Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 2477–2482
Bioorganic & Medicinal Chemistry

QSAR study on CA inhibitory activity of disulfonamides: effect of halogen substitution

Mona Jaiswal,[a] Padmakar V. Khadikar[a,*] and Claudiu T. Supuran[b]

[a]*Research Division, Laxmi Fumigation and Pest Control (P), 3, Khatipura, Indore 452007, India*
[b]*Universita degli Studi di Firenze, Laboratorio di Chiica Bioinorganica, Rm. 118, Via della Lastruccia 3, I-500 la Sesto Fiorentino (Firenze), Italy*

Received 25 November 2003; accepted 21 January 2004

Abstract—The paper deals with quantitative structure–activity relationship (QSAR) study on CA inhibitory activity ($\log IC_{50}$) of disulfonamides using a large series of distance-based topological indices. The study discusses effect due to halogen-substitution nearer (*o*-position) to $-SO_2NH_2$ groups. The results have shown that halogen substitution at R_3 has pronounced effect on the inhibitory activity. Predictive power of the proposed models is discussed on the basis of regression data and cross-validation parameters.

1. Introduction

The enzyme carbonic anhydrase (CA, carbonate hydrolase, EC 4.2.1.1.) is involved in a variety of physiological and physiopathological processes.[1] Among its inhibitors, sulfonamides are important clinical agents, used in the treatment of glaucoma, gastro-duodenal ulcers, certain neurological disorders, motion and altitude sickness etc.[1] One of us (CTS)[2,3] is very well recognized in the area of the CA inhibitors and has reviewed the potential use of sulfonamide CA inhibitors. Earlier, quantitative structure–activity relationship (QSAR) studies of sulfonamide CA inhibitors were mainly based on Hansch–Fujitu[4] and HOMO[5] methods using multiple regression analysis.

The study referred above[2,3] uses hydrophobic substituent constants and concerns with the investigation of the CA inhibitory ($\log IC_{50}$) activities of a series of 21 1,3-benzene-disulfonamides (Table 1, Fig. 1).

Table 1. Structural details of the disulfonamides used, their $\log IC_{50}$ activity and indicator parameter I_1 and I_2

S.No.	R_1	R_2	R_3	$\log IC_{50}$	I_1	I_2
1	H	H	H	−6.0000	0	0
2	Me	H	H	−7.4814	0	0
3	Et	H	H	−6.6020	0	0
4	*n*-Pr	H	H	−6.6020	0	0
5	F	H	H	−7.0000	1	0
6	Cl	H	H	−6.9586	1	0
7	Br	H	H	−6.7695	1	0
8	Me	Me	H	−7.3979	0	0
9	Me	H	Me	−7.6777	0	0
10	Me	H	F	−7.4437	0	1
11	Me	H	Cl	−7.6020	0	1
12	Me	Cl	H	−8.3010	0	0
13	Cl	Me	H	−7.8239	1	0
14	Me	H	Br	−7.6382	0	1
15	Et	H	Cl	−7.7447	0	1
16	*n*-Pr	H	Cl	−7.7447	0	1
17	Cl	Cl	H	−8.3010	1	0
18	Cl	H	Cl	−7.8239	1	1
19	Br	H	Br	−7.6020	1	1
20	Cl	H	F	−7.6020	1	1
21	Cl	H	Br	−7.5086	1	1

Using hydrophobic substituent constants π_i $(i = 1, 2, 3)$ of the group R_j $(j = 1, 2, 3)$ the following model was proposed for modeling CA inhibitory properties ($\log IC_{50}$):

Keywords: QSAR; CA inhibition; Disulfonamides; Topological indices; Regression analysis.
* Corresponding author. Tel.: +91-731-2531906; fax: +91-731-2701170; e-mail addresses: pvkhadikar@rediffmail.com; tomona@rediffmail.com

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.01.049

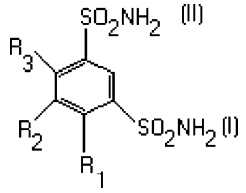

Figure 1. Disulfonamides used (for details see Table 1).

$$\log IC_{50} = 0.2125\pi_1 + 2.1814\pi_2 - 2.8731(\pi_3 - 0.6193)2 - 0.5674 \quad (1)$$
$$n = 19, \quad R = 0.9486, \quad F = 50.9414$$

The results indicated that substituents at R_1 and R_2 influence the inhibitory property ($\log IC_{50}$) in a positive manner while substituents at R_3 have negative influence. The negative influence of R_3 was attributed to the steric hindrance of the $-SO_2NH_2$ group.

