Acta Microbiologica et Immunologica Hungarica, 50 (4), pp. 385–393 (2003)

QSAR ANALYSIS OF ANTIBACTERIAL ACTIVITY OF SOME 4-AMINODIPHENYLSULFONE DERIVATIVES

V. K. AGRAWAL[1], S. BANO[1] and P. V. KHADIKAR[2]*

[1]Department of Chemistry, A. P. S. University, Rewa 486 003, India,
E-mail: vijay-agrawal@lycos.com
[2]Research Division, Laxmi Pest & Fumigation Pvt. Ltd.,
3 Khatipura, Indore 452 007, India

(Received: 13 January 2000; revised: 1 October 2001; accepted 15 October 2001)

QSAR studies on a set of 36 congeners of 4-aminodiphenylsulfone derivatives with measured inhibition potencies of dihydropterate synthase were made using multiple regression analysis. Conformational entropy in combination with indicator parameters gave excellent results.

Keywords: QSAR studies, molecular modelling, sulfones, 4-aminodiphenylsulfone derivatives, antibacterial activity

The inhibition potencies of 4-aminodiphenylsulfone antibacterial agents were studied earlier using linear free energy as well as molecular modelling methods [1].The earlier results showed that these antibacterials are quite flexible and possess multiple conformational energy minima. It was also observed that the application of molecular shape analysis (MSA) was not successful in generative quantitative structure-activity relationship (QSAR).

Compadre et al. [1] have shown that conformation of 4-aminodiphenylsulfone derivatives plays important role in exhibiting antibacterial activity and investigated which conformations were really active. In doing so they performed uniform conformational scanning at 30° increments for θ_1 and θ_2 (Figure 1) and for torsional rotations in flexible substituents. This conformational analysis of 36 4-aminodiphenylsulfone derivatives (Figure 1, Table I) indicates that these mole-

* Corresponding author.

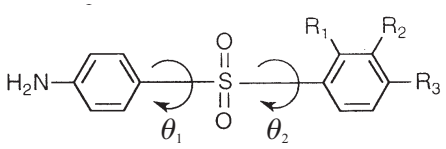

Figure 1. 4-aminodiphenylsulfone derivatives used in the present investigation

cules are quite f1exible with respect to θ_1 and θ_2. They observed that ortho-substituents (R_1) diminished conformational flexibility as compared to meta and/or para substituents. The details are given in Table I. We have therefore, used the conformation in the present study, shown in Figure 1.

We have undertaken the present investigation attempting multiple regression analysis using conformational entropy (S) in combination with indicator parameters for modelling antibacterial activity of a series of 4-aminodiphenylsulfone derivatives.

Materials and methods

A. Biological activity

The antibacterial activities as reported earlier [1], wherein the rate of folate production with respect to time was determined by periodic sampling of the incubation mixtures and quenching the reaction by addition of trichloroacetic acid was adopted in the present study. Thus, amounts of folate produced were determined with folate-requiring *Streptococcus faecium*.That is, the amount of dihydropteroic acid synthesized by the enzyme was measured microbiologically using *Streptococcus* strains that need pteroic acid to grow. Fifty percent inhibitory concentration (I_{50}) was calculated by using a curve-fitting procedure. Inhibition experiments were repeated until computed 95% confidence interval for the I_{50} value fell within 15%. The I_{50} values in micromoles per liter were transformed to pC units (log $1/I_{50}$).These pC values were used and presented in Table I.

B. QSAR descriptors

1. Free space intramolecular conformational entropy (S)

The free space intramolecular conformational entropy (S) [2, 3] was calculated using the following expression:

$$S = -R\sum_{i=1}^{N} P_i \ln P_1 \quad (1)$$

where R is the gas constant, N is the number of conformational states sampled, and P_i is Boltzmann probability of the i^{th} conformational state. The P_i, in term, was calculated from the fundamental statistical mechanics relationships [2]:

$$P_1 = \frac{\exp(-E_1 / RT)}{\sum_{i=1}^{N} \exp(-E_1 / RT)} \quad (2)$$

where E_i is the intramolecular conformational energy of the i^{th} state.

2. Indicator parameters

In an attempt to improve the QSAR models three structurally related dummy parameters [4, 5] (indicator parameters) Ip_1, Ip_2 and Ip_3 were introduced. When $R_1 = -OCH_3$ the parameter was Ip_1 with value of unity, otherwise it is zero. Similarly, Ip_2 is unity when $R_2 = -OCH_3$, otherwise it is also zero. Ip_3 is the indicator parameter whose value is unity when R_3 has a chain, otherwise it is zero.

C. Statistical analysis

Multiple regression analyses [6] for correlating antibacterial activities with the aforementioned molecular descriptors were carried out using Regress-1 software supplied by Prof. István Lukovits, Hungarian Academy of Sciences, Budapest, Hungary. Several multiple regressions were attempted using correlation matrix from this program and the best results are considered and discussed in developing QSAR and hence, for modelling the antibacterial activities of the compounds in the present study.

