Available online at www.sciencedirect.com
SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry 12 (2004) 1793–1798

Bioorganic & Medicinal Chemistry

QSAR study on ^{13}C NMR chemical shifts on carbinol carbon atoms

Mona Jaiswal and Padmakar Khadikar*

Research Division, Laxmi Fumigation & Pest Control (P), 3, Khatipura, Indore 452007, India

Received 4 October 2003; accepted 9 December 2003

Abstract—QSAR calculations of ^{13}C NMR chemical shifts (ppm, TMS = 0) on carbinol carbon atoms have been attempted using a large set of distance based topological indices: Wiener (W)-, Szeged (Sz)-, PI (Padmakar-Ivan) and Connectivity ($^m\chi$, $^m\chi^v$) indices. The regression analysis has shown that excellent results are obtained in multiparametric regression. The predictive power of the proposed models are discussed using cross-validated parameters.

1. Introduction

It is well known that many alcohols are specially useful as solvents and are much used in synthesizing several other derivatives. In addition, they (C_6–C_{10}) are used as plasticizers especially in vinyls. Alcohols (C_{12}–C_{16}) are used as synthetic detergents, evaporation retardants (C_{16}–C_{18}), lubricants (C_{20}–C_{20}) and mold release agents. Many of them exhibit narcosis and are toxicants.[1,2]

All the aforementioned properties/activities are molecular property/activity and are, therefore, related to the molecular structures of alcohols. The ^{13}C NMR chemical shifts on carbinol carbon atoms (ppm, TMS = 0) likewise is a molecular parameter and has direct relevance to the properties and activities mentioned above.

With the introduction of topology and graph theory in chemistry[3–10] it is established that molecular structures of the organic molecules can be considered as molecular graphs, obtained by deleting all the hydrogen atoms in the molecular structure. Numerical value assigned to such molecular graphs, under specified conditions, are called topological indices. These topological indices that is, graph theoretical descriptors are considered as numerical representative of molecular structures.

In the area of NMR spectroscopy, the graph-theoretical techniques, particularly topological indices are very useful. Randic and co-workers[11–13] have shown that graph theoretical techniques could also be used to obtain the chemical shifts of nucleus. They have developed[14] a computer code for listing equivalence classes of graphs works for most of the small, non transitive, and nonisospectral graphs. In addition, Duvenbeck[15] has discussed topological and geometrical approaches to develop models for the prediction of ^{13}C NMR chemical shifts and used Σ^{13}C NMR chemical shifts as a molecular property.

In one of our earlier reports,[16] we have used Wiener (W)[17]-, Szeged (Sz)[18–20]-indices for predicting ^{13}C NMR chemical shifts of 2,6- and 2,7-disubstituted naphthalenes. Our results showed that W and Sz indices can be successfully used for predicting ^{13}C NMR chemical shifts and that Σ^{13}C can be used as a molecular property which in turn can be modeled by both W and Sz indices successfully.

Prompted by the aforementioned success we have used PI (Padmakar-Ivan) index[21–23] for modeling ^{13}C NMR shifts in alkenes and cycloalkenes and obtained interesting results.[24] In another report, we observed that PI index can also be used successfully for modeling coupling constants.[25]

In view of the above, and in continuation of our earlier work, we have now used a large set of topological indices

---

Keywords: ^{13}C NMR; Carbinols; PI (Padmakar-Ivan) index; Regression analysis; Cross-validation; Pogliani Q factor.

* Corresponding author. Tel.: 91-731-2531906; fax: 91-731-2701170; e-mail: pvkhadikar@rediffmail.com

0968-0896/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.bmc.2003.12.017

(W, Sz, PI, $^0\chi$, $^1\chi$, $^2\chi$, $^0\chi^v$, $^1\chi^v$ and $^2\chi^v$) for the QSAR study of ^{13}C NMR chemical shifts (ppm, TMS = 0) on carbinol carbon atoms of the alcohols used in the present study (Table 1). Since the set of alcohols used are acyclic, by defination their W index coincides with Sz index.[19,20] The results, as discussed below, show that excellent model is a pentaparametric model containing PI,W and the connectivity indices[26–29] ($^0\chi$, $^1\chi$, $^2\chi$) for modeling ^{13}C NMR chemical shifts on carbinol carbon atoms. Our results, will, therefore, be useful for the organic chemists, pharmacists, and medicinal chemists interested in investigating/modeling properties/activities of alcohols (Table 1).

