QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIP MODELING OF BIOCONCENTRATION FACTORS OF POLYCHLORINATED BIPHENYLS

Manish Rao Ambedkar[1], Kushal Kumar[2], Madhu Gupta[3]

[1, 3] *Department of Chemistry, M. M. H. College Ghaziabad, C.C.S.University Meerut U.P. (India)*
[2] *IIMT College of Engineering Greater Noida , Uttar Pradesh Technical University Lucknow U.P. (India)*

ABSTRACT

Polychlorinated biphenyls (PCB) are now of concern as prevalent, persistent, and toxic pollutants. PCB contamination of water and soil and bioaccumulated in food chains due to their high hydrophobicity and chemical stability. PCB `s are chemicals that persist for a very long time in the environment and consequently may concentrate to a high level (10^6) in the food chain. They may also cause toxic effects on the animal and human reproduction, development and immunological function. That's why it is essential to determine the value of BCF for 58 PCB`s. The bioconcentration factors (BCFs) of 58 polychlorinated biphenyl (PCB) congeners were modeled by quantitative structure–activity relationship (QSAR) based on physicochemical & topological descriptors derived solely from molecular structure and calculated using Chem Sketch 12.0 & Dragon 6.0 software to calculate all kinds of descriptors. Descriptors utilized for the general model were selected by various statistical validation techniques. Multilinear models were developed using the best Multiple linear regression analysis (MLRA) followed by statistical evaluation by SPSS software (IBM) to corelate experimental BCF to a set of molecular descriptors. The proposed parameter model satisfactorily describes the relationship between observed and calculated values in terms of statistical parameters. Comparison with Alan R. Katritzky et al in their result have reported R^2 for best parameteric modeling is as R^2=0.931. We have observed that in our case R^2 for models with one, two, three, four five and six molecular descriptors are 0.8444, 0.9108, 0.9242, 0.9389, 0.9484 and 0.9503 Our results are much more superior then the result reported by Alan R. Katritzky et al. Therefore simple 2D QSAR reported by us is much betters then the 3D QSAR modeling of Alan R. Katritzky et al.

Keywords: Bioconcentration Factors (BCF) ; Multiple Linear Regression Analysis (MLRA) ; Polychlorinated Biphenyls (PCB); Quantitative Structure Activity Relationship (QSAR) ; SPSS.

I. INTRODUCTION

Polychlorinated biphenyls (PCB) formerly widely used industrially as dielectric fluids in transformers and capacitors, as hydraulic and heat transfer fluids, and as plasticizers, are now of concern as prevalent, persistent, and toxic pollutants.[1-2] Bioaccumulation of chemicals is quantitatively expressed in terms of BCF, defined as the equilibrium of its concentration inside an organism (or in a certain tissue of the organism, usually in the fat)

to that in the ambient environment.[3] The concentrations in tissues and in the environment are measured at steady-state after chronic exposure. However, the real test period may be too short to achieve steady-state. In addition, metabolism and chemical degradation may occur and large molecules may not permeate sufficiently through membranes into the organism, often lowering BCF values. Thus, experimental determination of BCF may underestimate the environmental risk.[4].The measured value of BCF should be strongly related to the high complexity of bioaccumulation process, taking into account such factors as metabolism, organ-specific bio concentration, irreversible binding onto proteins, incomplete depuration, and kinetic effects.[5] Bioaccumulation is a thermodynamically driven partitioning process between aquatic environment and the lipid tissues of fish, thus, n-octanol is often a satisfactory surrogate for biological lipids.[6] As demonstrated earlier, it is important to know the BCF of all PCB congeners. The literature data on experimental BCF of PCB are limited and their measurement is difficult and expensive. Thus, quantitative structure-property relationship (QSPR) methods based on the descriptors derived directly from the molecular structure are vital to supply the missing data independently of experimentation. The BCF of a chemical is most commonly estimated from established correlations between logBCF and logKOW[7-8]. molecular connectivity indices and polarity correction factors[9], theoretical molecular descriptors[10-11], and molecular electro negativity distance vector. The QSPR model development techniques utilized included genetic algorithms and artificial neural networks[12], fragment constant method[13] the heuristic method and support vector machine.[14] The present study was devoted to determining the bioaccumulation process (in terms of BCF) of 58 PCB congeners by means of multilinear QSAR approach in attempt to construct a statistically significant model. The main focus is on the selection of appropriate descriptors and investigation of how those selected descriptors are in fact related to the studied property elucidating the physical nature of the bioaccumulation phenomenon.

II. MATERIALS AND METHODS

2.1. Data Set

All data of the present investigation were obtained from the reference *(* Alan R. Katritzky *et al)*. The data set for this investigation consisted of 58 PCBs.The geometry of biphenyl template is depicted along with atom numbering (Fig. 1).

2.2. Molecular Descriptor Generation

To obtain a QSAR model, compounds are often represented by the molecular descriptors. The calculation process of the molecular descriptors was described as below: The two-dimensional molecular structures of 58 PCBs were drawn by Chem Sketch 12.0 then calculated some parameters. Then optimize molecular structure this optimize structure files were exported into software Dragon 6.0 to calculate all kinds of descriptors. The software Dragon 6.0 can calculate Physicochemical parameters, constitutional, topological, geometrical, descriptors and has been successfully used in various QSAR researches.Then value of all parameters put into SPSS statistical and data analysis software or NCSS (We can also use MSTAT instead of SPSS & NCSS) statistical and data analysis software to get data regression and correlation. Constitutional descriptors are related to the number of atoms and bonds in each molecule. Topological descriptors include valence and non-valence

molecular connectivity indices calculated from the hydrogen-suppressed formula of the molecule, encoding information about the size, composition, and the degree of branching of a molecule. The topological descriptors describe the atomic connectivity in the molecule. The geometrical descriptors describe the size of the molecule and require 3Dcoordinates of the atoms in the given molecule. The electrostatic descriptors reflect characteristics of the charge distribution of the molecule. The quantum chemical descriptors offer information about binding and formation energies, partial atom charge, dipole moment, and molecular orbital energy levels etc.

