Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 3035–3040

BIOORGANIC & MEDICINAL CHEMISTRY

QSAR Prediction of Toxicity of Nitrobenzenes

V.K. Agrawal[a] and P.V. Khadikar[b,*]

[a]*QSAR & Chemical Laboratories, A. P. S. University, Rewa-486 003, India*
[b]*Research Division, Laxmi Pest and Fumigation Pvt. Ltd., 3, Khatipura, Indore 452 007, India*

Received 9 April 2001; accepted 16 June 2001

Abstract—A QSAR analysis has been carried out on the toxicities of 40 mono-substituted nitrobenzenes using recently introduced PI and Sz indices, as well as older molecular redundancy (MRI) and Balaban indices (J). The results have shown that no statistically significant mono-parametric QSAR models are possible. Also, that along with PI, Sz, MRI and J indices are the appropriate parameters to be used in developing multiparametric QSAR models. The toxicities of nitrobenzenes are well predicted by a penta-parametric model consisting of PI, Sz, J, MRI and Ip_1 (an indicator parameter taking care of the effect of substitution at 2-position) as the correlating parameters. The predictive ability of the model is determined by a cross-validation method.

Introduction

Nitrobenzenes are widely used industrial chemicals and thus have a high potential for environmental pollution. These toxic substances are reported to be present in surface waters; and regarded as pro-electrophiles yielding the corresponding potentially highly toxic C-nitroso-compounds. The toxicity of nitrobenzenes has been quite extensively examined, and they have been the subject of number of quantitative structure–activity (QSAR) relationship.[1]

QSARs have provided a valuable approach in research into the toxicity of organic chemicals. Many investigators have used the 1-octanol/water partition coefficient (log P)-dependent QSAR as a basis for predicting the toxicity. Recently, some QSAR studies applying theoretical approaches to predict the toxicity have been reported.[2,3] In all of these studies, however, the reported models suffer from the lack of generacity. This is due to the fact that the correlations were based on a small number of compounds on the one hand or were not sufficiently validated on the other hand.

Among the different approaches employing computational chemistry, those based on chemical graph theory have been useful in establishing QSAR.[4–8] There are, however, difficulties in the use of these approaches when the QSARs are derived by means of multiple linear regression (MLR) techniques. Commonly, the topological descriptors are mutually inter-related by simple or multiple correlations and, therefore, fortuitous or artifactual MLR models may even be obtained. This difficulty in the use of the MLR method can be overcome through the cross-validation method; the probability of obtaining chance correlation is kept to minimum.[9–11]

In the present study, the MLR and cross-validation methods are used for modeling the toxicity, log ($1/IGC_{50}$) of 40 nitrobenzenes (Table 1), the toxicity of which were previously analyzed by Dearden et al.[1]

The proposed QSAR models were based on molecular descriptors (topological indices) that can be calculated for any compound utilizing only the knowledge of its molecular structure (molecular graph). The model included several topological constitutional descriptors: Szeged index, Sz,[12,13] Padmakar–Ivan index, PI,[14–16] Balaban index, J,[17] molecular redundancy index, MRI[18] along with three indicator parameters (Ip_1, Ip_2, Ip_3) related to substitution at 2-, 3-, and 4- positions respectively. The predictive ability of the models were assisted by the cross-validation method.[9,11]

The main objective of this work is to obtain QSAR models that could be used to predict toxicity of nitrobenzenes.

Results and Discussion

A set of 40 nitrobenzenes (I) used in the present investigation are recorded in Table 1, along with their

*Corresponding author. E-mail: vijay-agrawal@lycos.com

Abbreviations: QSAR, quantitative structure–activity relationship; Sz, Szeged index; PI, Padmakar–Ivan index; J, Balaban index; MRI, molecular redundancy index; IGC_{50}, 2-day (i.e., eight generations) static 50% inhibitory growth concentrations

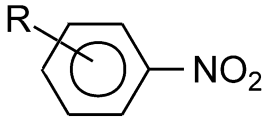
0968-0896/01/$ - see front matter
PII: S0968-0896(01)00211-5