The literature shows that till date no attempt has been made to investigate CA inhibitory activity $\log IC_{50}$ of the referred sulfonamides using distance-based topological indices. This was, therefore, primary objective of the present study. Careful examination of the structural details (Table 1) indicated that out of 20 substituted sulfonamides 15 sulfonamides contain halogen substitution. That is, in addition to the steric effect, halogen substitution also influences the activity. To investigate effect due to halogen substitution was the another objective of the present study. Thus, in continuation to our earlier study on sulfonamides,[6,7] we now report QSAR study and effect due to halogen substitution on carbonic anhydrase inhibitory activity of the disulfonamides (Table 1) using distance-based topological indices (Table 2).

Table 2. Distance-based topological indices (W, Sz, PI, $^0\chi$, $^1\chi$ and $^2\chi$) for the disulfonamides (Table 1)

S.No.	W	Sz	PI	$^0\chi$	$^1\chi$	$^2\chi$
1	289	420	176	10.9831	6.2103	7.2925
2	342	498	204	11.8534	6.6210	7.8258
3	410	591	234	12.5605	7.1590	8.0168
4	494	700	266	13.2676	7.6590	8.3973
5	342	498	204	11.8534	6.6210	7.8258
6	342	498	204	11.8534	6.6210	7.8258
7	342	498	204	11.8534	6.6210	7.8258
8	398	580	234	12.7236	7.0317	8.3433
9	399	582	234	12.7236	7.0317	8.3710
10	399	582	234	12.7236	7.0317	8.3710
11	399	582	234	12.7236	7.0317	8.3710
12	398	580	234	12.7236	7.0317	8.3433
13	398	580	234	12.7236	7.0317	8.3433
14	399	582	234	12.7236	7.0317	8.3710
15	472	682	266	13.4307	7.5697	8.5621
16	562	799	300	14.1378	8.0697	8.9425
17	398	580	234	12.7236	7.0317	8.3433
18	399	582	234	12.7236	7.0317	8.3710
19	399	582	234	12.7236	7.0317	8.3710
20	399	582	234	12.7236	7.0317	8.3710
21	399	582	234	12.7236	7.0317	8.3710

2. Results and discussion

The preliminary regression analysis has indicated that none of the distance-based topological indices, when used singly, are capable of modeling $\log IC_{50}$ activity for the entire set of 21 compounds. That is, no monoparametric regression equation could be obtainable. However, the correlation of the activity $\log IC_{50}$ with the topological indices indicated that the data set (21 compounds) split into two categories depending upon halogen substitution at R_1 (category I) and R_2 (category II), respectively. The results as discussed below indicated that excellent results are obtained under splitting of the data set. In both the cases we have used the parent compound ($R_1 = R_2 = R_3 = H$) in order to discuss the effect due to halogen substitution at R_1 and R_2, respectively, on the inhibitory activity ($\log IC_{50}$). The sulphonamides belonging to the category-I are: **1**, **5**, **6**, **7**, **13**, **17**, **18**, **19**, **20** and **21**; while the sulphonamides **1**, **10**, **11**, **14**, **15**, **16**, **18**, **19**, **20** and **21** belong to category II. Record that compounds **18**–**21**, have halosubstitution both at R_1 and R_3, while compounds **1**, **2**, **3**, **4**, **8** and **9** have no halogen at any of the positions R_1, R_2 and R_3. To account for presence/absence of halogen substitution at R_1 and R_3 we have used two indicator parameters I_1 and I_2, respectively. In presence of halogen they assume the value of 1 otherwise they are zero.