III. RESULTS AND DISCUSSION

By using the multiple linear regression analysis (MLRA) method of 2D-QSAR, regression models were developed for 58 PCBs. To select the sets of descriptors that are most relevant to logKBCF values and effectively show the relation between descriptors and logKBCF values of these compounds, six subsets with the descriptors from one to six were determined to establish the QSAR models. The initial decision concerning the optimal number of parameters describing the LogBCF of the PCB for the current set of structures was based on the application of the so-called breaking point criterion (Katritzky et al., 2006). An increase of the R^2 value less than 0.02 was chosen as a threshold. Multi-linear regression method for descriptor selection proceeds with a reselections of descriptors by sequentially eliminating descriptors which do not match any of the following criteria: (i) the F-test greater than one unit; (ii) R^2 value less than a value defined at the start (default 0.01); (iii) the student's t-test less than that defined (default 0.1); and (iv) duplicate descriptors having a higher squared inter-correlation coefficient than a predetermined level (usually 0.8). The next step involves correlation of the given property with (i) the top descriptor in the above list with each of the remaining descriptors, and (ii) the next one with each of the remaining descriptors, etc. The goodness of the correlation is tested by the correlation coefficient (R^2) and The stability of the correlations was tested against the cross-validated coefficient (R^2CV). Descriptors assigned a lower index number have a higher *t*-test value i.e. they are of higher significance in the derived model. Besides, it will demonstrate which descriptors have bad or missing values, which descriptors are insignificant, and which descriptors are highly intercorelated .This information will be helpful in reducing the number of descriptors involved in the search for the best QSAR/QSPR model. The statistical parameters of the derived model are as follows: (N= 58, Se=0.0420, R^2=0.9503, R^2A=0.9444, F-Ratio=162.4790, Q=23.2103) where R^2 is the squared correlation coefficient, N is the number of data points. Comparison to other models indicated advantages of the proposed model over previously reported ones. Derived parameter regression equation has improved statistics and is based on theoretical descriptors with a definite physicochemical meaning; it is easier to use and interpret due to the mathematical simplicity of the linear QSAR approach. Internal validation and scrambling procedure confirmed the stability and reliable predictive ability of the general model and indicated the absence of chance correlations. External validation demonstrated that the presented model can be applied to structurally similar sets of compounds, thus extending the domain of applicability of the model. A plot representing the observed vs predicted LogBCF values is given in Figure 2. Comparison with Alan R. Katritzky *et al* in their result have reported R^2 for best parameteric modeling is as R^2=0.931. We have observed that in our case R^2 for models with one, two, three, four five and six molecular descriptors are 0.8444, 0.9108, 0.9242, 0.9389, 0.9484 and 0.9503 Our results are much more superior then the result reported by Alan

R. Katritzky *et al*. Therefore simple 2D QSAR reported by us is much betters then the 3D QSAR modeling of Alan R. Katritzky *et al*.The data was subjected to regression analysis and the result obtained is discussed below.

3.1 Best Mono-Parametric Model

When topological indices were taken as independent parameter then we found that no single parameter is applicable is obtaining statistically significant mono-parametric model. The one which is the best model contain 2χ having R^2 value 0.8444. The model is as given below.

LogBCFexp= 0.7545(±0.0433)2χ -0.6967

N= 58, Se=0.0708, R^2=0.8444, F-Ratio=303.9296, Q=12.9789

3.2 Best Bi-Parametric Model

Many bi-parametric models have been obtained. The best model contains the R^2 value of two models with BAC and 0χ and BAC with 0χv has a similar value 0.9108. The model is as given below.

LogBCFexp=-0.0483(±0.0063)BAC 0.8016(±0.0611)0χ -3.2294

N= 58, Se=0.0541, R^2=0.9108, R^2A=0.9076, F-Ratio=280.9630,Q=17.6406

In above model the BAC have negative coefficient and 0χv has positive coefficient suggesting that the compound having high value of 0χv and low value of BAC will favor LogBCFexp activity.

3.3 Best Tri Parametric Model

When three parameters are taken together five tri- parametric models have been obtained. The best tri-parametric model contains BAC, 2χ, 3χv. The R^2 value of best model is 0.9242. The model is given below.

LogBCFexp=-0.0514(±0.0070)BAC 1.0786(±0.1117)2χ, 0.6405(±0.1519)3χv -4.1838

N= 58, Se=0.0504, R^2=0.9242, R^2A=0.9200, F-Ratio=219.3830, Q=19.0744

In above model the BAC have negative coefficient and 2χ and 3χ have positive coefficient suggesting that the compound having high value of 2χ and 3χv and low value of BAC will favor LogBCFexp activity.

3.4 Best Tetra-Parametric Model

When four parameters are taken together five tatra- parametric models have been obtained. The best tetra-parametric model contains W, J, Jhetz and 0χv. The R^2 value of best model is 0.9389. The model is given below.

LogBCFexp=-0.0453(±0.0064)W 93.1302(±20.1162)J, -60.4452(±12.4157)Jhetz 5.0308(±0.7117)0χv -26.7280

N= 58, Se=0.0457, R^2=0.9389, R^2A=0.9342, F-Ratio=202.4640, Q=21.2028

In above model the W and Jhetz have negative coefficient and J and 0χv have positive coefficient suggesting that the compound having high value of J and 0χv and low value of W and Jhetz will favor LogBCFexp activity.

3.5 Best Penta Parametric Model

When five parameters are taken together five penta- parametric models have been obtained. The best penta-parametric model contains W, J, Jhetz, 2χand 3χv.The R^2 value ranging from 0.9460 to 0.9484. The R^2 value of best model is 0.9484. The model is given below.

