INTERNATIONAL JOURNAL OF ADVANCES IN PHARMACY, BIOLOGY AND CHEMISTRY

Research Article

On Diuretic Activity of Benzene Sulfonamide Using ^{13}C NMR Chemical Shift as A Molecular Descriptor: Regular vs Ridge Regression

Asutosh Kumar Pandey

Chemistry Department, Oriental University, Indore, India.

ABSTRACT

Diuretic activity [p(1/C)] of benzene sulfonamides was modeled using ^{13} C NMR chemical shift (as a molecular descriptor. The regression analyses were carried out using regular as well as Ridge multiple regression analyses. Application of variety of statistics namely (statistics, Ridge regression and parameter derived there were used for modeling the diuretic activity. Results have shown that ^{13} C NMR chemical shift (yields an excellent model.

Keywords: Diuretic activity , benzene sulfonamide, topological index, regression analysis.

INTRODUCTION

A diuretic is an agent that promotes the secretion of urine, primarily by decreasing tubular respiration of sodium ions and their osmotic equivalent of water from the kidney tubule. Diuretics are used in edema, cardiac failure, renal disorders, and liver disorders.

Of-late it was proposed that the normal acidification of urine is used by secretion of hydrogen ion by the tubular cells of the kidney; these ions were provided by the active of the enzyme carbonic anhydrase. It was also observed that sulfonamide rendered the urine of dogs alkaline because of the inhibition of carbonic anhydrase which resulted in a lesser exchange of hydrogen ions for sodium ions in the kidney tubule. The sodium ions, along with bicarbonate ions, were then extracted, and a diuretic effect was noticed. The larger doses required and the side effects of sulfonamide prompted a search for more effective carbonic anhydrase inhibitors as diuretic drugs. It was then argued that the sulfonamide portion of an active diuretic molecule could not be mono- or di-substituted. It was also believed that a more acidic sulfonamide would bind more tightly to the carbonic anhydrase enzyme.

Recently we have carried out an extensive work on topological modeling of carbonic anhydrase inhibitors and observed that, use of distance-based topological indices resulted into excellent models.[1-12] In addition, we have also observed that information theoretical index and ad-hoc molecular descriptors can also be used for this purpose. Also, we observed that chemical shift in ^{13} C NMR chemical shift (can be used as a molecular descriptor [13-19] and that the use of the same was found very encouraging for modeling carbonic anhydrase inhibitory activity[19] This has prompted us to undertake the present investigation in that we have used ^{13} C NMR chemical shift(for modeling diuretic activity[p1 (1/C)] of the benzene sulfonamides (Fig 1, Table 1). The diuretic activities of the sulfonamides used were taken from the literature [20].

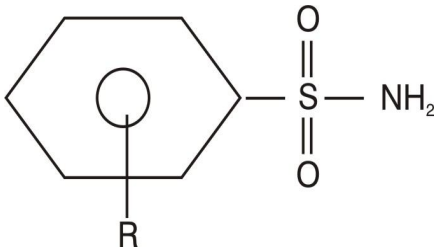


Fig. 1: General structure of benzene sulfonamide used in the present study

In addition to ^{13} C NMR chemical shift(δ), we have also used distance- based topological indices: Wiener (W) [21]; Szeged (Sz), [22-24] PI (Padmakar – Ivan) [25-27] -, first- order molecular connectivity ( ) [28-30] indices. In order to account for the structural feathers not contained in the topological

indices we have used four indicator parameters: IP_1, IP_2, IP_3, and IP_4 related to substitution at 4-, 3-, di- and 2- positions respectively. The modeling of diuretic activity[pl (1/C)] was done adopting step-wise regression analyses and have used both simple as well as robust multiple regression analyses.[31] Finally, data were subjected to Λ - statistic and we have also carried out ridge regression analyses and applied Mallow's Cp.- statistic. The result are summarized in Tables 1 to 6.The benzene sulfonamides used, their calculated topological indices, indicator parameters and the diuretic activities are given in Table 1.Table 2 records regression parameters and quality of correlation using regular multiple regression analysis. The other statistical parameters using this regular multiple regression analysis are summarized in Table 3.Such regression parameters for robust multiple regression analysis are summarized in Tables 4 and 5 respectively. Finally, the comparison of the results using regular and robust regression analyses are made in Table 6.