Results and discussion

The 4-aminodiphenylsulfone antibacterial agents used in the present study are presented in Table I and Figure 1.

The adopted antibacterial activities expressed in pC unit and the free space intramolecular conformational entropies (S) are given in Table I. In addition,

Table I

4-Aminodiphenylsulfone derivatives, their antibacterial activities (pC), conformational entropy (S, cal.) and indicator parameters (Ip_1, Ip_2, Ip_3)

Compound number	R_1	R_2	R_3	pC	S (cal.)	Ip_1	Ip_2	Ip_3
1	CI	H	NH_2	6.32	5.34	0	0	0
2	CH_3	H	NH_2	6.19	9.22	0	0	0
3	OH	H	NHC_2H_5	6.14	14.70	0	0	1
4	H	OH	NH	6.07	21.29	0	0	0
5	H	H	$NHCH_2COOH$	6.06	27.87	0	0	1
6	OCH_3	H	NHC_4H_6	5.99	33.01	1	0	1
7	NH_2	H	NH_2	5.99	36.66	0	0	0
8	OCH_3	H	NHC_3H_7	5.92	37.92	1	0	1
9	H	H	NH_2	5.92	44.59	0	0	0
10	NO_2	H	NH_2	5.87	47.20	0	0	0
11	H	H	$NHCH_3$	5.89	53.84	0	0	0
12	H	H	OH	5.82	60.51	0	0	0
13	OCH_3	H	NHC_2H_5	5.75	60.69	1	0	1
14	H	H	$N(CH_3)_2$	5.75	67.34	0	0	1
15	OCH_3	H	NH_2	5.73	72.23	1	0	0
16	H	OH	NHC_2H_5	5.70	78.83	0	0	1
17	OH	H	NHC_3H_7	5.71	84.38	0	0	1
18	H	OH	NHC_3H_7	5.64	91.00	0	0	1
19	H	H	$NHCH_2COOCH_3$	5.65	97.60	0	0	1
20	CN	H	NH_2	5.65	100.2	0	0	0
21	H	OCH_3	NH_2	5.56	106.5	0	1	0
22	NH_2	H	H	5.65	110.3	0	0	0
23	H	H	NHC_2H_5	5.56	116.9	0	0	1
24	H	OCH_3	NHC_2H_5	5.49	123.2	0	1	1
25	NO_2	H	H	5.61	126.3	0	0	0
26	H	H	COOH	5.44	133.0	0	0	0
27	H.	OCH_3	NHC_3H_7	5.40	139.0	0	1	1
28	H	H	$NHCOCH_3$	5.15	145.6	0	0	1
29	H	H	OCH_3	5.12	152.3	0	0	0
30	H	H	CH_3	5.09	158.9	0	0	0
31	NH_2	H	NO_2	4.93	162.1	0	0	0
32	H	H	$COOCH_3$	4.93	168.7	0	0	1
33	H	H	H	4.92	175.5	0	0	0
34	H	H	$CONHNH_2$	4.90	182.10	0	0	0
35	H	H	CI	4.89	188.8	0	0	0
36	H	H	NO_2	4.51	195.4	0	0	0

pC = activity; S = conformational entropy; and
Ip_1, Ip_2 and Ip_3 are indicator parameters as discussed in the text.

Table I also contains conformational entropy (S) and indicator parameters (Ip_1, Ip_2 and Ip_3).

The perusal of Table I shows that antibacterial activity expressed as pC follows from the sequence:

1>2>3>4>5>6=7>8=9>11>10>12>13=14>15>17>16>19=20=22>
18>25>21=23>24>26>27>28>29>30>31=32>33>34>35>36.

In an attempt to explain the above activity sequence and thus to develop QSAR models the data listed in Table I were used and QSAR was attempted to determine multiple regression analysis.

Before a multivariate analysis is undertaken, it is convenient to tailor the data in certain ways to make the calculations easier. Normally, it is sufficient to preprocess the data by means of auto-scaling and mean-centering the variables [6]. Auto-scaling [7, 8] gives each variable unit variance and hence the same chance to contribute to a calculated model, whereas mean-centering facilitates interpretation. In the present study we have checked for any auto-correlation by obtaining correlation matrix (Table II).

Table II

Correlation matrix for the correlation of antibacterial activities of 4-aminodiphenylsulfone derivatives with various descriptors used in the present study

	pC	Ip_1	Ip_2	Ip_3	S
pC	1.0000	0.2185	–0.0682	0.1473	–0.9705
Ip_1		1.0000	–0.1066	0.2391	–0.2812
Ip_2			1.0000	0.1529	0.1497
Ip_3				1.0000	–0.1437
S					1.0000

Abbreviations used as in Table I.

