Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 203–209
Bioorganic & Medicinal Chemistry Letters

QSAR study on pK_a vis-à-vis physiological activity of sulfonamides: a dominating role of surface tension (inverse steric parameter)

Mamta Thakur,[a] Abhilash Thakur,[b] Padmakar V. Khadikar[c,*] and Claudiu T. Supuran[d]

[a]*Department of Chemistry, Softvision College of Biotechnology and Science, 13, Green Park Colony, Indore, India*
[b]*Research Division, Upzon Drugs Pvt. Ltd, Sanwar Road, Indore, India*
[c]*Research Division, Laxmi Fumigation & Pest control, (P) Ltd, 3, Khatipura, Indore 452007, India*
[d]*University of Florence, Departmento di Chimica, Laboratoria di Chemica Bioinorganica, via della lestruccia polo scientifico 50019-Sest Firrentino Firenze, Italy*

Received 23 July 2004; accepted 4 October 2004

Abstract—The paper describes the dominating role of surface tension (ST) on the modeling, monitoring, and estimating pK_a for a large series of 43 substituted sulfonamides. Because of the direct correlation of ST with parachor (Pc) vis-à-vis molecular volume (MV), ST is considered as a steric parameter. Single as well as multi-parametric regressions have indicated that ST has a dominating role in QSAR of the set of sulfonamides used and that excellent results are obtained in multi-parametric regression analysis. The results are discussed critically on the basis of statistical parameters.

1. Introduction

The sulfonamides constitute an important class of drugs, with several types of pharmacological agents possessing antibacterial, anticarbonic anhydrase, diuretic, hypoglycemic, and antithyroid activity among others.[1–3] A large number of structurally novel sulfonamide derivatives have ultimately been reported to show substantial antitumor or antiviral activity in vitro and in vivo.[1–3] Although they have a common chemical motif of aromatic/heterocyclic or amino acid sulfonamide, there are a variety of mechanisms of their biological action, such as carbonic anhydrase inhibition, cell cycle perturbation in the G1 phase, disruption of microtubule assembly, functional suppression of the transcriptional activator NF-Y, and angiogenesis (matrix metalloproteinase, MMP) inhibition among others.[1–4] Some of these compounds selected via elaborate preclinical screenings or obtained through computer-based drug design, are currently being evaluated in clinical trials.[1–4]

The biological activity of the sulfonamides depends upon the way (strength) with which they bind to their receptor/enzyme. This ability of binding depends upon proton-ligand formation constants of the sulfonamides; more commonly expressed by pK_a of the sulfonamides.[5,6] Like many other organic compounds acting as drugs, the physiological activity of the sulfonamide also depends upon their pK_a. Prompted by these results, we have undertaken the present investigation in that we have carried out QSAR (quantitative structure activity relationship) study on a set of sulfonamides (Table 1) with a view to investigate the dominating role of surface tension (ST). This physicochemical parameter is directly related to parachor (Pc), which in turn is related to molar volume (MV). Thus, we can treat ST, as a steric parameter. Consequently the present study is based on the dominating role of steric effect on the exhibition of pK_a, that is for modeling, monitoring, and estimating pK_a of the sulfonamides. Such a study will be helpful to those interested in investigating physiological activity of sulfonamides. We have, therefore, considered a large set of 43 sulfonamides and adopted their pK_a as reported in the literature.[7,8] The results as discussed below

Keywords: QSAR; pK_a; Proton-ligand formation constant; Steric parameter; Surface tension; Regression analysis; Topological index; Sulfonamide.
* Corresponding author. Tel.: +91 731 2531906; fax: +91 731 2763618; e-mail addresses: abhilashthakur@yahoo.com; pvkhadikar@rediffmail.com

0960-894X/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.bmcl.2004.10.032

Table 1. Substituents, observed pK_a and indicator parameters for sulfonamides used in the present study

