Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 405–411
Bioorganic & Medicinal Chemistry Letters

QSAR study on the antibacterial activity of some sulfa drugs: building blockers of Mannich bases

Dheeraj Mandloi,[a] Sheela Joshi,[b] Padmakar V. Khadikar[c,*] and Navita Khosla[b]

[a]*Institute of Engineering and Technology, Devi Ahilya Vishwavidyalaya, Khandwa Road, Indore 452017, India*
[b]*School of Chemical Sciences, Devi Ahilya Vishwavidyalaya, Takshshila Campus, Khandwa Road, Indore 452017, India*
[c]*Research Division, Laxmi Fumigation and Pest Control Pvt. Ltd, 3, Khatipura, Indore 452007, India*

Received 20 July 2004; accepted 21 October 2004

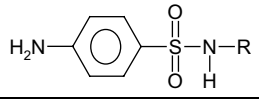
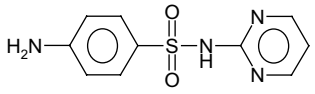
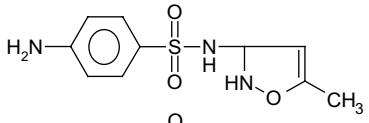
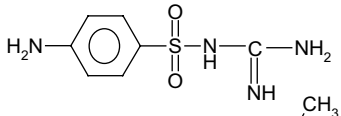
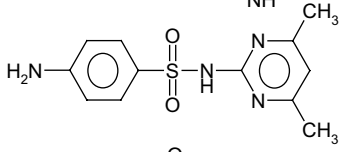
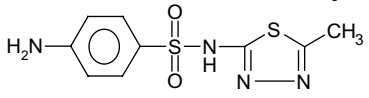
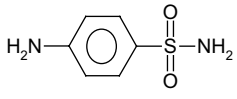
Abstract—Sulfa drugs are building blockers of several types of Mannich bases. Consequently, the antibacterial activities of sulfa drugs are reported in this paper, which will help in explaining and understanding antibacterial activities of Mannich bases. Reported QSAR is carried out using distance-based topological indices and discussed critically on the basis of statistical parameters.

1. Introduction

The importance of sulfa drugs (sulfonamides) is well established in pharmaceutical chemistry and drug design. This class of drugs is well known as antibacterial, carbonic anhydrase inhibitors, anti-cancerous and also as anti-inflammatory agents. Consequent to these physiological activities of sulfa drugs they are used as building blockers for making Mannich bases.[1,2]

In our earlier studies[3–16] several Mannich bases were synthesized from the sulfa drugs and have evaluated for their biological significance and toxicity, in particular antibacterial activity. However, to understand the biological potential of the derived Mannich bases, the same should be known for their building blockers, for example, sulfa drugs. This can be done more efficiently by investigating quantitative structure–activity relationship (QSAR) study using distance-based topological indices.[17–21] Such a study for the sulfa drugs used in the present study (Table 1) are not reported in the literature. Thus, the present work deals with the antibacterial activity of sulfa drugs (Table 1) against *E. coli*, *K. pneumonae* and *B. subtilis*. In doing so we have used distance-based topological indices: Wiener (W);[17] Szeged index (Sz),[20,21] first-order connectivity index ($^1\chi$)[18] and Balaban index (J).[19] The details of these indices are given in Section 4.

Table 1. Molecular structures of the sulfa drugs used in the present study

	General structure	
1	Sulfadiazine (S-1)	
2	Sulfamethoxazole (S-2)	
3	Sulfaguanidine (S-3)	
4	Sulfadimidine (S-4)	
5	Sulfamethiazole (S-5)	
6	Sulfanilamide (S-6)	

Keywords: QSAR; Mannich base; Topological index; Regression analysis; Topological drug design; Antibacterial activity.
* Corresponding author. Tel.: +91 731 2531906; fax: +91 731 2763618; e-mail addresses: dheeraj_m@sancharnet.in; pvkhadikar@rediffmail.com

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.10.076

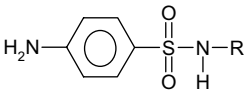


Figure 1. General structure of sulfa drugs used in the present study (see Table 1 for more details).

