IOSR Journal of Applied Chemistry (IOSR-JAC)
e-ISSN: 2278-5736.Volume 4, Issue 1 (Mar. – Apr. 2013), PP 26-33
www.iosrjournals.org

QSAR Studies of Some Bisphosphonates Compounds for their Anticancer Activities

*Anita K.[a] Shweta Sharma[a], Bashirulla Shaik[b], Sarita Shrivastava[c]
[a] *Dept. of Chemistry, Career College, Bhopal, Madhya Pradesh, India*
[b]*Dept. of Applied Science, National Institute of Technical Teachers Training and Research, Shamla Hills, Bhopal-462002, Madhya Pradesh, India*
[c]*Dept. of Chemistry, Govt. M. V. M., Madhya Pradesh, India*

Abstract: *The present Study deals with QSAR modeling of anticancer activities of some bisphosphonate compounds using Distance-based topological indices and Kier Hall Valance Connectivity indices. The multiple regression analysis (MLR) indicates that a tri-parametric model having an R^2 value of 0.9052 is the best model for predicting the log1/IC50 activity of bisphosphonate compounds. The predictive power of the models was tested using Leave-One-Out Cross validation method.*
Key words: *QSAR, topological indices, Multiple Regression (MLR), Cross validation*

I. Introduction

Studies show that cancer is most deadly disease in the world due to the uncontrolled growth and spread of abnormal cells anywhere in a body [1].In spite of all the efforts made so far, cancer is difficult to cure. The main reason for this difficulty is that cancer results from the uncontrolled multiplication of subtly modified normal human cells.

Most of the drugs used for the treatment of cancer are cytotoxic (cell-killing) drugs that work by interfering in some way with the operation of the cell's DNA. Unfortunately cytotoxic drugs have the potential to be very harmful to the body unless they are very specific to cancer cells which is difficult to achieve. Therefore a major challenge is to design new drugs that will be more selective for cancer cells, with lesser side effects.

Several studies have shown the effect of bisophosphonates on cancer cells and it is an established fact that their anti-tumor activity is related to interference with mevalonate pathway [2]. Bisphosphonate compounds are widely used since 1970 for disorders of bone metabolism like paget's disease and osteoporosis [3]. Researchers in this field have shown that nitrogen containing bisphosphonates are extensively used to block bone destruction in cancer patients with bone metastasis. They are effective inhibitors of osteoclast-mediated bone resorption. It has also been observed that nitrogen containing bisphosphonates have anti cancer activities through inhibition of tumor cell functions, enhancement of the cytotoxic activity of chemotherapy agents, inhibition of tumor angiogenesis and stimulation of antitumor immune reactions.

II. Methodology Used

The methodology used in the present study is to model the anti-cancer activities of bis phosphonates using topological indices. Quantitative Structure Activity Relationship (QSAR) modeling establishes a quantitative correlation between chemical structure and biological activity [4]. In the present study we took 24 compounds with $\log 1/IC_{50}$ activities reported in the literature [5]. In this work the biological activity of molecules has been taken as $\log 1/IC_{50}$ which is negative log of IC_{50} values and is known as the concentration of test compound required to inhibit 50% activity. The structural details of bisphosphonate compounds which act as anticancer compounds are given in Table 1.DRAGON software [6] has been used for calculation of topological descriptors. These calculated descriptors are reported in Table 2. They include: J, Jhetz, Jhetv, Jhete, Jhetp, ${}^0\chi^V$, ${}^1\chi^V$, ${}^2\chi^V$, and ${}^3\chi^V$. In fact topological parameters have been very successfully used by us in modeling different activities of drug molecules [7-10].From the descriptors calculated useful descriptors were generated by variable selection of descriptors in multiple regression analysis using NCSS software [11].Finally the proposed models obtained were subjected to cross-validation by leave-one-out procedure [12].