Compd no.	X	Y	pK_a	I_X	I_Y	I_N
1	H	H	9.10	0	0	0
2	H	4-OMe	9.42	0	1	0
3	H	4-Me	9.35	0	1	0
4	H	4-F	8.90	0	0	0
5	H	4-Cl	8.47	0	0	0
6	H	4-Br	8.50	0	0	0
7	H	3-Br	8.25	0	0	0
8	H	3-NO_2	7.50	0	0	0
9	3-NO_2	H	7.93	0	0	0
10	3-NO_2	4-OMe	8.44	0	1	0
11	3-NO_2	4-Me	8.27	0	1	0
12	3-NO_2	4-F	7.85	0	0	0
13	3-NO_2	4-Cl	7.51	0	0	0
14	3-NO_2	4-Br	7.42	0	0	0
15	4-Cl	H	8.75	0	0	0
16	4-Cl	4-OMe	9.19	0	1	0
17	4-Cl	4-Me	9.02	0	1	0
18	4-Cl	4-F	8.61	0	0	0
19	4-Cl	4-Cl	8.30	0	0	0
20	4-Cl	4-Br	8.24	0	0	0
21	4-Cl	3-NO_2	7.19	0	0	0
22	4-F	H	8.85	0	0	0
23	4-F	4-OMe	9.32	0	1	0
24	4-F	4-Me	9.20	0	1	0
25	4-F	4-F	8.73	0	0	0
26	4-F	4-Cl	8.41	0	0	0
27	4-F	4-Br	8.38	0	0	0
28	4-F	3-NO_2	7.27	0	0	0
29	4-Me	4-OMe	9.80	1	1	0
30	4-Me	4-Me	9.65	1	1	0
31	4-Me	H	9.34	1	0	0
32	4-Me	4-F	9.23	1	0	0
33	4-Me	4-Cl	8.78	1	0	0
34	4-Me	3-NO_2	7.80	1	0	0
35	4-NH_2	H	10.29	0	0	1
36	4-NH_2	4-Me	10.53	0	1	1
37	4-NH_2	3-Me	10.44	0	1	1
38	4-NH_2	4-Cl	9.76	0	0	1
39	H	4-N	6.66	0	0	0
40	3-NO_2	4-NO_2	5.51	0	0	0
41	4-Cl	4-NO_2	6.24	0	0	0
42	4-F	4-NO_2	6.45	0	0	0
43	4-Me	4-NO_2	6.78	1	0	0

I_X = accounting for alkyl substitution at X, I_Y = accounting for alkyl substitution at Y, I_N = accounting for amide substitution.

establishes the dominating role of ST and that excellent results are obtained for multi-parametric model for that some distance-based topological indices together with some indicator parameters are needed.

2. Results and discussion

In our earlier results we have shown that we can use distance-based topological indices for modeling inhibition activities of benzenesulfonamides.[9–13] In addition, some physicochemical parameters were also used successfully for modeling the inhibitory activity.[14,15] Our earlier studies[16–18] related to other type of organic compounds acting as drugs, indicated that one can use topological indices for modeling their pK_a. Consequently, our results herein relates to the use of surfaces tension (ST), distance based topological indices (W, $^1\chi$, J, Sz) along with indicator parameters (I_X, I_Y, I_I) for modeling, monitoring and estimating pK_a of the sulfonamides used. The details of these molecular descriptors are given in the experimental section. The set of 43 sulfonamides and their pK_a are same as reported in the literature[7,8] (Table 1). In doing QSAR study we have used maximum R^2 method and followed step-wise regression analysis[19] and used the correlation matrix (Table 3). This matrix shows that W, Sz, and $^1\chi$ are highly linearly correlated and that these topological indices along with ST can be used to yield a multi-parametric model for modeling pK_a of the sulfonamides used (Fig. 1).

The simple regression analysis has shown that among the molecular descriptors used (Table 2), surface tension (ST) gives better results. This, therefore, indicates the dominating role of ST in modeling pK_a of the sulfonamides used. The mono-parametric model is found as below:

$$\begin{aligned} \mathrm{p}K_a &= 15.0504 - 0.1120(\pm 0.0207)\mathrm{ST}, \\ n &= 43,\ \ \mathrm{Se} = 0.8739,\ \ R = -0.6450, \\ F &= 29.231,\ \ Q = -0.7400 \end{aligned} \tag{1}$$

Here and there after, n—number of compounds, Se—standard error of estimation, R—simple correlation coefficient, and R^2_A-adjustable R-square. Eq. 1 indicates that pK_a is inversely proportional to ST, that is, pK_a goes on increasing with decrease in the magnitude of ST. Since ST is related to parachor (Pc), which in turn is related to molar volume (MV) it (ST) accounts for steric effect. We can, therefore, say the decrease in pK_a is due to increase in the steric effect.

In successive regression analysis we have carried out several bi-parametric regression analyses in three ways:

(i) combination of ST with one of the indicator parameters,
(ii) combination ST with one of the nontopological parameters, and
(iii) combination ST with one of the topological indices.