Before discussing the results obtained in the present study, it is worthy to mention that QSAR methodology is very useful in screening a large library of possible drug candidates for selectivity and potency.[22–27] Mathematical models are formed that correlate molecular structure to an activity or property of interest. Molecular structure is encoded through the generation of descriptors, which are numerical values corresponding to topological, geometric or electronic structural features. The goal of QSAR methodology is to develop several models to predict activity/property/toxicity using correlation analysis employing statistical techniques. In the present study we have used simple as well as multiple regression analysis using maximum R^2 methods[28] for modelling antibacterial activity of sulfa drugs with the general structure as given in Figure 1.

2. Results and discussion

The details of molecular structures of sulfa drugs used in the present study are given in Table 1 and their antibacterial activities against *E. coli*, *K. pneumonae* and *B. subtilis* are present in Table 2. Here, we have used average zone of inhibition in mm at four different concentrations as antibacterial activity. The best should have been that the activity should have been used as log of ED_{50}. However, earlier also such type of activity were presented. In some cases the activity was reported as '+' or '++'. However, it become essential to make statistical analysis of the data related to the zone of inhibition. The best would be make Student *t*-test. The Student *t*-test can be performed using the following expression:

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

The data presented in Table 2 and the above expression indicated that for small sample we are 98% sure that the true mean potency level of the antibacterial activity against each of the bacteria used falls between 8.91 and 9.09. Therefore, *t*-test permits us to use effective zone of inhibition as the antibacterial activity.

The calculated values of distance-based topological indices: W, Sz, $^1\chi$, J, $\log RB$ are summarized in Table 3.

A perusal of Table 2 shows that only five sulfa drugs are effective against *E. coli*; however, against *K. pneumonae* only three viz., **1**, **2** and **5** are effective. In case of *B. Subtilis* sulfa drug **6** is not effective. In obtaining QSAR models we have used average value of zone of inhibition to account for their antibacterial activities against the three bacteria mentioned earlier. Based on these average values we obtained we can propose the following order of antibacterial activity:

Against *E. coli*

$$\mathbf{6} > \mathbf{2} > \mathbf{5} > \mathbf{1} > \mathbf{4} \quad (1)$$

Against *K. pneumonae*

$$\mathbf{5} > \mathbf{1} > \mathbf{2} \quad (2)$$

Against *B. subtilis*

$$\mathbf{2} > \mathbf{1} > \mathbf{4} > \mathbf{5} > \mathbf{3} \quad (3)$$

It is interesting to record that only sulfa drug **3** is active against *B. subtilis* and it is inactive against other two bacteria. Furthermore, these sequences (order) do not establish any quantitative structure–activity (QSAR) relationship. Therefore, we have made such study using topological indices, which encodes the molecular structures of sulfa drugs numerically. Since, different sulfa drugs are found effective against the three bacteria used we have obtained three different correlation matrices (Table 4) for preliminary investigation of correlatedness

Table 2. Antibacterial activity of the sulfa drugs against *E. coli*, *K. pneumonae* and *B. subtilis* adopted from our earlier work[6b]

No.	Compound	Zone of inhibition in mm				
		Concentration in μg/ml				Average
		10	20	40	80	
E. coli						
1	S-1	9.70	17.20	21.20	24.50	18.14
2	S-2	19.63	22.93	23.23	24.03	22.45
3	S-4	15.10	16.90	18.80	18.56	17.34
4	S-5	16.20	17.43	22.20	25.50	20.32
5	S-6	20.43	22.30	22.86	25.55	22.78
						$\bar{x}$ = 20.21
						σ = 2.4566
K. pneumonae						
1	S-1	27.03	29.23	29.53	25.63	27.83
2	S-2	11.63	14.92	22.43	26.46	18.86
3	S-5	30.33	29.86	29.60	27.76	29.39
						$\bar{x}$ = 25.36
						σ = 5.6828
B. subtilis						
1	S-1	20.86	27.23	26.36	25.93	25.10
2	S-2	22.80	25.46	27.43	27.90	25.90
3	S-3	11.23	10.90	17.90	21.76	15.45
4	S-4	15.40	19.96	21.76	22.08	19.80
5	S-5	18.06	19.13	19.46	20.50	19.29
						$\bar{x}$ = 21.11
						σ = 4.3566