It is necessary to mention that the distance-based topological indices used are: Wiener (W)-,[8] Szeged (Sz)-,[9,10] Padmakar–Ivan (PI)-[11,13] and Randic connectivity ($^0\chi$, $^1\chi$, $^2\chi$)-indices.[12,13] The calculations of these indices are available in the literature.[11,13] These indices, alongwith indicator parameters mentioned above, are used for obtaining satisfactory significant models. Maximum R^2 method[14] is used for obtaining statistically significant models. It is interesting to record that the topological indices used are highly linearly correlated. One may, argue that the used descriptors are redundant that is, duplicates of each other. However, such cases were critically examined and discussed by Randic.[18] He stated that one should particularly be aware of a common fit in all regression analysis, when the descriptors that are highly interrelated. He further mentioned that by discarding one of the descriptors that duplicates another (in one domain) we may discard a descriptor that may carry useful structural information in another domain, in which it does not parallel with other descriptors. Bearing this in mind, we conclude that the highly correlated topological indices (redundant descriptors) may be used and retained during simple or multiple regression and that their simultaneous presence need not be unjustified. Furthermore, the topological indices used by us are 2D descriptors and not 3D descriptors. Consequently one may argue that the 3D geometry, confirmation effects or mechanism of action at the receptor site at a biological level is difficult to interpret. However, the splitting of the data set into categories I and II, as mentioned above and the excellent statistics obtain thereafter, possibly indicates that the sulfonamides belonging to these categories have different type of mechanism, which may be attributed to their differential 3D geometries.

Table 3. Regression parameters and quality of correlation for disulfonamides belonging to category I

Model No.	Parameter used	Se	R	F	Q
1	W	0.2711	−0.9392	142.153	3.4644
2	Sz	0.2428	−0.9382	124.503	3.8641
3	PI	0.2363	−0.9416	262.260	3.9848
4	$^0\chi$	0.2354	−0.9421	150.036	4.0021
5	$^1\chi$	0.2354	−0.9421	150.036	4.0021
6	$^2\chi$	0.2495	−0.9347	131.329	3.7463
7	W, I_2	0.1732	0.9730	62.1502	5.6178
8	Sz, I_2	0.1735	0.9729	61.9317	5.6075
9	PI, I_2	0.1725	0.9732	62.7274	5.6414
10	$^0\chi$, I_2	0.1768	0.9718	59.5184	5.4966
11	$^1\chi$, I_2	0.1768	0.9718	59.5244	5.4966
12	$^2\chi$, I_2	0.1791	0.9711	57.8932	5.4221

The results of stepwise regression on the disulfonamides belonging to category I are presented in Table 3. Inspection of this Table 3 shows that all the six simple linear regressions involving single descriptor resulted into statistically significant models. Also, that among W, Sz and PI indices (Wiener type indices), PI index is found slightly better than both W and Sz indices for modeling $\log IC_{50}$ activity.

The linear model for modeling CA inhibitory activity ($\log IC_{50}$) of disulfonamides using PI index is found below:

$$\begin{aligned} &\log IC_{50} = -0.0300(\pm 0.0038)PI - 0.7564 \\ &n = 10,\ \ Se = 0.2623,\ \ R = -0.9416, \\ &F = 262.260,\ \ Q = 3.9848 \end{aligned} \tag{2}$$

Here 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 in the literature[15,16] as the ratio of correlation coefficient (R) to the standard deviation (Se) that is, $Q = R/Se$ and is used for accounting the predictive power of the model.

It is worthy to mention that W index is applicable to acyclic graphs alone, while Sz is the modification of W for cyclic compounds. For acyclic graphs W and Sz coincide. However, for cyclic molecules containing tree-like (acyclic) chains the coincidence of W and Sz indices is well known. The PI (Padmakar–Ivan) index on the other hand is applicable to both acyclic, cyclic graphs as well as cyclic graphs with tree like (acyclic) side chains. Since, Sz and PI indices are W-type indices they mainly account for size, shape and branching. The negative sign associated with the correlation coefficients of W, Sz and PI indicate that size, shape and branching have negative effect on the exhibition of $\log IC_{50}$ activity.

The connectivity indices, in addition to size, shape and branching also accounts for the presence of heteroatom. The simple regressions based on single descriptor (Table 3) show that models based on $^0\chi$ and $^1\chi$ indices are more or less similar to the model based on PI index. Here, also negative dependence of $^0\chi$ and $^1\chi$ is observed. Both $^0\chi$ and $^1\chi$ indices convey information about the number of atoms in the molecules, which in our case has negative influence on the activity $\log IC_{50}$.

In order to improve results derived from the simple linear equation 2 and to investigate effect due to halogen substitution at R_1 and R_3, we have attempted two variable regressions in that one of the variable is either I_1 or I_2 and the other one is one of the distance-based topological index used. All the two variable regression containing I_1 as one of the correlating parameter suffers from the defect in that the coefficients of I_1 were considerably smaller than the standard deviations. Such models are not allowed statistically. However, two variable models containing I_2 as one of the correlating parameter gave excellent results (Table 3). This shows that halogen substitution at R_1 does not have appreciable influence in the exhibition of $\log IC_{50}$ activity.

