Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 14 (2004) 4795–4801
Bioorganic & Medicinal Chemistry Letters

Topological estimation of electronic absorption bands of arene absorption spectra as a tool for modeling their toxicity and environmental pollution

Padmakar V. Khadikar,[a,*] Shalini Singh,[a] Mona Jaiswal[a] and Dheeraj Mandloi[b]

[a]*Research Division, Laxmi Fumigation and Pest, Control Pvt. Ltd, 3, Khatipura, Indore 452007, India*
[b]*Institute of Engineering and Technology, D.A. University, Indore 452017, India*

Received 12 May 2004; accepted 12 June 2004
Available online 22 July 2004

Abstract—A novel application of distance-based topological indices : Wiener (W)-, Szeged (Sz)-, Padmakar–Ivan (PI)-, and Sadhana (Sd)-indices in modeling electronic absorption bands of arene absorption spectra has been described. It is demonstrated that all these indices correlate linearly with the logarithm of β and *para* electronic absorption bands in several series of arene systems. The results have shown that our methodology is best suited for the estimation—$\ln \lambda_\beta$, while comparatively less significant results are obtained in case of the estimation of $\ln \lambda_p$. The statistical analysis of the data have shown that PI index gives better results for modeling $\ln \lambda_\beta$; while Sz index proved better for modeling $\ln \lambda_p$. The results are critically discussed on the basis of regression parameters and quality of correlation. Such a study will be useful as a tool for modeling toxicity of arene system as well as their environmental pollution.

1. Introduction

The introduction of the versatile mathematical concept viz., topology and graph theory in chemistry and other related disciplines resulted into very interesting results.[1] The development of quantitative structure–property–activity–toxicity-relationships (QSPR, QSAR, QSTR) are among such major results. In these methodologies properties/activities/toxicities are modeled using graph-theoretical indices also called as topological indices and have shown an explosive growth.[2–5]

A topological index is a single number, derived following a certain rule, which can be used to characterize the molecule. A plethora of topological indices are reported in the literature.[1–5] The Wiener index (W)[6] is the first, the oldest, and the widely used topological index for this purpose. It was introduced by Wiener in 1947. Even today this index finds very useful in modeling physicochemical properties and biological activities of organic molecules.[7] The Szeged index (Sz)[8,9] is the modification of Wiener index (W) to cyclic molecules. This index was introduced by Gutman in 1994.[8,9] For acyclic compounds Sz index coincides with W index. Consequently yet another index was introduced by one of the authors (P.V.K.) and was named as Padmakar–Ivan index (PI).[10,11] The PI index unlike Sz index, is different for acyclic and cyclic molecules. While attempting the derivation of recursion formula for PI index, Khadikar introduced yet another index called Sadhana index (Sd).[12,13] However, like Sz index, Sd index is applicable to cyclic compounds only.

The major use of topological indices is in modeling physicochemical properties and biological activity of organic molecules acting as drugs. The important properties/activities modeled being those related to carbonic anhydrase (CA) inhibition; and tadpole narcosis.[14–17] The other properties modeled being chiefly: boiling point (bp), molar volume (MV), molecular refraction (MR), polarizability (α), dipole molecules (μ), heat of formation (Hf), etc. However, comparatively very little work is done on topological modeling of spectroscopic properties. Some such spectroscopic properties modeled using topological indices are: infrared absorption bands,[18,19] Mossbaur parameters,[20] chemical shifts in NMR spectra,[21–24] g-tensors in ESR spectra[25,26] and

Keywords: Electronic spectra; Arene system; Arene absorption band; Molecular connectivity; Topological indices; Regression analysis.
* Corresponding author. Tel.: +91-731-2531906; fax: +91-766-2242175; e-mail addresses: pvkhadikar@rediffmail.com; dheeraj_m@sancharnet.in