LogBCFexp= -0.0360(±0.0044)W 71.7677(±15.2845)J, -44.4263(±9.0550)Jhetz 5.5883(±0.6915)2χ 2.3125(±0.2994)3χv -30.5554

N= 58, Se=0.0423, R^2=0.9484, R^2A=0.9434, F-Ratio=191.0870, Q=23.0226

In above model the W and Jhetz have negative coefficient and J , 2χ and 3χv have positive coefficient suggesting that the compound having high value of J , 2χ and 3χv and low value of W and Jhetz will favor LogBCFexp activity.

3.6 Best hexa Parametric Model

When six parameters are taken together two hexa- parametric models have been obtained. The best tetra-parametric model contains W, J, Jhetz, Jhetm, 2χ and 3χv. The R^2 value of best model is 0.9503. The model is given below.

LogBCFexp= -0.0358(±0.0044)W 78.1359(±15.8176)J, -144.3091(±72.0558)Jhet 95.6271(±68.4487)Jhetm 5.5205(±0.6870)2χ 2.3110(±0.2967)3χv -30.2199

N= 58, Se=0.0420, R^2=0.9503, R^2A=0.9444, F-Ratio=162.4790, Q=23.2103

In above model the W and Jhetz have negative coefficient and J , Jhetm, 2χ and 3χv have positive coefficient suggesting that the compound having high value of J , Jhetm, 2χ and 3χv and low value of W and Jhetz will favor LogBCFexp activity.

IV. CONCLUSION

The relationship of the BCF to molecular structure of 58 PCB congeners was investigated using molecular descriptors calculated by Chem Sketch 12.0 & Dragon 6.0 software to calculate all kinds of descriptors. Multiple linear regression analysis (MLRA) followed by statistical evaluation by SPSS software (IBM) software we can also use MSTAT & NCSS instead of SPSS. The 6-parameter regression equation provided insight into the structural features that influence BCF. In above best hexa parameteric model the W and Jhetz have negative coefficient and J , Jhetm, 2χ and 3χv have positive coefficient suggesting that the compound having high value of J , Jhetm, 2χ and 3χv and low value of W and Jhetz will favor LogBCFexp activity. The stability of the model was demonstrated by applying internal validation techniques. By comparison to other reported models, it was demonstrated that QSAR models utilizing descriptors of topological can be advantageously used for modeling of BCF of PCB. The ability of the proposed model to predict accurately BCF of structurally similar to PCB compounds was also demonstrated.

IV.1 CO-RELTION

(1.) The LogBFCexp has moderate correlation with Balaban and Balaban type's indices and good correlation with Randic and Kier and Hall connectivity indices.

(2.) 0χ, 0χv, and 1χv has very good correlation with all topological indices.

(3.) Balaban and Balaban type's indices have a very good correlation among themselves and with topological indices.

(4.) 0χ has 100 % correlation with 0χv and 1χv and very good correlation with all the indices used.

(5.) All the Randic and Kier Hall connectivity indices has good correlation among themselves and with all the others parameters used.

REFFERENCE

[1] Halsall,C.J. ; Lee,R.G.M. ; Coleman,P.J. ; Burnett,V. ; Hardingjones,P. ; Jones,K.C. ; *Environ. Sci. Technol.* 29(1995) 2368-2376.

[2] Herrice,R.F. ; Lefkowitz,D.J. ; Weymouth,G.A. ; *Environ. Health Persp*. 115 (2007) 173-175.

[3] Voutsas,E. ; Magoulas,K. ; Tassios,D. ; *Chemosphere* 48(2002) 645–651.

[4] Franke,C. ; *Chemosphere* 32(1996) 1897–1905.

[5] Franke, C. ; Studinger, G. ; Berger, G. ; Bohling, S. ; Bruckmann, U. ; Cohors-Fresenborg, D. ; Jtihnck, U. ; *Chemosphere* 29 (1994) 1501–1514.

[6] Barron, M. G. ; *Environ. Sci. Technol.* 24 (1990) 1612–1618.

[7] Bintein, S. ; Devillers, J. ; Karcher, W. ; *SAR QSAR Environ. Res.* 1 (1993) 29–39.

[8] Chiou, C. T. ; Freed, V. H. ; Schmedding, D. W. ; Kohnert, R. L. ; *Environ. Sci. Technol.* 11(1977) 475–478.

[9] Lu, X. X. ; Tao, S. ; Hu, H. Y. ; Dawson, R. W. ; *Chemosphere 41* (2000) 1675–1688.

[10] Gramatica, P. ; Papa, E. ; *QSAR Comb. Sci.* 24 (2005) 953–960.

[11] Gramatica, P. ; Papa, E. ; *QSAR Comb. Sci.* 22 (2003) 374–385.

[12] Fatemi, M. H. ; Jalali-Heravi, M. ; Konuze, E. ; *Anal. Chim. Acta* 486 (2003) 101–108.

[13] Lin, Z. ; Yu, H. ; Gao, S. ; Cheng, J. ; Wang, L. ; *Arch. Environ. Contam. Toxicol.* 41 (2001) 255–260.

[14] Liu, H. ; Yao, X. ; Zhang, R. ; Liu, M. ; Hu, Z. ; Fan, B. ; *Chemosphere* 63 (2006) 722–733.

[15] Katritzky, A. R.; Dobchev, D. A.; Tulp, I.; Karelson, M.; Carlson, D. A.. Pro. *Bioorg. Med. Chem. Lett.* 16 (2006) 2306-2311.

[16] Franke, C.; Studinger, G.; Berger, G.; Bohling, S.; Bruckmann, U.; Cohors-Fresenborg, D.; Jtihnck, U. *Chemosphere* 29 (1994) 1501–1514.

[17] Khadikar, P. V.; Singh, S.; Mandloi, D.; Joshi, S.; Bajaj, A. V. ; *Bioorg. Med. Chem.* 11 (2003) 5045–5050.

