Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Medicinal Chemistry 41 (2006) 360–366
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://france.elsevier.com/direct/ejmech

Original article

QSAR studies on benzopyran potassium channel activators

Vijay K. Agrawal [a], Jyoti Singh [a], Madhu Gupta [b], Yusuf Ali Jaliwala [c]
Padmakar V. Khadikar [d,*], Claudiu T. Supuran [e]

[a] *QSAR and Computer Chemical Laboratories, A.P.S. University, Rewa 486003, India*
[b] *Department of Chemistry, MMH P.G. College, Ghaziabad, India*
[c] *Risiraj College of Pharmacy, Sawer Road, Indore 452 0010, India*
[d] *Research Division, Laxmi Fumigation and Pest Control Pvt. Ltd., 3 Khatipura, Indore 452007, India*
[e] *University of Florence, Dipartmento di Chimica, Laboratorio di Chimica Bioinorganica, Via della Lastruccia 3, RM.188, Polo Scientifico, 50019 Sesto Fiorentino (Firenze), Italy*

Received 27 January 2005; received in revised form 27 September 2005; accepted 3 November 2005
Available online 03 February 2006

Abstract

QSAR studies on a series of benzopyrans as potassium channel activators have been carried out using a large set of distance-based topological indices. In addition, the molecular descriptors namely: negentropy and molecular redundancy indices have also been used. The relaxant potency in rat trachea, expressed as pEC_{50} was used for biological characterization of the benzopyrans. The results have shown that pEC_{50} can be modeled excellently in multiparametric model in that we have to include an indicator parameter. The predictive powers of the proposed models are discussed on the bases of cross-validation parameters.

Keywords: Potassium channel activators; QSAR; Negentropy; Topological indices; Benzopyrans

1. Introduction

Three families of potassium ion channels have been described to date [1]. Compounds interacting with these families of proteins may control various diseases, and research was reported for the synthesis of compounds which can activate or block K_{ATP}-dependent potassium channels [1]. Recently, much interest have been shown towards the K (potassium) channel activators (KCA) which have been found to be smooth muscle relaxants with their main utility in hypertension and bronchodilators [1].

In their paper Mannhold et al. [1] reported the synthesis, vasodilator properties and multivariate analysis of 6-substituted benzopyrans as KCAs. Initially they started with 34 compounds and in the end deleted two compounds from regression procedure as outliers (Fig. 1 and Table 1). Their regression analysis explained the variance in biological activity to 82% in the tracheal test system. They observed that low values of substituent size are favorable for high potency. However, they have not used topological indices in their analysis, have not mentioned statistical models along with their quality and also have not attempted estimation of predictive power of the proposed models.

Our earlier reports [2–11] indicated that the distance-based topological indices can be successfully used in such studies as mentioned above. Topological based studies gave better results than those in which indices other than topological indices are used for characterizing the biological activity.

We report in the present study a large set of molecular descriptors along with topological indices used to model pEC_{50} of the benzopyran K-channel activators reported in [1] (Fig. 1 and Tables 1 and 2). It is interesting to mention that in the Mannhold et al.'s paper [1], from which the data have been taken, the activity is recorded in two systems: aortic and tracheal. The former system showed much more dependence on

* Corresponding author. Tel.: +91 731 253 1906; fax: +91 7662 24 2175, +91 7662 23 0819.
E-mail address: pvkhadikar@rediffmail.com (P.V. Khadikar).

0223-5234/$ - see front matter © 2006 Elsevier SAS. All rights reserved.
doi:10.1016/j.ejmech.2005.11.004

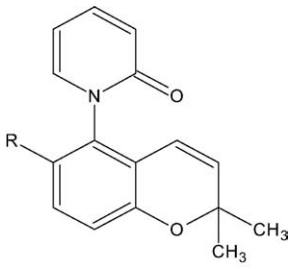

Fig. 1. Benzopyrans KCAs used in the present study.

the size of the substituents. The topological indices chosen in the present study will be much more useful for modeling aortic system. Consequently, the present study deals with the same.

2. Results and discussion

The initial regression analysis has shown that compounds **1**, **6**, **9**, **16**, **29** and **32** are outliers. Therefore, they are deleted from the process of regression. Thus, our methodology relates in 28 compounds only. At present we can't offer convincing proof for the deletion of six compounds. Maybe it is due to the ultimate result of regression and/or they have different type of mechanism than the remaining 28 compounds. It is worthy to mention that our set of compounds include compounds **13**, **34** and **35** which otherwise were deleted by Mannhold et al. [1] from their regression analysis.