2. Results and discussion

The alcohols, their ^{13}C NMR chemical shifts (ppm, TMS = 0) on carbinol carbon atoms and the values of Wiener (W)- and PI (Padmakar-Ivan) indices are presented in Table 1. This Table 1 shows that no degeneracy is present in ^{13}C NMR chemical shifts, while high degeneracy is present in W, Sz and PI indices. This is due to the fact that these indices belong to first-generation topological indices which inspite of their degeneracy are used successfully in Quantitative Structure–Property– Activity– Toxicity Relationships (QSPR/QSAR/QSTR).[30]

Table 1. ^{13}C NMR chemical shifts (ppm, TMS = 0) on carbinol carbon atoms in alcohols and their W and PI indices

Compd	Alcohol	^{13}C NMR chemical shifts (ppm, TMS = 0)	W	PI
1	Mehanol	49.0	1	0
2	Ethanol	57.0	4	2
3	Propanol	63.6	10	6
4	Butanol	61.4	20	12
5	Pentanol	61.8	35	20
6	Hexanol	61.9	56	30
7	Isopropanol	63.4	9	6
8	2-Butanol	68.7	18	12
9	2-Pentanol	67.0	32	20
10	2-Hexanol	67.2	52	30
11	3-Pentanol	73.8	31	20
12	3-Hexanol	72.3	50	30
13	3-Heptanol	72.6	76	42
14	4- Heptanol	70.6	75	42
15	4-Octanol	70.9	108	56
16	5-Nonanol	71.1	149	72
17	Isobutanol	68.9	18	12
18	Tert-Butanol	68.4	16	12
19	Neopentanol	72.6	28	20
20	2-Me-Pentanol	66.9	50	30
21	3-Me-Butanol	60.2	32	20
22	3-Me-2-Butanol	72.0	29	20
23	4-Me-2-Butanol	65.2	32	20
24	4-Me-3-Pentanol	77.3	46	30
25	3,3-Di-Me-Butanol	58.9	46	30
26	2,3-Di-Me-2-Butanol	72.2	42	30
27	3,3-Di-Me-2-Butanol	74.8	42	30
28	4,4-Di-Me-3-Butanol	80.9	46	30
29	2,4-Di-Me-3-Pentanol	80.4	65	42
30	2,3,3-Tr-Me-2-Butanol	74.1	58	42
31	2,4,4-TriMe-3-Pentanol	82.8	86	56
32	2,2,4,4-TetraMe-3-Pentanol	84.7	111	72

Table 2 contains calculated values of Randic[26,27] ($^0\chi$, $^1\chi$, $^2\chi$) and Kier and Hall[28,29] ($^0\chi^v$, $^1\chi^v$, $^2\chi^v$) connectivity indices. We observed that (Table 2) compared to W and PI indices lesser degeneracy is present in these indices. This is due to the fact that all of them are second generation topological indices.[30] Out of these indices, $^1\chi$ (first-order) is found very useful in QSAR/QSPR/QSTR studies.

A preliminary regression analysis[31] indicated that none of the topological indices (Tables 1 and 2), in one variable regression, are capable of modeling, monitoring, and estimating ^{13}C NMR chemical shifts of the alcohols used. That is, no monoparametric regression models are possible for modeling ^{13}C NMR chemical shifts of the alcohol using these indices.

In stepwise regression analysis several statistically significant biparametric expressions are obtained. The quality of these biparametric regression expressions are given in Table 3. The models (Table 3) are more or less of similar quality, one containing W and $^0\chi$ gave slightly better results. This model is found as:

$$
\begin{aligned}
&^{13}\text{C NMR chemical shift}\\
&\quad = 35.8437\ (\pm 4.9297) - 0.1643\ (\pm 0.0564)\ W\\
&\qquad + 7.5070\ (\pm 1.2946)\ {}^0\chi \qquad (1)
\end{aligned}
$$

$n = 32$, $Se = 4.6203$, $R = 0.8207$, $F = 29.946$, $Q = 0.1776$

Table 2. Connectivity indices of the alcohols used in the present investigation (Ref Table 1)