Experimental

Diuretic activities

The diuretic activity reported earlier as 1/C was converted into p (1/C) units and used in the present study.

Topological indices

All the topological indices (Table 1) were calculated from the hydrogen suppressed molecular grope of the benzene sulfonamides. These molecular graphs are obtained by transforming molecular structure into molecular graph deleting all the carbon – hydrogen as well as heteroatom - hydrogen bonds in the molecular structure; the calculation of these indices is well documented in the literature [36-45]. Therefore, the details of their calculations are not given here. However, below we give the corresponding expressions for their calculations.

Wiener indexes (W)

The Wiener index (W) is a widely used topological index[31]. It is based on the vertex-distances of the respective molecular graph.

Molecular graph can be denoted by G and having v_1, v_2, v_3, vn as its vertices. Let d (v_i,v_j|G) stands for the shortest distance between the vertices v_i and v_j. Then the Wiener index is defined as:

$$W = W(G) = 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} d(v_i, v_j | G) \qquad (6)$$

First-order connectivity index ($^1\chi$,)

The connectivity index $^1\chi = {}^1\chi(G)$ of a graph G is defined by Randic[32] as under:

$$^1\chi = {}^1\chi(G) = \sum_{ij} [\delta_i \delta_j]^{-0.5} \qquad (7)$$

Where δ_i and δ_j are the valence of a vertex i and j, equal to the number of bonds connected to the atoms i and j, in G.

Szeged indexes (Sz)

Let e be an edge of the molecular graph G. Let n_1 (e|G) be the number of vertices of G lying closer to one end of e; let n_2 (e|G) be the number of vertices of G lying closer to the other end of e. Then the Szeged index (Sz) is defined[35-37] as:

$$Sz = Sz(G) = \sum_{e} n_1(e|G)\, n_2(e|G) \qquad (8)$$

With the summation going over all the edges of G.In cyclic graphs, there are edges equidistant from both the ends of edge e; by definition of Sz such edges are not taken into account.

PI (Padmakar-Ivan) index

$$PI = PI(G) = \sum_{e} [\varepsilon\, n_1(e|G) + e\, n_2(e|G)] \qquad (9)$$

Cross Validation

As opposed to traditional regression methods, the cross-validation evaluates the validity of a model by how well it predicts data rather than how well it fits data. The analysis uses a "leave-one-out" scheme, a model is built with N-1 compounds and the nth compound is predicted. Each compound is left out of the model derivation and predicted in turn. As indication of the performance of the model is obtained from the cross-validated (or predictive) r^2_{CV} which is defined as:

$$r^2_{cv} = \frac{SD - PRESS}{SD} \qquad (10)$$

Where SD is the sum-of-squares deviation for each activity from the mean. PRESS (or predictive sum-of-squares) is the sum of squared difference between the actual and that predicted when the compound is omitted from the fitting process. Once a model is developed which has highest cross-validated r^2_{cv}, this method is used to derive the conventional QSAR equation and conventional r^2 and s values. The results of the final model are often visualized as contour maps of the coefficient. In addition to PRESS, SD, r^2_{cv}, S_{PRESS}, one also need to evaluate predictive-square-error (PSE) in an attempt to decide the predictive potential of the proposed models. The data's of calculation of cross-validated parameters are given in our publications.