The biological activity (pEC_{50}) of the present set of 34 compounds along with the indicator parameters used is given in Table 1. The data show that the biological activity (pEC_{50}) ranges between 4.63 and 7.84 and that little degeneracy is observed in the activity.

A perusal of Table 1 shows that pEC_{50} activity follow the following sequence in increasing order of pEC_{50}.

$$\begin{aligned} 29 > 33 > 30 > 34 > 3 > 19 > 11 > 15 > 18 > 17 > 13 \\ > 14 > 31 > 9 > 4 > 5 > 2 > 7 > 32 \\ > 10 > 12 > 6 > 16 > 21 > 27 > 20 \\ > 1 > 25 > 24 = 26 > 28 > 22 > 23 \end{aligned} \quad (1)$$

This shows that fluorosulfonyloxy substitution at R has highest effect on the exhibition of pEC_{50} while the substitution of 4-fluorophenacyloxy has most retarding effect. The above sequence (Eq. (1)), however, does not show any correlation between structure and activity.

Consequent to above, we have attempted regression analysis using the data recorded in Tables 1 and 2 shows that degeneracy (high to low) is present in these topological indices as they belong to first- and second-generation topological indices [11]. Balaban has shown that such indices in spite of their degeneracy can be used successful in QSPR/QSAR studies. This is the case in the present study also.

In order to arrive at the most significant model we have used maximum R^2 method [12]. The results have shown that

Table 1
Compounds used in the present study, their pEC_{50} and indicator values

Compound No	Compound	EC_{50}	I_{r1}
1[a]	H	5.43	0
2	Methoxy	6.55	1
3	Acetyl	7.37	0
4	Propionyl	6.63	0
5	Benzoyl	6.61	0
6[a]	4-Methoxybenzoyl	6.15	0
7	2-Thienoyl	6.40	0
8	2-Furoyl	6.49	0
9[a]	4-Hydroxybenzoyl	6.65	0
10	4-Nitrobenzoyl	6.20	0
11	2-Fluorobenzoyl	7.08	0
12	2-Nitrobenzoyl	6.17	0
13	2-Methylbenzoyl	6.83	0
14	2-(Trifluoromethyl)benzoyl	6.76	0
15	2,6-Difluorobenzoyl	6.97	0
16[a]	Aminohydroxyiminomethyl	5.99	0
17	Formyl	6.91	0
18	2,2-Dicyanoethenyl	6.95	0
19	2,5-Dimethyl-1-pyrrolyl	7.27	0
20	Hydroxy	5.44	1
21	Acetoxy	5.60	1
22	Phenacyloxy	5.05	1
23	4-Fluorophenacyloxy	4.63	1
24	Phenylcarbamoyloxy	5.30	1
25	2-Fluorophenylcarbamoyloxy	5.38	1
26	4-Fluorophenylcarbamoyloxy	5.30	1
27	2-(Trifluoromethyl) phenylcarbamoyloxy	5.58	1
28	4-(Trifluoromethyl) phenylcarbamoyloxy	5.10	1
29[a]	Fluorosulfonyloxy	7.95	1
30	Cyno	7.67	0
31	4-Pyridyl	6.68	0
32[a]	Thiocarboxamid	6.21	0
33	Bromo	7.84	0
34	Trifluoromethyl	7.61	0

$Ip_1 = 1$, when oxy group is present otherwise 0.

Balaban index [13] (J) is the promising index for this purpose. At this stage it is worthy to mention that compared to Wiener [14] (W) and Randic connectivity indices (Kier and Hall [15]), little attention is paid for the use of Balaban index [13] (J) in QSPR/QSAR studies. Same is the case with the molecular redundancy index MRI [16].

The perusal of Table 3 shows that except for MRI [16] all other topological indices are linearly correlated with each other. They are, therefore, autocorrelated. This means any combination of these indices in regression procedure may result into the defect due to co-linearity. However, such cases are well dealt with Randic [17] and we will follow his recommendation in those cases when auto correlated indices are involved in the model.

The data presented in Table 3 also show that W [14], Sz [18,19] and logRB [20] indices are best suited for monoparametric regression. The initial regression analysis indicated that the model containing J though of the low statistics, is better than the other models. Therefore, our methodology is centered on the use of J and related multiparametric models. As stated earlier another reason in favor of J is that very little attention is paid to this most discriminating index (Tables 4–6).