The analysis of the data (Table 3) for the compounds belonging to category I reveal that all the two variable models are more or less similar, the model based on PI and I_2 gave slightly better results. This model is found as:

$$\begin{aligned} &\log IC_{50} = -0.0361(\pm 0.0035)PI + 0.3993(\pm 0.1410)I_2 \\ &\qquad + 0.4214 \\ &n = 10,\quad Se = 0.1725,\quad R = 0.9732, \\ &F = 62.7274,\quad Q = 5.6414 \end{aligned} \tag{3}$$

The positive sign associated with I_2 indicates that halogen substitution at R_3 has dramatic influence on the exhibition of activity, the correlation coefficient (R) increased from 0.9416 to 0.9732, while the standard error of estimation (Se) is lowered down from 0.2363 to 0.1725 when we passed from mono- to bi-parametric regression.

Table 4. Regression parameters and quality of correlation for disulfonamides belonging to category II

Model No.	Parameter used	Se	R	F	Q
13	W	0.4049	−0.6926	17.517	1.7106
14	Sz	0.3896	−0.7198	21.025	1.8475
15	PI	0.3631	−0.7625	29.664	2.0100
16	$^1\chi$	0.3817	−0.7332	22.081	1.9209
17	$^2\chi$	0.2221	−0.9184	183.854	4.1351
18	W, I_1	0.3681	0.7898	5.8013	2.1456
19	Sz, I_1	0.3527	0.8095	6.6535	2.2952
20	PI, I_1	0.3255	0.8401	8.3923	2.5810
21	$^0\chi$, I_1	0.2680	0.8947	14.0700	3.3384
22	$^1\chi$, I_1	0.3422	0.8214	7.2601	2.4004
23	$^2\chi$, I_1	0.2030	0.9410	27.0014	4.6355
24	W, I_2	0.1129	0.9821	95.3164	8.6989
25	Sz, I_2	0.1128	0.9822	95.6111	8.7075
26	PI, I_2	0.1125	0.9823	96.1370	8.7316
27	$^0\chi$, I_2	0.1123	0.9823	96.4554	8.7471
28	$^1\chi$, I_2	0.1121	0.9824	96.8257	8.7636
29	$^2\chi$, I_2	0.1146	0.9816	92.4641	8.5655

We now discuss QSAR study on disulfonamides belonging to category II. The results of one- and two-variable regression analysis are summarized in Table 4.

The perusal of Table 4 shows that negative correlation is observed in all the five cases. Among W, Sz and PI, once again the model based on PI index is found slightly better:

$$\begin{aligned}&\log IC_{50} = -0.0131(\pm 0.0039)PI - 4.3617\\ &n = 10, \quad Se = 0.3631, \quad R = -0.7625,\\ &F = 29.664, \quad Q = 2.0100\end{aligned} \tag{4}$$

However, unlike results obtained for category I, no statistically significant model using ${}^0\chi$ could be obtained and that ${}^2\chi$ gave excellent results:

$$\begin{aligned}&\log IC_{50} = -1.1839(\pm 0.1803)^2\chi + 2.4023\\ &n = 10, \quad Se = 0.2221, \quad R = -0.9184,\\ &F = 183.854, \quad Q = 4.1351\end{aligned} \tag{5}$$

The ${}^2\chi$ index encodes more information about branching and the negative sign associated with this parameter indicates its negative influence on the exhibition of the activity ($\log IC_{50}$).

Once again, unlike disulfonamides belonging to category I, here we obtained statistically significant two variable models in that one of the correlating parameter is the indicator parameter I_1 and the other one is one of the distance-based topological indices used (Table 4).

The data presented in Table 4 show that the quality of simple regression involving a particular topological index is dramatically improved by the introduction of I_1 and that two variable model containing ${}^2\chi$ and I_1 gave the best results:

$$\begin{aligned}&\log IC_{50} = -1.1664(\pm 0.1653)^2\chi - 0.2105(\pm 0.1314)I_1\\ &\qquad + 2.3403\\ &n = 10, \quad Se = 0.2030, \quad R = 0.9410,\\ &F = 27.0014, \quad Q = 4.6355\end{aligned} \tag{6}$$

This shows that increase in the extent of branching and presence of halogen at R_1 have negative influence on the exhibition of the activity.