Compd	$^0\chi$	$^1\chi$	$^2\chi$	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$
1	2.0000	1.0000	0.0000	1.4472	0.4472	0.0000
2	2.7071	1.4142	0.7071	2.1543	1.0233	0.3162
3	3.4142	1.9142	1.0000	2.8614	1.5233	0.7236
4	4.1213	2.4142	1.3536	3.5685	2.0233	1.0772
5	4.8284	2.9142	1.7071	4.2756	2.5233	1.4307
6	5.5355	3.4142	2.0607	4.9827	3.0233	1.7843
7	3.5774	1.7321	1.7321	3.0246	1.4129	1.0937
8	4.2845	2.2701	1.8021	3.7317	1.9509	1.2573
9	4.9916	2.7701	2.1825	4.4388	2.4509	1.6377
10	5.6987	3.2701	2.5361	5.1459	2.9509	1.9912
11	4.9916	2.8081	1.9217	4.4388	2.4889	1.4703
12	5.6987	3.3081	2.3021	5.1459	2.9889	1.8507
13	6.4058	3.8081	2.6556	5.8530	3.4889	2.2043
14	6.4058	3.8081	2.6825	5.8530	3.4889	2.2312
15	7.1129	4.3081	3.0361	6.5601	3.9889	2.5847
16	7.8200	4.8081	3.3896	7.2672	4.4889	2.9383
17	4.2845	2.2701	1.8021	3.7317	1.8792	1.5764
18	4.5000	2.0000	3.0000	3.9472	1.7236	2.1708
19	5.2071	2.5607	2.9142	4.6543	2.1698	2.7188
20	5.6987	3.3081	2.3021	5.1459	2.9172	2.0764
21	4.9916	2.7701	2.1825	4.4388	2.3792	1.9061
22	5.1547	2.6427	2.4880	4.6019	2.3236	1.9846
23	4.9916	2.7701	2.1825	4.4388	2.4509	1.6377
24	5.8618	3.1807	2.6295	5.3090	2.8616	2.2196
25	5.9142	3.0607	3.3107	5.3614	2.6698	3.0343
26	6.0774	2.9434	3.5207	5.5246	2.6670	2.8084
27	6.0774	2.9434	3.5207	5.5246	2.6242	3.0420
28	5.8618	3.1807	2.6295	5.3090	2.8616	2.2196
29	6.7321	3.5534	3.3472	6.1793	3.2343	2.9786
30	7.0000	3.2500	4.5000	6.4472	2.9736	3.8090
31	7.6547	3.8541	4.3987	7.1019	3.5349	4.0549
32	8.5774	4.1547	5.4537	8.0246	3.8355	5.1346

It is interesting to record that when $^0\chi^v$ is used in place of $^0\chi$ similar results are obtained. The positive sign of $^0\chi$ and/or $^0\chi^v$ indicates that the nature of atom(s) has significant effect on ^{13}C NMR chemical shift; while negative coefficient of W indicates retarding effect of branching on the exhibition of ^{13}C NMR chemical shift of alcohols used.

Here and thereafter n- is the number of the compound, Se- is the the root mean square error, R- is the correlation coefficient, F- is the F-statistics and Q- is the quality factor.[32,33] The quality factor (Q) is defined[32,33] as the ratio of correlation coefficient (R) to the standard error of estimation (Se) that is, Q = R/Se and is a measure of predictive power of the model.

The successive regression resulted into several statistically significant triparametric models (Table 3). Out of these the models: 12, 13, 14–19 gave better results than the biparametric model [equation (1)] discussed above. Among the triparametric models the one containing $^0\chi$, $^1\chi$, and $^2\chi$ gave the best results:

$$\begin{aligned}
&^{13}\text{C NMR chemical shift}\\
&\quad = 27.9509\ (\pm 7.5240) + 31.9440\ (\pm 8.8760)\ ^0\chi\\
&\qquad - 31.8708\ (\pm 9.3597)\ ^1\chi\\
&\qquad - 15.2249\ (\pm 5.6145)\ ^2\chi \qquad (2)
\end{aligned}$$

$n = 32,\ \ Se = 4.3146,\ \ R = 0.8515,\ \ F = 24.6265,$
$Q = 0.1974$

The corresponding model containing $^0\chi^v$, $^1\chi^v$, and $^2\chi^v$ was of slightly lower quality and the one containing PI, $^0\chi$ and $^1\chi$ was of still lower quality.