Table 3. Distance-based topological indices calculated for sulfa drugs used in the present study

No.	Compound	Topological indices				
		W	Sz	$^1\chi$	J	$\log RB$
1	S-1	536	818	8.0773	1.8372	162.0572
2	S-2	535	731	7.9712	1.8479	161.6517
3	S-3	307	427	6.4155	2.4613	96.2437
4	S-4	722	1092	8.8650	1.8868	215.2262
5	S-5	535	731	7.9712	1.8479	161.6517
6	S-6	152	236	4.9990	2.3936	48.2757

Table 4. Correlation matrices for the sulfa drugs used

	Activity	*W*	*Sz*	$^1\chi$	*J*	log *RB*
E. coli (*n* = 5)						
Activity	1.0000					
W	−0.74316	1.0000				
Sz	−0.81869	0.98966	1.0000			
$^1\chi$	−0.71278	0.98908	0.96911	1.0000		
J	0.55374	−0.88960	−0.84188	−0.94645	1.0000	
log *RB*	−0.74081	0.99994	0.98875	0.99058	−0.89442	1.0000
K. pneumonae (*n* = 3)						
Activity	1.0000					
W	0.37815	1.0000				
Sz	0.37815	1.0000	1.0000			
$^1\chi$	0.37815	1.0000	1.0000	1.0000		
J	−0.37815	−1.0000	−1.0000	−1.0000	1.0000	
log *RB*	0.37815	1.0000	1.0000	1.0000	−1.0000	1.0000
B. subtilis (*n* = 5)						
Activity	1.0000					
W	0.40917	1.0000				
Sz	0.38494	0.98630	1.0000			
$^1\chi$	0.51942	0.98772	0.96965	1.0000		
J	−0.75239	−0.79654	−0.74201	−0.87764	1.0000	
log *RB*	0.41002	1.0000	0.98647	0.98790	−0.79698	1.0000

of topological indices against the antibacterial activity and also for investigating mutual correlatedness among the topological indices used. We will divide our discussion in the following three different headings based on the bacteria used.

2.1. Antibacterial activity of sulfa drugs against *E. coli*

The data presented in Table 4 show that the topological indices *W*, *Sz*, $^1\chi$ and log *RB* are significantly correlated with the antibacterial activity against *E. coli* and that *Sz* is the best topological index for this purpose. The correlation potential of Balaban index (*J*) is significantly lower than the other topological indices. This shows that we can obtain four mono-parametric models for modelling antibacterial activity against *E. coli* and that mono-parametric model based on *Sz* will be the best for this purpose. This Table 4 also shows that all the five topological indices are highly linearly correlated and thus any combination of these indices in multilinear regression analysis may result with a model suffering from the defect due to colinearity. However, such cases are nicely dealt with by Randic[29] and we will follow his recommendation to explain model containing highly correlated topological indices. Looking to the sample size and following 'Rule of Thumb' we can at the most go for bi-parametric regression analysis.

The regression parameters and quality of correlation for the different mono- and bi-parametric models are given in Table 5. This shows that among the mono-parametric models, the model based on *Sz* gives better results:

$$\begin{aligned}&\text{Antibacterial activity against } E.coli\\&\quad = 24.9010 - 0.0065(\pm 0.0026)Sz\\&\quad n = 5, \quad Se = 1.6295,\\&\quad r = -0.8187, \quad F = 6.098, \quad Q = 0.5024 \qquad (4)\end{aligned}$$

Table 5. Regression analysis and quality of correlation for modelling antibacterial activity of sulfa drugs against *E. coli*