In all such bi-parametric regression analysis better results than the mono-parametric model discussed above (Eq. 1) are obtained. However, the bi-parametric regression containing ST and I_N yielded excellent model according to the following equation:

$$\begin{aligned} \mathrm{p}K_a &= 16.4519 - 0.1401(\pm 0.0097)\mathrm{ST} \\ &\quad + 2.6977(\pm 0.2145)I_N, \\ n &= 43,\ \ \mathrm{Se} = 0.3971,\ \ R = 0.9394, \\ R^2_A &= 0.8765,\ \ F = 150,\ \ Q = 2.36 \end{aligned} \tag{2}$$

Looking to the sample size, this is an excellent model. Only a single physicochemical property viz ST is capable of modeling pK_a excellently. This model once again establishes the dominating role of ST in modeling, monitoring, and estimating pK_a of the benzene sulfonamides used. Like earlier case here also the coefficient ST term

Table 2. Topological indices and molecular parameters related to the sulfonamides used in the present study

Compd no.	W	$^1\chi$	J	Sz	MV	ST
1	447	7.68347	1.84162	681	178.5	54.6
2	646	8.61532	1.80081	976	202.5	52.0
3	538	8.07732	1.82718	820	194.8	52.1
4	538	8.07732	1.82718	820	182.7	52.8
5	538	8.07732	1.82718	820	190.5	56.5
6	538	8.07732	1.82718	820	194.7	57.5
7	528	8.07732	1.86027	800	194.7	57.5
8	726	8.98800	1.87293	1074	190.3	65.8
9	720	8.98800	1.89551	1068	190.3	65.8
10	986	9.91985	1.85781	1460	214.3	61.7
11	843	9.38185	1.88173	1254	206.6	62.3
12	843	9.38185	1.88173	1254	194.6	63.7
13	843	9.38185	1.88173	1254	202.3	63.7
14	843	9.38185	1.88173	1254	206.5	68.1
15	536	8.07732	1.83724	818	190.5	56.5
16	758	9.00917	1.80106	1148	214.5	53.8
17	638	8.47116	1.82474	974	206.7	54.0
18	638	8.47116	1.82474	974	194.7	54.8
19	638	8.47116	1.82474	974	202.4	58.3
20	638	8.47116	1.82474	974	206.6	59.3
21	847	9.38185	1.86872	1258	202.3	67.3
22	536	8.07732	1.83724	818	182.7	52.8
23	758	9.00917	1.80106	1148	206.7	50.5
24	638	8.47116	1.82474	974	199.0	50.6
25	638	8.47116	1.82474	974	186.9	51.1
26	638	8.47116	1.82474	974	194.7	54.8
27	638	8.47116	1.82474	974	198.9	55.8
28	847	9.38185	1.86872	1258	194.6	63.7
29	758	9.00917	1.80106	1148	218.8	50.1
30	638	8.47116	1.82474	974	211.1	50.2
31	536	8.07732	1.83724	818	194.8	52.1
32	638	8.47116	1.82474	974	199.0	50.6
33	638	8.47116	1.82474	974	206.7	54.0
34	847	9.38185	1.86872	1258	206.6	62.3
35	536	8.07732	1.83724	818	180.8	64.9
36	638	8.47116	1.82474	974	197.1	61.3
37	627	8.47116	1.85516	952	197.1	61.3
38	638	8.47116	1.82474	974	192.7	66.4
39	756	8.98800	1.79994	1134	190.3	65.8
40	756	8.98800	1.79994	1134	202.2	77.2
41	880	9.38185	1.80014	1324	202.3	67.3
42	880	9.38185	1.80014	1324	194.6	63.7
43	880	9.38185	1.80014	1324	206.6	62.3

W = Weiner index, $^1\chi$ = first order Randic connectivity index, J = Balaban index, Sz = Szeged index, MV = molar volume, ST = surface tension.

in Eq. 2 is negative and thus has the same meaning as discussed for Eq. 1. However, the coefficient of I_N term is positive thus accounting its positive role in the exhibition of pK_a.