Table 1. Nitrobenzenes, their molecular descriptors and toxicity

Compd no.	Substituent R	log 1/IGC$_{50}$	Ip$_1$	Ip$_2$	Ip$_3$	PI	Sz	J	MRI
1	H	0.350	0	0	0	66	142	2.2275	0.4735
2	2-NH_2	0.077	1	0	0	126	278	2.4000	0.3439
3	2-OH	0.770	1	0	0	94	228	2.4513	0.3074
4	2-CH_3	0.479	1	0	0	84	180	2.3945	0.3279
5	2-Cl	0.676	1	0	0	84	180	2.3945	0.3833
6	2-Br	0.863	1	0	0	84	180	2.3945	0.3833
7	2-CH_2OH	−0.155	1	0	0	126	287	2.4185	0.2667
8	2-C_6H_5	1.301	1	0	0	228	586	1.9566	0.5552
9	2-$CONH_2$	−0.721	1	0	0	176	404	2.6166	0.2667
10	2-NO_2	1.252	1	0	0	126	278	2.4000	0.4376
11	3-NH_2	0.026	0	1	0	126	278	2.0640	0.3439
12	3-OH	0.506	0	1	0	94	240	2.3270	0.3074
13	3-CH_3	0.572	0	1	0	84	192	2.2720	0.3279
14	3-Cl	0.836	0	1	0	84	192	2.2720	0.3833
15	3-Br	10.215	0	1	0	84	192	2.2720	0.3833
16	3-CN	0.451	0	1	0	104	240	2.3270	0.3074
17	3-CH_3OH	−0.220	0	1	0	126	305	2.2932	0.2667
18	3-C_6H_5	1.569	0	1	0	228	622	2.1555	0.5552
19	3-$CONH_2$	−0.193	0	1	0	176	434	2.4647	0.2667
20	3-CHO	0.140	0	1	0	104	240	2.3270	0.2814
21	3-NO_2	0.762	0	1	0	126	296	2.0640	0.4376
22	3-OCH_3	0.670	0	1	0	104	240	2.3270	0.3035
23	4-CH_3	0.796	0	0	1	84	192	2.4470	0.3279
24	4-C_2H_5	0.804	0	0	1	104	252	2.2401	0.2866
25	4-OCH_3	0.544	0	0	1	104	252	2.2401	0.3035
26	4-OC_2H_5	0.829	0	0	1	126	323	2.1948	0.3599
27	4-OC_4H_9	1.420	0	0	1	176	502	2.0888	0.1932
28	4-F	0.253	0	0	1	84	192	2.4470	0.3833
29	4-Cl	0.559	0	0	1	84	192	2.4470	0.3833
30	4-Br	0.461	0	0	1	84	192	2.4470	0.3833
31	4-CH_2CN	0.132	0	0	1	126	323	2.1948	0.2667
32	4-CH_2Cl	1.180	0	0	1	104	252	2.2401	0.2882
33	4-CH=NOH	0.678	0	0	1	150	406	2.1515	0.2400
34	4-NHC_6H_5	1.886	0	0	1	294	876	1.7814	0.3748
35	4-CH_2OH	0.101	0	0	1	126	323	2.1948	0.2667
36	4-$COOCH_3$	0.398	0	0	1	138	388	2.2958	0.2635
37	4-$COOC_2H_5$	0.710	0	0	1	176	475	2.2722	0.2381
38	4-$CONH_2$	−0.179	0	0	1	176	464	2.3478	0.2667
39	4-CHO	0.203	0	0	1	104	252	2.2401	0.2814
40	4-NO_2	1.301	0	0	1	126	314	2.2930	0.4376

log 1/ IGC$_{50}$, toxicity of nitrobenzenes; Ip$_1$, Ip$_2$, Ip$_3$, indicator parameters accounting for substitution at positions-2, -3, and –4, respectively; PI, Padmakar–Ivan index; Sz, Szeged index; J, Balaban index; MRI, molecular redundancy index.

toxicities, log (1/IGC$_{50}$), and calculated values of Sz, PI, J, and MRI indices. The adopted indicator parameters ((Ip$_1$, Ip$_2$, Ip$_3$) are also recorded in Table 1.



I

In the statistical process, the compound **1** was found as an outlier and was deleted from the original 40-compounds set. None of the descriptors used in the mono-parametric analysis can account for the reported log (1/IGC$_{50}$) values. The reason for this is not evident but it may be attributed to intervention of metabolic factors.