The connectivity index $^1\chi$ conveys more information about the nature of atoms and first-order branching, while $^2\chi$ encodes more information about branching. Negative coefficient of these parameters in eq 2 indicates retarding effect of branching on the exhibition of ^{13}C

Table 3. Attempted regression expressions, and their quality of correlations

(i) Monoparametric regressions

None of the topological indices resulted into statistically significant results for estimating ^{13}C NMR chemical shifts on carbinol carbon atoms

(ii) Biparametric regression

The following biparametric regressions gave statistically significant results for estimating ^{13}C NMR chemical shifts on carbinol carbon atoms

Model	Parameter used	Se	R	F	Q
1	W, PI	4.8530	0.7999	25.7638	0.1648
2	W, $^0\chi$	4.6203	0.8207	29.9216	0.1776
3	W, $^0\chi^v$	4.6203	0.8207	29.9216	0.1776
4	$^0\chi$, PI	4.8325	0.8018	26.1066	0.1659
5	$^0\chi^v$, PI	4.8325	0.8018	26.1066	0.1659

(iii) Triparametric regressions

The following triparametric regressions gave statistically significant results for estimating ^{13}C NMR chemical shifts on carbinol carbon atoms.

Model	Parameter used	Se	R	F	Q
6	PI, $^0\chi$, $^1\chi$	4.6850	0.8219	19.4697	0.1754
7	W, $^1\chi$, $^2\chi$	5.0359	0.7909	15.5949	0.1570
8	PI, $^0\chi^v$, $^1\chi^v$	4.7649	0.8154	18.5114	0.1711
9	W, $^1\chi^v$, $^2\chi^v$	5.0288	0.7916	15.6652	0.1574
10	PI, W, $^1\chi$	4.7727	0.8147	18.4199	0.1707
11	PI, $^1\chi^v$, $^2\chi^v$	5.2165	0.7735	13.8986	0.1483
12	$^0\chi$,$^1\chi$, $^2\chi$	4.3146	0.8515	24.6265	0.1974
13	$^0\chi^v$, $^1\chi^v$, $^2\chi^v$	4.4192	0.8436	23.0381	0.1910

(iv) Tetraparametric regressions

Model	Parameter used	Se	R	F	Q
14	PI, $^0\chi$,$^1\chi$,$^2\chi$	4.0529	0.8753	22.1157	0.2160
15	W, $^0\chi$,$^1\chi$,$^2\chi$	4.1653	0.8678	20.5783	0.2083
16	PI, $^0\chi^v$, $^1\chi^v$, $^2\chi^v$	4.2083	0.8648	20.0237	0.2055
17	W, $^0\chi^v$, $^1\chi^v$, $^2\chi^v$	4.2019	0.8652	20.1046	0.2059

(v) Pentaprametric regression

Model	Parameter used	Se	R	F	Q
18	PI,W, $^0\chi$, $^1\chi$, $^2\chi$	3.9149	0.8888	19.5494	0.2270
19	PI, W, $^0\chi^v$, $^1\chi^v$, $^2\chi^v$	4.2083	0.8648	20.0237	0.2055

NMR chemical shifts. On the other hand positive coefficient of $^{0}\chi$ once again indicates favourable effect of nature of atoms on the exhibition of ^{13}C NMR chemical shift.

The aforementioned results show that in both, bi- and tri-parametric regressions, $^{1}\chi$ was found superior than other topological indices. The results show that combination of $^{0}\chi$, $^{1}\chi$, and $^{2}\chi$ is the most favourable combination for estimating ^{13}C NMR chemical shifts of the alcohols used.

With a hope of obtaining regression expression (model) with still improved statistics, we attempted tetraparametric regression analysis in that combination W or PI with $^{0}\chi$, $^{1}\chi$, $^{2}\chi$ on one hand and $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$ on the other hand (Table 3) were considered. The results show that a tetraparametric model containing PI, $^{0}\chi$, $^{1}\chi$, and $^{2}\chi$ gave better results:

$$\begin{aligned}{}^{13}\text{C NMR chemical shift} &= 12.3400\ (\pm 1.0717) - 0.3476\ (\pm 0.1598)\ \text{PI} \\ &\quad + 37.8727\ (\pm 8.7717)\ {}^{0}\chi \\ &\quad - 32.4372\ (\pm 8.7958)\ {}^{1}\chi \\ &\quad - 17.2943\ (\pm 5.3590)\ {}^{2}\chi \end{aligned} \tag{3}$$