[18] Kubinyi, H.. *QSAR in drug design, in: handbook of chemoinformatics. From Data to Knowledge, Vol. 4*, J. Gasteiger, Ed. Wiley-VCH, Weinheim, (2003) pp 1532-1554.

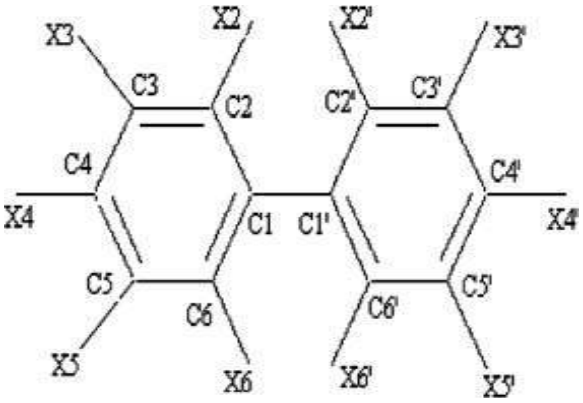

Fig.1 Geometry of biphenyl template with atom numbering

Table-1 Details of compounds with their activity used in the present study.

S.N.	COMPOUND NAME	LogBCFexp
1	Biphenyl	2.64
2	Biphenyl, 4-chloro-	2.77
3	Biphenyl, 2,2'-dichloro-	3.38
4	Biphenyl, 2,3-dichloro-	4.11
5	Biphenyl, 2,3'-dichloro-	3.80
6	Biphenyl, 2,4-dichloro-	3.55
7	Biphenyl, 2,4'-dichloro-	3.57
8	Biphenyl, 2,5-dichloro-	3.89
9	Biphenyl, 3,5-dichloro-	3.78
10	Biphenyl, 4,4'-dichloro-	3.28
11	Biphenyl, 2,2',5-trichloro-	4.11
12	Biphenyl, 2,4,4'-trichloro-	4.20
13	Biphenyl, 2,4,5-trichloro-	4.26
14	Biphenyl, 2,4',5-trichloro-	4.23
15	Biphenyl, 2,2',3,3'-tetrachloro-	4.23
16	Biphenyl, 2,2',3,5'-tetrachloro-	4.84
17	Biphenyl, 2,2',4,4'-tetrachloro-	4.85
18	Biphenyl, 2,2',4,5-tetrachloro-	5.00
19	Biphenyl, 2,2',4,5'-tetrachloro-	4.84
20	Biphenyl, 2,2',5,5'-tetrachloro-	4.63
21	Biphenyl, 2,2',6,6'-tetrachloro-	3.85
22	Biphenyl, 2,3,4',6-tetrachloro-	4.60
23	Biphenyl, 2,3',4',5-tetrachloro-	4.77
24	Biphenyl, 3,3',4,4'-tetrachloro-	4.59
25	Biphenyl, 2,2',3,4,5'-pentachloro-	5.38
26	Biphenyl, 2,2',3,4',5-pentachloro-	5.00
27	Biphenyl, 2,2',3',4,5-pentachloro-	5.43
28	Biphenyl, 2,2',4,4',5-pentachloro-	5.00
29	Biphenyl, 2,2',4,5,5'-pentachloro-	5.40
30	Biphenyl, 2,3,3',4,4'-pentachloro-	5.00
31	Biphenyl, 2,3,3',4,6-pentachloro-	5.00
32	Biphenyl, 2,3',4,4',5-pentachloro-	5.00
33	Biphenyl, 3,3',4,4',5-pentachloro-	5.81
34	Biphenyl, 2,2',3,3',4,4'-hexachloro-	5.77
35	Biphenyl, 2,2',3,3',6,6'-hexachloro-	5.43
36	Biphenyl, 2,2',3,4,4',5-hexachloro-	5.88
37	Biphenyl, 2,2',3,4,4',5'-hexachloro-	5.39
38	Biphenyl, 2,2',3,4,5,5'-hexachloro-	5.81

39	Biphenyl, 2,2',3,4',5,6'-hexachloro-	5.39
40	Biphenyl, 2,2',3,5,5',6-hexachloro-	5.54
41	Biphenyl, 2,2',4,4',5,5'-hexachloro-	5.65
42	Biphenyl, 2,2',4,4',6,6'-hexachloro-	4.93
43	Biphenyl, 2,3,3',4,4',5-hexachloro-	5.39
44	Biphenyl, 2,3,3',4,4',5'-hexachloro-	5.39
45	Biphenyl, 3,3',4,4',5,5'-hexachloro-	5.97
46	Biphenyl, 2,2',3,3',4,5,6'-heptachloro-	5.80
47	Biphenyl, 2,2',3,4,4',5,5'-heptachloro-	5.80
48	Biphenyl, 2,2',3,4,4',5,6'-heptachloro-	5.80
49	Biphenyl, 2,2',3,4,4',5',6-heptachloro-	5.84
50	Biphenyl, 2,2',3,4',5,5',6-heptachloro-	5.80
51	Biphenyl, 2,3,3',4,4',5',6-heptachloro-	5.84
52	Biphenyl, 2,2',3,3',4,4',5,5'-octachloro-	5.81
53	Biphenyl, 2,2',3,3',4,4',5,6-octachloro-	5.92
54	Biphenyl, 2,2',3,3',4,4',5,6'-octachloro-	5.92
55	Biphenyl, 2,2',3,3',4,5,5',6-octachloro-	5.88
56	Biphenyl, 2,2',3,3',5,5',6,6'-octachloro-	5.82
57	Biphenyl, 2,2',3,3',4,5,5',6,6'-nonachloro-	5.71
58	Decachlorobiphenyl	5.44

Table-2 Values of calculated topological parameters for the compounds used in the present study.