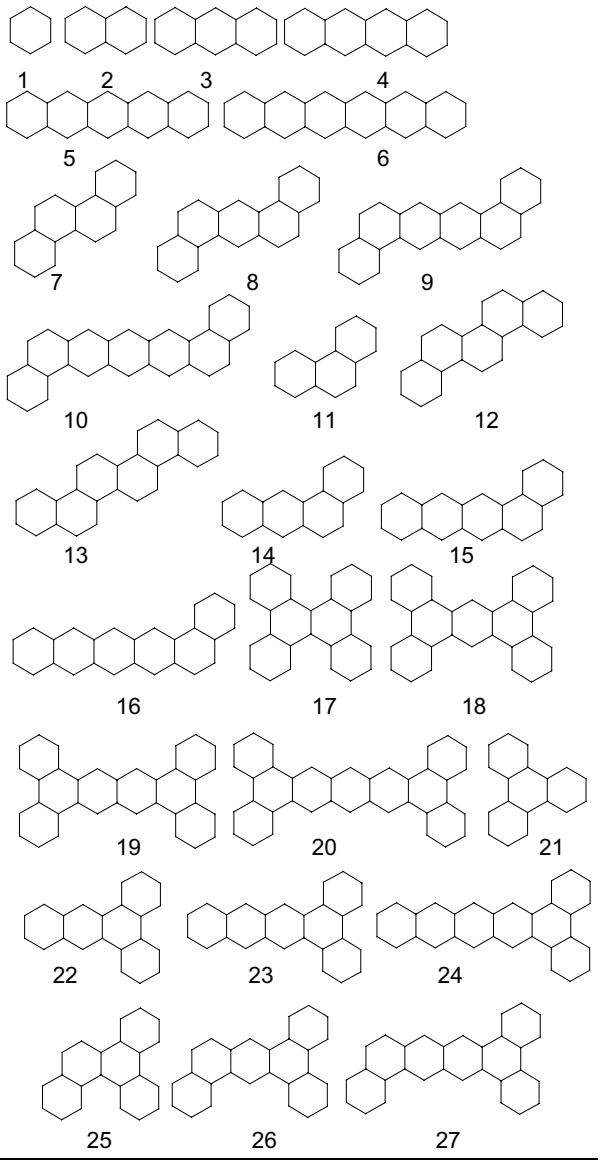
0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.06.094

X-ray absorption fine-structure parameters.[27–33] However, very little work is done on the application of topological indices to electronic spectra (both UV and visible spectra). The only reference which we could trace out is that of the application of Randic index in the correlation of arene absorption spectra.[34] Rouvary and El-Basil[34] demonstrated that, for a several series of arene system χ index (first order Randic connecting index)[35] correlated linearly with the logarithm of the β and *para* electronic absorption bands. They have shown that the correlation splits into two categories depending upon the nature of the member of the arene series. Accordingly, linearity in correlations exist only for the members of the given series, and that such members must have closely related topologies and geometries. Such series were named by them as 'family' and relationship was called 'familial' relationship. They were of the opinion that chemical graphs can belong to families in the same way that plants belong to genera. Furthermore, the methodology adopted by Rouvary and El-Basil[34] was very much qualitative and no statistics was reported. The mentioned spectral parameters are successfully used as molecular descriptors in developing QSPR/QSAR/QSTR models for modeling physiological activities.

In view of the above, and in continuation of our earlier work related to the topological modeling of spectroscopic properties, we have undertaken the present work, in that we have used *W*, Sz, PI, and Sd indices for modeling, monitoring, and estimation of $\ln\lambda_\beta$ and $\ln\lambda_p$ electronic absorption bands for the same arene series used by Rouvary and El-Basil.[34] The results are discussed below.

Another objective of the present study is that arene system is well known for their toxicity as well as for their environmental pollution. Also, that electronic spectral data can be used as molecular descriptors for modeling toxicity and pollution. The methodology described herein for the topological estimation of electronic absorption band will be useful for those who are interesting in modeling toxicity and pollution of arenes.

Table 1. Arene Systems Used in the Present Study



2. Results and discussion

Before discussing our results it is worthy to mention that catacondensed arene molecules usually exhibit three principal bands in the near ultraviolet and visible region of their spectra. These three bands have been designated as the alpha (α), beta (β), and *para* (*p*) bands.[34] For the latter two bands there are adequate data available, whereas there are only limited data on the α band. In certain instances the α bands are completely subsumed by the *para* bands and cannot be measured at all. Consequent to this, and following Rouvary and El-Basil,[34] we have also chosen β and *p* bands for our study and have used their natural logarithm. The experimental observations[36] and the four reasons given by Rouvary and El-Basil[34] support that β and *para* bands are sensitive to the topological structure of the arene molecules from which they are originated. Hence, the use of topological indexes for modeling, monitoring and estimating $\ln\beta$ and $\ln p$ is well justified and required to be investigated.

The set of arene systems used in the present investigation is given in Table 1. Table 2 records the adopted values of $\ln\beta$ and $\ln p$ along with the calculated values of *W*, Sz, PI, and Sd indices. For statistical correlation of $\ln\beta$ and $\ln p$ with each of the topological indices used (*W*, Sz, PI, Sd) we have followed simple regression as well as curvilinear methodology.[37] The derived statistical parameters and the quality of correlations are given in Table 3. Tables 4 to 6 record the comparison of observed and estimated $\ln\beta$ and $\ln p$, respectively.