III. Results And Discussions

The biological activities of these compounds in terms of log $1/IC_{50}$ are reported in Table 2. Table 2 also contains calculated parameters viz- Balaban and Balaban type indices and zero to three order valence-connectivity indices. The correlation matrix among all the topological indices and biological activity is reported in Table 3. A close look at this table clearly indicates that for modeling $\log 1/IC_{50}$ activity connectivity indices play a dominant role. However, it is interesting to observe that higher order valence-connectivity indices show a decreasing trend. The 1st and 2nd order connectivity-indices show good correlation. The data was subjected to regression analysis and the best obtained correlations are summarized in Table 4. On the basis of R^2 following models are best for modeling the anticancer activity of these compounds:

One variable model (model no. 4, Table 4)

$$\log 1/IC_{50} = -0.6139(\pm 0.0644)\,{}^{1}\chi^{v} + 6.9457 \quad (1)$$

$N=24$, $R^2 = 0.8050$, $R^2A = 0.7961$, $Se = 0.3077$, $F = 90.821$, $Q = 2.9159$

Two variable model (model no. 7, Table 4)

$$\log 1/IC_{50} = -0.1939\,(\pm 0.1079)\,J_{hetp} - 0.6631\,(\pm 0.0672)\,{}^{1}\chi^{v} + 8.1470 \quad (2)$$

$N = 24$, $R^2 = 0.8310$, $R^2A = 0.8149$, $Se = 0.2932$, $F = 51.626$, $Q = 3.1091$

Three variable model (model no. 8, Table 4)

$$\log 1/IC_{50} = -0.2564\,(\pm 0.1022)\,J_{hetp} - 0.7133\,(\pm 0.0651)\,{}^{1}\chi^{v} + 0.3063\,(\pm 0.1342)\,{}^{3}\chi^{v} + 7.0165 \quad (3)$$

$N = 24$, $R^2 = 0.8659$, $R^2A = 0.8141$, $Se = 0.2677$, $F = 43.049$, $Q = 3.4760$

However during the regression analysis it has been found that compound 4 is a serious outlier. Its behavior may be because of its geometry of side chain attached to nitrogen. When compound 4 is deleted drastic improvement in quality of regression model is obtained and the new models are as shown below:

One variable model (model no. 12, Table 4)

$$\log 1/IC_{50} = -0.6248\,(\pm 0.0570)\,{}^{1}\chi^{v} + 6.9891 \quad (4)$$

$N = 23$, $R^2 = 0.8513$, $R^2A = 0.8442$, $Se = 0.2777$, $F = 120.215$, $Q = 3.3225$

Two variable model (model no. 15, Table 4)

$$\log 1/IC_{50} = -0.2232\,(\pm 0.0900)\,J_{hetp} - 0.6823\,(\pm 0.0561)\,{}^{1}\chi^{v} + 8.3756 \quad (5)$$

$N = 23$, $R^2 = 0.8863$, $R^2A = 0.8749$, $Se = 0.2489$, $F = 77.930$, $Q = 3.7824$

Three variable model (model no. 16, Table 4)

$$\log 1/IC_{50} = -0.2661\,(\pm 0.0871)\,J_{hetp} - 0.7171\,(\pm 0.0555)\,{}^{1}\chi^{v} + 0.2287\,(\pm 0.1173)\,{}^{3}\chi^{v} + 7.5016 \quad (6)$$

$N = 23$, $R^2 = 0.9052$, $R^2A = 0.8903$, $Se = 0.2331$, $F = 60.498$, $Q = 4.0816$

The best one-parametric model contains 1st order valence-connectivity index for which the R^2 value comes to be 0.8513. The systematic addition of topological indices gave three bi-parametric correlations which were found statistically better than the mono-parametric correlations. However the best bi-parametric model among these contains ${}^{1}\chi^{v}$ and $J_{hetp.}$ The R^2 value for this bi-parametric model comes out to be 0.8863. The adjusted R^2 for this model is 0.8749, which clearly indicates that addition of J_{hetp} is justified .Further improvement in R^2 is observed when ${}^{3}\chi^{v}$ is added to the bi-parametric model the R^2 changes from 0.8863 to 0.9052. The R^2A also show significant improvement from 0.8749 to 0.8903. On the basis of Pogliani's quality factor [13-14] we infer that the three-parametric model is the best for modeling $\log 1/IC_{50}$ activity of present set of compounds. Further confirmation is obtained by estimating the activity using model 16 which is reported in Table 5. The estimated values are in good agreement with observed values. When observed activity values are plotted against estimated values, we obtained a graph which is reported in Fig. 1. The predictive power of the model comes out to be 0.905.