Table 2
Values of topological indices calculated for compounds used in the present study

Compound No	*N*	*W*	[1]χ	MRI	*J*	Sz	logRB
1	38.8512	838	9.9991	0.3527	1.7975	1558	258.9346
2	40.8720	955	10.3719	0.3412	1.8019	1748	292.3170
3	46.3974	1094	10.9099	0.3194	1.7882	1960	330.1472
4	41.6300	1756	12.9443	0.4557	1.4864	3130	502.2575
5	41.9164	1564	12.4442	0.4032	1.5018	2696	454.0554
6	41.9164	1564	12.4442	0.4032	1.5018	2696	454.0554
7	48.9072	2462	14.2488	0.3939	1.4030	4271	667.1037
8	44.6936	1930	13.3549	0.4156	1.4966	3407	548.5343
9	48.3168	2336	14.2656	0.4012	1.5016	4019	652.1990
10	45.6729	1930	13.3549	0.4484	1.4966	3407	548.5342
11	49.5880	2542	14.5663	0.4012	1.5126	4328	705.0710
12	46.4784	2108	13.7656	0.3923	1.5115	3690	596.1970
13	36.2916	838	9.9991	0.3522	1.7975	1558	258.8346
14	42.5480	1438	11.9301	0.3360	1.7272	2450	417.9004
15	47.4230	1668	12.3929	0.3965	1.6697	2753	470.3089
16	34.9055	838	9.9991	0.3540	1.7975	1558	258.9346
17	42.9160	1112	10.8549	0.3302	1.7610	1978	333.0909
18	51.7300	2186	13.8381	0.3910	1.4566	3784	607.6320
19	54.7200	2393	14.2319	0.3558	1.4638	4122	660.3986
20	48.5688	2494	14.3381	0.4135	1.4010	4114	672.6647
21	51.6264	2720	14.7488	0.3766	1.4068	4446	728.0526
22	51.6264	2768	14.7319	0.3766	1.3853	4542	733.6471
23	59.0672	3497	15.9601	0.4224	1.4178	5541	912.1655
24	59.0672	3689	15.9433	0.4224	1.3492	5925	932.5504
25	27.7060	838	9.9991	0.4873	1.7975	1558	258.9346
26	38.4506	1396	12.0336	0.4525	1.5082	2606	409.9767
27	32.3340	723	9.4611	0.3789	1.8073	1370	226.2454
28	36.8483	1074	10.6724	0.3649	1.8178	1940	326.3927

The monoparametric regression containing *J* yielded the following regression expression (model):

$$pEC_{50} = 3.1471(\pm 0.8480)J + 1.3840 \quad (2)$$

$n = 28$, S.E. = 0.7267, $R = -0.5918$, $F = 14.011$, $Q = 0.8143$.

Here and hereafter *n* is the number of compounds used, S.E. is the standard error of estimation, *R* is the multiple correlation coefficient, *F* is the *F*-ratio, *Q* is the quality factor [21,22]. This quality factor is defined as the ratio of correlation coefficient to the standard error of estimation, $Q = R/\text{S.E.}$

In the above model (Eq. (2)) the *J* involved is a highly discriminating distance-based topological index, whose value do not substantially increase with the molecular size and the number of rings present The positive sign of its coefficient indicates its favorable contribution for modeling pEC_{50}.

In several biparametric regressions attempted we observed that the biparametric regression containing *J* and the indicator parameter $^{I}p_1$ gave better results. This model is found as below:

$$pEC_{50} = 2.0694(\pm 0.4385)J + 1.3247(\pm 0.1495)IP_1 + 3.5646 \quad (3)$$

$n = 28$, S.E. = 0.3641, $R = 0.9182$, $F = 67.167$, $Q = 2.5218$.

Here, the positive coefficient of *J* has the same influence for the exhibition of pEC_{50} as mentioned above. The negative coefficient of $^{I}p_1$ indicates that 'Oxy' substitution at R has retarding effect on the exhibition of pEC_{50}.

