Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 1731–1736
Bioorganic & Medicinal Chemistry

QSAR study on tadpole narcosis using PI index: a case of heterogenous set of compounds

Mona Jaiswal and Padmakar Khadikar*

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

Received 10 October 2003; accepted 8 January 2004

Abstract—QSAR Study on Tadpole narcosis of heterogenous set of compounds has been carried out using a large set of distance-based topological indices, including logP as molecular descriptors. Excellent results are obtained in biparametric regression models. It was found that PI index, is superior to ${}^{0}\chi$,${}^{2}\chi$ and logP indices for modeling Tadpole narcosis.

1. Introduction

Overton[1] found that a large number of compounds of widely diverse chemical structures exhibited varying degrees of narcosis on frog tadpoles. Leo et al.[2,3] have found a good correlation between the effective concentration (log1/c), and the partition coefficients (logP) of the molecules in octanol and water. Besides hydrophobicity/lipophilicity (logP), the steric factors are also found important in the narcotic effect of drugs. Thus, the narcotic activity is not only governed by lipid solubility of drug molecules, but steric factor is also equally important in constituting the activity of these drugs.

Abraham and Rafols[1] have reported factors that influence Tadpole narcosis using LFER (linear free-energy relatiponship) approach. According to them Tadpole narcosis is a linear function of solvametric parameters. We have arranged 'marriage' of topological and solvametric parameters to yield still better results.[4] Consequently, we thought perhaps PI (Padmakar-Ivan) index[5–7] alone can give better results for modeling Tadpole narcosis. Our earlier report[4] has shown that use of PI index, except for first-order connectivity index (${}^{1}\chi$),[8,9] results into better QSPR/QSAR models. Also, that in some cases combination of PI index with other molecular descriptors yielded excellent results. Record that PI index, like Szeged index (Sz)[10,11] is a modification of Wiener index (W).[12]

At this stage it is worthy to mention that in our earlier reports[13–28] we have examined variety of cases in that the PI index was found more useful than many of the other distance-based topological indices in QSAR/QSTR analysis. The success of PI index, is, therefore, established with rigorous examples. However, the utility of PI index in describing the mechanism of drug receptor-interactions is yet to be established. The efforts in this direction is underway and will be published elsewhere. Also, we observed that in many cases[13–28] the biological activity is well correlated with PI index and not logP. The present report is, therefore, an extension of our earlier work in demonstrating the applicability of PI index to model tadpole narcosis and to investigate relative potential of PI index in such studies.

In a recent review[29] Hansch et al. have examined QSAR lacking positive hydrophobic terms and advocated to examine those instances where hydrophobic terms are not significant.

This is, therefore, another objective of present study in that we have to examine potential of distance-based topological indices over logP for modeling tadpole narcosis. In this regard Hansch et al.[29] have discussed the cases in that parameters other than logP were found more fruitful.

Keywords: Tadpole narcosis; QSAR; PI index; Regression analysis; Distance based topological indices; Lipophilicity.
* Corresponding author. Tel.: +91-732-2531906; fax: +91-731-2531906; e-mail addresses: tomona@rediffmail.com (M. Jaiswal); pvkhadikar@rediffmail.com (P. Khadikar).

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.01.009

In view of the above, our objective of the present study is two-fold: (i) to propose QSAR model for modeling Tadpole narcosis using distance-based topological indices: Wiener (W),[12] PI (Padmakar-Ivan)[5-7], Randic connectivity(($^0\chi$,$^1\chi$,$^2\chi$)[8,9] and Kier and Hall valence connectivity ($^{0v}\chi$,$^{1v}\chi$,$^{2v}\chi$)[30,31] indices for modeling tadpole narcosis and (ii) to examine the use of logP for modeling tadpole narcosis. For this purpose we have used a heterogenous set of compounds and adopted their tadpole activity.[3]

2. Results and discussion

The heterogenous set of 22 compounds, their tadpole narcosis activity expressed as log1/c, and lipophilicity/hydrophobicity (logP), as adopted from the literature[3] are presented in Table 1.