Further confirmation is obtained by calculating cross-validated parameters from 8-16. Such values are given in Table 6. The PSE value close to0.4 for model 16 suggests that this model is the best. Also cross-validation R^2 value for model 16 comes to be 0.8953 which is the highest among all the discussed models. For any kind of possible defect we have calculated variance inflation factor, tolerance and condition number for various parameters using VIF plot which is given in Table 7. All the parameters show the value within the permissible limit. Therefore the model is free from any kind of defect. Ridge trace suggests that there is no co- linearity in the model.

IV. Conclusions

A close look at model 16 gives following conclusions:

1. J_{hetp} is a negative coefficient

2. $^{1}\chi^{v}$ also shows negative coefficient
3. $^{3}\chi^{v}$ has a positive coefficient

Since $^{3}\chi^{v}$ has a positive coefficient its higher value will give a improved value of log $1/IC_{50}$.The 1st order branching is not favourable and similarly negative coefficient of Balaban type index J_{hetp} suggest that the lower value of these parameters will enhance biological activity. While designing a new drug molecule these factors should be considered.

Acknowledgement

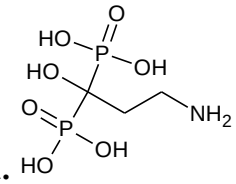
The authors are very thankful to Prof. Vijay Kumar Agrawal, a well known scientist in the field of Computer Assisted Drug Designing for his guidance.

References

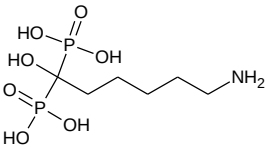
[1]. Cancer Facts and Figures 2007(American Cancer Society, Atlanta, 2007).
[2]. M.T. Velenti, L.D. Carbnare, F. Bertoldo, L.Donatelli, V. Lo Cascio The effects on h-TERT gene expression is an additional mechanism of amino bisphosphonates in prostatic cancer cells, European Journal of Pharmacology, 1-2, 2008, 36-42.
[3]. R. Graham, G. Russel, Determinants of structure function relationships among bisphosphonates, Bone, 40, (5), 2007, S21-S25.
[4]. X. Y.Yao, J. P. Panaye,J.P. Doucet, R.S. Zang, H.F.Chen, N.C. Liu, Z.D. Hu,B.T. Fan, Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks and multiple linear regression. J. Chem. Inf. Comput. Sci., 44, 2004, 1257-1266.
[5]. M.Talebi, G.H.Ghasemi, H.Kefayati, A QSAR study on biological activities of bisphosphonates compounds as anticancer drugs, World Academy of Science, Engineering and Technology, 81, 2011.
[6]. DRAGON : Software for Calculation of Topological Indices, www.disat.unimib.it
[7]. B. Shaik, J. Singh, S.Singh ,N. Sohani, V.K. Agrawal, P.V. Khadikar**,** Mutagenicity of Nitrated Polycyclic Aromatic Hydrocarbons: A QSAR Investigation Chemical Biology Drug Design, 71(3), 2008, 230-243.
[8]. B.Louis, J. Singh, B. Shaik, V.K. Agrawal, P.V. Khadikar, QSPR Study on the Estimation of Solubility of Drug-like Organic Compounds: A Case of Barbiturates Chemical Biology Drug Design, 74, 2009, 190-195.
[9]. J. Singh, I. Ahmad, B. Shaik, V.K. Agrawal, P.V. Khadikar, Modeling of Clostridium histolyticum collagenase inhibitory activity of some sulfony1-1-amine hydroxamates : a molecular connectivity approach. J. Indian Chem. Soc., 86, 2009, 1197-1203.
[10]. J. Singh, I. Ahmad, B. Shaik, R. Kumar, M. Gupta, V. K. Agrawal, Modeling of protease inhibitory activity of some hydroxamates using physic-chemical and topological parameters, International J Pure Appl. Chem., 4(3), 2009,195-203.
[11]. NCSS, Kaysville Utah, www.ncss.com
[12]. S. Chaterjee, A.S. Hadi, Price B, Regression Analysis by Examples, (3rd Ed. Wiley: New York 2000.)
[13]. L. Pogliani , Structure property relationships of amino acids and some dipeptides, Amino Acids, 6, 1994,141.
[14]. L.Pogliani , Modeling with Special Descriptors Derived from a Medium-Sized Set of Connectivity Indices, J. Phys. Chem., **100**, 1996, 18065-18077.

Table 1 . Structures of compounds used in the study

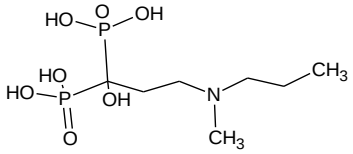
1.