The step-wise regression resulted into several triparametric models out of which the model containing *J*, MRI and $^{I}p_1$ was found to be the best. This model is as follows:

$$pEC_{50} = 2.7548(\pm 0.4761)J + 4.9137(\pm 1.8983)\text{MRI} - 1.1747(\pm 0.1468)I_{P1} + 0.5055 \quad (4)$$

$n = 28$, S.E. = 0.3286, $R = 0.9367$, $F = 57.220$, $Q = 2.8505$.

Here, the contributions of *J* and $^{I}p_1$ are same as discussed earlier. In addition, positive coefficient of MRI indicates that pEC_{50} is favored by the information content.

Looking to the sample size (28 compounds) we have also can finally go for tetraparametric analysis. Out of the several tetraparametric models, the model containing *J*, MRI, *W* and $^{I}p_1$ was found to be the best. This model is found as under:

$$pEC_{50} = 3.7015(\pm 0.9099)J + 5.2400(\pm 1.8987)\text{MRI} + 2.184\,0 \times 10^{-4}(\pm 1.795\,5 \times 10^{-4})\text{W} - 1.2294(\pm 0.1522)I_{p1} - 1.4982 \quad (5)$$

$n = 28$, S.E. = 0.3254, $R = 0.9406$, $F = 44.143$, $Q = 2.8905$.

At this stage it is interesting to record that in both the models (Eqs. (4) and (5)) MRI term is involved. Its high positive coefficient value in these models indicates its dominating role in the exhibition of pEC_{50}. This molecular descriptor MRI is an

Table 3
Correlation matrix for the inter-correlation of structural descriptors and their correlation with the activity

	pEC_{50}	*N*	*W*	$^1\chi$	MRI	*J*	Sz	log(RB)	Ip_1
pEC_{50}	1.0000	–0.6841	–0.6183	–0.5947	0.1203	0.5918	–0.6132	–0.6103	–0.8387
N		1.0000	0.9069	0.8973	–0.0138	–0.7624	0.9022	0.9085	0.4670
W			1.0000	0.9791	0.3042	–0.8868	0.9974	0.9984	0.4087
$^1\chi$				1.0000	0.3368	–0.9346	0.9864	0.9882	0.3218
MRI					1.0000	–0.4655	0.3222	0.3102	–0.2067
J						1.0000	–0.9044	–0.8981	–0.2773
Sz							1.0000	0.9986	0.3830
log(RB)								1.0000	0.3849
Ip_1									1.0000

Table 4
Regression parameters and quality of the proposed models

Compound number	Parameter used	Ai[a] *I* = 1, 2, 3, 4	(B)	(R^2)	(*R*)	(S.E.)	*F*-Ratio	*Q* = *R*/S.E.	Probability
1	*J*	3.1471(± 0.8408)	1.3840	–0.3502	–0.5918	–0.7267	14.011	0.8143	$9.10\ 2 \times 10^{-4}$
2	*J*	2.0694(± 0.4385)	3.5646	–0.8431	–0.9182	–0.3641	67.167	2.5218	8.819×10^{-11}
	Ip_1	–1.3247(± 0.1495)							
3	MRI	4.9137(± 1.8983)	0.5055	–0.8773	–0.9367	–0.3286	57.220	2.8505	4.407×10^{-11}
	J	2.7548(± 0.4761)							
	Ip_1	–1.1747(± 0.1468)							
4	*W*	$2.1840 \times 10^{-4}(\pm 1.7955 \times 10^{-4})$	–1.4982	–0.8848	–0.9406	–0.3254	44.143	2.8905	1.807×10^{-10}
	MRI	5.2400(± 1.8987)							
	J	3.7015(± 0.9099)							
	Ip_1	–1.2294(± 0.1522)							

[a] Ai stands for the coefficient of I^{th} term.

Table 5
Various correlation models and their qualities

Model number	Regression expression
1	$pEC_{50} = 3.1471\ (\pm 0.8408)J + 1.3840$
2	$pEC_{50} = 2.069(\pm 0.4385)J - 1.3247(\pm 0.1495)Ip_1 + 3.5646$
3	$pEC_{50} = 4.9137(\pm 1.8983)$MRI + $2.7548(\pm 0.4761)J - 1.1747$ $(\pm 0.1468)Ip_1 + 0.5055$
4	$pEC_{50} = 2.1840 \times 10^{-4}(\pm 1.7955 \times 10^{-4})W + 5.2400$ (± 1.8987)MRI + $3.7015(\pm 0.9099)J - 1.2294(\pm 0.1522)Ip_1 -$ 1.4982

information theoretic index. The results, therefore, show that the information content of the benzopyrans used has a dominating effect on the exhibition of pEC_{50}. The Wiener index (*W*) on the other hand accounts for the effect due to shape, size and branching. Its positive coefficient in expressed by equation (5) favors the recommendation of Mannhold et al. [1] that low values for substituent size is favorable for high potency.