Looking to such an excellent result there was no need to attempt for further regression analysis. However, with a hope of obtaining still better results we have carried out several tri-parametric regression analyses and the one containing ST, I_N and J yielded significantly improved statistics:

$$\begin{aligned} \mathrm{p}K_a &= -0.3152 - 0.1546(\pm 0.0082)\mathrm{ST} \\ &\quad + 2.7653(\pm 0.1706)I_N + 9.4458(\pm 1.9088)J, \\ n &= 43,\ \ \mathrm{Se} = 0.3152,\ \ R = 0.9632, \\ R_A^2 &= 0.9227,\ \ F = 166.926,\ \ Q = 3.06 \end{aligned} \quad (3)$$

Here, the coefficient of ST and I_N have the same sign as that of Eq. 2 and thus carry the same meaning. In addition, the coefficient of added Balaban index (J) is positive. This index is the most discriminating topological index and is actually extended connectivity index introduced by Randic. In spite of it being most discriminating index very little work is done of QSAR using J. However, our attempt of using of J indicates it to be the best for modeling pK_a of benzohydroxamic acids.[20] It is worthy to mention that the values of Balaban index (J) do not substantially increase with the molecular size and the number of rings present. Also, it is a variant of Randic connectivity index, represents extended connectivity and is a good descriptor for the shape of the molecule. Thus, the positive coefficient of J in Eq. 3 indicates that all such factors are useful for the exhibition of pK_a of the sulfonamides used.

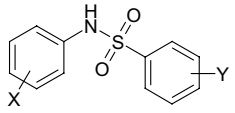

Figure 1. General structure of the sulfonamides used in the present study.

Table 3. Correlation matrix

	pK_a	W	$^1\chi$	J	Sz	MV	ST	I_X	I_Y	I_N
pK_a	1.000									
W	−0.5847	1.000								
$^1\chi$	−0.5738	0.9956	1.000							
J	−0.0765	0.1880	0.2447	1.000						
Sz	−0.5755	0.9991	0.9918	0.1553	1.000					
MV	−0.0646	0.5690	0.5603	−0.1404	0.5844	1.000				
ST	−0.6451	0.5335	0.5449	0.3476	0.5171	−0.02971	1.000			
I_X	0.1232	0.0566	0.0392	−0.1468	0.0664	0.3874	−0.2972	1.000		
I_Y	0.5177	0.0823	0.0901	−0.1804	0.0927	0.5149	−0.3732	0.0065	1.000	
I_N	0.5157	−0.1976	−0.2012	0.0069	−0.1922	−0.2237	0.2304	−0.1412	0.1577	1.000

For symbols used see Tables 1 and 2.

Addition of the indicator parameter I_Y during the stepwise regression analysis yielded a tetra-parametric regression expression with improved statistics; no other tetra-parametric model was found better than this model, which contains ST, I_N, J and I_Y as the correlating parameters. This model is found as below:

$$\begin{aligned} pK_a &= -1.1240 - 0.1442(\pm 0.0080)ST \\ &\quad + 2.6263(\pm 0.1581)I_N + 9.6607(\pm 1.7063)J \\ &\quad + 0.3537(\pm 0.1072)I_Y, \\ n &= 43,\ \ Se = 0.2818,\ \ R = 0.9715, \\ R^2_A &= 0.9379,\ \ F = 159.634,\ \ Q = 3.45 \end{aligned} \quad (4)$$

The physical significance associated with the coefficients of ST, I_N and J is the same as discussed for Eq. 3. The added indicator parameter I_Y has a positive coefficients indicating its favorable role in the exhibition of pK_a. When Szeged index (Sz) is added to Eq. 4 only a slight improved in the statistics is observed. No other topological index yields such an improvement in the statistics; the resulted penta-parametric model containing ST, I_N, J, I_Y, and Sz is found as below:

$$\begin{aligned} pK_a &= -0.9988 - 0.12481(\pm 0.01045)ST \\ &\quad + 2.3857(\pm 0.1728)I_N + 9.4371(\pm 1.5879)J \\ &\quad + 0.5171(\pm 0.1163)I_Y - 8.4817 \\ &\quad \times 10^{-4}(\pm 3.2059 \times 10^{-4})Sz, \\ n &= 43,\ \ Se = 0.2617,\ \ R = 0.9702, \\ R^2_A &= 0.9464,\ \ F = 149.269,\ \ Q = 3.71 \end{aligned} \quad (5)$$

Since, the Szeged index (Sz) is basically for cyclic compounds its negative coefficient (Eq. 5) indicates that decrease in the number of cycle favors the exhibition of pK_a. However, for monocyclic compounds containing tree-like acyclic side chain the variation in the magnitude of Sz and its consequent effect on the biological activity is dooerhed by changes in the side chain for which coincidences of Sz and W indices is well known.