2.1. Estimation of $\ln\lambda_\beta$

The results obtained for estimating $\ln\lambda_\beta$ by simple regression analysis using *W*, Sz, PI, and Sd indices are given in Table 3. A perusal of this Table 3 shows that out of the four topological indices used the PI index gives best statistics. This model based on PI index is found as below:

Table 2. $\ln \lambda_\beta$, $\ln \lambda_p$ for arene series and their topological indices (W, Sz, PI, Sd)

Arene no.	$\ln \lambda_\beta$	$\ln \lambda_p$	W	Sz	PI	Sd	$^1\chi$[a]
1	7.524	7.634	27	54	24	12	1.000
2	7.700	7.955	109	243	96	44	1.414
3	7.830	8.228	279	656	216	96	1.732
4	7.916	8.457	569	1381	384	168	2.000
5	8.016	8.658	1011	2506	600	260	2.236
6	—	8.844	1637	4119	864	372	2.449
7	7.890	8.068	545	1301	390	168	2.414
8	7.996	—	971	2354	610	260	2.808
9	8.088	8.384	1581	3897	878	372	3.121
10	8.162	8.562	2407	5964	1194	504	3.388
11	7.828	7.981	271	632	218	96	1.914
12	7.962	8.099	963	2330	612	260	2.914
13	7.990	8.131	1557	3799	884	372	3.414
14	7.962	8.177	553	1325	388	168	2.270
15	8.058	8.417	987	2410	606	260	2.561
16	8.155	8.014	1605	3975	872	372	2.812
17	7.692	—	1333	3447	884	372	7.718
18	7.981	—	2079	5364	1204	504	10.637
19	8.088	—	3073	7913	1572	656	12.075
20	8.149	—	4347	11,182	1988	828	13.242
21	7.852	7.951	513	1269	390	168	3.914
22	7.962	8.158	907	2290	610	260	4.684
23	8.049	8.393	1485	3783	878	372	5.293
24	8.140	8.591	2279	5836	1194	504	5.812
25	7.958	8.114	883	2202	612	260	4.828
26	8.019	8.164	1453	3647	882	372	5.722
27	8.095	8.350	2239	5644	1200	504	6.391

[a] Taken from Ref. 34.

Table 3. Regression parameters and quality of correlations for modeling $\ln \lambda_\beta$ and $\ln \lambda_p$ using W, Sz, PI, Sd, $^1\chi$ indices and employing simple and curvilinear regression analysis

Absorption band	Topological index	Se	R	Q
Simple regression analysis				
$\ln \lambda_\beta$ ($\ln \lambda_p$)	W	0.0978 (0.2254)	0.7523 (0.6182)	7.6922 (2.7426)
	Sz	0.0996 (0.2249)	0.7421 (0.6201)	7.4508 (2.7572)
	PI	0.0910 (0.2274)	0.7904 (0.6090)	8.6857 (2.6816)
	Sd	0.0937 (0.2266)	0.7758 (0.6126)	8.2796 (2.7034)
	$^1\chi$	0.1307 (0.2784)	0.4746 (0.2393)	3.6312 (0.8599)
Curvilinear regression analysis (second polynomial)				
$\ln \lambda_\beta$ ($\ln \lambda_p$)	W	0.0725 (0.2257)	0.8785 (0.6412)	12.1172 (2.8409)
	Sz	0.0747 (0.2257)	0.8705 (0.6413)	11.6532 (2.8413)
	PI	0.0668 (0.2283)	0.8978 (0.6307)	13.4401 (2.7625)
	Sd	0.0720 (0.2256)	0.8803 (0.6417)	12.2263 (2.8444)
	$^1\chi$	0.1221 (0.2784)	0.5934 (0.3233)	4.8599 (1.1612)
Curvilinear regression analysis (third polynomial)				
$\ln \lambda_\beta$ ($\ln \lambda_p$)	W	0.0604 (0.2219)	0.9211 (0.6791)	15.2500 (3.0603)
	Sz	0.0621 (0.2219)	0.9163 (0.6788)	14.7552 (3.0590)
	PI	0.0575 (0.2257)	0.9287 (0.6651)	16.1513 (2.9468)
	Sd	0.0604 (0.2231)	0.9210 (0.6745)	15.2483 (3.0233)
	$^1\chi$	0.0960 (0.2520)	0.7853 (0.5523)	8.1802 (2.1916)

Values in the parenthesis are for modeling $\ln \lambda_p$.

$$\ln \lambda_\beta = 7.7900 + 0.00025\ \mathrm{PI}$$
$$n = 26, \quad \mathrm{Se} = 0.0910, \quad R = 0.7904,$$
$$F = 39.9595, \quad Q = 8.6857 \qquad (1)$$

Similarly, in curvilinear analysis, best results are obtained using PI index. The second and third polynomials are found below:

Table 4. Statistical data supporting Rouvary–El-Basil 'Familial' behavior in the exhibition of $\ln\lambda_\beta$ and $\ln\lambda_p$