In view of the failure of obtaining a simple regression model, we have attempted MLR analysis. Before a multivariate analysis is undertaken, it is convenient to tailor the data in certain ways to make the calculations easier. Normally, it is sufficient to pre-process the data by means of auto-scaling and mean-centering the variables.[9] Auto-scaling,[9] gives each variable unit variance and, hence, the same chance to contribute to a calculated model, whereas mean scaling facilitates interpretation. The best way to achieve this is to obtain a correlation matrix (Table 2). Thus, all the variables used were initially auto-scaled to zero mean and unit variance to give each descriptor equal importance in the MLR analysis. The statistical significance of the screened models was judged by the multiple correlation coefficient (R), standard error of estimation (Se), adjusted R-squared (R^2_A), the F statistics and quality factor (Q). The quality factor (Q) is defined[19,20] as the ratio of R to Se (R/Se).

The predictive ability was evaluated by cross-validation coefficient (R^2_{CV}), which is based on the predictive error sum of squares (PRESS).

The MLR analysis for remaining 39 compounds resulted in the most significant five parametric model (Table 3)

Table 2. Correlation matrix showing the correlation of molecular descriptors and toxicity of nitrobenzenes

	$\log 1/IGC_{50}$	Ip_1	Ip_2	Ip_3	PI	Sz	J	MRI
$\log 1/IGC_{50}$	1.0000							
Ip_1	−0.1415	1.0000						
Ip_2	−0.0074	−0.3873	1.0000					
Ip_3	0.1272	−0.4851	−0.6216	1.0000				
PI	0.2580	−0.0211	−0.0644	0.0791	1.0000			
Sz	0.3292	−0.0924	−0.0877	0.1619	0.9894	1.0000		
J	−0.3584	0.3486	−0.1729	−0.1319	−0.5485	−0.5862	1.0000	
MRI	0.6597	0.1085	0.1002	−0.1874	0.2176	0.2030	−0.3085	1.0000

For symbols see Table 1.

Table 3. Proposed statistically significant models for predicting toxicity of nitrobenzenes

(1)	$\log 1/IGC_{50} = -0.0417\ (\pm 0.0062)PI + 0.0141(\pm 0.0020)Sz + 4.7509(\pm 0.5496)MRI\ -0.2084$
(2)	$\log 1/IGC_{50} = -0.0384\ (\pm 0.0064)PI + 0.0128(\pm 0.0022)Sz - 0.5637(0.03459)J + 4.4785(\pm 0.5905)MRI\ + 1.1924$
(3)	$\log 1/IGC_{50} = -0.0429\ (\pm 0.0069)PI + 0.0141(\pm 0.0023)Sz - 0.7443(\pm 0.3565)J\ + 4.3125(\pm 0.5865)MRI + 0.1833(\pm 0.1144)Ip_1 + 1.7469$
(4)	$\log 1/IGC_{50} = -0.0381\ (\pm 0.0067)PI + 0.0126\ (\pm 0.0023)Sz - 0.5884(\pm 0.3705)J + 4.4799(\pm 0.5990)MRI - 0.0195(\pm 0.0941)Ip_2 + 1.2531$
(5)	$\log 1/IGC_{50} = -0.0433\ (\pm 0.0077)PI + 0.0144\ (\pm 0.0026)Sz - 0.5309(\pm 0.3454)J + 4.3641(\pm 0.5961)MRI - 0.1176(\pm 0.1024)Ip_3 + 1.3003$

For symbols see Table 1.

with the following statistics: $Se = 0.2475$; $R_A^2 = 0.7714$; $R = 0.8952$; $F = 26.645$, and $Q = 3.6170$. The model accounted for about 80% of the variance in $\log(1/IGC_{50})$ according to following regression equation (model):

$$\log(1/IGC_{50}) = -0.0429\ (\pm 0.0069)PI + 0.0141 \times (\pm 0.0023)Sz - 0.7443(\pm 0.3565)J + 4.315(\pm 0.5865)MRI + 0.1833 \times (\pm 0.1144)Ip_1 + 1.7469 \quad (1)$$

The positive sign of the indicator parameter Ip_1 suggests that the toxicity is dependent on the substitution at the 2-position and that substituents at the 2-position play a dominant role in deciding mechanism of toxic effect of nitrobenzenes used. Furthermore, the molecular descriptor Sz also has a positive coefficient in eq (1). This topological index accounts for the shape and size of the molecule. Hence, the presence of bulky groups at the 2-position accounts for the enhancement of toxicity. Also the positive coefficient of MRI indicates that the increase in the information content increases the toxicity. On the other hand, the negative signs associated with the coefficients of PI and J indicates their negative contribution in the exhibition of toxicity.