The perusal of Table III indicates that excellent correlation exists between pC and S. Also, that combinations of S with indicator parameters may improve the quality of the model based solely on S.

From the aforementioned discussion it is clear that a highly significant mono-variate QSAR model exists between pC and S, thus, giving the following linear correlation:

$$pC = 6.2969 - 0.0074\ (\pm 3.2823 \times 10^{-4})\ S \qquad (3)$$

$$n = 36,\quad R = -0.9705,\quad Se = 0.1070,\quad F = 551.248.$$

Table III

Observed and estimated antibacterial activity of 4-aminodiphenylsulfone derivative

Compound number	Observed activity	Estimated (Est.) antibacterial activity using equation (7)	
		Est.	Res.
1	6.320	6.245	0.075
2	6.190	6.216	–0.026
3	6.140	6.176	–0.036
4	6.070	6.125	–0.055
5	6.060	6.078	–0.018
6	5.990	5.980	0.010
7	5.990	6.013	–0.023
8	5.920	5.944	–0.024
9	5.920	5.955	–0.035
10	5.870	5.935	–0.065
11	5.890	5.886	0.004
12	5.820	5.837	–0.017
13	5.750	5.775	–0.025
14	5.750	5.786	–0.036
15	5.730	5.690	0.040
16	5.700	5.701	–0.001
17	5.710	5.660	0.050
18	5.640	5.611	0.029
19	5.650	5.562	0.088
20	5.650	5.543	0.107
21	5.560	5.605	–0.045
22	5.650	–	–
23	5.560	5.605	–0.045
24	5.490	5.419	0.071
25	5.610	–	–
26	5.440	5.481	–0.041
27	5.400	5.300	–0.100
28	5.150	5.364	–0.214
29	5.120	5.207	–0.087
30	5.090	5.108	–0.018
31	4.930	–	–
32	4.930	5.036	0.106
33	4.920	4.985	–0.065
34	4.900	4.937	–0.037
35	4.890	4.887	0.003
36	4.510	–	–

Res. = Residue = difference between observed and estimated activity.

This QSAR indicates that, as entropy (S) decreases, inhibition potency (pC) increases.

Equation (3) accounts for over 93% of the variance in the inhibition measured as reported in Table I. This QSAR may be too good in that the standard devi-

ation of the fit might be less than the uncertainty in the experimental measurements.

In spite of the fact that the model expressed by equation (3) accounts for 93% variance in the inhibition it has four outliers (compounds 22, 25, 31 and 36). When they are deleted from the data set the quality of the model [equation (3)] improved significantly: the correlation coefficient increased from –0.9705 to –0.9847 and the standard error of estimation decreased from 0.1070 to 0.0716. This improve.d model is found as follows:

$$pC = 0.6285 - 0.0072\ (\pm 2.3287 \times 10^{-4})\ S \qquad (4)$$

$$n = 32,\quad R = -0.9846,\quad Se = 0.0716,\quad F = 953.248.$$

Step-wise regression analysis indicated that the quality of the model expressed by equation (4) goes on increasing as we pass from mono- to tri-parametric models and that no statistically significant tetra-parametric models are possible.

Only two bi-parametric models containing S and Ip_1 and S and Ip_2 are found to be statistically significant. The bi-parametric model containing S and Ip_3 results in a model, in which the coefficient of Ip_3 term was much lower than the standard deviation. Such models are not allowed statistically. Furthermore, the correlation coefficient of this model remained the same and the standard error is comparatively increased (Se = 0.072).

The aforementioned bi-parametric models were found as:

$$pC = 6.2863 - 0.0073\ (\pm 2.3343 \times 10^{-4})\ S - 0.0670\ (\pm 0.0380)\ Ip_1 \qquad (5)$$

$$n = 32,\quad R = -0.9861,\quad Se = 0.0693,\quad F = 510.679$$

and

$$pC = 6.2688 - 0.0073\ (\pm 2.1398 \times 10^{-4})\ S + 0.1133\ (\pm 0.0399)\ Ip_2 \qquad (6)$$

$$n = 32,\quad R = -0.9880,\quad Se = 0.0644,\quad F = 592.793.$$

The latter model [equation (6)] containing Ip_2 is slightly better and has positive coefficient for the Ip_2 term. This means that the presence of $–OCH_3$ at R_2 is favorable for the inhibition effect.

Finally, the most excellent tri-parametric model consisting of S, Ip_1 and Ip_2 is shown below:

$$pC = 6.2848 - 0.0074\ (\pm 2.1339 \times 10^{-4})S - 0.0600\ (\pm 0.0346)Ip_1 + 0.1084\ (\pm 0.0387)Ip_2 \qquad (7)$$

$$n = 32,\quad R = -0.9892,\quad Se = 0.0623,\quad F = 423.496.$$

As stated earlier, no other higher parametric model is found statistically significant. It means that the model expressed by equation (7) is the most appropriate model for modelling antibacterial activity of 4-aminodiphenylsulfones.