$n = 32$, $Sd = 4.0529$, $R = 0.8753$, $F = 22.116$, $Q = 0.2160$

The successive regression analysis ultimately resulted into a pentaparametric model containing PI, W, $^{0}\chi$, $^{1}\chi$, and $^{2}\chi$ and having slightly better statistics:

$$\begin{aligned}{}^{13}\text{C NMR chemical shift} &= -9.8149\ (\pm 1.6179) - 1.6192\ (\pm 0.7578)\ \text{PI} \\ &\quad + 0.5402\ (\pm 0.3152)\ \text{W} \\ &\quad + 59.3994\ (\pm 15.1512)\,{}^{0}\chi \\ &\quad - 54.1402\ (\pm 15.2322)\,{}^{1}\chi \\ &\quad - 25.3112\ (\pm 6.9769)\ {}^{2}\chi \end{aligned} \tag{4}$$

$n = 32$, $Sd = 3.9149$, $R = 0.8888$, $F = 19.549$, $Q = 0.2270$

Looking to the sample size and considering recommendations of Thumb's rule[34] we can't attempt still higher parametric regressions. However, the attempts made resulted into slight improvement in R, but the resulting models contains one or more parameters whose coefficients were considerably smaller than their corresponding standard deviations. Such models are not allowed statistically.

At this stage it is worthy to comment on the occurrence of highly correlated topological indices in the models discussed above. Normally when such is the case, the model suffers from the defect due to colinearity. However, such cases were dealt by Randic[35] who made certain recommendations to interpret such models. We will use the recommendations of Randic in our case also.

Before discussing Randic recommendations,[35] it is interesting to mention that the multicolinearity problem results into one of the following problems: (i) coefficients of the parameters involved have unexpected signs, (ii) incorrect signs of the coefficient; (iii) a change in the values of the previous coefficient when a new variable is added to the model; (iv) change to insignificance of the previously significant variable when a new variable is added to the model and (v) increase in the standard error of the estimate when a variable is added to the model. The mathematical recommendation is to eliminate the variables that are chief cause of the multicolinearity problem. Dealing with multicolinearity problem requires a great deal of experience.

We now discuss Randic recommendations[35] accounting the presence of highly correlated parameters in the regression equations such as those discussed above. Though our proposed models contain such highly correlated topological indices, the corresponding regression equations are statistically significant as in each case the coefficients of the topological indices involved are significantly larger than their standard deviation. Further evidence is obtained from the recommendations made by Randic,[35] in that he stated that one should be particularly aware of common fit in regression analysis in describing descriptors that are highly intercorrelated. Randic[35] further stated that, by discarding one of the descriptors which commonly duplicate another, may carry useful structural information in that part in which it does not parallel with another descriptors. In fact highly correlated topological indices in our models do have different structural information. Hence, our models containing highly correlated topological indices are statistically significant.

We now discuss the predictive power of our proposed models. In a simple case this can be done using Pogliani's quality factor, Q.[32,33] This factor is defined as the ratio of correlation coefficient (R) to the standard error of estimations (Se) that is, Q = R/Se. This means Q is directly proportional to R and inversely to Se. Hence, larger the Value of R, the smaller the value of Se, the higher will be Q, and the better will be the predictive power of the models.

A persual of Table 3 shows that the value of Q go on increasing as we pass from bi- to penta-parametric models and is highest for such model containing PI, W, $^{0}\chi$, $^{1}\chi$, $^{2}\chi$ as the correlating parameters (model 18).

Though we observed that the quality factor Q is quite useful in the present case, its use was highly criticized by Todeschini.[36] In view of this we have used leave-one-out methodology adopting cross-validation method.[31–34] The cross-validated parameters so calculated are given in Table 4.

It is worthy to record that as opposed to traditional regression models, cross-validation evaluates the validity of a model by how well it predicts the data rather than how well it fits the 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 deviation and predicted in turn.

PRESS (predictive sum of squares) is one of the most important cross-validated parameters. PRESS smaller than SSY (sum of squares of deviation for each activity) and the ratio PRESS/SSY smaller than 0.4 indicates statistically significant model. If the ratio goes smaller than 0.1 the model is considered having excellent predictive power. In our case (Table 4), the models 12, 14, 15 and 18 have this ratio smaller than 0.4 and is the least for the pentaparametric model-18 indicating to be the best model having highest statistical quality and predictivity.

Another cross-validated parameter of interest is the predictive correlation which is usually named as cross-validation r^2_{cv}. The highest value of r^2_{cv} indicates the best predictive power. Once again r^2_{cv} is in favour of model-18.