Successive regression analysis yielded a hexa-parametric model having the best statistic than those described above. This model was found to contain ST, I_N, Sz, $^1\chi$, and MV as the correlating parameters. It is observed that during the regression procedure if we use $^1\chi$ and MV there is improvement in the statistics. This is found to be case in present study also and is made clear by the following model:

$$\begin{aligned} pK_a &= -14.7461 - 0.1424(\pm 0.0076)ST \\ &\quad + 2.8341(\pm 0.1285)I_N 5.1420(\pm 1.8217)J \\ &\quad - 0.0098(\pm 0.0020)Sz + 3.0909(\pm 0.7198)^1\chi \\ &\quad + 0.0299(\pm 0.0051)MV, \\ n &= 43,\ \ Se = 0.2140,\ \ R = 9845, \\ R^2_A &= 0.9641,\ \ F = 189.190,\ \ Q = 4.60 \end{aligned} \quad (6)$$

Here, in this model expressed by Eq. 6 both ST and MV are involved. It is worthy to comment upon their mutual relationships. This relationship could only be expressed through their expressions for calculation. These three parameters are calculated by the following equations:

$$\text{Molar Volume} = MV = MW/d \quad (a)$$

$$Pc = (MW/d)ST^{1/4} \quad (b)$$

$$\text{Surface tension} = ST = (Pc/MV)^4 \quad (c)$$

Thus, the ST is inversely proportional to MV. This, therefore, provided physicochemical significant to the model expressed in Eq. 6.

The residual plot has indicated the presence of four outliers in the model expressed by Eq. 6, the deletion of these outliers gave excellent statistics:

$$\begin{aligned} pK_a &= 18.3014 - 0.1416(\pm 0.0057)ST \\ &\quad + 3.0096(\pm 0.1097)I_N 6.1887(\pm 1.3733)J \\ &\quad - 0.0103(\pm 0.0016)Sz + 3.2040(\pm 0.5583)^1\chi \\ &\quad + 0.0316(\pm 0.0039)MV, \\ n &= 39,\ \ Se = 0.1570,\ \ R = 0.9919, \\ R^2_A &= 0.9808,\ \ F = 324.991,\ \ Q = 6.32 \end{aligned} \quad (7)$$

In order to confirm our results we have estimated pK_a for the benzene sulfonamides using Eqs. 5–7 and compared them with the observed values of pK_a. Comparison is shown in Table 4. The residue, that is the difference between observed and calculated pK_a confirms our results.

It is worth to mention that a model with good statistics may not necessary have a good predictive power. The best model is one, which has the best statistics as well as best predictive power. Thus, we have obtained predictive correlation coefficient (R^2_{Pred}) for the models expressed or (Eqs. 5–7) by correlating observed pK_a with the calculated ones. The R^2_{Pred} obtained are presented in Figures 2–4. This shows that the model based on Eq. 7 is the most suitable for modeling pK_a.

The predictive power or the models can also be justified by calculating quality factor Q.[21,22] These Q values fixed as: 3.71, 4.60 and 6.32 for the models expressed by Eqs. 5–7, respectively. This again shows that the model expressed by Eq. 7 is the most appropriate model for modeling pK_a.

3. Conclusions

From the results and discussion made above we concluded that the surface tension (ST) is a dominating parameter for modeling, monitoring, and estimating pK_a for the set of sulfonamides used in the present study. Its reciprocal correlation with molar volume indicates it to be an inverse steric parameter. The results obtained herein also suggest that ST can be used as a parameter for modeling physicochemical activity of sulfonamides vis-à-vis carbonic anhydrase.

Table 4. Experimental and calculated pK_a values using the models expressed by Eqs. 5–7