Regression Analysis

Maximum R^2 improvement method to identify prediction models[39-41]. This method finds the "best" one variable model, the "best" two variable model and so forth for the prediction of property /activity. Several models (combinations of variables) were examined to identify combinations of variables with good prediction capabilities. In all regression models developed a variety of statistics associated with residues, i.e. the Wilks - Shapiro test for normality and Cooks D-statistics for outliners, to obtain the most reliable results were examined. Multiple regression analyses for correlating antiviral activity of the present set of compounds with the aforementioned molecular descriptors were carried out using Regress-1 software as supplied by Professor I. Lukovits, Hungarian Academy of Sciences, Budapest, Hungary. Several multiple regressions were attempted using correlation matrix from this program and the best results were considered and discussed in developing QSAR.

RESULTS AND DISCUSSION

A perusal of Table 1 shows that degeneracy is present in the topological indices as well as in diuretic activity. The degeneracy in the former can be understood from the fact that some of them (W, Sz, PI) belong to first -generation topological indices , while ^{1} belongs to second -generation topological indices [32-33] Balaban has stated [32-33] that such indices in spite of their observed degeneracy can be successfully used in developing QSAR models. The correlation of the data set is presented in the dendrogram [34]

Multiple linear regression

With the number of variables (topological indices, indicator parameters and ^{13}C NMR chemical shift) used, several models were possible. Consequently, we have used step- wise regression analyses for obtaining model of statistical best quality. None of the parameters (Table 1) when used singly yielded one-variable model with significant statistics. The results presented in Table 2 indicate that 2- variable model containing IP_4 and δ gave statistically significant model .This 2-variable model is given below:

$$p(I/C) = -5.0958\ (\pm 1.1639) - 0.7940\ (\pm 0.1317)\,IP_4 + 0.8698\ (\pm 0.1865)\,\delta \qquad (1)$$

$$n = 19,\ R^2 = 0.7426,\ R^2_A = 0.7105,\ CV = 0.8358,\ F = 23.083$$

Here and there after, n-is the number of compounds used, R^2 is the multiple- R^2, $R2_A$ is the adjustable R^2, CV is the coefficient of variation and F-is the F-statistics. The positive coefficient of indicates its favorable contribution for modeling the diuretic activity [p (1/C)} On the other hand the negative

coefficient of IP_4 indicates that substitution at 2-position is not favorable for the exhibition of the activity. Successive regression analysis gave still better model which contains δ, IP_3 and IP_4 as the correlating parameters:

$$p(1/C) = -6.2066\ (\pm 1.1874) - 0.2756\ (\pm 0.1329)\ IP_3 - 0.8817\ (\pm 0.1272)\ IP_4 + 1.0563\ (\pm 0.1922)\ \delta \quad (2)$$

$$n = 19,\ R^2 = 0.8000,\ R^2_A = 0.7600,\ CV = 0.8065,\ F = 19.999$$

The improved statistics is due to the added indicator parameter IP_3 that is di- substitution. The negative coefficient of IP_3 indicates that di-substitution is also not favorable for the exhibition of the diuretic activity. The physical significance of the other two correlating parameters is the same as discussed under model expressed by the equation (1). Further step- wise regression analyses failed to yield any model with still better statistics than the model expressed by equation (2). Also, none of the topological indices alone or in combination with other topological indices gave a model with better statistics than those discussed above. These results, therefore, indicates better modeling ability of ^{13}C NMR chemical shift compared to other molecular descriptors used.

Ridge Multiple Regression

It is interesting to record that regular multiple regression is optimum when all of its assumptions are valid. When some of these assumptions are invalid, least- square regression can perform poorly. Ridge regression provides an alternative to least- squares regression that works with less restrictive assumptions. The Ridge regression conducts its own residual analyses and down-weight or completely removes various observations. Huber's method is currently the most frequently recommended method for the Ridge regression. The results obtained using this method for the present study are given in Tables 4 and 5. The data presented in these tables indicate that better results are obtained in Ridge-multiple regression analyses. A perusal of these tables also indicates that like regular multiple regression analyses, statistics goes on improving as we proceed from 2- variable to 3 -variable models. Obviously meaning that the 3-variable model, is also the best in Ridge analysis for modeling the diuretic activity. Since in regular multiple regression analysis the best model is also the 3-variable model containing δ, □□IP_3, IP_4 as the correlating parameters we discuss this model obtained from Ridge analysis as below. The results on other models are available in Tables 4 and 5.