The uncertainty in prediction is determined by S_{PRESS}. However, it is not useful in our case as it coincides with standard error of estimation (Se). In such cases the parameter called predictive square error (PSE) is used to account for the uncertainty of prediction; the smallest value of which (PSE) favour highest predictive power. Once again PSE (Table 4) is in favour of model-18.

3. Conclusion

From the aforementioned results and discussion we conclude that PI index in combination with other topological indices (W, $^0\chi$, $^1\chi$, $^2\chi$) can be used successfully for modeling ^{13}C NMR chemical shift vis-a-vis physiochemical and biological activity of organic compounds. In lower parametric models, $^1\chi$ is found better than PI index. The branching has significant retarding effect on the exhibition of ^{13}C NMR chemical shift of alcohols.

4. Experimental

^{13}C NMR chemical shift — obtained from ^{13}C NMR spectra of the compounds were used.

Topological indices — Hydrogen-suppressed molecular graphs of aromatic hydrocarbons and heterocycles are used for the computation of the topological indices. The details of calculations are available in the literature.[4–10] However, below we give the expressions used for their calculations.

Wiener index (W)[17] — The Wiener index W = W(G) of a graph G is defined as the half-sum of the elements of the distance matrix:

$$W = W(G) = 1/2\sum_{i=1}\sum_{j=1}(D)_{ij} \qquad (5)$$

where $(D)_{ij}$ is the *ij*-th element of the distance matrix which denotes the shortest graph theoretic distance between sites *i* and *j* in G. This index is not applicable to cyclic graphs.

Szeged index (Sz)[18,20] — It is supposed to be the modification of W index for cyclic graphs and is defined as:

$$Sz = Sz(G) = \sum_{e} n_u . n_v \qquad (6)$$

where n_u is the number of vertices nearer to the vertex u than the vertex v. The meaning of n_v is analogous. The vertices equidistant to edge e = uv are not taken into account for the calculation of Sz. For acyclic graphs Sz coincides with W index.

PI (Padmakar-Ivan) index[21–23] — The PI (Padmakar-Ivan) index is considered to be the modification of Sz index such that unlike Sz index; PI index for cyclic graphs do not coincide with PI index for acyclic graphs. It is defined as:

$$PI = PI(G) = \sum_{e}(n_{eu} + n_{ev}) \qquad (7)$$

where n_{eu} is the number of edges nearer to u than v (u and v being both ends of an edge e = uv). The meaning n_{ev} is analogous. Like Sz index, here also the edges equidistant from both ends of an edge (e = uv) are not considered for the calculation of PI index.

Randic Connectivity indice ($^0\chi$, $^1\chi$, $^2\chi$)[26,27] — The general expression used for the calculation of Randic connectivity indices ($^0\chi$, $^1\chi$, $^2\chi$) is:

$$^m\chi^R = \sum_{(i,j,k)}\left(\delta_i, \delta_j, \delta_k\right)^{-0.5} \qquad (8)$$

Table 4. Cross-validation parameters for the most significant regressions under particular category for estimating ^{13}C NMR chemical shifts on carbinol carbon atoms

Model (eq)	Parameters used	PRESS	SSY	PRESS/SSY	r^2_{cv}	S_{PRESS}	PSE
2	W, $^0\chi$ (1)	619.0767	1277.5021	0.4846	0.5154	4.6203	4.3984
12	$^0\chi$, $^1\chi$, $^2\chi$ (2)	521.2447	1375.3340	0.3840	0.6160	4.3146	4.0360
14	PI, $^0\chi$, $^1\chi$, $^2\chi$ (3)	443.4992	1453.0796	0.3052	0.6948	4.0529	3.7228
18	PI, W, $^0\chi$, $^1\chi$, $^2\chi$ (4)	398.4830	1498.0958	0.2660	0.7340	3.9149	3.5188

Kier-Hall valence connectivity index (${}^0\chi^v$, ${}^1\chi^v$, ${}^2\chi^v$)[28,29] — Like the above general expression used for the calculation of Kier and Hall valence connectivity indices (${}^0\chi^v$, ${}^1\chi^v$, ${}^2\chi^v$) is as:

$$ {}^{m}\chi^{v} = {}^{m}\chi^{v}(G)\sum_{(i,j,k)}\left(\delta_i^v, \delta_j^v, \delta_k^v\right)^{-0.5} \tag{9} $$

where δ^v the valence delta values given by:

$$ \delta_i^v = \frac{Z_i^v - H_i}{Z_i = Z_j - 1} \tag{10} $$

where Z_i is the atomic number of atom i, Z_i^v is the number of the valence electron of the atom i and H_i is the number of the hydrogen atoms attached to atom i.