The estimated values of distance-based topological indices are recorded in Table 2. The details of the calculation of these indices are available in the literature.[30–36]

The perusal of data in Table 2 shows that low to high degeneracy is present in all the topological indices used. This is due to the fact that either they belong to first-generation (PI, W) or second generation ($^0\chi$,$^1\chi$,$^2\chi$;$^{0v}\chi$,$^{1v}\chi$ and $^{2v}\chi$) topological indices. Balaban[37] has shown that such indices inspite of their observed degeneracy can be used successfully in QSPR/QSAR studies. This is found to be the case in the present study also.

Table 1 shows that no degeneracy is present in the activity (log 1/c) as well as lipophilicity (logP). However, some of their values are negative.

Table 1. Heterogenous set of compounds used, their tadpole activity (log 1/c) and lipopholicity/hydrophobocity (logP)

Compd	Compound	logP	log1/c
1	Ethyl acetate	0.730	1.410
2	Ethyl propionate	1.230	2.100
3	Ethyl butyrate	1.730	2.620
4	Ethyl valerate	2.230	3.050
5	Acetone	−0.210	0.490
6	2-butanone	0.290	1.020
7	2-pentanone	0.790	1.570
8	Nitromethane	0.330	0.850
9	Ethylether	0.770	1.350
10	Ethanol	0.160	0.260
11	Propanol	0.340	0.980
12	Butanol	0.840	1.770
13	Hexanol	1.840	3.030
14	Heptanol	2.340	3.600
15	Octanol	2.840	4.050
16	Methyl carbamate	−0.650	0.590
17	Ethyl carbamate	−0.150	1.460
18	Propyl carbamate	0.350	2.330
19	Isobutyl carbamate	0.650	2.490
20	Isoamyl carbamate	1.150	3.000
21	Chloroform	1.970	3.120
22	Methanol	−0.660	−0.190

The regression analysis[38] (Table 3) have shown that except $^2\chi$ all other topological indices are found to be well correlated with the activity. Furthermore, the relative correlation power of logP is the least and that $^1\chi^v$ is the most useful index for this purpose. Thus, Tadpole narcosis can be modeled more efficiently without using logP parameter. The best model obtained using first-order valence connectivity index $^1\chi^v$ is as below:

$$\log 1/c + -0.7839\ (\pm 0.1289) + 1.1859\ (\pm 0.0544)\ {}^1\chi^v \quad (1)$$

$$n = 21,\ \mathrm{Se} = 0.2737,\ \mathrm{R} = 0.9806,\ \mathrm{F} = 475.7429,\ \mathrm{Q} = 3.5828$$

In obtaining the above results we have deleted compound-**21** as an outlier. The same is found to be the case in other regressions also.

In the above (eq 1) and thereafter n is the number of compounds, Se is standard error of estimation, R is the correlation coefficient, F is F-statistics and Q is the quality factor. This quality factor Q is defined[39,40] as the ratio of correlation coefficient (R) to the standard

Table 2. Distance based topological indices for the heterogenous set of compounds used (ref Table 1)

Compd	W	$^0\chi$	$^1\chi$	$^2\chi$	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$	PI
1	32	4.9916	2.7701	2.1825	4.0236	1.9040	0.9246	20
2	50	5.6987	3.3081	2.3021	4.7307	2.4647	1.1586	30
3	75	6.4058	3.8081	2.6825	5.4378	2.9647	1.5550	42
4	102	7.2760	4.1639	3.5234	6.3081	3.3205	2.4030	56
5	9	3.5774	1.7321	1.7321	2.9082	1.2041	0.9082	6
6	18	4.2845	2.2701	1.8021	3.6154	1.7648	1.0556	12
7	32	4.9916	2.7701	2.1825	4.3225	2.2648	1.4520	20
8	9	3.5774	1.7321	1.7321	2.2637	0.8124	0.4397	6
9	20	4.1213	2.4142	1.3536	3.8225	1.9916	0.7815	12
10	4	2.7071	1.4142	0.7071	2.1543	1.0233	0.3162	2
11	10	3.4142	1.9142	1.0000	2.8614	1.5233	0.7236	6
12	20	4.1213	2.4142	1.3536	3.5685	2.0233	1.0772	12
13	56	5.5355	3.4142	2.0607	4.9827	3.0233	1.7843	30
14	84	6.2426	3.9142	2.4142	5.6899	3.5233	2.1378	42
15	120	6.9497	4.4142	2.7678	6.3970	4.0233	2.4914	56
16	20	4.1213	2.4142	1.3536	3.3938	1.2879	0.5676	12
17	35	4.8284	2.9142	1.7071	4.1010	1.8754	0.7873	20
18	56	5.5355	3.4142	2.0607	4.8081	2.3754	1.2027	30
19	79	6.4058	3.7701	2.8896	5.6783	2.7313	2.0591	42
20	114	7.1129	4.2701	3.2432	6.3854	3.2313	2.3852	56
21	9	3.5774	1.7321	1.7321	1.7112	0.6547	0.2474	6
22	1	2.0000	1.0000	0.0000	1.4472	0.4472	0.0000	00