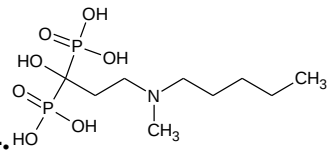
2.



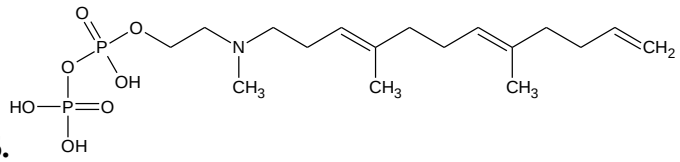
3.



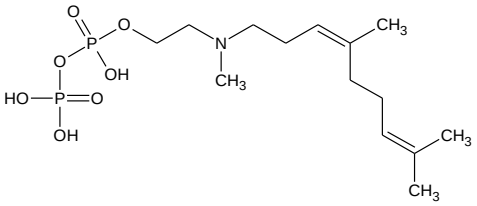
4.



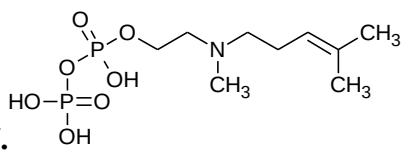
5.



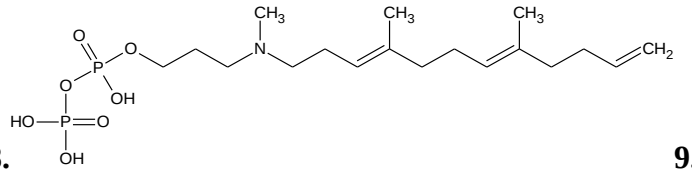
6.



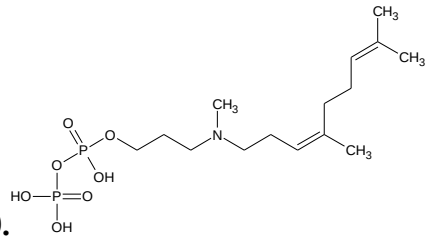
7.



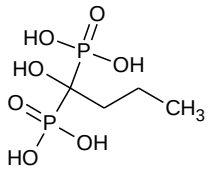
8.



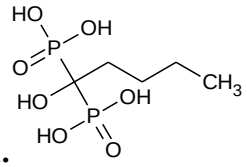
9.



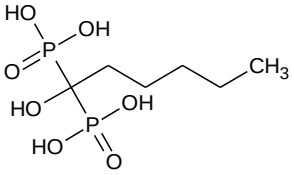
10.



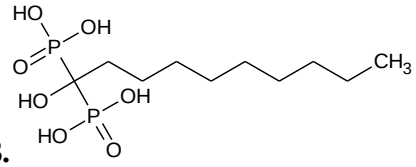
11.



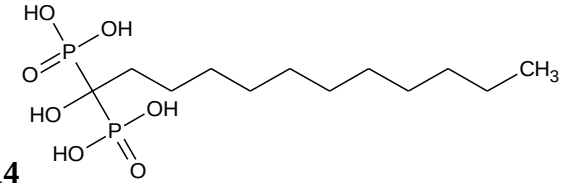
12.



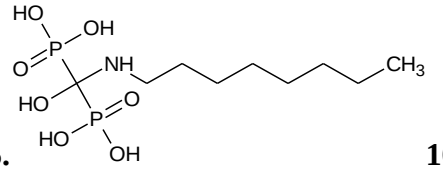
13.



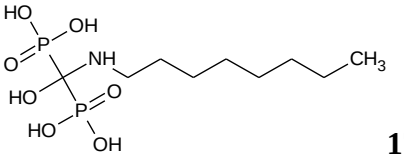
14



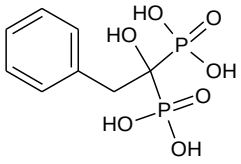
15.



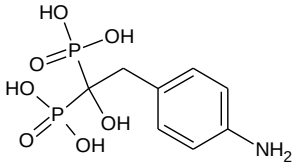
16.



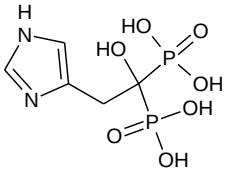
17.