Regression analysis — All the regressions is carried out by maximum R^2 method while 'leave-one-out' methodology is used in deriving cross-validated parameters.

Softwares — The topological indices were calculated using the software prepared by Raj Singh Sisodia, while the regression analysis were performed using Regress-1 program provided by Professor Istvan Lukovits, Hungarian Academy of Sciences, Budapest, Hungary.

Acknowledgements

One of the authors (PVK) is thankful to Prof. Ivan Gutman for introducing him (PVK) to the fascinating field: Chemical Topology and Graph Theory.

References and notes

1. Roberts, J. B.; Weigert, F. J.; Kroschcoister, J. I.; Reich, H. J. *J. Am. Chem. Soc.* **1970**, *92*, 1338.
2. Kier, L. B.; Hall, L. H. *Report in Molecular Theory* **1990**, *1*, 121.
3. Trinajstic, N. *Chemical Graph Theory*, 2nd ed.; CRC: Boca Raton: FL, 1993.
4. Devillers, J., Balaban, A. T., Eds. *Topological Indices and Related Descriptors in QSAR and QSPR*; Garden & Breach Scientific: Amsterdam, 1999.
5. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, 2000.
6. Devillers, J. *Comparative QSAR*; Tailor & Francis: Philadelphia, 1998.
7. Diudea, M. V., Ed. *QSPR/QSAR Studies by Molecular Descriptors*; Nova Science, 2000.
8. Diudea, M. V.; Khadikar, P. V. Molecular Topology and its Applications, New Age Int: New Delhi, 2003, in press.
9. Gutman, I.; Polansky, O. E. *Mathematical Concept in Organic Chemistry*; Springer-Verlag: Berlin, 1986.
10. Trinajstic, N. *Chemical Graph Theory*, Vol. II, Chapter V; CRC Press INC: Boca Raton: Florida, 1983.
11. Randic, M. *Int. J. Quantum Chem.* **1983**, *23*, 1707.
12. Randic, M. *J. Mathematical Chem.* **1984**, *59*, 34.
13. Randic, M.; Trinajstic, N. *Theor. Chem. Acta.* **1988**, *73*, 233.
14. Khadikar, P. V., Bajaj, A. V., Mandloi, D. *Bioorg. Med. Chem. Lett.*, in press.
15. Duvenbeck, C. *Topological and Geometrical Approach to Develop Models for Prediction of* ^{13}*C NMR Shifts*, Bochum: GER, 1995.
16. Khadikar, P. V.; Pathre, S.; Shrivastava, A. *Bioorg. Med. Chem.* **2002**, *12*, 2673.
17. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
18. Gutman, I. *Graph Theory Notes New York* **1994**, *27*, 9.
19. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
20. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *Commun. Math. Comput. Chem. (MATCH)* **2001**, *43*, 7.
21. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 934.
22. Khadikar, P. V.; Karmarkar, S.; Varma, R. G. *Acta. Chem. Slov.* **2002**, *49*, 755.
23. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2001**, *29*, 134.
24. Khadikar, P. V.; Bajaj, A. V.; Mandloi, D. *Indian J. Chem.* **2002**, *41A*, 2065.
25. Khadikar, P. V.; Mandloi, D.; Bajaj, A. V. *Oxid. Commun.*, in press.
26. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
27. Randic, M. *J. Mol. Graphis & Modelling* **2001**, *20*, 19.
28. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; Wiley: New York, 1986.
29. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
30. Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 23.
31. Chatterjee, S.; Hadi, A.; Price, B. *Regression Analysis by Examples*, 3rd ed.; John Wiley & Sons: New York, 2000.
32. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
33. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
34. Khadikar, P. V.; Agrawal, V. K.; Karmarkar, S. *Bioorg. Med. Chem.* **2002**, *10*, 3499 (and references therein).
35. Randic, M. *Croat. Chem. Acta* **1993**, *66*, 289.
36. Todeschini, R. *Chemometrics Web News, Milano Chemometrics & QSAR Research Group,* File: ///CI/WINDOWS/DESKTOP/Web News on Chemometrics,html. Feb **2001**.