Table 3. Regression data for one-variable models attempted

Model No.	Parameter	Se	R	Q
1	PI	0.4515	0.9242	2.0470
2	W	0.4477	0.9255	2.0672
3	$^0\chi$	0.4500	0.9226	2.0233
4	$^1\chi$	0.3690	0.9501	2.5748
5	$^2\chi$	0.6813	0.8173	1.1996
6	$^0\chi^v$	0.4034	0.9400	2.3302
7	$^1\chi^v$	0.2317	0.9806	4.2322
8	$^2\chi^v$	0.4314	0.9311	2.1583
9	logP	0.5041	0.9046	1.7945

deviation (Se) i.e., Q = R/Se and is used for accounting predictive power of the model.

It is interesting to record that $^{1}\chi^{v}$ distinguishes the degree of unsaturation and the presence of heteroatoms. The positive coefficient of this term in (eq 1) indicates that degree of unsaturation and the presence of heteroatom are favourable for the exhibition of tadpole narcosis of the heterogeneous set of compounds under present study.

The statistical characteristics of the models presented in Table 3 indicate that the next best model is the one containing first-order connectivity index ($^{1}\chi$):

$$\log 1/c = -1.3620\ (\pm 0.2510) + 1.1028\ (\pm 0.0831)\ ^{1}\chi \quad (2)$$

$$n = 21,\ \mathrm{Se} = 0.3690,\ \mathrm{R} = 0.9503,\ \mathrm{F} = 175.8718,\ \mathrm{Q} = 2.5748$$

The first-order connectivity index ($^{1}\chi$) conveys more information about the number of atoms in a molecule. Hence, the positive coefficient $^{1}\chi$ in (eq 2) shows that in addition to degree of unsaturation and the heteroatoms, the number of atoms also plays a dominating role in the exhibition of tadpole narcosis activity of the compound used.

The statistical characteristics (Table 3) shows that the model based on PI index is inferior to both the models discussed above:

$$\log 1/c = 0.3938\ (\pm 0.1659) + 0.0578\ (\pm 0.0054)\ \mathrm{PI} \quad (3)$$

$$n = 21,\ \mathrm{Se} = 0.4515,\ \mathrm{R} = 0.9242,\ \mathrm{F} = 111.2392,\ \mathrm{Q} = 2.0470$$

However, PI index gives better results than $^{0}\chi$ and $^{2}\chi$. The inferior quality of the PI index is due to the fact that it accounts for size, shape and branching and does not account for the presence of heteroatom.

As stated earlier, Hansch et al.[29] have reviewed the role of positive hydrophobic term in QSAR study. Their objective was to investigate whether or not logP is needed for the better exhibition of biological activity. In view of this, we have used logP as one of the correlating parameters and obtained the following result:

$$\log 1/c = 0.9230\ (\pm 0.1453) + 1.089\ (\pm 0.1178)\ \log\mathrm{P} \quad (4)$$

$$n = 21,\ \mathrm{Se} = 0.5041,\ \mathrm{R} = 0.9046,\ \mathrm{F} = 180.1614,\ \mathrm{Q} = 1.7945$$

The results (Table 3) shows that logP is a better parameter than $^{2}\chi$ but it is worsed parameter compared to other topological indices including PI index.

In view of the point raised by Hansch et al.[29] we have also investigated the role of logP in multiparameric regression analysis for the exhibition of tadpole narcosis activity in that we have used several two variable regressions with logP as one of the correlating parameters. The statistical characteristics of these regressions are given in Table 4.