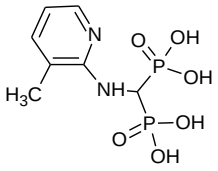
18.



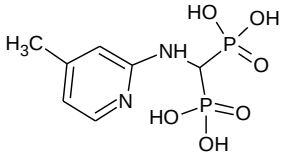
19.



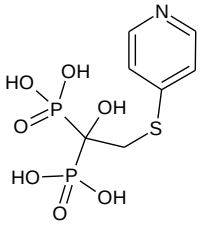
20.



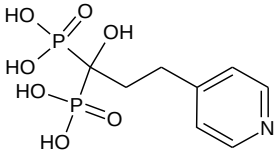
21.



22.



23.



24.

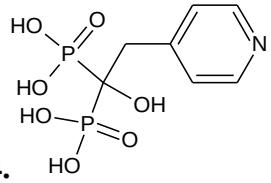


Table 2.Calculated values of topological parameters for the compounds used in the present study along with observed log1/IC_{50} values

Compd. no	log1/IC_{50}	J	J_{hetZ}	J_{hetv}	J_{hete}	J_{hetp}	$^{0}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$	$^{3}\chi^{v}$
1	2.16	4.96	9.44	4.47	5.36	5.86	8.23	6.32	6.41	5.29
2	2.36	4.59	6.57	4.30	4.80	5.10	10.35	7.82	7.47	6.03
3	2.31	4.76	6.82	4.15	5.10	4.65	11.51	8.20	7.92	6.24
4	2.46	4.54	5.98	3.98	4.83	4.30	12.93	9.20	8.63	6.78
5	-0.52	3.74	4.83	3.21	4.33	3.36	18.20	11.82	9.26	5.90
6	-0.14	3.74	5.02	3.05	4.36	3.23	16.50	10.65	8.67	5.22
7	1.55	3.77	5.64	2.75	4.46	3.02	13.01	8.60	7.08	4.19
8	-0.47	3.69	4.67	3.21	4.23	3.34	18.91	12.32	9.62	6.11
9	-0.34	3.69	4.80	3.05	4.24	3.21	17.21	11.15	9.02	5.43
10	2.52	4.96	9.34	4.54	5.33	6.03	8.65	6.62	6.62	5.40
11	2.66	4.87	8.18	4.52	5.17	5.76	9.36	7.12	6.98	5.68
12	1.30	4.74	7.26	4.45	4.98	5.46	10.06	7.62	7.33	5.93
13	0.93	4.18	5.18	4.04	4.29	4.52	12.89	9.62	8.74	6.93
14	0.15	3.97	4.67	3.87	4.05	4.22	14.31	10.62	9.45	7.43
15	1.02	4.18	5.42	3.65	4.44	3.90	12.69	9.37	8.24	6.21
16	1.96	4.44	6.30	3.80	4.79	4.17	11.27	8.37	7.53	5.71
17	2.37	2.84	4.46	2.94	3.24	3.52	10.33	7.68	7.53	5.96
18	2.53	2.81	4.29	2.94	3.24	3.43	10.83	7.88	7.79	6.10
19	2.42	2.87	4.65	2.73	3.22	3.25	9.55	7.12	7.11	5.65
20	2.58	2.73	4.58	2.49	3.32	2.73	10.55	7.77	7.16	5.25
21	2.79	2.67	4.47	2.46	3.26	2.69	10.55	7.76	7.21	5.17
22	2.18	2.68	5.12	2.80	3.06	3.62	11.42	8.65	8.39	7.36
23	2.39	2.68	3.89	2.74	3.00	3.16	10.91	8.03	7.73	6.31
24	1.91	2.84	4.48	2.91	3.25	3.47	10.20	7.53	7.40	5.87

Table 3. Correlation matrix

	log1/IC_{50}	J	J_{hetz}	J_{hetv}	J_{hete}	J_{hetp}	$^{0}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$	$^{3}\chi^{v}$
log1/IC_{50}	1.00									
J	-0.09	1.00								
J_{hetz}	0.27	0.83	1.00							
J_{hetv}	0.05	0.94	0.83	1.00						
J_{hete}	-0.13	0.98	0.84	0.87	1.00					
J_{hetp}	0.22	0.85	0.90	0.95	0.78	1.00				
$^{0}\chi^{v}$	-0.90	-0.06	-0.45	-0.24	0.00	-0.44	1.00			
$^{1}\chi^{v}$	-0.90	-0.05	-0.47	-0.20	-0.02	-0.41	0.98	1.00		
$^{2}\chi^{v}$	-0.79	-0.09	-0.52	-0.13	-0.11	-0.34	0.88	0.94	1.00	
$^{3}\chi^{v}$	-0.09	-0.02	-0.21	0.21	-0.16	0.14	0.10	0.25	0.55	1.00