Model	TI	*Se*	R_A^2	*R*	*F*	*Q*
1	*W*	1.8987	—	−0.7432	3.701	0.3914
2	*Sz*	1.6295	—	−0.8187	6.098	0.5024
3	$^1\chi$	1.9903	—	−0.7128	3.098	0.3581
4	*J*	2.3629	—	0.5537	1.327	0.2343
5	log *RB*	1.9061	—	−0.7408	3.649	0.3886
6	*W*, *J*	2.1772	0.2151	0.7795	1.548	0.3580
7	*Sz*, *J*	1.7948	0.4666	0.8563	2.750	0.4771
8	$^1\chi$, *J*	2.0613	0.2964	0.8051	1.843	0.3906
9	*J*, log *RB*	2.1758	0.2161	0.7798	1.551	0.3584
10	*W*, *Sz*	1.1584	0.7778	0.9428	8.001	0.8139
11	*W*, $^1\chi$	2.2654	0.1502	0.7584	1.354	0.3347
12	*Sz*, $^1\chi$	1.6408	0.5542	0.8815	3.487	0.5372

Here and thereafter *n* is the number of compound, *Se*—standard error of estimation, *r*—simple correlation coefficient, *F*—Fisher's statistics and *Q*—quality factor, which is defined[30,31] as the ratio of correlation coefficient to the standard error of estimation, that is $Q = r/Se$.

The coefficient of *Sz* in the mono-parametric model represented by Eq. 1 is negative indicating thereby that antibacterial activity of sulfa drugs against *E. coli* is inversely proportional to the magnitude of *Sz*. This index *Sz* precisely accounts for the cyclic nature of the compound, whose negative coefficient in Eq. 1 indicates that the activity decreases with the increase in number of cycles present in the sulfa drugs. Furthermore, if the sulfa drugs are mono-cyclic containing tree-like acyclic side chain then the variation in *Sz* will be due to variation in side chains for which coincidence with *Sz* and *W* is well known. Thus, the overall interpretation of negative *r* values is that decrease in the magnitude of *Sz* increases the antibacterial activity against *E.coli*.

As stated earlier we have attempted several bi-parametric regressions and the results obtained are presented in Table 5. These results show that the bi-parametric model containing W and Sz gave excellent results in accordance with the following expression:

$$\begin{aligned}&\text{Antibacterial activity against } E.\ coli\\&= 24.32538 + 0.0384(\pm 0.0194)W\\&\quad - 0.0324(\pm 0.0131)Sz\\&n = 5, \quad Se = 1.1584, \quad r = 0.9428,\\&R_A^2 = 0.8889, \quad F = 8.001, \quad Q = 0.8139 \qquad (5)\end{aligned}$$

The physical significance of the negative coefficient of Sz term in the Eq. 5 is the same as discussed for Eq. 4. The positive coefficient of W indicates that tree-like (acyclic) side chain is favourable for the exhibition of antibacterial activity of sulfonamides against *E. coli*. This bi-parametric model though violets 'Rule of Thumb' and be considered statistically good as the coefficients of W and Sz are larger than their standard deviation. Furthermore, bi-parametric model Eq. 5 contains two highly linearly correlated topological indices, viz., W and Sz and, therefore, needs further explanation. The multicolinearity (auto-correlation) occurs when two independent variables are correlated with each other and this problem continue to be of prime concern to theoretical statistician. From a decision makers view point, one has to recognize the following problems and indication of severe multicolinearity:

(1) Incorrect signs of the coefficients,
(2) A change in the values of the previous coefficient, when a new variable is added for the model,
(3) Change to insignificant of a previously significant variable when a new variable is added to the model and
(4) An increase in the standard error of the estimate when a variable is added to the model.

In addition, decision makers should also consider the recommendations of Randic,[29] who stated that selection of descriptors to be used in QSAR studies should not be delegated solely to the computers, although the statistical criteria will continue to be useful for preliminary screening of the descriptors taken from a large pool. Often in a automated selection of descriptors a descriptor will be discarded because it is highly correlated with another descriptor already selected, what is more important is not descriptor parallel to one another, but whether they differ in those parts that are important to QSAR correlations. If they differ in the domain, which is important for the QSAR both descriptors should be retained. If they differ in parts that are not relevant for the correlation, one of them can be discarded. Thus, we observed that none of the four problems/indications of multi-colinearity are present in the model expressed by Eq. 5. Also that both W and Sz indices carry different information content. The W index basically accounts for effect due to acyclic side chain (tree-like structure), while Sz index deals primarily with cyclic nature of the molecule. Hence, the model Eq. 1 though not very good for theoretical statistician, it is excellent from chemical point of view. The coefficients of both W and Sz terms are significantly larger than their respective standard deviation, further supporting that model Eq. 1 is an excellent model for monitoring, modelling and estimating antibacterial activity of sulfa drugs against *E. coli*.