The results (Table 4) show dramatic change in the quality of correlations when each of the topological indices independently is combined with logP. This shows that in addition to the topology of the molecule, the lipophilicity/hydrophobicity (logP) supports the exhibition of the Tadpole narcosis activity.

The data presented in Table 4 shows that the results obtained in two variable models containing: (i) $^{1}\chi$ and logP; (ii) $^{1}\chi^{v}$ and logP are more or less similar:

$$\log 1/c = -0.6821\ (\pm 0.1980) + 0.4908\ (\pm 0.0878)\ \log\mathrm{P} + 0.7278 \times (\pm 0.0846)\ ^{1}\chi, \quad (5)$$

$$n = 21,\ \mathrm{Se} = 0.2291,\ \mathrm{R} = 0.9821,\ \mathrm{F} = 243.9518,\ \mathrm{Q} = 4.2868$$

$$\log 1/c = -0.6168\ (\pm 0.1912) + 0.1437\ (\pm 0.1224)\log\mathrm{P} + 1.056\ (\pm 0.1230)\ ^{1}\chi^{v} \quad (6)$$

$$n = 21,\ \mathrm{Se} = 0.2295,\ \mathrm{R} = 0.9820,\ \mathrm{F} = 243.1467,\ \mathrm{Q} = 4.2789$$

Eqs 5 and 6 indicate that the number of atoms, degree of unsaturation, presence of heteroatom and lipophilicity/hydrophobicity are the important parameters affecting tadpole narcosis of the set of compounds used.

The results (Table 4) also show that the two variable model containing PI and logP is quite inferior to above models but is better than the two variable models containing: (i) $^{2}\chi$ and logP; (ii) $^{2}\chi^{v}$ and logP. The two variable model containing PI and logP yielded the following results.

Table 4. Regression data for two-variable models in that logP is one of the correlating parameters

Model No.	Parameter	Se	R	F	Q
10	W,logP	0.3144	0.9659	125.308	3.0722
11	PI,logP	0.3186	0.9650	121.822	3.0289
12	$^{0}\chi$,logP	0.3024	0.9685	136.181	3.2027
13	$^{1}\chi$,logP	0.2291	0.9821	243.952	4.2868
14	$^{2}\chi$,logP	0.4176	0.9390	67.137	2.2486
15	$^{0}\chi^{v}$,logP	0.2678	0.9754	176.104	3.6023
16	$^{1}\chi^{v}$,logP	0.2295	0.9820	243.147	4.2789
17	$^{2}\chi^{v}$,logP	0.3401	0.9600	105.773	2.8227

$$\log 1/c = 0.5059\ (\pm 0.1196) + 0.5539\ (\pm 0.1234)\log P + 0.0348\ (\pm 0.0064)PI \quad (7)$$

$$n = 21,\ Se = 0.3186,\ R = 0.9650,\ F = 121.8215,\ Q = 3.0289$$

In order to confirm the positive role of logP we have attempted two variable models in that only binary combinations of topological indices were used. Such results are presented in Table 5.

Out of all possible combinations, the combination of PI index with connectivity indices were interesting. Two such models are those containing PI and $^{0}\chi$ and the other containing PI and $^{2}\chi^{v}$; the latter is found better and is as below.

$$\log 1/c = 0.1027 + 0.0262(\pm 0.0147)PI + 0.8492\ (\pm 0.3718)\ ^{2}\chi^{v} \quad (8)$$

$$n = 21,\ Se = 0.4084,\ R = 0.9418,\ F = 70.596,\ Q = 2.3061$$

However, this model (eq 8) is worse than all the earlier two variable models containing logP as one of the correlating parameters (Table 4).

The best biparametric model containing topological indices alone is the one containing $^{1}\chi$ and $^{2}\chi^{v}$ as the correlating parameters:

$$\log 1/c = -0.9642 + 0.7170(\pm 0.02049)\ 1c + 0.5680\ (\pm 0.2796)\ ^{2}\chi^{v} \quad (9)$$

$$n = 21,\ Se = 0.3419,\ R = 0.9596,\ F = 104.5983,\ Q = 2.8067$$

The statistical characteristics show that even this model is worsed (except model containing $^{2}\chi$ and logP) than the two variable models in which logP is one of the correlating parameters. The results, therefore, favours the positive contribution of logP in the exhibition of tadpole activity.