In the aforementioned models we observed that the quality factor Q goes on increasing as we pass from the mono- to tetraparametric models. This means that the quality of regression and the predictive power of the model run parallel to each other. However, the use of *Q* factor is criticized recently [23], therefore, we attempted cross-validation method [24,25] for discussing predictive power of the proposed models. The various cross-validated parameters calculated for the proposed models are given in Table 7. The meaning of PRESS, SSY, r^2_{CV}, S_{PRESS} and PSE is given as a footnote in Table 7.

PRESS is an important cross-validation parameter. Its value smaller than SSY indicates that the proposed models are significant and they predict better than chance.

Furthermore, the ratio PRESS/SSY is found smaller than 0.4 indicating that the proposed models are reasonable QSAR models. The ratio of PRESS/SSY smaller than 0.1 indicates an excellent model. In our case this ratio ranges between 0.13 and 0.19. Thus we can safely say that these models are excellent. The parameter r^2_{CV} is cross-validated predictive correlation coefficient. The values for our models favor this finding.

The other important cross-validated parameters are S_{PRESS} and PSE i.e. uncertainty of prediction and predictive square error. Unfortunately, S_{PRESS} is of no use in the present case as it is found to be the same as that of S.E. Under such circumstances one has to make use of PSE, the smallest value of which indicates highest predictive power of the model. Hence, the tetraparametric model possesses highest statistical quality as well as highest predictive power.

In support of our findings, we have estimated pEC_{50} using tri- and tetraparametric models [Eqs. (4) and (5)) and obtained graphs between observed and estimated pEC_{50} (Figs. 2 and 3). The predictive correlation coefficient (R^2_{pred}) for these models are found to be 0.8773 and 0.8848, respectively, which are in favor of tetraparametric model; thus confirming our earlier results.

3. Conclusion

From the results and discussion made above we conclude that combination of distance-based topological indices with in-

Table 6
Comparison of estimated biological activity (pEC_{50}) with their observed values

Compound number	Observed pEC_{50}	Estimated pEC_{50} using			
		Model-3	Residue	Model-4	Residue
1	6.550	6.016	0.534	5.957	0.593
2	7.370	7.146	0.224	7.168	0.202
3	6.630	7.001	–0.371	7.033	–0.403
4	6.610	6.839	–0.229	6.775	–0.165
5	6.400	6.624	–0.224	6.515	–0.115
6	6.490	6.624	–0.134	6.515	–0.025
7	6.200	6.306	–0.106	6.297	–0.097
8	7.080	6.670	0.410	6.641	0.439
9	6.170	6.613	–0.443	6.672	–0.502
10	6.830	6.832	–0.002	6.813	0.017
11	6.760	6.644	0.116	6.758	0.002
12	6.970	6.597	0.373	6.613	0.357
13	6.910	7.188	–0.278	7.184	–0.274
14	6.950	6.915	0.035	6.970	–0.020
15	7.270	7.053	0.217	7.124	0.146
16	5.440	6.022	–0.582	5.964	–0.524
17	5.600	5.805	–0.205	5.764	–0.164
18	5.050	5.265	–0.215	5.190	–0.140
19	4.630	5.112	–0.482	5.078	–0.448
20	5.300	5.222	0.078	5.170	0.130
21	5.380	5.057	0.323	5.047	0.333
22	5.300	4.998	0.302	4.978	0.322
23	5.580	5.312	0.268	5.497	0.083
24	5.100	5.123	–0.023	5.285	–0.185
25	7.670	7.852	–0.182	7.892	–0.222
26	6.680	6.884	–0.204	6.760	–0.080
27	7.840	7.346	0.494	7.335	0.505
28	7.610	7.306	0.304	7.377	0.233

Table 7
Cross-validation parameters for the proposed models

Model (Eq.)	Parameters used	PRESS	SSY	PRESS/SSY	R^2cv	S_{PRESS}	PSE
2. (3)	J, Ip_1	3.314-9	17.8120	0.1861	0.813-9	0.364-2	0.3440
3. (4)	J, MRI, Ip_1	2.588-8	18.5380	0.1397	0.860-4	0.328-4	0.3040
4. (5)	W, MRI, J, Ip_1	2.434-8	18.6921	0.1303	0.869-7	0.325-4	0.2949

formation theoretic indices are more fruitful in QSAR studies. Also, that less attended J is found more useful than the other distance based topological indices used. In addition our results show that this less attended J index gives even better results than the most widely used W and 1_χ indices.