The aforementioned three variable Ridge model is found as below:

$$p(I/C) = -5.8006\ (+-1.0868) - 0.2360\ (+-0.1161)\ IP_3 - 0.8468\ (+-0.1106)\ IP_4 - 09877\ (+-0.1759)\ \delta \quad (3)$$

$$n = 19,\ R^2 = 0.8244,\ R^2_A = 0.7893,\ CV = 0.7405,\ F = 23.481$$

The physical significance of the parameters involved in this Ridge model, equation (3), is the same as that mentioned under regression multiple regression models. The comparison of the statistics of regular and Ridge multiple regression analyses (Table 2) indicates that the latter analysis yields better results. From the results discussed above observed that the co-efficient of δ term is the highest in all the proposed models. This, therefore, indicates that ^{13}C NMR chemical shift (δ□□ plays a dominating role in the modeling the diuretic activity. In all the cases (Tables 2 and 3) the magnitude of R^2_A parameter goes on increasing as we pass from 2- to 3-variable models. This means that the added parameter in each case has favorable contribution for developing the model and thus for modeling the activity. In order to confirm our results we have calculated p(I/ C) from the best model and compared the same with the observed values of p(I/C). Such a comparison is shown in Table 6 and demonstrated in Figure. 2 and 3. From these figures we have calculated predicative correlation coefficient (R^2pred), which are found as 0.7992 and 0.7998 for the models using regular and Ridge multiple regression analyses.

Λ - Statistics

We have also attempted to investigate the presence or otherwise of co-linearity in the proposed models. The Fisher variance ratio F and Λ statistics are the measure of the seriousness of co- linearity in the regression equation [31-35]. The latter is defined as:

$$\Lambda = 1/n \sum_{i=1}^{n} 1/\lambda I \qquad (4)$$

Where n is the number of descriptors and the □□ are the eigen values of the correlation matrix of descriptors. A value of /\ greater than 5 is taken to indicate that a co-linearity problem exists in the equation. /\ values recorded in Table 4 are all smaller than 5 indicating absence of problem due to collinearly. In addition to /\, the condition number k and sum of the reciprocals of eigen values are also used to investigate multi-collinearly. A large condition number k and the summation of reciprocals of the eigenvalues grater that five times the number of predictor variables indicates the malti-collinearity Problem does exist in the proposed models.

Variance Inflation Factor

Ridge regression is yet another useful technique for analyses of co-linearity data [31.] Application of this analysis to the data and the eq (2) and (3) yields Ridge traces as show in Fig. 5 supporting again the absence of collinearly. Furthermore, a variance inflation factor :

$$VIF = 1 / (1 - r_i^2)) \qquad (5)$$

I was also calculated for each of the descriptors in the model by regressing it against the other and are given in Tables 3 and 5. In all the three proposed models VIF was found smaller than 10, thus again indicating absence of multi-co-linearity.

CONCLUSIONS

The results and discussion made above indicates that ^{13} C NMR chemical shift (δ□□) is a useful molecular descriptor for successful modeling of diuretic activity of benzene sulfonamides used. The results also show that combination of ^{13}C NMR chemical shift (δ)and the distance- based topological indices (W, Sz, PI, and ^{1}□) are not that useful for yielding statistically significant multi-variable models for modeling the diuretic activity. Also, that 2- and di- substation on the benzene nucleus of the sulfonamide is responsible for the exhibitive of the activity. Furthermore ,results show that there is no need to separate out 2 - substituted sulfonamides due to ortho- effect.