Table 5. Regression data for two-variable models without the involvement of logP

Model no.	Parameter	Se	R	F	Q
18	PI, $^{0}\chi$	0.4471	0.9298	57.4303	2.0796
19	PI, $^{1}\chi$	0.4084	0.9418	70.5960	2.3060
20	$^{1}\chi$,$^{2}\chi$	0.3611	0.9548	92.8113	2.6441
21	$^{1}\chi$,$^{2}\chi^{v}$	0.3419	0.9596	104.5948	2.8067
22	$^{0}\chi^{v}$, $^{2}\chi^{v}$	0.3886	0.9475	78.9457	2.4382
23	$^{0}\chi$,$^{1}\chi$	0.3506	0.9574	99.0251	2.7308
24	$^{0}\chi$,$^{2}\chi^{v}$	0.3926	0.9463	77.1692	2.4103
25	$^{0}\chi$,$^{2}\chi^{v}$	0.4041	0.9430	72.3181	2.3336

The physical significance of the models expressed by eqs 7–9 is the same as discussed above.

We now discuss the predictive power of the proposed models. In doing so we have used Pogliani's quality factor Q.[39,40]

Among the one variable regressions (Table 3), the highest value for Q is found for the model containing $^{1}\chi^{v}$ as the correlating parameter. This model, therefore, has highest predictive power also.

In case of two variable regression models containing logP as one of the correlating parameters, the highest Q value is observed for the model based on $^{1}\chi$ and logP. The second best model is the one containing $^{1}\chi^{v}$ and logP. Hence, in addition to logP, $^{1}\chi$ and $^{1}\chi^{v}$ play dominating role in deciding predictive power. The model containing $^{2}\chi$ and logP is the worsed in this respect.

Though the two variable models containing topological indices were of less importance compared to those in which logP is involved, we nevertheless found that the model based on $^{1}\chi$ and $^{2}\chi^{v}$ has the highest-predictive power.

We, therefore, observed that Pogliani's quality factor Q is quite useful in the present case. However, its use is highly criticized.[41] Consequently, we have used 'leave-one-out' cross-validation method[38,42–44] to discuss predictive potential of the models. Since the methodology is based on 'leave-one-out' validation there is no need to consider the learning and test sets. The cross-validated parameter, for the best model under a particular category are given in Table 6.

PRESS (predictive residual sum of squares) $\ll$SSY (sum of squares of deviation for each activity) and the ratio PRESS/SSY < 0.1(Table 6) indicates that all the models are statistically excellent and have also excellent predictive power. The predictive correlation coefficient i.e., cross-validated r^{2}_{cv} (Table 6) also favours this finding. All the three cross-validated parameters taken together are in favour of models 13 and 16.

The parameter S_{PRESS} is the uncertainty in prediction. However, it is of no use in the present case as it is similar to standard error of estimation (Se). In such cases predictive square error PSE is used to account for uncertainty of prediction; the lowest value of which is in favour of models 13 and 16 (Table 4).

Inspite of the success of cross-validation methodology it was argued that the selected model- a combination of descriptors, could be examined for prediction of the activity of a set of compounds which are not in the learning set (the set used for building the model). According to them a good prediction would be the best proof for the reliability of the model, much better than using some statistical factors. In view of this and in view of the data set (22 compounds) we have chosen every fourth compound that is, 1,5,9,13 and 17 as a member of the training set and examined the predictive power of

Table 6. Cross-validated parameters for the best models under a particular category

Model No.	Parameter	PRESS	SSY	PRESS/SSY	r^2_{cv}	S_{PRESS}	PSE
11	PI,logP	1.8270	24.7295	0.0739	0.9761	0.3186	0.2950
13	$^1\chi$,logP	0.9449	25.1158	0.0376	0.9624	0.2291	0.2121
16	$^1\chi^v$,logP	0.9479	25.6086	0.0370	0.9630	0.2295	0.2125
21	$^1\chi$,$^2\chi^v$	2.1040	24.4524	0.0860	0.9140	0.3420	0.3165

Table 7. Results for the test set (compounds **1**,**5**,**9**,**13**,**17**) for modeling log 1/c (tadpole activity)

Topological index	Se	R	F	Q
$^1\chi$	0.9195	0.9327	66.692	1.0144
$^1\chi^v$	0.9195	0.9929	169.427	1.0798
PI	0.9195	0.9328	66.699	1.0145
logP	0.9195	0.8801	30.436	0.9572

this training set (compounds different from the learning set) and observed that the proposed model is excellent statistically as well as have excellent predictivity (Table 7).