Family category	Absorption band	Regression parameters using the indices									
		$^1\chi$		W		Sz		PI		Sd	
		R	Se	R	Se	R	Se	R	Se	R	Se
I_B	$\ln\lambda_\beta$	0.9900	0.0096	0.9121	0.0907	0.9072	0.0931	0.9441	0.0729	0.9461	0.0717
II_B	$\ln\lambda_\beta$	0.9998	0.0025	0.9754	0.0317	0.9744	0.0323	0.9874	0.0003	0.9874	0.0227
III_B	$\ln\lambda_\beta$	0.9863	0.0147	0.9498	0.0279	0.9491	0.0281	0.9678	0.0224	0.9686	0.0222
IV_B	$\ln\lambda_\beta$	0.9966	0.0047	0.9698	0.0416	0.9684	0.0425	0.9841	0.0303	0.9843	0.0301
V_B	$\ln\lambda_\beta$	0.9057	0.0460	0.9799	0.0216	0.9799	0.0216	0.9775	0.0228	0.9776	0.0228
VI_B	$\ln\lambda_\beta$	0.9988	0.0071	0.9806	0.0294	0.9807	0.0294	0.9918	0.0191	0.9918	0.0191
VII_B	$\ln\lambda_\beta$	0.9448	0.0127	0.9675	0.0317	0.9671	0.0319	0.9837	0.0226	0.9837	0.0226
I_p	$\ln\lambda_p$	0.9997	0.0107	0.9173	0.2006	0.9128	0.2058	0.4521	0.1540	0.9538	0.1513
II_p	$\ln\lambda_p$	0.9833	0.0234	0.6840	0.2433	0.6290	0.2434	0.9346	1.2417	0.9346	0.2418
III_p	$\ln\lambda_p$	0.9954	0.0336	0.9955	0.0335	0.9931	0.0349	0.9993	0.0130	0.993	0.0130
IV_p	$\ln\lambda_p$	0.9999	0.0004	0.9958	0.0040	0.9957	0.0041	0.9987	0.0021	0.9988	0.0021
V_p	$\ln\lambda_p$	0.9699	0.0489	0.9742	0.0453	0.9740	0.0455	0.9799	0.0400	0.9599	0.0400
VI_p	$\ln\lambda_p$	0.9999	0.0010	0.9902	0.0427	0.9901	0.0428	0.9953	0.0296	0.9953	0.0295

Table 5. Observed and estimated $\ln\lambda_\beta$ using W, Sz, PI, and Sd indices

Arene no.	$\ln\lambda_\beta$ (Ob.)	$\ln\lambda_\beta$ estimated using							
		W		Sz		PI		Sd	
		Est.	Res.	Est.	Res.	Est.	Res.	Est.	Res.
1	7.524	7.832	−0.308	7.842	−0.318	7.796	−0.272	7.796	−0.272
2	7.700	7.842	−0.142	7.850	−0.150	7.814	−0.114	7.815	−0.115
3	7.830	7.860	−0.030	7.867	−0.037	7.844	−0.014	7.845	−0.015
4	7.916	7.893	0.023	7.898	0.018	7.886	0.030	7.887	0.029
5	8.016	7.942	0.074	7.946	0.070	7.940	0.076	7.941	0.075
6	—	8.011	—	8.014	—	8.006	—	8.006	—
7	7.890	7.890	0.000	7.895	−0.005	7.887	−0.003	7.887	−0.003
8	7.996	7.937	0.059	7.939	0.057	7.942	0.054	7.941	0.055
9	8.088	8.005	0.083	8.004	0.084	8.009	0.079	8.006	0.082
10	8.162	8.097	0.065	8.092	0.070	8.088	0.074	8.083	0.079
11	7.828	7.860	−0.032	7.866	−0.038	7.844	−0.016	7.845	−0.017
12	7.962	7.936	0.026	7.938	0.024	7.943	0.019	7.941	0.021
13	7.990	8.002	−0.012	8.000	−0.010	8.011	−0.021	8.006	−0.016
14	7.962	7.891	0.071	7.896	0.066	7.887	0.075	7.887	0.075
15	8.058	7.939	0.119	7.941	0.117	7.941	0.117	7.941	0.117
16	8.155	8.008	0.147	8.008	0.147	8.008	0.147	8.006	0.149
17	7.692	7.977	−0.285	7.985	−0.293	8.011	−0.319	8.006	−0.314
18	7.981	8.060	−0.079	8.066	−0.085	8.091	−0.110	8.083	−0.102
19	8.088	8.171	−0.083	8.174	−0.086	8.183	−0.095	8.172	−0.084
20	8.149	8.312	−0.163	8.312	−0.163	8.287	−0.138	8.272	−0.123
21	7.852	7.886	−0.034	7.893	−0.041	7.887	−0.035	7.887	−0.035
22	7.962	7.930	0.032	7.936	0.026	7.942	0.020	7.941	0.021
23	8.049	7.994	0.055	8.000	0.049	8.009	0.040	8.006	0.043
24	8.140	8.082	0.058	8.086	0.054	8.088	0.052	8.083	0.057
25	7.958	7.928	0.030	7.933	0.025	7.943	0.015	7.941	0.017
26	8.019	7.991	0.028	7.994	0.025	8.015	0.004	8.006	0.013
27	8.095	8.078	0.017	8.078	0.017	8.090	0.005	8.083	0.012