Table 4 also shows that among the two variables regression based on W, Sz, PI, the regression based on PI and I_1 gave better results.

It is interesting to mention that simple regression involving ${}^0\chi$ was statistically insignificant, however, the combination ${}^0\chi$ and I_1 resulted into statistically significant model (Table 4).

All two variable regressions involving I_2 and one of the topological indices gave statistically excellent models (Table 4). All of them, more or less, have similar correlation potential. However, the two variable model containing ${}^2\chi$ and I_2 was comparatively less significant; and the model with ${}^1\chi$ and I_2 as the correlating parameters was slightly better:

$$\begin{aligned}&\log IC_{50} = -0.1593(\pm 0.1073)^1\chi - 1.4757(\pm 0.1593)I_2\\ &\qquad - 5.0107\\ &n = 10, \quad Se = 0.1121, \quad R = 0.9824,\\ &F = 96.8257, \quad Q = 8.7636\end{aligned} \tag{7}$$

This shows that for the disulfonamides belonging to category II, the number of atoms and halogen substitution at R_3 have retarding effect on the exhibition of the activity.

The above results show that the disulfonamides used in the present investigation have dual characteristics. They behave differently when they belong to category I or category II. This may be attributed to different types of mechanism for the sulfonamides belonging to the categories I and II.

Prompted by the encouraging results obtained above, we have finally attempted effect due to halogen-substitution considering all the 21 disulfonamides. Out of he several regressions attempted the tri-parameric regression containing ${}^1\chi$, PI and I_2 gave excellent results:

$$\begin{aligned}&\log IC_{50} = 11.4514(\pm 1.7713)^1\chi - 0.1874(0.0279)PI\\ &\qquad + 0.2918(\pm 0.1635)I_2 - 44.5688\\ &n = 21, \quad Se = 0.2917, \quad R = 0.8814,\\ &F = 19.7259, \quad Q = 3.0216\end{aligned} \tag{8}$$

The negative sign of the coefficient of PI in the above Eq. 8 is due to its high collinearity with ${}^1\chi$. The physical significance of the parameters involved are the same as discussed above.

From the above results an interesting conclusion can be made that halogen substitution at R_3 has a dominating effect on the exhibition of $\log IC_{50}$. Record that R_1 is at *ortho*-position to SO_2NH_2 (I), likewise R_3 is at *ortho*-position to $-SO_2NH_2$ (II). The latter *ortho* substitution plays a dominating role in the exhibition of $\log IC_{50}$.

Simply getting good statistics is not the best way to deal with 'biological studies' when no sound predictivity is possible. In view of this we now discuss the predictive power of the proposed models. Fortunately, a quality factor (Q) is available[15,16] for this purpose. Use of quality factor Q indicates that all the models under the referred condition have the best predictive power too. However, the use of Q for deciding predictive power is criticized recently.[17] In view of this we have used cross-validated method[14] and obtained various cross-validation parameters (Table 5). The meaning of the

Table 5. Cross-validation parameters for the best models

Mode (eq)	Parameter used	Category	PRESS	SSY	PRESS/SSY	r^2_{cv}	S_{PRESS}	PSE
9 (3)	PI, I_2	I	0.2802	3.7316	0.0558	0.9442	0.1725	0.1443
23 (6)	$^2\chi$, I_1	II	0.2886	2.9313	0.1293	0.8707	0.2030	0.1699
28 (7)	$^1\chi$, I_2	II	0.0879	2.4320	0.0361	0.9639	0.1140	0.0938

PRESS—predicted residual sum of squares; SSY—sum of the squares of the residual values; r^2_{cv}—cross-validated correlation coefficient; S_{PRESS}—uncertainty of prediction; PSE—predictive squared error.

cross-validation parameters used is given in the footnote of Table 5.

A perusal of Table 5 shows that in all the three cases PRESS<<SSY and the ratio of PRESS/SSY is smaller than 4; and the same for models 9 (Eq. 3) and 28 (Eq. 7) is very much smaller than 0.1. All these results indicate that the models are free from chance and have excellent predictive power. The values of predictive correlation coefficients (r^2_{cv}) supports this finding.