Table 1: Sulfonamide, their distance –base topological indices ^{13} C NMR chemical shift, parameter are diuretic action

Comp No.	R	W	SZ	PI	^{1}□	IP_1	IP_2	IP_3	IP_4	□	p(1/C)
1	4-MeNH	201	306	126	5.5370	1	0	0	0	5.58	-0.04
2	4-NH_2	152	236	104	4.9990	1	0	0	0	5.85	-0.30
3	4-MeO	201	306	126	5.5370	1	0	0	0	6.01	0.19
4	4-Me	152	236	104	4.9999	1	0	0	0	6.06	0.18
5	3-Me	148	228	104	4.9999	0	1	0	0	6.07	0.18
6	H	114	177	84	4.6052	0	0	0	0	6.12	0.13
7	4-Cl	152	236	104	4.9999	1	0	0	0	6.26	0.29
8	4-Br	152	236	104	4.9999	1	0	0	0	6.25	0.27
9	3-Cl	148	228	104	4.9999	0	1	0	0	6.30	0.29
10	4-Ac	325	472	176	6.3929	1	0	0	0	6.34	0.40
11	4-CN	201	306	126	5.5370	1	0	0	0	6.42	1.00
12	3-NO_2	240	354	150	5.9097	0	1	0	0	6.51	0.68
13	4-NO_2	252	378	126	5.9097	1	0	0	0	6.48	0.82
14	3,4-Cl_2	189	291	126	5.4097	0	0	1	0	6.37	0.24
15	3NO2 ,4-Cl	289	427	176	6.3204	0	0	1	0	6.50	0.32
16	3-CF3 ,4NO2	484	694	266	7.5317	0	0	1	0	6.62	0.58
17	2-Me	144	220	104	5.0159	0	0	0	1	6.30	-0.30
18	2-Cl	144	220	104	5.0159	0	0	0	1	6.39	-0.30
19	2-$N0_2$	228	334	150	5.9265	0	0	0	1	6.50	-0.30

□□□□□□- ^{13} C NMR, chemical shift, W – Wiener Index, Sz – Szeged Index, PI – Padmakar – Ivan Index;
^{1}□ – First- order molecular connective Index,
IP_1 – indicator parameter for 4- substitution,
IP_2 – indicator parameter for 3- substitution,
IP_3- indicator parameter for di-substitution,
IP_4- indicator parameter for 2- substitution,
p(1/C)- logarithmic value of diuretic activity.

Table 2: Regression parameters and quality corrections are using regular multiple regression analysis

Model No.	Parameters used	Regular Regression Analysis			
		R^2	R^2_A	CV	F
1	□ , IP_4	0.7420	0.7105	0.8858	23.083
2	□□ IP_2, IP_4	0.7775	0.7330	0.8506	17.471
3	□□ IP_1, IP_4	0.7866	0.7440	0.8329	18.434
4	□ ,IP_3, IP_4	0.8000	0.7600	0.8065	19.999

Table 3: Variance inflation factor (VIF), eigen values (□), condition number (k) and Λ- statistic and other parameters using regular multiple regression analysis

Model	Parameter used	VIF	□i	□□□□i	□□□□□i	Λ
1	IP_4	1.0807	1.2732	0.7854	2.1612	1.0806
	□	1.0807	0.7268	1.3758		
2	IP_4	1.3261	1.3470	0.7423	4.0940	1.3646
	□	1.3751	1.2623	0.7922		
	IP_2	1.3930	0.3907	2.5595		
3	IP_4	1.2160	1.7690	0.5652	3.9306	1.3102
	□	1.2847	0.7298	1.3702		
	IP_1	1.4301	0.5012	1.9952		
4	IP_3	1.3270	1.4045	0.7119	3.9253	1.3084
	IP_4	1.2149	1.1718	0.8533		
	δ	1.3836	0.4237	2.3601		

Table 4: Regression parameters and quality corrections using robust multiple regression analysis

Model No.	Parameters used	Robust Regression Analysis			
		R^2	R^2_A	CV	F
5	□ , IP_4	0.7972	0.7718	0.7785	31.446
6	□□ IP_2, IP_4	0.7907	0.7489	0.8227	18.895
7	□□ IP_1, IP_4	0.8027	0.7633	0.7836	20.347
8	□ ,IP_3, IP_4	0.8244	0.7893	0.7405	23.481