3. Conclusions

From the results and discussion made above we conclude that tadpole narcosis of the heterogenous set of compounds can be model excellently even in one variable regression and that $^1\chi^v$ is best for this purpose and is superior to PI index. However, PI index gave better results than $^0\chi$, $^2\chi$ and logP. In multiparametric models the connectivity indices play dominating role in accounting tadpole narcosis. Although use of topological indices is quite useful for modeling Tadpole narcosis, the use of logP enhances the quality of the models based on topological indices.

4. Experimental

Tadpole narcosis activity—The tadpole narcosis activity (log 1/c) as well as hydrophobicity (logP) were adopted from the literatures.[3]

Topological indices—All the topological indices used in the present study are calculated from hydrogen-suppressed molecular graphs obtained by deleting all the carbon–hydrogen and heteroatom-hydrogen bonds from the molecular structures. The details of the calculations of topological indices are available in the literature. However, below we give expression for their calculations.

Wiener index (W)[12]—The Wiener index $W = W(G)$ of a graph G is defined as the half-sum of the elements of the distance matrix:

$$W = W(G) = 1/2\sum_{i=1}\sum_{j=1}(D)_{ij} \tag{10}$$

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.

PI (Padmakar-Ivan) index[5–7]—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:

$$PI = PI(G) = \sum_e (n_{eu} + n_{ev}) \tag{11}$$

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 ($^0\chi$, $^1\chi$, $^2\chi$)[8,9]—The general expression used for the calculation of Randic connectivity indices ($^0\chi$, $^1\chi$, $^2\chi$) is:

$$^m\chi^R = \sum_{(i,j,k)} (\delta_i, \delta_j, \delta_k)^{-0.5} \tag{12}$$

Kier-Hall valence connectivity index ($^0\chi^v$, $^1\chi^v$, $^2\chi^v$)[30,31]—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:

$$^m\chi^v = {}^m\chi^v(G) = \sum_{(i,j,k)} \left(\delta_i^v, \delta_j^v, \delta_k^v\right)^{-0.5} \tag{13}$$

where δ^v the valence delta values given by:

$$\delta_i^v = \frac{Z_i^v - H_i}{Z_i - Z_{j-1}} \tag{14}$$

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.

Pogliani Q Quality index—The Pogliani quality index Q is defined as: Q = R/S, where R is the multiple correlation cefficient and S the standard deviation, defined as.

$$S = (RSS/n - p)^{1/2} \tag{15}$$

where RSS is the residual sum of squares, n is the total number of objects and p is the total number of model parameters.

Regression analysis[38-40]—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

Authors are thankful to the referees for their excellent suggestions which has improved the quality of the paper. They are also thankful to Prof. Istvan Lukovits for providing the softwares. One of the authors (PVK) is thankful to Prof. Ivan Gutman for introducing him to the fascinating field: Chemical Topology and Graph Theory.