Second polynomial

$$\ln\lambda_\beta = 7.6600 + 0.000643\ \mathrm{PI} - 2.16\times10^{-7}(\mathrm{PI})^2$$
$$n = 26,\quad \mathrm{Se} = 0.0668,\quad R = 0.8978,\qquad (2)$$
$$F = 99.7112,\quad Q = 13.4401$$

Third polynomial

$$\ln\lambda_\beta = 7.5700 + 0.00115\ \mathrm{PI} - 8.83\times10^{-7}(\mathrm{PI})^2 + 2.27\times10^{-10}(\mathrm{PI})^3$$
$$n = 26,\quad \mathrm{Se} = 0.0575,\quad R = 0.9287,$$
$$F = 150.5454,\quad Q = 16.1513\qquad (3)$$

2.2. Estimation of $\ln\lambda_p$

The results obtained for estimating $\ln\lambda_p$, both by simple as well as curvilinear regression analysis, were of comparatively poor statistics (Table 3). Unlike the estimation of $\ln\lambda_\beta$, here Sz index is found slightly better than the other three indexes. The model based on Sz index is found as,

$$\ln\lambda_p = 7.9800 + 9.7\times10^{-5}\ \mathrm{Sz}$$
$$n = 22,\quad \mathrm{Se} = 0.2249,\quad R = 0.6201,$$
$$F = 12.4959,\quad Q = 2.7572\qquad (4)$$

In curvilinear analysis, W, Sz, and Sd indices gave similar results, both in second as well as third polynomial analysis (Table 3). The PI index gave worse results in simple as well as curvilinear analysis.

It is important to mention that all the four topological indices used are less successful in modeling $\ln \lambda_p$. Also, that at present we cannot give convincing proof as to why PI index is better for modeling $\ln \lambda_\beta$ and why Sz index is so for modeling $\ln \lambda_p$. Out of all possibilities this might be due either to the different origins for $\ln \lambda_\beta$ and $\ln \lambda_p$ and/or different information contents in PI and Sz indices.

It is interesting to record that Rouvary and El-Basil[34] used entirely different methodology for calculating $^1\chi$ index. In contrast to the usual procedure involving characterization of the full chemical graph of the species they used the caterpillar tree of the molecule under consideration. Using such $^1\chi$ index they observed that the correlation of $^1\chi$ with $\ln \lambda_\beta$ splits into seven categories, I–VII, and these categories contain the compounds as below:

$$\begin{aligned} &\mathrm{I_B}: \mathbf{1},\mathbf{2},\mathbf{3},\mathbf{4},\mathbf{5}; \quad \mathrm{II_B}: \mathbf{7},\mathbf{8},\mathbf{9},\mathbf{10};\\ &\mathrm{III_B}: \mathbf{11},\mathbf{7},\mathbf{12},\mathbf{13}; \quad \mathrm{IV_B}: \mathbf{11},\mathbf{14},\mathbf{15},\mathbf{16};\\ &\mathrm{V_B}: \mathbf{17},\mathbf{18},\mathbf{19},\mathbf{20}; \quad \mathrm{VI_B}: \mathbf{21},\mathbf{22},\mathbf{23},\mathbf{24};\\ &\mathrm{VII_B}: \mathbf{21},\mathbf{25},\mathbf{26},\mathbf{27} \end{aligned} \tag{5}$$

Note that compound **11** was common for categories $\mathrm{III_B}$ and $\mathrm{IV_B}$. Similarly, compound **21** was common for categories $\mathrm{VI_B}$ and $\mathrm{VII_B}$ while compound **6** was deleted due to no availability of $\ln \lambda_\beta$ value.

Likewise Rouvary and El-Basil[34] demonstrated that correlation $\ln \lambda_p$ with $^1\chi$ resulted into splitting of the data set into six categories, that is, each of these category I_p–VI_p contains the following compounds:

$$\begin{aligned} &\mathrm{I}_p: \mathbf{1},\mathbf{2},\mathbf{3},\mathbf{4},\mathbf{5},\mathbf{6}; \quad \mathrm{II}_p: \mathbf{11},\mathbf{14},\mathbf{15},\mathbf{16};\\ &\mathrm{III}_p: \mathbf{7},\mathbf{9},\mathbf{10}; \quad \mathrm{IV}_p: \mathbf{7},\mathbf{12},\mathbf{13};\\ &\mathrm{V}_p: \mathbf{21},\mathbf{25},\mathbf{26},\mathbf{27}; \text{ and } \mathrm{VI}_p: \mathbf{22},\mathbf{23},\mathbf{24} \end{aligned} \tag{6}$$

Here also compound **7** was common for categories III_p and IV_p. Also, that compounds **8**, **17**, **18**, **19**, **20** were not considered due to unavailability of the data. Thus, Rouvary and El-Basil[34] used **26** compounds for modeling $\ln \lambda_\beta$ and **22** for modeling $\ln \lambda_p$.