Table 4 . Regression parameters and quality of correlation

Model No.	Parameters Used	Ai=(1…….3)	B	Se	R^2	R^2A	F-ratio	Q=R/Se
1	$^{3}\chi^{v}$	-0.1390(±0.3247)	2.4513	0.6940	0.0083	0.0000	0.183	0.1313
2	J_{hete}	-0.1860(±0.2966)	2.4064	0.6907	0.0176	0.0000	0.393	0.1921
3	J_{hetp}	0.2394(±0.2286)	0.6706	0.6801	0.0475	0.0042	1.097	0.3205
4	$^{1}\chi^{v}$	-0.6139(±0.0644)	6.9457	0.3077	0.8050	0.7961	90.821	2.9159
5	$^{3}\chi^{v}$ $^{1}\chi^{v}$	0.2160(±0.1447) -0.6380(±0.0647)	5.8751	0.2995	0.8237	0.8069	49.062	3.0303
6	J_{hete}	-0.2075(±0.1275)	7.8284	0.2968	0.8269	0.8109	50.141	3.0638

	$^{1}\chi^{v}$	-0.6156(±0.0621)						
7	J_{hetp} $^{1}\chi^{v}$	-0.1939(±0.1079) -0.6631(±0.0672)	8.1470	0.2932	0.8310	0.8149	51.626	3.1091
8	J_{hetp} $^{1}\chi^{v}$ $^{3}\chi^{v}$	-0.2564(±0.1022) -0.7133(±0.0651) 0.3063(±0.1342)	7.0165	0.2677	0.8659	0.8458	43.049	3.4760

After deletion of compound no. 4

Model No.	Parameters Used	Ai=(1…….3)	B	Se	R^2	R^2A	F-ratio	Q=R/Se
9	$^{3}\chi^{v}$	-0.2143(±0.3373)	2.8533	0.7133	0.0189	0.0000	0.404	0.1927
10	J_{hete}	-0.2319(±0.3028)	2.5556	0.7103	0.0272	0.0000	0.587	0.2322
11	J_{hetp}	0.2293(±0.2318)	0.6778	0.7039	0.0445	0.0000	0.979	0.2997
12	$^{1}\chi^{v}$	-0.6248(±0.0570)	6.9891	0.2777	0.8513	0.8442	120.215	3.3225
13	$^{3}\chi^{v}$ $^{1}\chi^{v}$	0.1382(±0.1351) -0.6392(±0.0586)	6.2998	0.2774	0.8587	0.8446	60.766	3.3405
14	J_{hete} $^{1}\chi^{v}$	-0.2709(±0.1052) -0.6287(±0.0506)	8.1479	0.2466	0.8883	0.8772	79.547	3.8220
15	J_{hetp} $^{1}\chi^{v}$	-0.2232(±0.0900) -0.6823(±0.0561)	8.3756	0.2489	0.8863	0.8749	77.930	3.7824
16	J_{hetp} $^{1}\chi^{v}$ $^{3}\chi^{v}$	-0.2661(±0.0871) -0.7171(±0.0555) 0.2287(±0.1173)	7.5016	0.2331	0.9052	0.8903	60.498	4.0816

Table 5 Observed and estimated values of log1/IC_{50} before and after deletion (Using model no. 8 and model 16)