4. Experimental

4.1. Biological activity (pEC_{50})

The biological activities for the set of compounds used in the present study are adopted from the literature [1].

4.2. Topological indices

All the topological indices were calculated using hydrogen suppressed graph of the compounds used. Such molecular graphs are obtained by deleting all the hydrogen atoms present

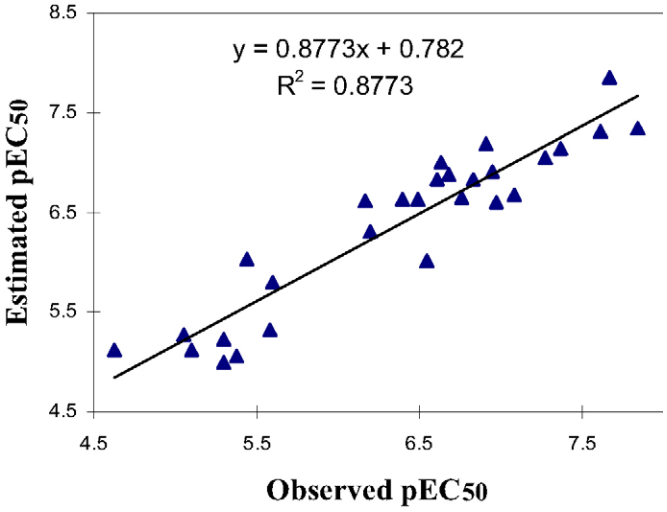

Fig. 2. Correlation of observed vs. estimated pEC_{50} using model-3 (Eq. 4).

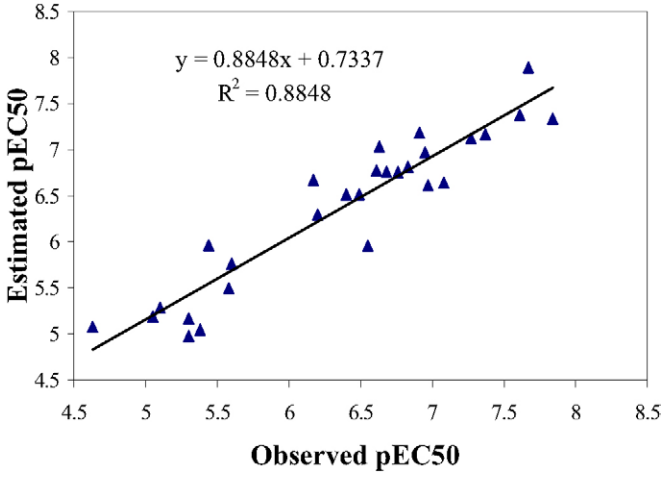

Fig. 3. Correlation of observed vs. estimated pEC_{50} using model-4 (Eq. 5).

in the structure. Below we give only the final expressions for their calculations.

4.3. Wiener index (W)

The Wiener index (W) is a widely used topological index [14]. It is based on the vertex-distances of the respective molecular graph.

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

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

4.4. Szeged index (Sz)

Let e be an edge of the molecular graph G. Let n_1(e|G) be the number of vertices of G lying closer to one end of e; let n_2(e|G) be the number of vertices of G lying closer to the other

end of e. Then the Szeged index (Sz) is defined [18,19] as:

$$\mathrm{Sz} = \mathrm{Sz(G)} = \sum \mathrm{e}\, n_1(\mathrm{e|G}) n_2(\mathrm{e|G}) \quad (7)$$

with the summation going over all the edges of G.

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

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

The connectivity index $^1\chi = {}^1\chi(\mathrm{G})$ of a graph G is defined by Randic (Kier and Hall [15]) as under:

$$^1\chi = {}^1\chi(G) = \sum_{ij} [\delta_i \delta_j]^{-1/2} \quad (8)$$

where δ_i and δ_j are the valences of vertices i and j, equal to the number of bonds connected to the atoms i and j, in G.