Since, Rouvary and El-Basil[34] have not attempted statistical analysis we did so using their new type $^1\chi$ calculated by them. In both the cases very poor statistics were obtained:

2.3. Simple regression for modeling $\ln \lambda_\beta$ and $\ln \lambda_p$

$$\begin{aligned} &\ln \lambda_\beta = 7.8800 + 0.0213\,{}^1\chi\\ &n = 26, \quad \mathrm{Se} = 0.1307, \quad R = 0.4746,\\ &F = 6.9766, \quad Q = 3.6312 \end{aligned} \tag{7}$$

$$\begin{aligned} &\ln \lambda_p = 8.1000 + 0.0434\,{}^1\chi\\ &n = 22, \quad \mathrm{Se} = 0.2784, \quad R = 0.2393,\\ &F = 1.2135, \quad Q = 0.8595 \end{aligned} \tag{8}$$

2.4. Second polynomial for modeling $\ln \lambda_\beta$ and $\ln \lambda_p$

$$\begin{aligned} &\ln \lambda_\beta = 7.7300 + 0.0853\,{}^1\chi - 0.00471({}^1\chi)^2\\ &n = 26, \quad \mathrm{Se} = 0.1221, \quad R = 0.5934,\\ &F = 13.0447, \quad Q = 4.8599 \end{aligned} \tag{9}$$

$$\begin{aligned} &\ln \lambda_p = 7.7800 + 0.2540\,{}^1\chi - 0.0282({}^1\chi)^2\\ &n = 22, \quad \mathrm{Se} = 0.2784, \quad R = 0.3233,\\ &F = 2.3341, \quad Q = 1.1612 \end{aligned} \tag{10}$$

2.5. Third polynomial for modeling $\ln \lambda_\beta$ and $\ln \lambda_p$

$$\begin{aligned} &\ln \lambda_\beta = 7.3700 + 0.3500\,{}^1\chi - 0.0536({}^1\chi)^2 + 0.00241({}^1\chi)^3\\ &n = 26, \quad \mathrm{Se} = 0.0960, \quad R = 0.7873,\\ &F = 39.13, \quad Q = 8.2010 \end{aligned} \tag{11}$$

$$\begin{aligned} &\ln \lambda_p = 6.5600 + 1.5700\,{}^1\chi - 0.0438({}^1\chi)^2 + 0.0358({}^1\chi)^3\\ &n = 22, \quad \mathrm{Se} = 0.2520, \quad R = 0.5523,\\ &F = 8.7782, \quad Q = 2.1916 \end{aligned} \tag{12}$$

Thus, in both the cases improved results were obtained in curvilinear regression analysis. Though, the statistics using $^1\chi$ index of Rouvary and El-Basil[34] was of low quality, the results did show that there $^1\chi$ is better suited for modeling $\ln \lambda_\beta$. Hence, the methodology used by us was effectively better than that of Rouvary and El-Basil[34].

It will be interesting to extend our study considering the 'familial' behavior recommended by Rouvary and El-Basil.[34] That is, to perform regression analysis for modeling $\ln \lambda_\beta$ under the seven families and that of $\ln \lambda_p$ under six families mentioned above.

It is important to state that Rouvary and El-Basil[34] have not performed any statistical analysis. Using $^1\chi$ they have qualitatively demonstrated 'familial' corrections of $\ln \lambda_\beta$, $\ln \lambda_p$, and $^1\chi$. Using these 'families' we have