In order to confirm our results we have estimated the antibacterial activity of sulfa drugs against *E. coli* using model expressed by Eq. 5 and compared them with the observed values. The data presented in Table 6 show that the observed and the estimated activities are very close to each other.

The predictive power of the models can be judged from quality factor Q. The Q values are recorded in Table 5. The highest $Q = 0.8139$ for the model expressed by Eq. 5 indicates that it has highest predictive power. Further, conformation regarding predictive power is made by calculating predictive correlation coefficient, R^2_{pred}, which is obtained from the correlation between the observed and the estimated activity. The $R^2_{\text{pred}} = 0.8906$ confirms that the predictive power of the proposed model Eq. 5 is highest.

In support of our results we have also calculated three important statistical parameters: probable error of the coefficient of correlation (PE), least square error (LSE) and Friedman's lack of fit measure (LOF). These parameters are calculated from the following equations and summarized in Table 7.

Table 6. Found and estimated antibacterial activity of sulfa drug against *E. coli* using the best model containing W and Sz indices

No.	Compound	Exp. activity	Estimated activity	Residue	(Residue)^2
1	S-1	18.14	18.64	−0.50	0.25
2	S-2	22.45	21.40	1.05	1.1025
3	S-4	17.34	16.98	0.36	0.1296
4	S-5	20.32	21.40	−1.08	1.1664
5	S-6	22.78	22.60	0.18	0.0324
					$\sum$ 2.6809

Table 7. PE, LSE and LOF values calculated for the derived models for modelling antibacterial activity of sulfa drugs against *E. coli*

Model	TI	PE	LSE	LOF
1	W	0.1335	10.8149	67.5931
2	Sz	0.0983	7.9653	49.7831
3	$^1\chi$	0.1466	11.8836	74.2725
4	J	0.2067	16.7492	104.6825
5	$\log RB$	0.1345	10.8994	68.1212
6	W, J	0.1170	9.4800	14.8125
7	Sz, J	0.0795	6.4424	10.0662
8	$^1\chi$, J	0.1049	8.4976	13.2775
9	J, $\log RB$	0.1168	9.4683	14.7942
10	W, Sz	0.0331	2.6809	4.1889
11	W, $^1\chi$	0.1266	10.2637	16.0370
12	Sz, $^1\chi$	0.0665	5.3842	8.4128

$$\mathrm{PE} = \frac{2}{3}\frac{1-r^2}{\sqrt{n}} \quad (6)$$

where, r—coefficient of correlation and n—number of compounds used.

$$\mathrm{LSE} = \sum (Y_{\mathrm{obs}} - Y_{\mathrm{calc}})^2 \quad (7)$$

where, Y_{obs} and Y_{calc} are the observed and calculated activities. In our case antibacterial activity of the sulfa drugs against *E. coli*

$$\mathrm{LOF} = \frac{\mathrm{LSE}}{\{1-(C+d\cdot p)/n\}^2} \quad (8)$$

where, LSE—least square error, C—number of descriptors + 1, p—number of independent parameters, n—number of compounds used, d—smoothing parameter which controls the bias in the scoring factor between equations with different number of terms and was kept 1.0.

It is argued that if

$r <$ PE, r is not significant;
$r >$ PE, several times at least three times greater correlation is indicated;
$r >$ 6PE, correlation is definitely good.

The values of PE (Table 7) indicates that all the proposed correlations are definitely good and the one expressed by Eq. 5 is the best. The lowest value of both LSE and LOF are also in favour of the proposed model. It is important to mention that one should use LOF directly rather than LSE, the reason being LOF does not decrease with increased number of descriptors and the lowest value is found for an equation with the optimum number of parameters.