In the case of hetero-systems the connectivity is given in terms of valence delta values δ_i^{v} and δ_j^{v} of atoms i and j and is denoted by 1_χ^{v}. This version of the connectivity index is called the valence connectivity index and is defined [15] as under:

$$^1\chi^v = {}^1\chi^v(G) = \sum_{ij} [\delta_i^v \delta_j^v]^{-1/2} \quad (9)$$

where the sum is taken over all bonds i–j of the molecule. Valence delta values are given by the following expression:

$$\delta_i^v = \frac{Z_i^v - H_i}{Z_i - Z_j - 1} \quad (10)$$

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

4.6. Balaban index (J)

The Balaban index, J (the average distance sum connectivity index) is defined [13] by:

$$\mathrm{J} = \mathrm{J(G)} = \frac{\mathrm{M}}{\mu + 1} \sum_{\mathrm{bond}} (\mathrm{d_i d_j})^{-1/2} \quad (11)$$

where M is the number of bonds in a graph G, μ is the cyclomatic number of G and d_i's($i = 1, 2, 3,\ldots,N$) are the distance sums (distance degrees) of atoms in G such that

$$d_i = \sum_{j=1}^{N} (D)_{ij}. \quad (12)$$

The cyclomatic number μ of G indicates the number of independent cycles in G and is equal to the minimum number of cuts (removal of bonds) necessary to convert a polycyclic structure into an acyclic structure:

$$\mu = M - N + 1. \quad (13)$$

One way to compute J for a hetero-system is to modify the elements of the distance matrix for hetero-system as follows:

(i) The diagonal elements:

$$(D)_{ij} = 1 - (Z_c/Z_i) \quad (14)$$

where $Z_\mathrm{c} = 6$ and Z_i = atomic number of the given element.

(ii) The off-diagonal elements:

$$d_i = \sum rR_r \quad (15)$$

where the summation is over all bonds. The bond parameter k_r is given by:

$$k_r = 1/b_r(Z_c^2/Z_i Z_j) \quad (16)$$

where b_r is the bond weight with values: 1 for single bond, 2 for double bond, 1.5 for aromatic bond and 3 for triple bond.

4.7. logRB

The logRB has been calculated by the method as described in literature [20].

4.8. Molecular negentropy (N)

The method of calculation of molecular negentropy is the same as described by Kier [16]. First, the negentropy per atom i, is computed for the drug molecule using Shannon's formula [16]:

$$i = -K \sum j P_j \log P_j \quad (17)$$

where K is constant depending on the logarithmic base, j is the set, and P_j is the complete array of probabilities.

Multiplying i with the total number of vertices (atoms), n, gives the molecular negentropy (N):

$$N = i.n. \quad (18)$$

4.9. Molecular redundancy index (MRI)

A message or information in the form of electron probability fields distributed around in space in a frame work of atomic radii has lead to introduction of molecular descriptor referred to as molecular redundancy index [16] (MRI), derived from information theory and molecular graph theory [20]. MRI is a molecular symmetry descriptor and indicates the capacity and symmetry of a molecule and can be computed as:

$$MRI = \frac{\sum n_i \log n_i}{N \log N} \quad (19)$$

where n is the number of atoms of the same kind in the i^{th} atom set, i is the number of different atoms sets and $N = \Sigma n_i$ is the total number of atoms in the molecule.

Eq. (19) shows that calculation of MRI leads to quantification of the information content. It encodes the salient steric properties of the molecules in cases where biological activity is non-specific. It ranks them correctly according to non-specific biological potency and thus, provides mechanistic interpre-

tation of drugs at molecular level based on probability consideration.

4.10. Regression analysis

Maximum R^2 improvement method was used to identify prediction models [12]. This method finds the "best" one variable model, the "best" two variable models and so forth for the prediction of property/activity. Several models (combinations of variables) were examined to identify combinations of variables with good prediction capabilities. In all regression models developed a variety of statistics associated with residues, i. e. the Wilks–Shapiro test for normality and Cook's D-statistics for outliers, to obtain the most reliable results were examined.

Multiple regression analyses for correlating activity of the present set of compounds with the aforementioned molecular descriptors were carried out using Regress-1 software as supplied by Professor I. Lukovits, Hungarian Academy of Sciences, Budapest, Hungary.