Compd	pK_a	pK_a[a]	Residual	pK_a[b]	Residual	pK_a[c]	Residual
1	9.10	8.99	0.11	9.36	−0.26	8.61[d]	0.49
2	9.42	9.19	0.22	10.23	−0.81	9.43	−0.01
3	9.35	9.56	−0.21	9.98	−0.63	9.22	0.13
4	8.90	8.96	−0.06	9.52	−0.62	8.74	0.16
5	8.47	8.50	−0.03	9.23	−0.76	8.46	0.01
6	8.50	8.37	0.13	9.21	−0.71	8.45	0.05
7	8.25	8.70	−0.45	9.58	−1.33	8.86[d]	−0.61
8	7.50	7.55	−0.05	8.46	−0.96	7.72	−0.22
9	7.93	7.77	0.16	8.63	−0.70	7.92	0.01
10	8.44	8.11	0.33	8.78	−0.34	7.98[d]	0.46
11	8.27	8.44	−0.17	8.94	−0.67	8.19	0.08
12	7.85	7.75	0.10	8.39	−0.54	7.62	0.23
13	7.51	7.30	0.21	8.10	−0.59	7.35	0.16
14	7.42	7.20	0.22	8.12	−0.70	7.37	0.05
15	8.75	8.59	0.16	9.30	−0.55	8.54	0.21
16	9.19	8.83	0.36	9.86	−0.67	9.05	0.14
17	9.02	9.17	−0.15	9.77	−0.75	8.99	0.03
18	8.61	8.56	0.05	9.29	−0.68	8.49	0.12
19	8.30	8.12	0.18	9.02	−0.72	8.24	0.06
20	8.24	7.99	0.24	9.01	−0.77	8.23	0.01
21	7.19	7.17	0.02	8.00	−0.81	7.23	−0.04
22	8.85	9.06	−0.21	9.59	−0.74	8.82	0.03
23	9.32	9.24	0.08	10.10	−0.78	9.27	0.05
24	9.20	9.60	−0.40	10.02	−0.82	9.22	−0.02
25	8.73	9.02	−0.29	9.59	−0.86	8.77	−0.04
26	8.41	8.56	−0.15	9.29	−0.88	8.49	−0.08
27	8.38	8.43	−0.05	9.28	−0.90	8.48	−0.10
28	7.27	7.62	−0.35	8.28	−1.01	7.49	−0.22
29	9.80	9.29	0.51	10.52	−0.72	9.71	0.09
30	9.65	9.65	0.00	10.44	−0.79	9.66	−0.01
31	9.34	9.14	0.20	10.06	−0.72	9.30	0.04
32	9.23	9.08	0.15	10.02	−0.79	9.22	0.01
33	8.78	8.66	0.12	9.77	−0.99	8.99	−0.21
34	7.80	7.79	0.01	8.84	−1.04	8.07	−0.27
35	10.29	9.93	0.36	10.65	−0.36	10.06	0.23
36	10.53	10.65	−0.12	11.27	−0.74	10.66	−0.13
37	10.44	10.95	−0.51	11.64	−1.20	11.07[d]	−0.63
38	9.76	9.49	0.27	10.42	−0.66	9.80	−0.04
39	6.66	6.81	−0.15	7.50	−0.84	6.65	0.01
40	5.51	5.39	0.12	6.23	−0.72	5.41	0.10
41	6.24	6.47	−0.23	7.00	−0.76	6.12	0.12
42	6.45	6.92	−0.47	7.28	−0.83	6.39	0.06
43	6.78	7.09	−0.31	7.84	−1.06	6.97	−0.19

[a] Calculated pK_a from Eq. 5.
[b] Calculated pK_a from Eq. 6.
[c] Calculated pK_a from Eq. 7.
[d] Data points are not included in calculation from Eq. 7.

4. Experimental

4.1. pK_a

The pK_a values for the set of 43 benzene sulfonamides were taken from the literature.[7,8]

4.2. Topological indices

The topological indices (W, Sz, $^1\chi$, J) used in the present investigation were calculated from the molecular graphs of benzene sulfonamides obtained by deleting carbon–hydrogen as well as heteroatom hydrogen bonds from the respective molecular structures of the benzene sulfonamides. The calculations are of those topological indices are available in the literature,[23] therefore, the details of their calculations are not given here. However, below we give the expressions for their calculations:

4.2.1. Wiener index (W). Wiener index $W = W(G)$ of G is defined[23] as the half-sum of the elements of the distance matrix:

$$W = W(G) = 1/2 \sum_{i=1}, \sum_{j=1} (D)_{ij} \quad (8)$$

where, $(D)_{ij}$ is the ijth element of the distance matrix, which denotes the shortest graph-theoretical distance between sites i and j of G.