2.2. Antibacterial activity of sulfa drugs against *K. penumonae*

As stated earlier only three sulfa drugs are effective against *K. pneumonae* and, therefore, it is not a good example for QSAR study. The data presented in Tables 4, 8–10 also indicate that this is not a good example for drug modelling. However, data did show that all the topological indices are equally worked and qualitatively the antibacterial activity follow the sequence:

$$\mathbf{5} > \mathbf{1} > \mathbf{2} \quad (9)$$

Table 8. Regression analysis and quality correlation for modelling antibacterial activity of sulfa drugs against *K. pneumonae*

No.	TI (S)	*Se*	*r* (*R*)	*F*	*Q*
1	*W*	7.4458	0.3781	0.167	0.0508
2	*Sz*	7.4458	0.3781	0.167	0.0508
3	$^1\chi$	7.4458	0.3781	0.167	0.0508
4	*J*	7.4458	−0.3781	0.167	−0.0508
5	log *RB*	7.4458	0.3781	0.167	0.0508

Table 9. Found and estimated antibacterial activity of sulfa drugs against *K. pneumonae* using the best model containing *W* index

No.	Compound	Exp. activity	Estimated activity	Residue	(Residue)^2
1	S-1	27.85	27.85	0.00	0.00
2	S-2	18.86	24.12	−5.26	27.6676
3	S-5	29.39	24.12	5.27	27.7729
					$\sum$ 55.4405

Table 10. PE, LSE and LOF values calculated for the derived models for modelling antibacterial activity of sulfa drugs against *K. pneumonae*

Model	TI	PE	LSE	LOF
1	*W*	0.3299	55.4405	124.5042
2	*Sz*	0.3299	55.4405	124.5042
3	$^1\chi$	0.3299	55.4405	124.5042
4	*J*	0.3299	55.4405	124.5042
5	log *RB*	0.3299	55.4405	124.5042

2.3. Antibacterial activity of sulfa drugs against *B. subtilis*

Like the case of *E. coli* here also five sulfa drugs are found active against *B. subtilis*. A perusal of Table 2 shows that the antibacterial activity of these five sulfa drugs follow the following sequence:

$$\mathbf{2} > \mathbf{1} > \mathbf{4} > \mathbf{5} > \mathbf{3} \quad (10)$$

However, this sequence doesn't establish any structure–activity relationship. Consequently, we have undertaken simple and multiple linear regression analysis. The correlation matrix (Table 4) indicates that here Balaban index (*J*) is capable of giving statistically significant mono-parametric model. The regression parameters of *J* is the only topological index capable of yielding statistically significant mono-parametric model. This model is found as below:

$$\begin{aligned}&\text{Antibacterial activity against } B.\ subtilis\\ &= 44.9398 - 12.0593(\pm 6.0957)J\\ &n = 5, \quad Se = 3.3139, \quad r = -0.7524,\\ &F = 3.913, \quad Q = 0.2270\end{aligned} \quad (11)$$

This shows that antibacterial activities of sulfa drugs against *B. subtilis* are inversely proportional to the magnitude of *J* index. This *J* index is the extended connectivity index, indicating there by that extended connectivity is not favourable for the exhibition of the activity against the bacteria used. That is, negative r values indicates that the decrease in the magnitude of Balaban index **J** increases antibacterial activity against *B. subtilis*.

Once again following the 'Rule of Thumb' we have ultimately made bi-parametric regression analysis. The results recorded in Table 11 show that all the attempted bi-parametric regressions resulted into statistically significant models and that the model containing *W* and $^1\chi$ is the best model. This model is found as below:

Table 11. Regression analysis and quality of correlation for modelling antibacterial activity of sulfa drugs against *B. subtilis*

Model	TI	*Se*	R_A^2	*R*	*F*	*Q*
1	*W*	4.5904	—	0.4092	0.603	0.0891
2	*Sz*	4.6432	—	0.3849	0.522	0.0829
3	$^1\chi$	4.2990	—	0.5194	1.1080	0.1208
4	*J*	3.3139	—	−0.7524	3.914	0.2270
5	log *RB*	4.5885	—	0.4100	0.606	0.0893
6	*W*, *J*	3.5663	0.3300	0.8155	1.985	0.2287
7	*Sz*, *J*	3.7330	0.2659	0.7956	1.724	0.2131
8	$^1\chi$, *J*	3.6322	0.3050	0.8078	1.878	0.2224
9	*J*, log *RB*	3.5681	0.3293	0.8153	1.982	0.2285
10	*W*, *Sz*	5.5789	−0.6397	0.4245	0.220	0.0761
11	*W*, $^1\chi$	3.3067	0.4240	0.8438	2.472	0.2552
12	*Sz*, $^1\chi$	4.3323	0.0112	0.7111	1.023	0.1641