Table 5: Variance inflation factor (VIF), eigen values (□), condition number (k) and Λ - statistic and other parameters using robust multiple regression analysis

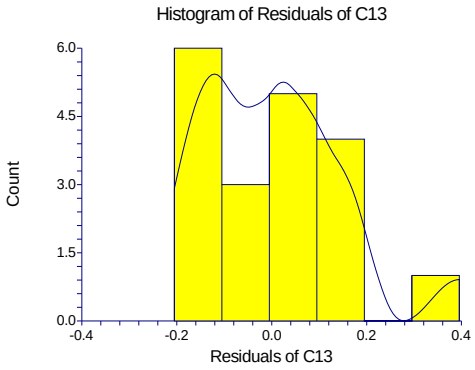
Model	Parameter used	VIF	□i	□□□□i	□□□□□i	Λ
5	IP_4	1.0929	1.2915	0.7742	2.1856	1.0928
	□	1.0929	0.7085	1.4114		
6	IP_4	1.3603	1.3567	0.7370	4.2018	1.4006
	□	1.4080	1.2698	0.7875		
	IP_2	1.4337	0.3735	2.6773		
7	IP_4	1.2080	1.7747	0.5634	3.9498	1.3166
	□	1.3039	0.7289	1.3719		
	IP_1	1.4380	0.4964	1.0145		
8	IP_3	1.3664	1.4148	0.7068	4.0297	1.3432
	IP_4	1.2405	1.1813	0.8465		
	δ	1.4226	0.4038	2.4764		

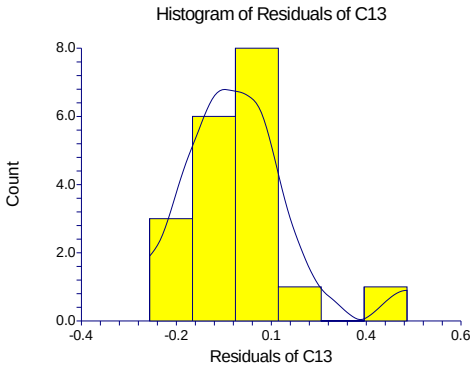
Table 6: Comparison of observed and calculated diuretic activity using three variable models containing IP_3. IP_4 and □

Compound No.	p (1/C) (Obs.)	Calculated p(1/C)			
		Regular Multiple regression		Robust Multiple regression	
		Cal.	Res.	Cal	Res.
1	-0.040	-0.313	0.273	-0.289	0.249
2	-0.300	-0.027	-0.273	-0.023	-0.277
3	0.190	0.142	0.048	0.135	0.055
4	0.180	0.194	-0.014	0.185	-0.005
5	0.180	0.205	-0.025	0.195	-0.015
6	0.130	0.258	-0.128	0.244	-0.114
7	0.290	0.406	-0.116	0.382	-0.092
8	0.270	0.395	-0.125	0.372	-0.102
9	0.290	0.448	-0.158	0.422	-0.132
10	0.390	0.490	-0.100	0.461	-0.071
11	1.000	0.575	0.425	0.540	0.460
12	0.680	0.670	0.010	0.629	0.051
13	0.820	0.638	0.182	0.599	0.221
14	0.240	0.246	-0.006	0.255	-0.015
15	0.320	0.384	-0.064	0.383	-0.063
16	0.580	0.510	0.070	0.502	0.078
17	-0.300	-0.434	0.134	-0.425	0.125
18	-0.300	-0.339	0.039	-0.336	0.036
19	-0.300	-0.127	-0.173	-0.139	-0.161

Res- Residual i.e. difference between observed and calculated diuretic activity.