Acknowledgments

Authors are thankful to Professor Istvan Lukovits, Hungarian Academy of Sciences, Budapest, Hungary for providing software to carryout regression analysis. Authors (V.K.A. and J. Singh) are thankful to Professor A.D.N. Bajpai, Vice-chancellor of APS University, Rewa, India for providing research facilities and encouragement. Authors are also thankful to CSIR New Delhi, India for providing financial support through project No *01(1785)/02/EMR-II*.

References

[1] R. Mannhold, G. Cruciani, H. Weber, H. Lemoine, A. Derix, C. Weichel, M. Clementi, J. Med. Chem. 42 (1999) 981–991.
[2] V.K. Agrawal, J. Singh, P.V. Khadikar, Bioorg. Med. Chem. 10 (2002) 3981–3996.
[3] V.K. Agrawal, R. Sharma, P.V. Khadikar, Bioorg. Med. Chem. 10 (2002) 2993–2999.
[4] V.K. Agrawal, R. Sohgaura, P.V. Khadikar, Bioorg. Med. Chem. 9 (2001) 3295–3299.
[5] V.K. Agrawal, S. Bano, C.T. Supuran, P.V. Khadikar, Eur. J. Med. Chem. 39 (2004) 593–600.
[6] V.K. Agrawal, S. Shrivastva, P.V. Khadikar, Moecular Div. 8 (2004) 413–419.
[7] V.K. Agrawal, S. Karmarkar, P.V. Khadikar, S. Shrivastva, I. Lukovits, Ind. J. Chem. 42A (2003) 1426–1435.
[8] P.V. Khadikar, I. Lukovits, V.K. Agrawal, S. Shrivastva, M. Jaiswal, I. Gutman, S. Karmarkar, A. Shrivastava, Ind. J. Chem. 42A (2003) 1436–1442.
[9] V.K. Agrawal, P.V. Khadikar, Bioorg. Med. Chem. 10 (2002) 3517–3522.
[10] V.K. Agrawal, S. Joseph, P.V. Khadikar, S. Karmarkar, Acta Pharm. 50 (2000) 329–338.
[11] V.K. Agrawal, K. Mishra, P.V. Khadikar, Oxidation Commun. 26 (2003) 14–21.
[12] S. Chaterjee, A.S. Hadi, B. Price, Wiley, New York, 2000 (3rd Ed).
[13] A.T. Balaban, Chem. Phys. Lett. 89 (1982) 399–404.
[14] H. Wiener, J. Am. Chem. Soc. 69 (1947) 17–20.
[15] L.B. Kier, L.H. Hall, Molecular connectivity in structure-activity relationship, Wiley, New York, 1986, in: L.B. Kier, L.H. Hall (Eds.), Molecular structure description, Academic Press, New York, 1999.
[16] L.B. Kier, J. Pharm. Sci. 69 (1980) 807–810.
[17] M. Randic, Croat. Chem. Acta 66 (1993) 289–299.
[18] P.V. Khadikar, S. Sharma, V. Sharma, S. Joshi, I. Lukovits, M. Kaveeshwar, Bull. Soc.Chem.Belg. 106 (1997) 767–772.
[19] P.V. Khadikar, N.V. Deshpandey, P.P. Kale, A. Dubrynin, I. Gutman, G. J. Domotor, Chem. Inf. Comput. Sci. 35 (1995) 547–550.
[20] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim GER, 2000. M. Karelson, Molecular Descriptors in QSAR/QSPR, J. Wiley & Sons, New York, 2000. M.V. Diudea, Ed. QSPR/QSAR Studies by Molecular Descriptors, Babes-Bolyai University, Cluj. Romania, 2000. N. Trinajstic, Chemical Graph Theory, 2nd revised ed, CRC Press, Boca Raton, Fl, 1992.
[21] L. Pogliani, Amino Acids J. Phys. Chem. 98 (1994) 1494–1499.
[22] L. Pogliani, Recent trends in graph theoretical descriptors, Chem. Rev. 100 (2000) 3827–3858.
[23] J. Devillers, A.T. Balaban, Topological Indices and Related Descriptors in QSAR and QSPR, Gorden & Breach, Williston, VT, 2000.
[24] V.K. Agrawal, S. Bano, P.V. Khadikar, Bioorg. Med. Chem. 13 (2003) 4039–4047.
[25] P.V. Khadikar, A. Srivastava, V.K. Agrawal, S. Shrivastava, Bioorg. Med. Chem. Lett. 13 (2003) 3014–3020.