$$\begin{aligned}&\text{Antibacterial activity against } B.\ subtilis\\&= 94.2415 + 23.1249(\pm 11.8920)^1\chi\\&\quad - 0.1250(\pm 0.0719)W\\&n = 5, \quad Se = 3.3067, \quad r = 0.8438,\\&R_A^2 = 0.4240, \quad F = 2.472, \quad Q = 0.2552 \quad (12)\end{aligned}$$

The positive coefficient of first-order connectivity ($^1\chi$) indicates that first-order branching is favourable for the exhibition of the activity. The Wiener index (*W*) accounts for the shape and size, the negative coefficient of which indicates their unfavourable contribution in the exhibition of the activity.

As in the case of *E. coli*, here also we have used *Q* and R^2_{pred} to estimate predictive power of the model. Table 11 shows that *Q* value is the highest for the model expressed by Eq. 12. Also, the value of 0.8280 for R^2_{pred} supports this finding. Furthermore, the comparison of found and calculated activity (Table 12), and the lowest values of LSE and LOF (Table 13) are in favour of the proposed model Eq. 12.

Table 12. Found and estimated antibacterial activity of sulfa drug against *B. subtilis* using the best model containing *W* and $^1\chi$ indices

No.	Compound	Exp. activity	Estimated activity	Residue	(Residue)^2
1	S-1	25.10	25.00	0.10	0.01
2	S-2	25.90	22.67	3.23	10.4329
3	S-3	15.45	15.43	0.02	0.0004
4	S-4	19.80	19.77	0.03	0.0009
5	S-5	19.29	22.67	−3.38	11.4244
					$\sum$ 21.8686

Table 13. PE, LSE and LOF values calculated for the derived models for modelling antibacterial activity of sulfa drugs against *B. subtilis*

Model	TI	PE	LSE	LOF
1	*W*	0.2482	63.2164	395.1025
2	*Sz*	0.2540	64.6773	404.2331
3	$^1\chi$	0.2177	55.4429	346.5181
4	*J*	0.1294	32.9464	205.9150
5	log *RB*	0.2480	63.1637	394.7731
6	*W*, *J*	0.0998	25.4372	39.7456
7	*Sz*, *J*	0.1094	27.8700	43.5468
8	$^1\chi$, *J*	0.1036	26.3851	41.2267
9	*J*, log *RB*	0.0999	25.4627	39.7854
10	*W*, *Sz*	0.2444	62.2481	97.2626
11	*W*, $^1\chi$	0.0858	21.8686	34.1696
12	*Sz*, $^1\chi$	0.1474	37.5382	58.6534

3. Conclusions

From the results and discussion made above we conclude that the distance-based topological indices can be used successfully for modelling antibacterial activity of sulfa drugs against *E. coli* and *B. subtilis*. However, the set of topological indices used are not good for modelling the activity against *K. pneumonae*. The results also show that different topological indices are responsible for giving statistical significant QSAR models for different bacteria used.

4. Experimental

Antibacterial activity: The antibacterial activity used here are those that are reported by us earlier.[6]

Topological indices: All the distance-based topological indices are calculated from the hydrogen suppressed molecular graphs. These molecular graphs are obtained by dealing all the carbon–hydrogen as well as heteroatom–hydrogen bond from the molecular structures of the sulfa drugs. The details of the calculations of these indices are available in the literature[17–21,32–34] and, therefore, they are not mentioned here.

Statistical analysis: The regression analysis is made using maximum R^2 method.[28]

Acknowledgements

Authors thanks are due to Prof. Istvan Lukovits, Hungarian Academy of Sciences, Budapest, Hungary for providing softwares for calculating topological indices as well as for making regression analysis.