Compd. No.	Observed log1/IC_{50}	Estimated log1/IC_{50}	Residual log1/IC_{50}	Estimated log1/IC_{50}	Residual log1/IC_{50}
1	2.16	2.59	-0.43	2.62	-0.46
2	2.36	1.96	0.40	1.91	0.44
3	2.31	2.39	-0.08	1.81	0.49
4	2.46	1.95	0.51		
5	-0.52	-0.57	0.05	-0.52	0.00
6	-0.14	0.17	-0.31	0.20	-0.33
7	1.55	1.60	-0.05	1.49	0.06
8	-0.47	-0.94	0.47	-0.83	0.35
9	-0.34	-0.22	-0.12	-0.11	-0.22
10	2.52	2.20	0.32	2.38	0.13
11	2.66	2.06	0.60	2.16	0.50
12	1.30	1.89	-0.59	1.94	-0.63
13	0.93	1.05	-0.12	0.98	-0.05
14	0.15	0.56	-0.41	0.46	-0.31
15	1.02	1.47	-0.45	1.16	-0.14
16	1.96	2.06	-0.10	1.69	0.26
17	2.37	2.25	0.12	2.42	-0.05
18	2.53	2.24	0.29	2.34	0.19
19	2.42	2.78	-0.36	2.83	-0.40
20	2.58	2.45	0.13	2.40	0.17
21	2.79	2.39	0.40	2.40	0.39
22	2.18	2.08	0.11	2.02	0.16
23	2.39	2.30	0.10	2.35	0.04
24	1.91	2.38	-0.47	2.52	-0.61

Table 6 . Cross validation parameters of proposed models before and after deletion

Model No.	Parameters used	PRESS/ SSY	R^2_{CV}	S_{PRESS}	PSE
4	$^1\chi^v$	0.24	0.76	0.50	0.48
7	$J_{hetp}, ^1\chi^v$	0.20	0.80	0.48	0.45
8	$J_{hetp}, ^1\chi^v, ^3\chi^v$	0.15	0.85	0.44	0.40
12	$^1\chi^v$	0.17	0.83	0.44	0.42
15	$J_{hetp}, ^1\chi^v$	0.13	0.87	0.40	0.37
16	$J_{hetp}, ^1\chi^v, ^3\chi^v$	0.10	0.90	0.37	0.34