Comp.no.	W	J	JhetZ	Jhetm	Jhetv	Jhete	Jhetp	BAC
1	198.0000	1.8000	2.7000	2.7000	2.7000	2.7000	2.7000	0.0000
2	252.0000	1.7890	2.7280	2.7300	2.6400	2.6680	2.6650	2.0000
3	287.0000	1.9630	3.0470	3.0520	2.8470	2.9070	2.9020	5.0000
4	291.0000	1.9330	2.9990	3.0040	2.8050	2.8640	2.8590	5.0000
5	294.0000	1.9110	2.9610	2.9650	2.7760	2.8320	2.8270	5.0000
6	298.0000	1.8880	2.9250	2.9290	2.7430	2.7990	2.7940	5.0000
7	301.0000	1.8670	2.8890	2.8930	2.7150	2.7680	2.7640	5.0000
8	293.0000	1.9210	2.9770	2.9820	2.7890	2.8460	2.8410	5.0000
9	298.0000	1.8820	2.9140	2.9180	2.7350	2.7900	2.7850	5.0000
10	315.0000	1.7800	2.7460	2.7500	2.5950	2.6420	2.6380	5.0000
11	346.0000	1.9940	3.1300	3.1370	2.8610	2.9420	2.9350	10.0000
12	368.0000	1.8720	2.9250	2.9310	2.6960	2.7650	2.7590	10.0000
13	354.0000	1.9470	3.0530	3.0590	2.7960	2.8730	2.8670	10.0000
14	362.0000	1.9030	2.9760	2.9820	2.7380	2.8100	2.8040	10.0000
15	408.0000	2.0420	3.2370	3.2450	2.9020	3.0010	2.9930	17.0000
16	410.0000	2.0330	3.2210	3.2290	2.8910	2.9890	2.9800	17.0000
17	426.0000	1.9580	3.0930	3.1010	2.7910	2.8810	2.8730	17.0000
18	414.0000	2.0140	3.1880	3.1970	2.8650	2.9610	2.9530	17.0000
19	419.0000	1.9910	3.1480	3.1560	2.8350	2.9280	2.9200	17.0000
20	412.0000	2.0240	3.2050	3.2130	2.8800	2.9760	2.9680	17.0000
21	394.0000	2.1250	3.3790	3.3880	3.0110	3.1190	3.1100	17.0000
22	418.0000	1.9980	3.1610	3.1690	2.8440	2.9380	2.9300	17.0000
23	426.0000	1.9520	3.0810	3.0890	2.7840	2.8720	2.8650	17.0000
24	440.0000	1.8840	2.9670	2.9740	2.6940	2.7760	2.7690	17.0000
25	486.0000	2.0500	3.2690	3.2780	2.8960	3.0050	2.9960	26.0000
26	488.0000	2.0420	3.2560	3.2650	2.8860	2.9950	2.9860	26.0000
27	486.0000	2.0490	3.2670	3.2760	2.8940	3.0040	2.9940	26.0000
28	496.0000	2.0090	3.1980	3.2070	2.8420	2.9470	2.9380	26.0000

29	488.0000	2.0410	3.2540	3.2630	2.8850	2.9940	2.9840	26.0000
30	502.0000	1.9790	3.1470	3.1560	2.8030	2.9050	2.8960	26.0000
31	482.0000	2.0700	3.3050	3.3150	2.9210	3.0340	3.0240	26.0000
32	504.0000	1.9710	3.1330	3.1420	2.7940	2.8940	2.8850	26.0000
33	510.0000	1.9430	3.0840	3.0930	2.7560	2.8530	2.8450	26.0000
34	569.0000	2.0690	3.3180	3.3290	2.9080	3.0280	3.0170	37.0000
35	537.0000	2.2020	3.5510	3.5640	3.0800	3.2160	3.2050	37.0000
36	568.0000	2.0750	3.3290	3.3400	2.9160	3.0360	3.0260	37.0000
37	571.0000	2.0620	3.3050	3.3160	2.9000	3.0180	3.0080	37.0000
38	559.0000	2.1080	3.3860	3.3970	2.9590	3.0830	3.0720	37.0000
39	555.0000	2.1270	3.4190	3.4310	2.9840	3.1100	3.0990	37.0000
40	545.0000	2.1680	3.4910	3.5030	3.0360	3.1680	3.1570	37.0000
41	573.0000	2.0550	3.2930	3.3030	2.8910	3.0080	2.9980	37.0000
42	555.0000	2.1320	3.4280	3.4400	2.9890	3.1160	3.1060	37.0000
43	577.0000	2.0380	3.2640	3.2740	2.8680	2.9840	2.9740	37.0000
44	578.0000	2.0320	3.2520	3.2620	2.8600	2.9750	2.9650	37.0000
45	587.0000	1.9960	3.1890	3.1990	2.8140	2.9240	2.9150	37.0000
46	632.0000	2.1870	3.5380	3.5510	3.0510	3.1910	3.1790	50.0000
47	652.0000	2.1170	3.4130	3.4250	2.9610	3.0920	3.0810	50.0000
48	642.0000	2.1540	3.4790	3.4920	3.0080	3.1440	3.1330	50.0000
49	644.0000	2.1480	3.4680	3.4800	3.0000	3.1350	3.1240	50.0000
50	636.0000	2.1750	3.5160	3.5290	3.0360	3.1740	3.1630	50.0000
51	652.0000	2.1160	3.4120	3.4230	2.9600	3.0910	3.0800	50.0000
52	738.0000	2.1730	3.5210	3.5340	3.0250	3.1680	3.1560	65.0000
43	726.0000	2.2150	3.5970	3.6110	3.0780	3.2270	3.2140	65.0000
54	729.0000	2.2040	3.5780	3.5910	3.0650	3.2120	3.2000	65.0000
55	717.0000	2.2430	3.6470	3.6610	3.1140	3.2670	3.2540	65.0000
56	702.0000	2.2950	3.7390	3.7540	3.1800	3.3390	3.3260	65.0000
57	800.0000	2.3230	3.7980	3.8130	3.2090	3.3760	3.3620	82.0000
58	907.0000	2.3480	3.8480	3.8640	3.2360	3.4100	3.3950	101.0000

Table-3 Values of calculated connectivity indices for the compounds used in the present study.