References and notes

1. Tramotini, M.; Angiolini, L. *Mannich Bases, Chemistry and Uses*; CRC: Boca Rota, FL, 1994.
2. Tramontini, M.; Angiolini, L.; Ghedini, N. *Polymer* **1998**, *29*, 771.
3. Joshi, S.; Khosla, N. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3747.
4. Joshi, S.; Khosla, N.; Tiwari, P. *Bioorg. Med. Chem.* **2004**, *12*, 571.
5. Joshi, S.; Khosla, N.; Khare, D.; Tiwari, P. *Acta Pharm.* **2002**, *52*, 197.
6. (a) Khosla, N.; Joshi, S. *Acta Pharm.* **1998**, *48*, 55; (b) Khosla, N. Synthesized and Characterization of Some Mannich Bases. Ph.D. Thesis. D.A. University, Indore, 1996.

7. Joshi, S.; Maskar, S.; Khosla, N.; Bhandari, V. *J. Indian Chem. Soc.* **1997**, *74*, 156.
8. Joshi, S.; Khosla, N. *Indian Drugs* **1995**, *32*, 398.
9. Joshi, S.; Khosla, N. *Indian Drugs* **1994**, *35*, 548.
10. Khosla, N.; Joshi, S. *Indian J. Pharm. Sci.* **1993**, *55*, 198.
11. Dhaneshwar, S. R.; Khadikar, P. V.; Katiyer, J. C.; Dhawan, B. N.; Chaturvedi, S. C. *Indian J. Pharm. Sci.* **1991**, *53*, 207.
12. Dhaneshwar, S. R.; Khadikar, P. V.; Katiyer, J. C.; Dhawan, B. N.; Chaturvedi, S. C. *Indian J. Pharm. Sci.* **1990**, *52*, 261.
13. Dhaneshwar, S. R.; Khadikar, P. V.; Chaturvedi, S. C. *Indian Drugs* **1990**, *28*, 21.
14. Dhaneshwar, S. R.; Khadikar, P. V.; Katiyer, J. C.; Dhawan, B. N.; Chaturvedi, S. C. *Indian Drugs* **1990**, *28*, 24.
15. Dhaneshwar, S. R.; Khadikar, P. V.; Chaturvedi, S. C. *Indian Drugs* **1990**, *27*, 431.
16. Dhaneshwar, S. R.; Khadikar, P. V.; Chaturvedi, S. C. *Indian Drugs* **1990**, *27*, 625.
17. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
18. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
19. Balaban, A. T. *Chem. Phys. Lett.* **1982**, *89*, 399.
20. Gutman, I. *Graph Theory Notes*, New York, **1994**, *27*, 9.
21. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Dumotor, G. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
22. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Research Studies Press: Letchworth (UK), 1976.
23. Martin, Y. C. *Quantitative Drug Design: A Critical Introduction*; Marcel Dekker: New York (NY), 1978.
24. *3D QSAR in Drug Design: Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden, The Netherlands, 1993.
25. Kubinyi, H. QSAR: Hansch Analysis and Related Approach. In *Methods and Principles in Medicinal Chemistry*; Kanch, S., Ed.; VCH: Weinheim (GER), 1994; p 240.
26. *QSPR/QSAR Studies by Molecular Descriptors*; Diuden, M. V., Ed.; Nova Science, 2000.
27. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; John Wiley & Sons: New York, 2000.
28. Chaterjee, S.; Hadi, A. S.; Price, B. *Regression Analysis by Examples*, 3rd ed.; Wiley: New York, 2000.
29. Randic, M. *Croat. Chem. Acta* **1993**, *66*, 289.
30. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
31. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
32. Thakur, A.; Thakur, M.; Khadikar, P. V.; Supuran, C. T.; Sudole, P. *Bioorg. Med. Chem.* **2004**, *12*, 789.
33. Khadikar, P. V.; Agrawal, V. K.; Karmarkar, S. *Oxid. Commun.* **2004**, *27*, 17.
34. Khadikar, P. V.; Sharma, S.; Sharma, V.; Joshi, S.; Lukovits, I.; Kaveeshwar, M. *Bull. Soc. Chem. Belg.* **1997**, *106*, 767.