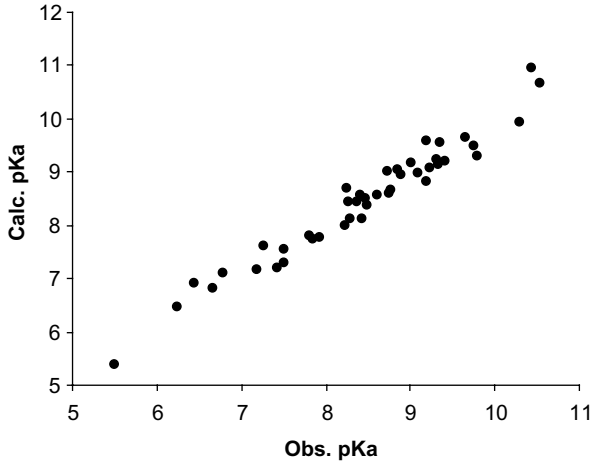

Figure 2. Showing relationship between obs. pK_a and calc. pK_a values using Eq. 5.

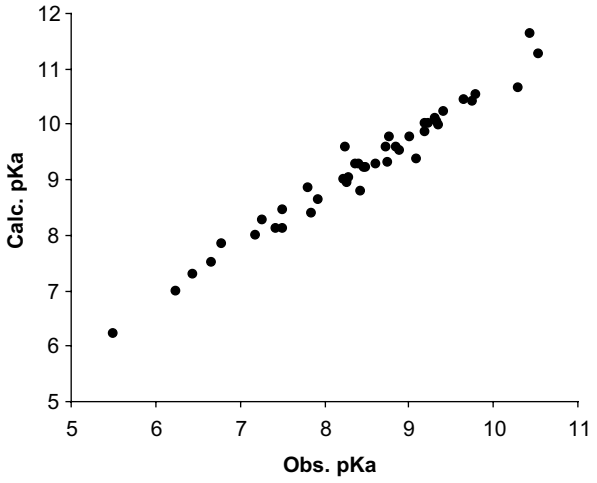

Figure 3. Showing relationship between obs. pK_a and calc. pK_a values using Eq. 6.

4.2.2. Szeged index (Sz). The Szeged index, Sz = Sz(G), is calculated[24–26] according to the following expression:

$$\mathrm{Sz} = \mathrm{Sz}(G) = \sum_{\mathrm{Edges}} n_u \cdot n_v \tag{9}$$

Here n_u is the number of vertices lying closer to one end of the edge $e = uv$; the meaning of n_v is analogous. Edges equidistance from both the ends of an edges, $e = uv$ are not taken into account.

4.2.3. The connectivity index ($^1\chi$). The connectivity index $^1\chi = {}^1\chi(G)$ of G is defined by Randic[27] as:

$$^1\chi = {}^1\chi(G) = \sum_{i,j} [d(i) \cdot d(j)]^{-0.5} \tag{10}$$

4.2.4. Balaban index (J). The Balaban index $J = J(G)$ of G is defined[28] as:

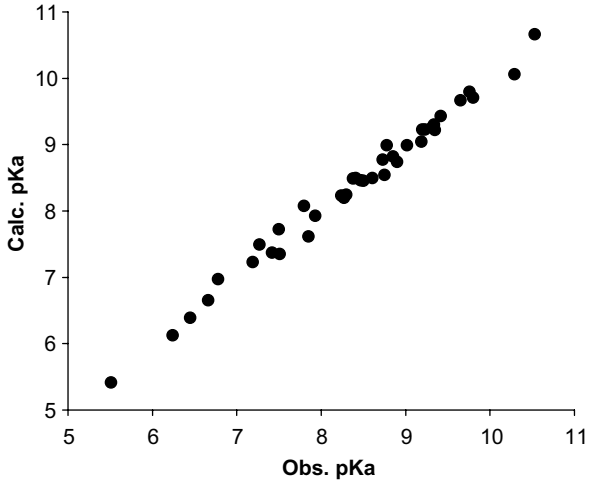

Figure 4. Showing relationship between obs. pK_a and calc. pK_a values using Eq. 7.

$$J = M/\mu + 1 \sum_{\mathrm{bonds}} (d_i \cdot d_j)^{-0.5} \tag{11}$$

where M is the number of bonds in G, μ is the cyclomatic number of G, and d_i ($i = 1, 2, 3, \ldots, N$; N is the number of vertices in G) is the distance sum. The cyclomatic number $\mu = \mu(G)$ of a cyclic graph G is equal to the minimum number of edges necessary to be erased from G in order to transform it into the related acyclic graph. In case of monocyclic graph $\mu = 1$ otherwise it is calculated by means of the following expression

$$\mu = M - N + 1 \tag{12}$$

4.2.5. Molecular descriptors. In present study, molar volume (MV), surface tension (ST), (Table 2) are tested and calculated from computer software acdlabs.[29]

4.3. Indicator parameters

In present study three different indicator parameters[30] are used for the understanding of positional significance of substituents in their activity/property. Indicator parameter I_X accounting for alkyl substitution at X, I_Y accounting for alkyl substitution at Y position and I_N accounting for amide substitution. They assume only two values 1 (in presence) and 0 (in absence) of the structural future.