Comp.no.	***0 χ***	***1χ***	***2χ***	***3χ***	***0χ*v**	***1χ*v**	***2χ*v**	***3χ*v**
1	8.2260	5.9660	4.7960	3.9660	6.7740	4.0710	2.7320	1.8810
2	9.0960	6.3600	5.4180	4.3770	7.8300	4.5490	3.3090	2.2030
3	9.9660	6.7880	5.8510	4.9060	8.8870	5.0390	3.7740	2.6970
4	9.9660	6.7880	5.8290	5.0250	8.8870	5.0390	3.7530	2.8550
5	9.9660	6.7710	5.9570	4.7600	8.8870	5.0330	3.8340	2.5740
6	9.9660	6.7710	5.9570	4.7530	8.8870	5.0330	3.8340	2.5640
7	9.9660	6.7710	5.9460	4.8440	8.8870	5.0330	3.8300	2.6110
8	9.9660	6.7710	5.9570	4.7760	8.8870	5.0330	3.8340	2.5780
9	9.9660	6.7540	6.0760	4.5310	8.8870	5.0270	3.8970	2.4060
10	9.9660	6.7540	6.0400	4.7880	8.8870	5.0270	3.8870	2.5260
11	10.8370	7.1820	6.4850	5.2480	9.9430	5.5160	4.3550	2.9870
12	10.8370	7.1650	6.5790	5.1640	9.9430	5.5100	4.4110	2.8860
13	10.8370	7.1820	6.4650	5.3450	9.9430	5.5160	4.3340	3.1300
14	10.8370	7.1650	6.5790	5.1860	9.9430	5.5100	4.4110	2.9000
15	11.7070	7.6090	6.8620	6.0900	11.0000	6.0060	4.7730	3.8310
16	11.7070	7.5920	6.9910	5.8400	11.0000	6.0000	4.8540	3.5540
17	11.7070	7.5750	7.1190	5.5460	11.0000	5.9940	4.9360	3.2480

18	11.7070	7.5920	6.9930	5.8180	11.0000	6.0000	4.8550	3.5380
19	11.7070	7.5750	7.1190	5.5680	11.0000	5.9940	4.9360	3.2620
20	11.7070	7.5750	7.1190	5.5900	11.0000	5.9940	4.9360	3.2760
21	11.7070	7.6090	6.9260	5.7110	11.0000	6.0060	4.8220	3.4350
22	11.7070	7.5920	6.9880	5.8440	11.0000	6.0000	4.8540	3.5480
23	11.7070	7.5750	7.0870	5.7640	11.0000	5.9940	4.9110	3.4630
24	11.7070	7.5750	7.0550	5.9440	11.0000	5.9940	4.8870	3.6510
25	12.5770	8.0030	7.4960	6.4280	12.0560	6.4830	5.3540	4.1120
26	12.5770	7.9860	7.6360	6.0690	12.0560	6.4770	5.4390	3.7820
27	12.5770	8.0030	7.4980	6.4100	12.0560	6.4830	5.3540	4.1050
28	12.5770	7.9860	7.6260	6.1380	12.0560	6.4770	5.4360	3.8140
29	12.5770	7.9860	7.6260	6.1600	12.0560	6.4770	5.4360	3.8280
30	12.5770	8.0030	7.4640	6.6020	12.0560	6.4830	5.3290	4.2980
31	12.5770	8.0030	7.5180	6.2560	12.0560	6.4830	5.3600	4.0230
32	12.5770	7.9860	7.5950	6.3340	12.0560	6.4770	5.4110	4.0140
33	12.5770	7.9860	7.5730	6.4500	12.0560	6.4770	5.3900	4.1730
34	13.4470	8.4300	7.8730	7.2650	13.1130	6.9730	5.7720	4.9480
35	13.4470	8.4300	7.9370	6.9230	13.1130	6.9730	5.8210	4.5760
36	13.4470	8.4140	8.0140	6.9230	13.1130	6.9670	5.8570	4.6230
37	13.4470	8.4140	8.0040	6.9970	13.1130	6.9670	5.8540	4.6640
38	13.4470	8.4140	8.0140	6.9460	13.1130	6.9670	5.8570	4.6370
39	13.4470	8.3970	8.1850	6.3780	13.1130	6.9610	5.9660	4.1040
40	13.4470	8.4140	8.0450	6.7680	13.1130	6.9670	5.8810	4.4490
41	13.4470	8.3970	8.1340	6.7300	13.1130	6.9610	5.9360	4.3800
42	13.4470	8.3970	8.2170	6.1690	13.1130	6.9610	5.9910	3.8930
43	13.4470	8.4140	7.9820	7.1200	13.1130	6.9670	5.8320	4.8240
44	13.4470	8.4140	7.9820	7.1070	13.1130	6.9670	5.8320	4.8200
45	13.4470	8.3970	8.0900	6.9550	13.1130	6.9610	5.8920	4.6950
46	14.3170	8.8410	8.4230	7.6090	14.1690	7.4570	6.3000	5.2870
47	14.3170	8.8240	8.5210	7.5150	14.1690	7.4510	6.3570	5.1890
48	14.3170	8.8240	8.5630	7.2320	14.1690	7.4510	6.3840	4.9450
49	14.3170	8.8240	8.5530	7.3190	14.1690	7.4510	6.3810	4.9890
50	14.3170	8.8240	8.5530	7.3370	14.1690	7.4510	6.3810	5.0010
51	14.3170	8.8240	8.5310	7.4230	14.1690	7.4510	6.3600	5.1440
52	15.1880	9.2520	8.9080	8.3010	15.2260	7.9400	6.7780	5.9980
43	15.1880	9.2690	8.8100	8.3840	15.2260	7.9460	6.7210	6.0850
54	15.1880	9.2520	8.9400	8.1050	15.2260	7.9400	6.8020	5.7980
55	15.1880	9.2520	8.9500	8.0480	15.2260	7.9400	6.8050	5.7690
56	15.1880	9.2520	8.9720	7.9500	15.2260	7.9400	6.8270	5.6230
57	16.0580	9.6790	9.3590	8.7290	16.2820	8.4300	7.2480	6.4230
58	16.9280	10.1070	9.7460	9.5080	17.3390	8.9190	7.6690	7.2230