Regular MLR **Robust MLR**





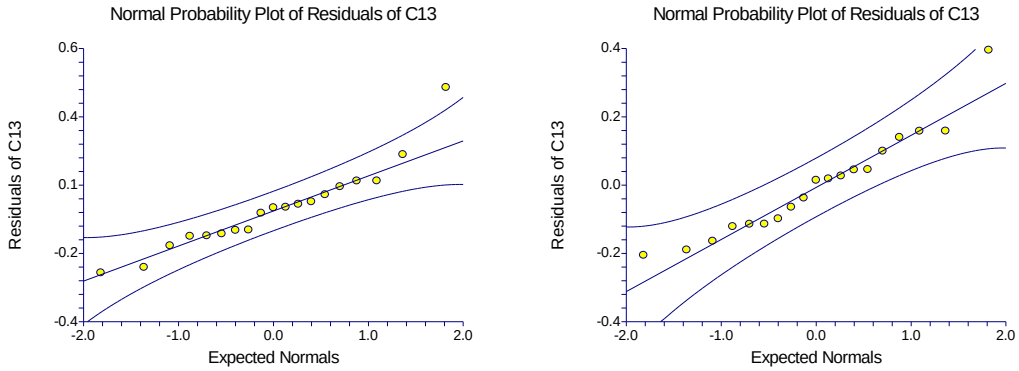


Figure2: calculation of predicative correlation coefficient (R^2pred),

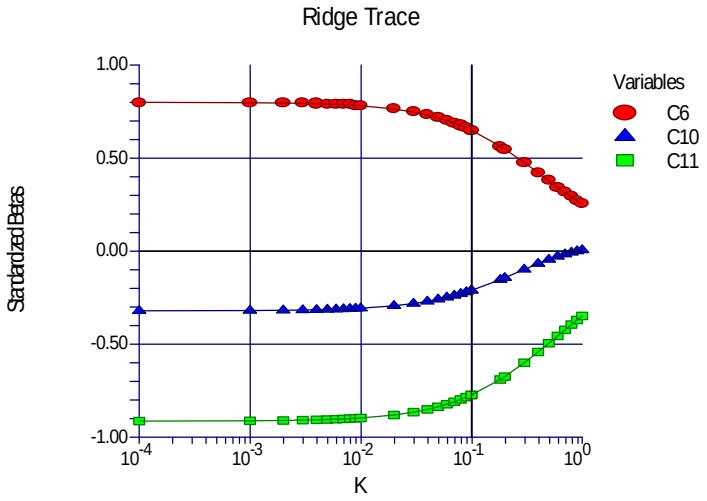


Fig. 3: Ridge multiple regression analyses

REFERENCES

1. Clare, B.W.; Supuran, C.T.; QSAR Studies of Sulfonamide Carbonic Anhydrase Inhibitors, in Supuran, C.T.; Scozzatava, A.; Conway, J. (Eds.), Carbonic Andydrase, its Inhibitors and Activators, CRC Press, Boca Raton: FI, USA, 2004.
2. Saxena, A.; Khadikar, P.V. Acta Pharma 1999, 49, 171.
3. Agrawal, V.K.; Sinha, S.; Bano, S.; Khadikar, P, V. Acta Mitrobin et. Immina. Hung. 2001, 48, 17.
4. Agrawal, V.K.; Shrivastava, R.; Khadikar, P.V. Bioorg.Med Chem. 2001, 9, 3287.
5. Agrawal, V.K; Sharma R.; Khadikar, P.V. Bioorg. Med. Chem. 2002, 10, 2993.
6. Agrawal, V.K.; Shrivastava, S.; Khadikar, P.V.; Supuran, C. T. Bioorg.Med. Chem. 2003, 11, 5353.
7. Thakur, A.; Thakur; Khadikar, P.V.; Supuran, C.T.; Sudele, P. Bioorg.Med.Chem. 2004, 12, 789.
8. Jaiswal, M.; Khadikar, P.V.; Supuran, C.T. Bioorg.Med.Chem.Letter. 2004, 12, 2477.
9. Jaiswal, M.; Khadikar, P. V.; Scozzafava, A.; Supuran, C.T. Bioorg.Med. Chem.Letter. 2004, 14, 3283.
10. Agrawal, V. K.; Bano, S.; Supuran, C.T.; Khadikar, P.V. Eur. J.Med. Chem. 2004, 39, 593.
11. Agrawal, V. K.; Khadikar, P.V. Bioorg. Med. Chem. Letter.2003, 13, 447.
12. Khadikar, P.V; Sharma, V; Karmarkar.S; Supuran, C.T. Bioorg. Med. Chem. Lett. 2005 (In press)
13. Khadikar, P. V; Pathak, S.; V.; Shrivastava, A.Bioorg, Med. Chem. Lett. 2002, 12,2673.