4.4. Statistical analysis

The regression analysis was performed using maximum R^2 method using the software Provided by Lukovits.

Acknowledgements

The authors thank Prof. Istvan Lukovits for providing the software for making calculations of topological indices as well as for making regression analysis.

References and notes

1. Abbate, F.; Supuran, C. T.; Scozzafava, A.; Orioli, P.; Stubbs, M. T.; Klebe, G *J. Med. Chem.* **2002**, *45*, 3583.
2. Scozzafava, A.; Owa, T.; Mastrolorenzo, A.; Supuran, C. T. *Curr. Med. Chem.* **2003**, *10*, 925–953.
3. Supuran, C. T.; Scozzafava, A. *Exp. Opin. Ther. Pat.* **2000**, *10*, 575.
4. Supuran, C. T.; Scozzafava, A. *Curr. Med. Chem., Immunol. Endocr. Metab. Agents* **2001**, *1*, 61.
5. Khadikar, P. V.; Pol, B.; Ali, S. M. *Rev. Microbiol.* **1986**, *12*, 291.
6. Khadikar, P. V.; Pol, B.; Ali, S. M. *Bacteriology* **1987**, *22*, 81.
7. Dauphin, G.; Kergowrd, A. *Bull. Soc. Chem. Fr.* **1961**, 468.
8. Hansch, C.; Kurup, A.; Garg, R.; Gao, H. *Chem. Rev.* **2001**, *101*, 819.
9. Agrawal, V. K.; Shrivastava, S.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem.* **2003**, *11*, 3553.
10. Thakur, A.; Thakur, M.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem.* **2004**, *12*, 789.
11. Jaiswal, M.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem.* **2004**, *12*, 2477.
12. Jaiswal, M.; Khadikar, P. V.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3283.
13. Agrawal, V. K.; Bano, S.; Supuran, C. T.; Khadikar, P. V. *Eur. J. Med. Chem.* **2004**, *39*, 593.
14. Agrawal, V. K.; Khadikar, P. V. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 447.
15. *Carbonic Anhydrase, its Inhibitors and Activators*; Supuran, C. T., Scozzafava, A., Conway, J., Eds.; CRC Press: Boca Raton, FL (USA), 2004; pp 1–363.
16. Karmarkar, S.; Khadikar, P. V.; Mandloi, M.; Joshi, S.; Agrawal, V. K. *Indian J. Chem.* **2001**, *40A*, 12.
17. Karmarkar, S.; Khadikar, P. V.; Agrawal, V. K.; Mathur, K. C.; Mandloi, M.; Joshi, S. *Proc. Indian Acad. Sci. (Chem. Sci.)* **2000**, *112*, 43.
18. Khadikar, P. V.; Karmarkar, S.; Joshi, S.; Gutman, I. *J. Serb. Chem. Soc.* **1996**, *61*, 89.
19. Chaterjee, S.; Hadi, A. S.; Price, B. *Regression Analysis by Examples*, 3rd ed.; Wiley VCH: New York, 2000.
20. Khadikar, P. V.; Sharma, S.; Sharma, V.; Joshi, S.; Lukovik, L.; Kaveeshwar, M. *Bull. Soc. Chem. Belg.* **1997**, *106*, 767.
21. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
22. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
23. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
24. Gutman, I. *Graph Theory Notes*; New York, 1994; Vol. 27, p 9.
25. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
26. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *Commun. Math. Comput. Chem. (MATCH)* **2001**, *43*, 7.
27. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
28. Balaban, A. T. *Chem. Phys. Lett.* **1982**, *89*, 399.
29. ACD-Lab software for calculating the referred physicochemical parameters; Chemsketh S.O, www.Acdlabs.com.
30. Todeschini, R.; Consonno, V. *Handbook of Molecular Descriptors*; Wieley-VCH: Weinheim (GER), 2000.