Table 6. Observed and estimated $\ln \lambda_p$ using W, Sz, PI, and Sd indices

Arene no.	$\ln \lambda_p$ (Ob.)	$\ln \lambda_p$ Estimated using							
		W		Sz		PI		Sd	
		Est.	Res.	Est.	Res.	Est.	Res.	Est.	Res.
1	7.634	7.986	−0.352	7.985	−0.351	7.941	−0.307	7.953	−0.319
2	7.955	8.006	−0.051	8.003	−0.048	7.976	−0.021	7.991	−0.036
3	8.228	8.048	0.180	8.043	0.185	8.035	0.193	8.051	0.177
4	8.457	8.119	0.338	8.113	0.344	8.117	0.340	8.134	0.323
5	8.658	8.227	0.431	8.223	0.435	8.223	0.435	8.241	0.417
6	8.844	8.381	0.463	8.379	0.465	8.352	0.492	8.371	0.473
7	8.068	8.113	0.045	8.106	−0.038	8.120	−0.052	8.134	−0.066
8	—	8.217	—	8.208	—	8.228	—	8.241	—
9	8.384	8.367	0.017	8.356	0.028	8.359	0.025	8.371	0.013
10	8.562	8.569	0.007	8.558	−0.004	8.513	0.049	8.526	0.036
11	7.981	8.046	−0.065	8.041	−0.060	8.036	−0.055	8.051	−0.070
12	8.099	8.215	−0.116	8.206	−0.107	8.229	−0.130	8.241	−0.142
13	8.131	8.361	−0.230	8.348	−0.217	8.362	−0.231	8.371	−0.240
14	8.177	8.115	0.062	8.108	0.069	8.119	0.058	8.134	0.043
15	8.417	8.221	0.196	8.213	0.204	8.226	0.191	8.241	0.176
16	8.014	8.373	−0.363	8.365	−0.351	8.356	−0.342	8.371	−0.357
17	—	8.306	—	8.314	—	8.362	—	8.371	—
18	—	8.489	—	8.500	—	8.518	—	8.524	—
19	—	8.732	—	8.747	—	8.698	—	8.700	—
20	—	9.045	—	9.064	—	8.902	—	8.900	—
21	7.951	8.105	−0.154	8.103	−0.152	8.120	−0.169	8.134	−0.183
22	8.158	8.202	−0.044	8.202	−0.044	8.228	−0.070	8.241	−0.083
23	8.393	8.343	0.050	8.346	0.047	8.359	0.034	8.371	0.022
24	8.591	8.538	0.053	8.546	0.045	8.513	0.078	8.524	0.067
25	8.114	8.196	−0.082	8.193	−0.079	8.229	−0.115	8.241	−0.127
26	8.164	8.335	−0.171	8.333	−0.169	8.361	−0.197	8.371	−0.207
27	8.350	8.528	−0.178	8.527	−0.177	8.516	−0.166	8.524	−0.174

now carried out regression analysis using their $^1\chi$ and the other topological indices (W, Sz, PI, Sd) used by us in the present study. The results are summarized in Table 4. The regression parameters and quality correlation for the 'familial' behavior shows that excellent results are obtained under the splitted categories (families). From the data presented in Table 4 we conclude that under splitted category (Family: Rouvary and El-Basil):

(1) The $^1\chi$ index as calculated by Rouvary and El-Basil[34] gives the best results as compared to other indices (W, Sz, PI, Sd) used.
(2) Correlation power of W and Sz indices are identical.
(3) Correlation power of PI and Sd indices are identical.
(4) Among the topological indices used, W and Sz indices on one hand and PI and Sd indices on the other hand, the latter pair of indices gave slightly better results than the former pair of indices.
(5) The results show that compared to Sz index (introduced as modification of W index for cyclic compound), the PI, and Sd indices are better suited for the cyclic compounds used.

The statistical analysis made by us is in support of Rouvary and El-Basil[34] recommendations in that it was agreed that the arene systems studied correlate according to the family to which they belong and that excellent results obtained demonstrate convincingly that the data reduction technique employed is an effective one. Also, that the number of data points available was probably too small to make meaningful regression analysis, when attempted to do so the correlation coefficients obtained appeared to be remarkably good. Hence, the approach should be viewed as a viable additional new technique in the growing armory of data reduction technique applicable to multiorganic systems.

The excellent statistical results obtained by us for the entire set of arene systems (26/22) employing W, Sz, PI, and Sd indices are superior to the Rouvary and El-Basil[34] findings.

From the discussion made above we observed that the methodology used by us as well as the one used by Rouvary and El-Basil[34] are more successful for modeling $\ln \lambda_\beta$ as compared to the modeling of $\ln \lambda_p$. This may be attributed to the different types of origins of β and p bands of the arene series used (Table 6).

3. Conclusions

The methodology used by us is superior to that used by Rouvary and El-Basil.[34] In our case the methodology used is applicable to the entire set of arene systems for modeling monitoring, and estimation of electronic bands viz., $\ln \beta$ and $\ln p$. In accordance with our methodology, there is no need to consider 'familial' behavior

that is, to split the data set into different categories. However, in the splitted data $^{1}\chi$ as calculated by Rouvary and El-Basil[34] gave better results than the topological indices used by us. The study reported herein is, therefore, well suited for the estimation of electronic spectral bands vis-à-vis toxicity and environmental pollution of arene system.

4. Experimental

4.1. Electronic absorption bands

The electronic absorption bands viz., β and *p* of the arene system are adopted from the work of Rouvary and El-Basil[34] and used by converting them into their natural logarithmic units.

4.2. Topological indices

All the topological indices were calculated using Luko-1 program of Lukovits, Hungarian Academy of Sciences, Budapest, Hungary.