References and notes

1. Abrahm, M. H.; Rafols, C. J. *J. Chem. Soc. Perkin Trans* **1995**, 1843 and references therein..
2. Hansch, C.; Leo, L. *Exploring QSAR Fundamentals and Applicability in Chemistry and Biology*; American Chemical Society: Washington DC, 1996.
3. Chaturvedi, S. K. *QSAR Studies on Non-Polar Narcosis*, Ph. D. Thesis, A. P. S. Univ; Rewa, India 2001 and references therein.
4. Agrawal, V. K.; Chaturvedi, S.; Abrahm, M. H.; Khadikar, P. V. *Bioorg. Med. Chem.,* in press.
5. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 934.
6. Khadikar, P. V.; Karmarkar, S.; Varma, R. G. *Acta. Chem. Slov.* **2002**, *49*, 755.
7. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2001**, *29*, 134.
8. Randic, M. J. *Am. Chem. Soc.* **1975**, *97*, 6609.
9. Randic, M. J. *Mol. Graphis & Modelling* **2001**, *20*, 19.
10. Gutman, I.; Polansky, O. E. *Mathematrical Concepts in Organic Chemistry*; Springer-Verlag: Berlin, 1986.
11. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. J. *Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
12. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
13. Khadikar, P. V.; Joshi, S.; Bajaj, A. V.; Mandloi, D. *Bioorg. Med. Chem. Lett*., in press.
14. Jaiswal, M.;Khadikar, P. V. *Bioorg. Med. Chem*., in press.
15. Khadikar, P. V.; Singh, S.; Mandloi, D.; Joshi, S.; Bajaj, A. V. *Bioorg. Med. Chem.* **2003**, *11*, 5045.
16. Khadikar, P. V.; Mandloi, D.; Bajaj, A. V.; Joshi, S. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 419.
17. Khadikar, P. V.; Phadnis, A.; Shrivastava, A. *Bioorg. Med. Chem.* **2002**, *10*, 1181.
18. Khadikar, P. V.; Singh, S.; Shrivastava, A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1125.
19. Khadikar, P. V.; Karmarkar, S.; Singh, S.; Shrivastava, A. *Bioorg. Med. Chem.* **2002**, *10*, 3163.
20. Khadikar, P. V.; Agrawal, V. K.; Karmarkar, S. *Bioorg. Med. Chem.* **2002**, *10*, 3499.
21. Khadikar, P. V.; Bajaj, A. V.; Mandloi, D. *Indian. J. Chem.* **2002**, *41A*, 2065.
22. Singh, S.; Joshi, S.; Shrivastava, A.; Khadikar, P. V. *J. Sci. & Ind. Res.* **2002**, *61*, 961.
23. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2001**, *29*, 143.
24. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *Nat. Acad. Sci. Lett.* **2000**, *23*, 165.
25. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *Nat. Acad. Sci. Lett.* **2000**, *23*, 124.
26. Khadikar, P. V. *Nat. Acad. Sci. Lett.* **2000**, *23*, 113.
27. Khadikar, P. V.; Mandloi, D.; Bajaj, A. V. *Oxid. Commun.,* in press.
28. Khadikar, P. V.; Mandloi, D.; Bajaj, A. V. *Oxid. Commun.,* in press.
29. Hansch, C.; Kurup, A.; Garg, R.; Gao, H. *Chem. Rev.* **2001**, *101*, 619.
30. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity–Relationship*; Wiley: New York, 1986.
31. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
32. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, 2000.
33. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; J Wiley & Sons: New York, 2000.
34. Devillers, J., Balaban, A. T., Eds. *Topological Indices and Related Descriptors in QSAR and QSPR*; Garden & Breach: Amsterdam, The Netherlands, 1999.
35. Diudea, M. V.; Khadikar, P. V. *Molecular Topology and its Applications*; New Age Int: New Delhi, 2003; in press.
36. Diudea, M. V.; Ivancine, O. *Molecular Topology*, Complex: Cluj **1995**.
37. Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 23.
38. Chatterjee, S.; Hadi, A.; Price, B. *Regression Analysis by Examples*, 3rd ed; Wiley VCH: New York, 2000.
39. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
40. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
41. Todeschini, R. Some observations About the Pogliani Q Quality Factor, *Chemometrics Web* News, Milano Chemometrics & QSAR Research Group, File: ///CI/WINDOWS /DESKTOP/Web News on Chemometrics, html.
42. Khadikar, P. V.; Lukovits, I.; Agrawal, V. K.; Shrivastava, S.; Jaiswal, M.; Gutman, I.; Shrivastava, A.; Indian, J. *Chem.* **2003**, *42A*, 1436.
43. Agrawal, V. K.; Sharma, R.; Khadikar, P. V. *Bioorg. Med. Chem.* **2002**, *10*, 357.
44. Agrawal, V. K.; Sharma, R.; Khadikar, P. V. *Bioorg. Med. Chem.* **2002**, *10*, 2965.