The indicator parameter Ip_2, used for accounting toxic effect of substitution at the 3-position, was not found to be statistically significant (Table 3). The same is found to be the case for the toxic effect of substitution at the 4-position (Ip_3). However, the data presented in Table 1 show that most compounds with high toxicity are 3- and 4-substituents within the same substituents. It is worth recording that the proposed model must be able to describe structural or biological meaning, not just mathematical correlation.

The aforementioned results, therefore, indicate that 2-, 3-, and 4-substituted nitrobenzenes have different mechanisms of action. The parameters responsible for their different mechanisms are molecular size, shape, complexity connectivities, information content as reflected by Sz, MRI, J and PI indices. We postulate that 4-substitution hinders binding to the same receptor site as that involved with the 2-, 3-derivatives, thereby causing other mechanisms to come into play. However, further study with a wide range of compounds and the use of a variety of other molecular descriptors would be needed to elucidate the mechanisms of toxicity. This, however, is not the objective of the present study.

It is interesting to record that PI and Sz are highly auto-correlated (0.9894, Table 2). This means that the proposed model [eq (1)], as given above, should suffer from the defect of collinearity. In fact, all the proposed models presented in Table 3 should suffer from collinearity defects. However, in such cases, R_A^2 values will help the interpretation of the proposed models, as it takes into account the adjustment of R^2.

If a variable is added that does not contributes its fare share, then R_A^2 will actually decline. However, in the proposed models, R_A^2 goes on increasing from 0.7495 to 0.7714, indicating that the collinearity defect is not serious. Furthermore, the coefficients of both PI and Sz terms in all the proposed models are higher than their corresponding standard deviations; such parameters are considered statistically significant. Hence, the high collinearity between Sz and PI indices does not make the correlation involving them an insignificant correlation. This is due to fact that, in spite of high auto-correlation, the information content in Sz and PI indices is different, making their occurrence a statistically allowed model.

One will think that just increasing the value of R_A^2 may not solve the problem of multi-collinearity. When we

Table 4. Regression parameters and quality of correlations of the proposed models (refer Table 3)

Model no.	Parameters used	Se	R^2_A	R	F	Q
1	PI, Sz, MRI	0.2591	0.7495	0.8771	38.907	3.3852
2	PI, Sz, J, MRI	0.2532	0.7609	0.8856	31.225	3.4976
3	PI, Sz, J, MRI, Ip_1	0.2475	0.7714	0.8952	36.645	3.6170
4	PI, Sz, J, MRI, Ip_2	0.2568	0.7539	0.8867	24.286	3.4529
5	PI, Sz, J, MRI, Ip_3	0.2520	0.7631	0.8912	35.479	3.5635

tried to add and analyze only PI except Sz to other descriptors or vice versa we obtained similar results as mentioned above. Also, a high correlation between PI and Sz made the correlation coefficient increase.

The observed high collinearity among PI, Sz, MRI, and J indicates that the new PI index duplicates much of the information content in Sz, MRI, and J. Randic[21] has stated that one should be particularly aware of a common fit in regression analysis in describing descriptors that are highly inter-correlated. He further stated that, by discarding one of the descriptors which commonly duplicates another, we may be discarding a descriptor that nevertheless may carry useful structural information in the parts in which it does not parallel with the another descriptors. Thus, following Randic,[21] we may safely say that the PI index caries the usual structural information not present in other molecular descriptors in that PI exhibited high multi-collinearity. A detailed study on this aspect is underway.

It is worth mentioning that, when Ip_2 is added to a tetra-parametric model (Table 4) containing PI, Sz, J, and MRI, the R^2_A (0.7609) declines to 0.7539 indicating that the added parameter (Ip_2) behave differently giving different mechanism of action. Further, here the coefficient of Ip_2 is much less than its standard error, making it statistically insignificant.

An addition of Ip_3 to the aforementioned tetra-parametric model results in increase in both R^2_A and R as well as decrease in the standard error of estimation (Se). Thus, this resulting penta-parametric model is definitely better than the tetra-parametric model mentioned above. However, this new model is statistically less significant than the model proposed by eq (1).

On application of the cross-validation method, we obtained cross-validated parameters for the proposed model [eq (1)] as: $R^2_{CV} = 0.7523$, $S_{PRESS} = 0.2475$, $PSE = 0.2277$, $PRESS/SSY = 0.2477$.