Table-4 Regression parameters and quality of correlation with topological Parameters

Model No.	**Parameter used**	**Ai=(1…6)**	**B**	**Se**	**R^2**	**R^2a**	**F Ratio**	**Q=R/Se**
1	**2χ**	0.7545(±0.0433)	-0.6967	0.0708	0.8444	-	303.9296	12.9789
2	**BAC 0χv**	-0.0483(±0.0063) 0.8016(±0.0611)	-3.2294	0.0541	0.9108	0.9076	280.9630	17.6406
3	**BAC 2χ**	-0.0514(±0.0070) 1.0786(±0.1117)	-4.1838	0.0504	0.9242	0.9200	219.3830	19.0744

	3χv	0.6405(±0.1519)						
4	**WJ** **Jhetz** **0χv**	-0.0453(±0.0064) 93.1302(±20.1162) -60.4452(±12.4157) 5.0308(±0.7117)	-26.7280	0.0457	0.9389	0.9342	202.4640	21.2028
5	**WJ** **Jhetz** **2χ** **3χv**	-0.0360(±0.0044) 71.7677(±15.2845) -44.4263(±9.0550) 5.5883(±0.6915) 2.3125(±0.2994)	-30.5554	0.0423	0.9484	0.9434	191.0870	23.0226
6	**WJ** **Jhetz** **Jhetm** **2χ** **3χv**	-0.0358(±0.0044) 78.1359(±15.8176) -144.3091(±72.0558) 95.6271(±68.4487) 5.5205(±0.6870) 2.3110(±0.2967)	-30.2199	0.0420	0.9503	0.9444	162.4790	23.2103

Table-5 Observed and calculated activity for the compounds using model-6 (Table N.- 4)

Comp.no.	Observed LogBCFexp	Calculated LogBCFexp	Residual
1	2.6400	2.7210	-0.0810
2	2.7700	2.9360	-0.1660
3	3.3800	3.5680	-0.1880
4	4.1100	3.6610	0.4490
5	3.8000	3.6460	0.1540
6	3.5500	3.4360	0.1140
7	3.5700	3.4880	0.0820
8	3.8900	3.7900	0.1000
9	3.7800	3.7940	-0.0140
10	3.2800	3.4730	-0.1930
11	4.1100	4.2000	-0.0900
12	4.2000	4.0500	0.1500
13	4.2600	4.1140	0.1460
14	4.2300	4.2360	-0.0060
15	4.2300	4.6500	-0.4200
16	4.8400	4.7260	0.1140
17	4.8500	4.5240	0.3260
18	5.0000	4.7750	0.2250
19	4.8400	4.7080	0.1320
20	4.6300	4.7950	-0.1650
21	3.8500	4.2570	-0.4070
22	4.6000	4.5960	0.0040
23	4.7700	4.9600	-0.1900
24	4.5900	4.8570	-0.2670
25	5.3800	5.1710	0.2090
26	5.0000	5.1180	-0.1180
27	5.4300	5.1850	0.2450
28	5.0000	5.0950	-0.0950
29	5.4000	5.1880	0.2120
30	5.0000	5.2430	-0.2430
31	5.0000	5.1360	-0.1360
32	5.0000	5.2950	-0.2950
33	5.8100	5.5240	0.2860
34	5.7700	5.5050	0.2650
35	5.4300	5.3840	0.0460
36	5.8800	5.5010	0.3790

37	5.3900	5.5860	-0.1960
38	5.8100	5.6590	0.1510
39	5.3900	5.4880	-0.0980
40	5.5400	5.5690	-0.0290
41	5.6500	5.5170	0.1330
42	4.9300	5.1300	-0.2000
43	5.3900	5.6450	-0.2550
44	5.3900	5.7150	-0.3250
45	5.9700	5.9540	0.0160
46	5.8000	5.7710	0.0290
47	5.8000	5.8900	-0.0900
48	5.8000	5.6900	0.1100
49	5.8000	5.6360	0.1640
50	5.8000	5.8180	-0.0180
51	5.8400	5.7160	0.1240
52	5.8100	6.0320	-0.2220
43	5.9200	5.7990	0.1210
54	5.9200	5.7160	0.2040
55	5.8800	5.9180	-0.0380
56	5.8200	5.9180	-0.0980
57	5.7100	5.7120	-0.0020
58	5.4400	5.4830	-0.0430