In order to confirm our findings we have evaluated pC values from equation (7) and compared them with the observed values. Such a comparison is shown in Table III and demonstrated in Figure 2. The difference between observed and calculated pC values and the predictive correlation coefficient ($R^2 = 0.975$) confirms that the model given by equation (7) is the most appropriate model.

Again, in order to further confirmation of the findings, we have calculated the quality factor Q [9] given by the ratio of correlation coefficient (R) to the standard error of estimation (Se) i.e. Q = R/Se. The Q values for the proposed QSAR models [equations (3) to (7)] are found to be –9.0701, 13.7514, 14.1645, 15.3416 and l5.8780, respectively. This shows that our proposed model [equation (7)] is the best model for estimating, monitoring, and modelling the antibacterial activities of the compounds used in the present study. Note that our proposed model is better than that of Compadre and coworkers [1].

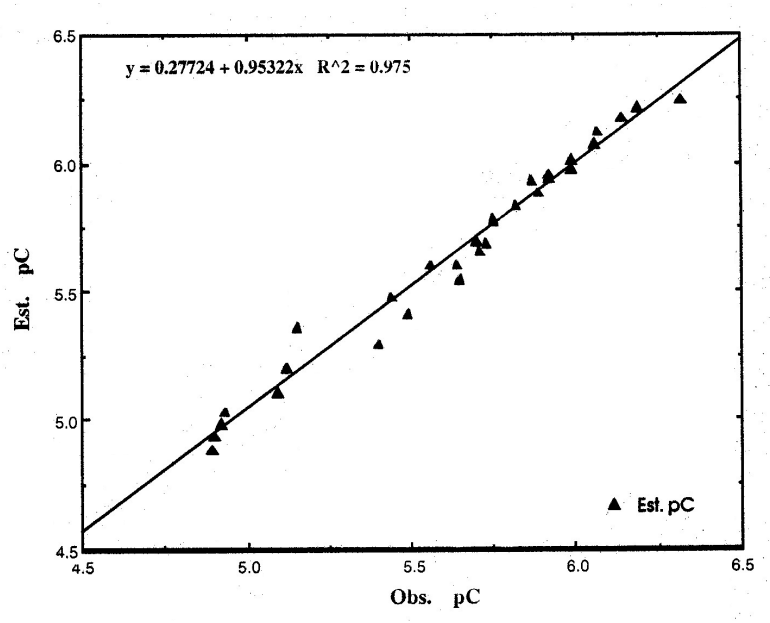

Figure 2

Acknowledgements. One of the authors, V. K. Agrawal is thankful to the AICTE, New Delhi, India for sanctioning a major Research Project on "Computer Aided Drug Designing and Calculation of some Physico-chemical Parameters". Authors are also thankful to Prof. István Lukovits, Hungarian Academy of Sciences, Budapest, Hungary for providing Regress-1 software.

References

1. Compadre, Lopez de R. L, Pearlstein, R. A., Hopfinger, A. J., Seydel, J. K.: A quantitative structure-activity relationship: Analogues of some 4-amino-diphenylsulfone antibacterial agents using linear free energy and molecular modeling methods. J Med Chem **30,** 900 (1987).
2. Hopfinger, A. J.: Conformational properties of macromolecules. Academic Press, New York, 1973.
3. Flory, P. G. J.: Statistical mechanics of chain molecules. Wiley Inter Science, New York, 1969.
4. Wold, S., Alballo, C., Dunn, W. J., Esbeesen, K., Helberg, S., Johansson, E., Sjostrom, M., In: Martend, M. M., Russwurm H., Jr. (eds): Food Research and Data Analysis. Proceeding of the IUFOST Symposium on Data Analysis in Food Research, Oslo, 1982. Applied Sci Publ Ltd. London, 1983, pp.147–188.
5. Stahle, L., Wold, S., In: Ellis, G. P., West, G. B. (eds): Progressive Medicinal Chemistry, Vol. 25, Elsevier Sci Publ, Amsterdam, 1988, pp. 2192–2238.
6. Chaterjee, S., Hadi, A. S., Price, B.: Regressiom analysis by examples, 3rd ed., Wiley, New York, 2000.
7. Topliss, I. G., Edwards, R. P.: Chance factors in studies of quantitative structure-activity relationships. J Med Chem **22,** 1238 (1979).
8. Groebner, D. F., Shanon, P. W.: Business Statistics: A decision making approach. 3rd ed. Merril Pub Company, New York, 1990.
9. Pogliani, L.: Structure property relationships of amino acids and some dipeptides. Amino Acids **6,** 141 (1994).