The ratio PRESS/SSY can be used to calculate the approximate confidence interval of prediction. To be a reasonable QSAR model, this ratio should be smaller than 0.4, and the value of this ratio smaller than 0.1 indicates an excellent model.[9] The observed value of PSE (0.2477) proposes that the model expressed by eq (1) is an excellent model.

It is worth mentioning that both Se and S_{PRESS} have similar values. Therefore, they cannot be used as parameters describing the uncertainty of prediction. We have, therefore, used predictive squared error (PSE) to account for the uncertainty of the prediction of the model proposed by us [eq. (1)]. The value of $PSE = 0.2277$ indicates that compared to Se and S_{PRESS} is directly related to the uncertainty of predictions.

Finally, using the model expressed by eq (1), we have calculated the toxicities of nitrobenzenes and compared them with their experimental values. The agreement between the observed (experimental) and calculated (estimated) values is very satisfactory, as indicated by predictive $R^2_{Predictive} = 0.802$ (90%) (Fig. 1).

Conclusion

The aforementioned results and discussion lead us to conclude that distance-based topological indices can be used successfully for modeling, monitoring and estimating toxicity of nitrobenzenes and related compounds.

Experimental

Toxicity data

Toxicities, for example 2-day (eight-generations) static 50% inhibitory growth concentration (IGC_{50}) for the set of 40 nitrobenzenes used were taken from the studies reported by Dearden et al.[1] in that toxicities against the aquatic ciliate *Tetrahymena pyriformis*, strain GL-C, were measured. Axenic cultures of the ciliate were used

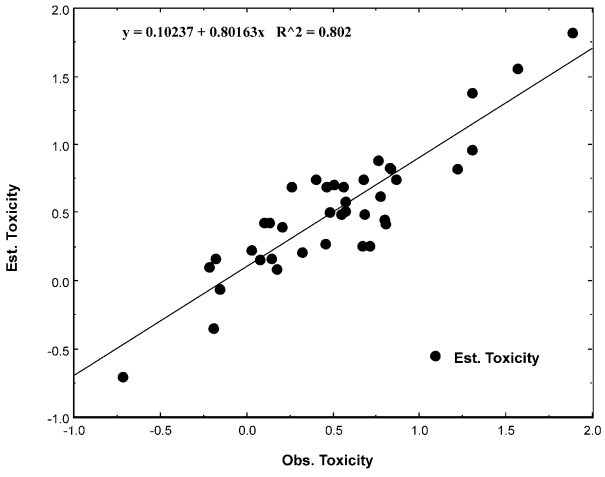


Figure 1. Correlation of observed versus estimated toxicities of nitrobenzenes.

and 2-day (i.e., 8–9 generations) static 50% inhibitory growth concentrations (IGC_{50}) were measured. Assessments were made in 250 mL Erlenmeyer flasks containing 50 mL of proteose peptone-based medium. Each assay was set up in triplicate and compared in at least a five-step concentration series. Cultures without test chemicals served as controls.

Molecular descriptors (topological indices)

The following indices based on molecular topology were considered: Szeged index, Sz,[12,13] Padmakar–Ivan index, PI,[14–16] Balaban index, J,[17] and molecular redundancy index, MRI.[18] In addition, three indicator parameters (Ip_1, Ip_2, Ip_3) related to substitution at the 2-, 3-, and 4-positions, respectively, were also used.

Statistical methods

The MLR and cross-validation methods were employed to search for relationships between the toxicity, log ($1/IGC_{50}$), and the aforementioned structural descriptors. The predictive ability of the proposed models is finally confirmed using the cross-validation method.[9]

Computer software, as supplied by Professor Lukovits, Hungarian Academy of Sciences, Budapest, Hungary were used for the calculations of Sz, and J, and for making MLR analysis.

Calculations of molecular descriptors

Szeged index (Sz).[12,13] 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 as

$$\mathrm{Sz}(G) = \mathrm{Sz} = \sum_e n_1(e|G) n_2(e|G)$$

with the summation giving over all 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 index.[14–16] Let G be a simple molecular graph without directed and multiple edges and without loops, the vertex and edge-shapes of which are represented by $V(G)$ and $E(G)$, respectively. If e is an edge of G, connecting the vertices u and v then we write $e = uv$. The number of vertices of G is denoted by n.

The distance between a pair of vertices u, w of G is denoted by $\mathrm{d}(u, w|G)$.

Let e be an edge of G connecting the vertices u and v, $e = uv \in E(G)$.