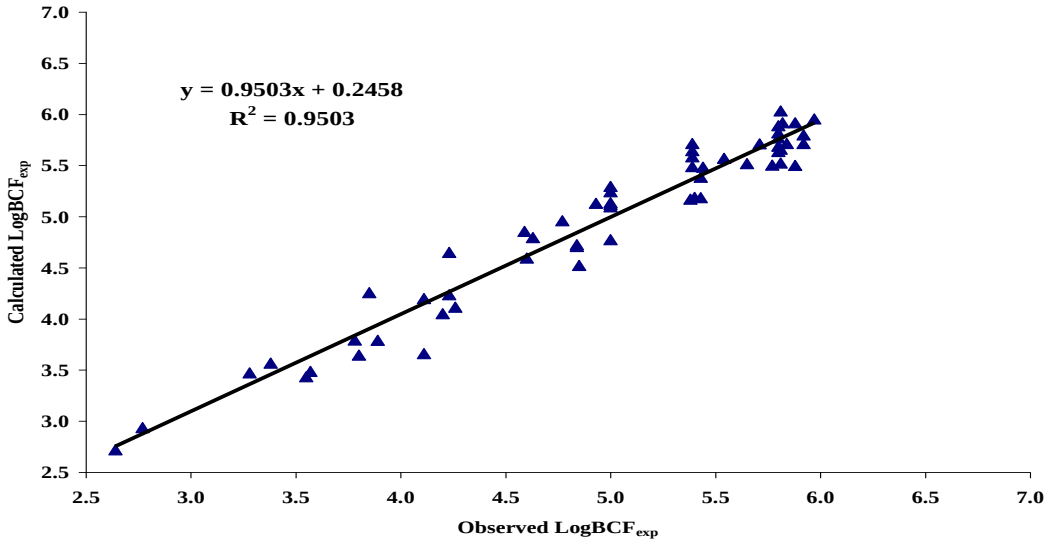

Fig-2 Correlation between Observed and Calculated activity using model No.-6 (Table- 4).

Table-6 Cross validated values for Topological parameters

S.no.	Parameters used	PRESS	SSY	PRESS/SSY	R^2cv	PSE	S_{PRESS}
1.	**2χ**	6.8667	37.2679	0.1842	0.8158	0.3440	0.3501
2.	**BAC,0χv**	3.9346	40.2000	0.0978	0.9022	0.2604	0.2674
3.	**BAC,2X,3χv**	3.3465	40.7881	0.0820	0.9180	0.2402	0.2489
4.	**W,J,Jhetz,0χv**	2.6984	41.4362	0.0651	0.9349	0.2156	0.2256
5.	**W,J,Jhetz,2χ,3χv**	2.2780	41.8566	0.0544	0.9456	0.1981	0.2093
6.	**W,J,Jhetz,Jhetm,2χ,3χv**	2.1940	41.9406	0.0523	0.9477	0.1944	0.2074

Table-7 Correlation matrix showing inter-correlation among all the parameters with the activity.

	LogBCFexp	*IR*	*ST*	*D*	*POL*	*W*	*J*	*JhetZ*	*Jhetm*
LogBCFexp	1.0000								
IR	0.9328	1.0000							
ST	0.9308	0.9996	1.0000						
D	0.9359	0.9998	0.9993	1.0000					
POL	0.9037	0.9907	0.9936	0.9887	1.0000				
W	0.8775	0.9723	0.9772	0.9693	0.9935	1.0000			
J	0.7702	0.8974	0.9006	0.8934	0.9152	0.8884	1.0000		
JhetZ	0.8082	0.9275	0.9299	0.9242	0.9401	0.9135	0.9972	1.0000	
Jhetm	0.8098	0.9286	0.9310	0.9253	0.9410	0.9145	0.9970	1.0000	1.0000
Jhetv	0.6940	0.8339	0.8385	0.8287	0.8618	0.8362	0.9915	0.9791	0.9785
Jhete	0.7475	0.8790	0.8826	0.8746	0.8999	0.8734	0.9992	0.9933	0.9930
Jhetp	0.7439	0.8760	0.8797	0.8715	0.8974	0.8709	0.9989	0.9926	0.9922
BAC	0.7950	0.9236	0.9317	0.9181	0.9667	0.9827	0.9062	0.9199	0.9203
0χ	0.9038	0.9907	0.9936	0.9888	1.0000	0.9935	0.9152	0.9401	0.9410
1χ	0.8999	0.9892	0.9923	0.9871	0.9999	0.9939	0.9186	0.9428	0.9436
2χ	0.9189	0.9942	0.9961	0.9933	0.9957	0.9853	0.8979	0.9265	0.9276
3χ	0.8731	0.9640	0.9683	0.9608	0.9840	0.9877	0.8939	0.9167	0.9175
0χv	0.9038	0.9907	0.9936	0.9888	1.0000	0.9935	0.9152	0.9401	0.9410
1χv	0.9027	0.9903	0.9932	0.9883	1.0000	0.9936	0.9163	0.9410	0.9418
2χv	0.9143	0.9942	0.9963	0.9929	0.9982	0.9888	0.9060	0.9332	0.9342
3χv	0.8636	0.9571	0.9620	0.9536	0.9806	0.9867	0.8932	0.9148	0.9155

	Jhetv	*Jhete*	*Jhetp*	*BAC*	**0χ**	**1χ**	**2χ**	**3χ**	**0χv**	**1χv**	**2χv**	**3χv**
Jhetv	1.0000											
Jhete	0.9960	1.0000										
Jhetp	0.9966	1.0000	1.0000									
BAC	0.8746	0.8976	0.8962	1.0000								
0χ	0.8617	0.8998	0.8973	0.9665	1.0000							
1χ	0.8666	0.9038	0.9013	0.9692	0.9999	1.0000						
2χ	0.8380	0.8805	0.8777	0.9470	0.9957	0.9942	1.0000					
3χ	0.8451	0.8800	0.8776	0.9723	0.9839	0.9858	0.9655	1.0000				
0χv	0.8617	0.8998	0.8973	0.9665	1.0000	0.9999	0.9957	0.9839	1.0000			
1χv	0.8632	0.9011	0.8986	0.9674	1.0000	0.9999	0.9953	0.9845	1.0000	1.0000		
2χv	0.8485	0.8894	0.8867	0.9545	0.9982	0.9972	0.9994	0.9729	0.9982	0.9979	1.0000	
3χv	0.8469	0.8801	0.8779	0.9758	0.9806	0.9828	0.9602	0.9994	0.9806	0.9813	0.9681	1.0000