4.3. Regression analysis

Regression analysis were performed using maximum R^2-method[37] using REGRESSION-1 program of Lukovits. The predictive power of the model was investigated using quality factor (*Q*).[38–40]

Acknowledgements

Authors thanks are due to Prof. Istvan Lukovits, Hungarian Academy of Sciences, Hungary for providing softwares for the calculations of topological indices as well as for performing regression analysis. One of the author (P.V.K.) is thankful to Prof. Ivan Gutman for introducing him (P.V.K.) to this fascinating field viz., Chemical Topology and Chemical Graph Theory.

References and notes

1. Trinajstic, N. *Chemical Graph Theory*, 2nd ed.; CRC: Boca Raton, FL, 1992.
2. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic: New York, 1976.
3. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Relationship*; Wiley: New York, 1986.
4. Diudea, M. V.; Khadikar, P. V. *Molecular Topology and its Applications*; Gilotia: New Delhi, India, in press.
5. Todechini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley, VCH: Weinheim, Germany, 2000.
6. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
7. Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. *Indian J. Chem.* **1993**, *32A*, 651.
8. Gutman, I. *Graph Theory Notes*; New York 1994; Vol. 27, p 9.
9. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
10. Khadikar, P. V. *Natl. Acad. Sci. Lett.* **2000**, *23*, 113.
11. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2000**, *29*, 143.
12. Khadikar, P. V.; Agrawal, V. K.; Karmarkar, S. *Bioorg. Med. Chem.* **2002**, *10*, 3499.
13. Khadikar, P. V.; Joshi, S.; Bajaj, A. V.; Mandloi, D. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1187.
14. Thakur, A.; Thakur, M.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem.* **2004**, *12*, 789.
15. Jaiswal, M.; Khadikar, P. V. *Bioorg. Med. Chem.* **2004**, *12*, 1793.
16. Jaiswal, M.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem.* **2004**, *12*, 2477.
17. Jaiswal, M.; Khadikar, P. V.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3283.
18. Khadikar, P. V.; Mandloi, M.; Shrivastava, A.; Phadnis, A. *Oxid. Commun.* **2003**, *26*, 161.
19. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 934.
20. Sharma, A., *Structural Studies of Some Metal Chelates*, Ph.D. Thesis, Indore University: Indore, India, 1999.
21. Khadikar, P. V.; Pathre, S.; Shrivastava, A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2673.
22. Jaiswal, M.; Khadikar, P. V. *Bioorg. Med. Chem.* **2004**, *12*, 1731.
23. Khadikar, P. V.; Bajaj, A. V.; Mandloi, D. *Indian J. Chem.* **2002**, *41A*, 2065.
24. Khadikar, P. V.; Mandloi, D.; Bajaj, A. V. *Oxid. Commun.* **2004**, *27*, 23, **2004**, *27*, 29.
25. Khadikar, P. V.; Sapre, N. S.; Kumar, S.; Patel, R. N. *Indian J. Chem.* **2000**, *39A*, 386.
26. Patel, R. N.; Kumar, S.; Khadikar, P. V. *Natl. Acad. Sci. Lett.* **2000**, *23*, 46.
27. Mishra, A.; Dave, M.; Khadikar, P. V.; Joshi, A. *Indian J. Pure Appl. Chem.* **2000**, *38*, 263.
28. Mishra, A.; Dave, M.; Khadikar, P. V.; Joshi, A. *Asian J. Phys.* **2000**, *9*, 414.
29. Joshi, S.; Singh, S.; Agrawal, V. K.; Mathur, K. C.; Karmarkar, S.; Khadikar, P. V. *Natl. Acad. Sci. Lett.* **1999**, *22*, 189.
30. Khadikar, P. V.; Joshi, S. In *X-ray Spectroscopy and Allied Areas*; Joshi, S. K., Shrivastava, B. D., Deshpande, A. P., Eds.; Narosa: New Delhi, India, 1998; pp 198–203.
31. Joshi, A.; Karmarkar, S.; Khadikar, P. V.; Joshi, S. In *X-ray Spectroscopy and Allied Areas*; Joshi, S. K., Shrivastava, B. D., Deshpande, A. P., Eds.; Narosa: New Delhi, India, 1998; pp 207–211.
32. Khadikar, P. V.; Mandloi, M.; Joshi, S.; Karmarkar, S. *Chem. Acta Turc.* **1998**, *26*, 11.
33. Khadikar, P. V.; Joshi, S. *X-ray Spectrom.* **1995**, *24*, 201.
34. Rouvary, D. H.; El-Basil, S. *J. Mol. Struct. (Theochem)* **1988**, *165*, 9.
35. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
36. Clar, E. *The Aromatic Sextent*; Wiley: London, 1972.
37. Chatterjee, S.; Hadi, A. S.; Price, B. *Regression Analysis by Examples*, 3rd ed.; Wiley: New York, 2000.
38. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
39. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
40. Khadikar, P. V.; Shrivastava, A.; Agrawal, V. K.; Shrivastava, S. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3009.