We define for $e = uv$ two quantities $n_{eu}(e|G)$ and $n_{ev}(e|G)$. $n_{eu}(e|G)$ is the number of edges lying closer to the vertex u than the vertex v, and $n_{ev}(e|G)$ is the number of edges lying closer to the vertex v than the vertex u. Edges equidistant from both ends of the edge uv are not counted.

The PI index, $\mathrm{PI} = \mathrm{PI(G)}$ of the graph G is defined as:

$$\mathrm{PI} = \mathrm{PI(G)} = \sum_{e \in E(G)} [n_{eu}(e|G) + n_{ev}(e|G)]$$

The summation goes over all edges of G.

Balaban index (J)..[17] The Balaban index,[17] $J = J(G)$, was introduced by Balaban in 1982 as the average distance sum connectivity index. It is defined as:

$$J = \frac{M}{\mu + 1} \sum_{\text{all edges}} (d_i \; d_j)^{-0.5}$$

where M is the number of edges of G; μ is the cyclomatic number of G; and d_i is the distance sum where $i = 1,2,3 \ldots N$. The cyclomatic number = the minimum number of edges that must be removed from G in order to transform it to the related acyclic graph.

Molecular redundancy index (MRI).[18] The MRI is derived from information theory and molecular graph theory and is defined as:

$$\mathrm{MRI} = \frac{\sum n_i \log n_i}{N \log N}$$

where n_i is the number of atoms of the same kind in the ith atom set, i is the number of different atoms in the molecule.

Indicator parameters. Indicator variables (parameters), sometimes called dummy variables or de novo constants, are used in linear multiple regression analysis to account for certain features which cannot be described by continuous variables. In QSAR equations they normally describe a certain structural element, be it a substituent or another molecular fragment.

In the present study, we have used three indicator parameters (variables), namely Ip_1, Ip_2, and Ip_3 which are taken as unity when substituents are present at the 2-, 3-, and 4-positions, respectively, otherwise their values are zero.

Acknowledgements

The authors are thankful to Professor I. Lukovits, Hungarian Academy of Sciences, Budapest, Hungary for providing Wiener-1 and Regress-1 Software for computation of some of the topological indices and for making regression analysis.

References and Notes

1. Dearden, J. C.; Cronin, M. T. D.; Schultz, T. W.; Lin, D. T. *Quant. Struct.-Act. Relat.* **1995**, *14*, 427.
2. Repetto, M. *Rev. Toxicol.* **1992**, *9*, 3.
3. Domenech, R. G.; Alvarez, J. G.; Anton, G. M.; Mari Buigues, J. I. *Rev. Toxicol.* **1992**, *9*, 15.

4. Wold, S. In *Chemometric Methods in Molecular Design*; van de Waterbeemd, H., Ed.; VCH: Weinheim, 1995; Vol. 2, p 195.
5. Luco, J. M.; Sosa, M. E.; Cesco, J. C.; Tonn, C. E.; Giordano, O. S. *Pestic. Sci.* **1994**, *41*, 1.
6. Trinajstic, N. *Chemical Graph Theory*, 2nd ed.; CRC: Boca Raton, FL, 1992.
7. Devillers, J.; Balaban, A. T. Eds. *Topological Indices and Related Descriptors in QSAR an QSPR*; Gorden and Breach Scientific: Amsterdam, 2000.
8. Devillers, J., Ed. *Comparative QSAR*; Tailer and Francis: Philadelphia, 1998.
9. Chaterjee, S.; Hadi, A. S.; Price, B. *Regression Analysis by Examples*, 3rd ed.; Wiley: New York, 2000.
10. Clark, M.; Cramer, R. D., III *Quant. Struct.-Act. Relat.* **1993**, *12*, 137.
11. Luco, J. M.; Ferretti, H. F. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 392.
12. Gutman, I. *Graph Theory Notes New York* **1994**, *27*, 9.
13. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 934.
14. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *Nat. Acad. Sci. Lett.* **2000**, *23*, 113.
15. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2001**, *29*, 143.
16. Khadikar, P. V. *Nat. Acad. Sci. Lett.* **2000**, *23*, 113.
17. Balaban, A. T. *Chem. Phys. Lett.* **1982**, *89*, 399.
18. Kier, L. B. *J. Pharm. Sci.* **1980**, *69*, 807.
19. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
20. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
21. Randic, M. *Croat. Chem. Acta* **1993**, *66*, 289.