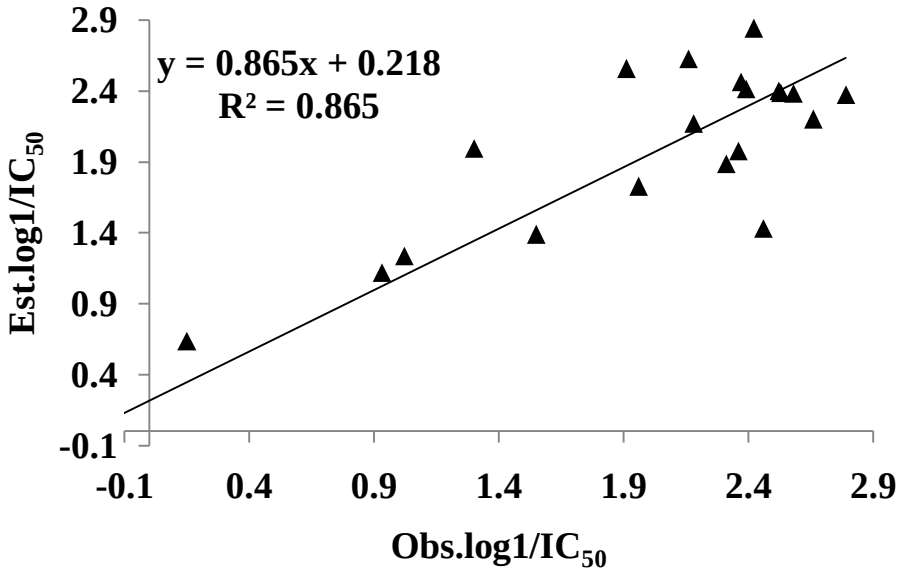


Fig. 1 Correlation between observed and estimated log1/IC_{50} using model 8

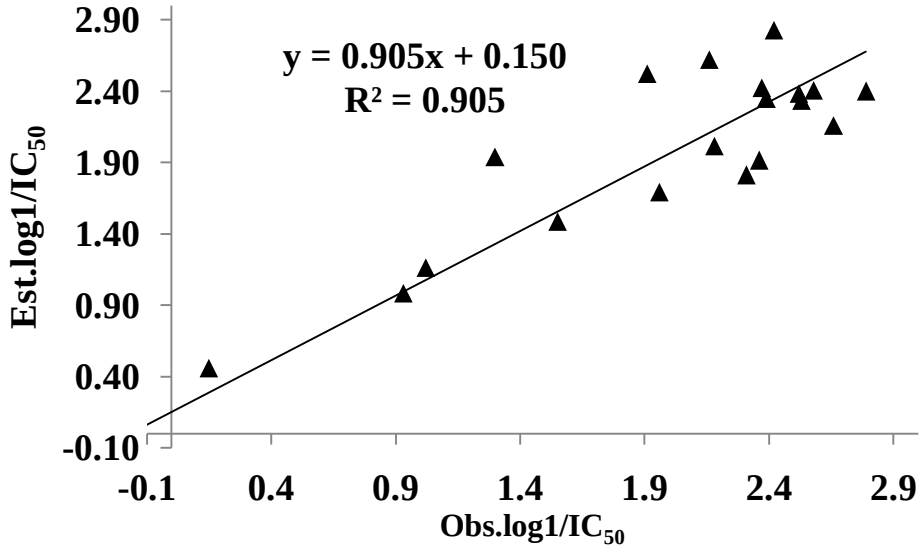


Fig. 2 Correlation between observed and estimated log1/IC_{50} using model 16

Table 7 . Ridge analysis for the three variable model (model 16, Table 4)

Parameters used	VIF	T	λ	k
J_{hetp}	1.29	0.77	1.43	1.00
$^{1}\chi^{v}$	1.34	0.74	1.10	1.29
$^{3}\chi^{v}$	1.13	0.88	0.46	3.10

VIF =Variance inflation factor
λ = Eigen Values
T = Tolerance
k = Condition number

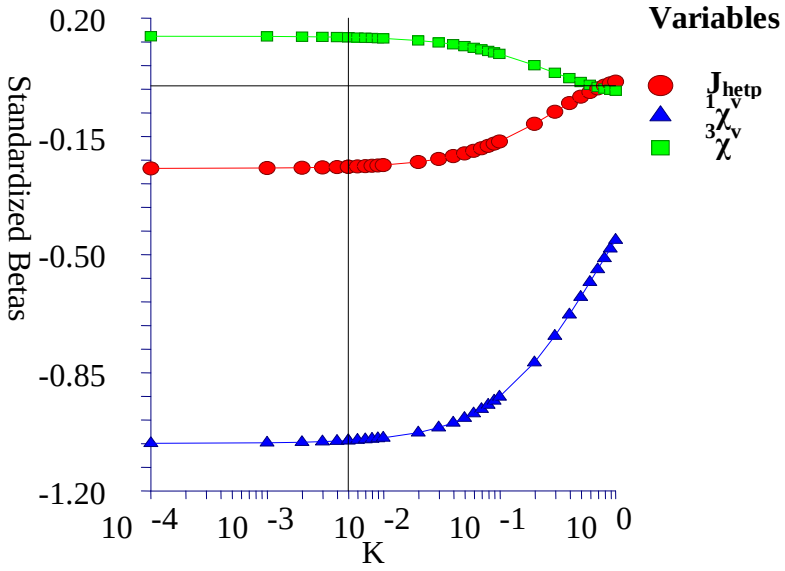


Fig. 3 Ridge Trace for three variable model (model 16, Table 4)

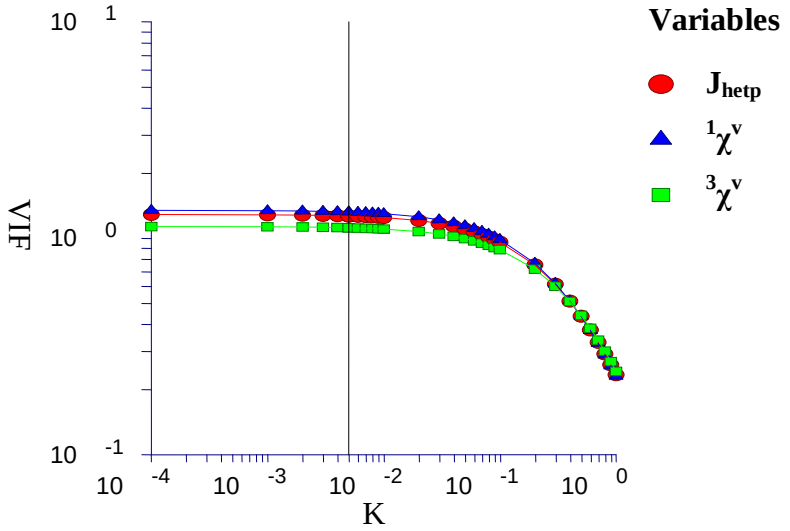


Fig.4 VIF plot for three variable model (model 16, Table 4)