Another important cross-validation parameter is S_{PRESS} that is, uncertainty of prediction. However, this parameter is of no use in the present case. The reason being S_{PRESS} coincides with Se that is, standard error of estimation. Under such situation we have another interesting cross-validation parameter namely PSE that is, predictive sum of errors. This parameter that is, PSE is more directly related to the uncertainty of prediction. The lowest values of which records uncertainty of prediction. The values of PSE (Table 5) finally supports the above mentioned results.

3. Conclusions

From the above results and discussion we conclude that CA inhibitory properties of the disulfonamides used can be modeled excellently using distance-based topological indices. Based on halogen substitution at R_1 and R_3, the data set splits into two categories. In both the categories excellent results are obtained in two variable models. Two variable models containing I_2 gave better results. The splitting of the data further indicates that the sulfonamides have different mechanism of action.

In case of the disulfonamides belonging to category I, PI index is found more useful in both simple as well as in two variable models. This is not the case with the sulfonamides belonging to category II. Here, $^2\chi$ is found more useful for modeling the activity.

The sulfonamides **18**–**21**, have dual characteristics. They behave differently when belong to category I or II. Such a differential behaviour is attributed to di-halogen substitution that is, halogen substitution at both R_1 and R_3 dominates the activity more efficiently.

4. Experimental

CA inhibitory properties—The CA inhibitory activity reported as IC_{50} were adopted[2,3] in the present by converting them to log units that is, as $\log IC_{50}$.

Topological indices—All the topological indices were calculated from the hydrogen-suppressed graphs obtained by deleting all the C–H and C-hetero atom bond in the molecular structure of disulfonamides. The details of calculations are available in the literature.[13]

Regression analysis—Statistical analysis were performed by stepwise regression analysis employing maximum R^2 method.[14]

Computation—The topological indices were calculated using the software prepared by Raj Singh Sisodia, while the regression analysis were performed using regress-1 program[19,20] provided by Prof. Istvan Lukovits, Hungarian Academy of Sciences, Budapest, Hungary.

Acknowledgements

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

References and notes

1. (a) Supuran, C. T. In *Carbonic Anhydrase and Modulation of Physiologic and Pathologic Processes in the Organism*; Puscas, I., Ed.; Helicon: Timisoara (Romania), 1991; pp 29–111; (b) Supuran, C. T.; Scozzafava, A.; Cassini, A. *Med. Res. Rev.* **2003**, *23*, 146–189; (c) Scozzafava, A.; Supuarn, C. T. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1551; (d) Casini, A.; Winum, J. V.; Montero, J. L.; Scozzafava, A.; Supuarn, C. T. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1551.
2. Supuarn, C. T. *Roum. Chem. Quart. Rev.* **1993**, *1*, 77.
3. Maren, T. H.; Clare, B. W.; Supuarn, C. T. *Roum. Chem. Quart. Rev.* **1994**, *2*, 259.
4. Kubiny, H. *QSAR: Hansch Analysis and Related Approaches*; VCH: Weinheim (Germany), 1994.
5. Dias, J. R. *Molecular Orbital Calculations Using Chemical Graph Theory*; Springer: Berlin, 1993.
6. Agrawal, V. K.; Khadikar, P. V. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 447.
7. Saxena, A.; Khadikar, P. V. *Acta Pharm.* **1999**, *49*, 171.
8. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
9. Gutman, I. *Graph Theory Notes New York* **1994**, *27*, 9.
10. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. *J. Cem. Inf. Comput. Sci.* **1995**, *35*, 547.
11. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 934.
12. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.

13. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, 2000.
14. Chatterjee, S.; Hadi, A.; Price, B. *Regression Analysis by Examples*. 3rd ed.; Wiley: New York, 2000.
15. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
16. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
17. Todeschini, R. *Chemometrics Web News, Milano Chemometrics & QSAR Research Group*, File: ///I/WINDOWS/DESKTOP/Web News on Chemometrics, html. Feb **2001**.
18. Randic, M. *Croat. Chem. Acta* **1993**, *66*, 289.
19. Khadikar, P. V.; Lukovits, I.; Agrawal, V. K.; Shrivastava, S.; Jaiswal, M.; Gutman, I.; Shrivastava, A. *Indian J. Chem.* **2003**, *42A*, 1436.
20. Agrawal, V. K.; Karmarkar, S.; Khadikar, P. V. *Bioorg. Med. Chem.* **2002**, *10*, 2885.