14. Khadikar, P.V.; Bajaj, A. V.; Mandloji, D. Indian . J. Chem. 2002, 41A, 2065.
15. Khadikar, P.V.; Mandloji D.; Bajaj, A. V. Oxid. Commun. 2004, 27, 23.
16. Jaswal, M.; Khadikar, P.V. Bioorg. Med. Chem. 2004, 12, 1793.
17. Jaiswal, M.; Khadikar, P. V. J. Indian. Chem. Soc.2004 (In press)
18. Khadikar, P. V; Sharma, V.; Supuran, C. T. Bioorg. Med. Chem. Lett. 2004 (In press)
19. Khadikar, P.V; Sharma, V.; Karmarkar, S., Supuran, C. T. Bio org Med. Chem. Lett. 2005 in press this ref.'s different from rq12.)
20. Kakeya, N.; Yata, N. Kamada, A. Aoki, M. Chem.Pharm.Bull. 1969,17,2000; Kakeya, N.; Yata, N.; Kamada, A.; Aoki, M. Chem.Pharm.Bull. 1969,17,2558; Kakeya, N.; Yata,N.; Kamada,A.; Aoki,M. Chem Pharm.Bull. 1970, 18, 191.
21. Wiener, H. J. Am. Chem .Soc., 1947, 69.17.
22. Gutman, I., Graph Theory Notes, New York, 1994, 27, 9.
23. Khadikar, P.V.; Deshpande, N.V.; Kale, P.P.; Dobrynin, A.; Gutman, I.; Domotor, G., J. Chem. Inf. Comput. Sci., 1975, 35, 547.
24. Khadikar, P.V.; Kale,P.P.; Deshpande,N.V.; Karmarkar,S. ;Agrawal,V.K., Commun, Math. Comput. Chem. (METCH), 2001,43,7.
25. Khadikar, P.V., Nat. Acad, Sci, Lett. 2000, 23, 113.
26. Khadikar, P.V.; Karmarkar, S; Agrawal, V.K., J. Chem.Inf.Comput.Sci, 2001, 41, 934.
27. Khadikar, P.V.; Kale, P.P.; Deshpande, N.V.; Karmarkar, S.; Agrawal, V.K., J. Meth.Chem. 2001, 29, 134.
28. Randic, M. J.Am.Chem.Soc, 1975,97,6609.
29. Kier,L.B., Hall,L.H., Molecular Connectivity in Chemistry and Drug Research, Academic: New York, 1976.
30. Personal communition
31. Chaterjee, S, Hadi, A.S.; Price, B., Regression Analyses by Examples , 3rd ed., Wiley: New York, 2000.
32. Balaban A.T., J. Chem. Inf. Comput. Sci. , 1992, 32, 23.
33. Balaban A.T., Math. Chem., 1994, 21, 115.
34. Sokal, R.R.; Sneath, P.H. Principles of Numerical Taxonomy: San Francisco, 1963; P- 182.
35. Kikuchi, O. Quant. Struct. -Act Relat . 1987, 6, 179.
36. Todeschini, R., Consonni, V., HandBook of Molecular Descriptors. Wiley-VCH: Weinheim (GER). 2000