TWO NEW TOPOLOGICAL INDICES TO DIFFERENTIATE HIGH AND LOW INHIBITORY ACTIVITY OF DIPHENYL SULFONE DERIVATIVES

*Edward Cornwell1 and Gianni Cordano 2

[1]Departamento de Química Inorgánica y Analítica
[2]Departamento de Química Orgánica y Fisicoquímica
Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Olivos 1007, Independencia , Santiago, Chile.
ecornwel@abello.dic.uchile.cl

SUMMARY

This work presents two new topological indices, C(u) and C^A(u), that permit their application to organic molecules containing oxygen, nitrogen, sulfur and halogens. These indices are obtained by modifying Schultz molecular topological index (2-MTI) and are validated by correlating the inhibitory activity of a group of diphenyl sulfone derivatives (SU).

The proposed indices show that:

a.- The use of two types of distance matrices to calculate the proposed indices does not produce significant differences in the calculated biological activity values obtained by using the corresponding regression equations.

b.- In a set of nineteen 4-amino-4'-substituted diphenylsulfones, the new index C(u) permits to differentiate between two SU subsets by correlating it with the biological activity of these derivatives .

c.- By applying the C(u) index to the set of the nineteen-4-aminodiphenylsulfone derivatives and to a subset of fourteen of them, it is possible to demonstrate that the correlation of biological activity with this index is in both cases better than the original Schultz index and the Kier chemical connectivity index (χ_1^v) order-1 valence type.

Key Words: Topological indices; diphenylsulfone, biological activity; electronegativity; actual atomic distance.

INTRODUCTION

4-amino-4'-substituted diphenylsulfone derivatives (SU), are important antibacterial, antimalarial, and antileprotic agents. Some SU compounds exert their biological action by inhibiting competitively with respect to the substrate p-aminobenzoate (PAB), the enzyme dihydropteroate synthase (DHPS), which catalyzes the formation of dihydropteroate from PAB and (hydroxymethyl) dihydropteridine pyrophosphate[4]).

Early studies have shown that SU inhibition is largely modulated by differences in net charges on 4-NH_2 group in SU molecules with respect to PAB (Dq (NH_2)), by differences in net charges on an oxygen atom in SO_2 of SU with respect to an oxygen atom in COO^- of PAB (Dq (O)) and shift differences of the proton measured in ppm for the 4-NH_2 group with respect to the same proton in PAB ($\Delta\delta$ (NH_2))[1]). Although there are other types of charges and factors involved in the modulation of biological activity[1]), they do not contribute significantly to the correlation between biological activity and independent variables treated in this study, so they were not considered.

The inhibitory effect of SU derivatives was taken from literature[1]). This effect on Escherichia Coli was expressed as apE = log(1/Ell_{50}). The SU concentration values corresponding to 50% enzyme activity inhibition were calculated by interpolating linear regression (1/dpm) vs. SU concentration. These values were divided by PAB concentration. They gave origin to the enzyme inhibition index, Ell_{50}. Since apE was not significant in our study, the Ell_{50} value was the unique expression used to indicate inhibitory activity.

From the original work[1]), we conclude that the inhibitory activity, Ell_{50}, of some SU derivatives is modulated by the electronic differences due to resonance and inductive effect of the 4'-substituents and, in addition, this work shows that there is also an important stereochemical factor[1]) that influences the inhibitory activity.

By modifying the Schultz molecular topological index[2-7]), two new, universal topological indices, C(u) and C^A(u), of wider range were developed. They overcome certain restrictions of Schultz topological index and permit their use with every organic substance, e.g., those containing atoms of O, N, S and halogens in our study. Similar modifications to Schultz topological index are supported in a reported work[8]).

Linear correlations of the type y = a + b * X, using Ell_{50} as dependent variable and C(u) as independent variable were performed for a group of nineteen 4'-substituted SU[1]). By checking the distribution of selected, ordered pairs that will permit regression analysis, two groups of SU derivatives were detected that responded to two independent linear regressions with identical mathematical pattern and both having high correlation and Fisher indices (r, F). These groups responded to quite different levels of Ell_{50} . This may be due to the existence of dissimilar inhibitory mechanisms for both SU derivative subsets.

THEORETICAL PROCEDURE

The Schultz topological index, 2-MTI[2-7]) is defined by the following expression,

$$\mathbf{V}([\mathbf{A}]_{(nxn)} + [\mathbf{B}]_{(nxn)}) = [e1, e2,....,en]$$

$$2\text{-MTI} = \sum_{i=1}^{n} e_i$$

Definition of Schultz multiplicative vector (V).

The vector V defined by Schultz contains the d_i^v elements of the L. B. Kier chemical connectivity index χ_1^v calculated according to equation.

$$\mathbf{d}_j^v = (Z^v - h) / (Z - Z^v - 1)$$

where Z^v is the atom valence electron number; Z, the total number of electrons of the atom; and h, the number of hydrogen atoms linked to the carbon atom to which the d_i^v value is calculated[9])

Definition of H. P Schultz adjacency matrix $[A]_{(nxn)}$

The vertex-adjacency matrix $[A]_{(nxn)} = A(G)$, of labeled conected graph G with n vertices, is a square symmetric matrix of order n .

It is defined below.

$a_{ij} = \{1$ if vertices i and j are adjacent, 0 otherwise$\}$

Definition of H. P Schultz distance matrix $[D]_{(nxn)}$

The distance matrix $[D]_{(nxn)} = D(G)$ of a labeled connected graph G with n vertices is a aquare symmetric matrix of order n.

It is defined below.

$d_{ij} = \{ l_{ij}$ if $i \neq j$. 0 otherwise$\}$

where l_{ij} is he length of the shortest path (i.e., the distance) between the vertices i and j in G

The $[e_1, e_2,...,e_n]$ are the elements of the product row (1xn) V vector

Authors' modifications to the adjacency matrix $[A_c]_{(nxn)}$

The diference is that in $[A_c]_{(nxn)}$ matrix, the adjacency of two carbon (i, j) in the the graph G (molecule structure withought H) is indicated by the ratio (>1) between the electronegativity value of any element with respect to the bonded atom[10]), usually carbon (Table 1). This originates a_{ij} elements in the modified adjacency matrix $[A_c]_{(nxn)}$. The value for each term a_{ii} in the adjacency matrix is 0.

Authors' modifications to the distance matrix [BA](nxn)

The modification consist in determining the elements d_{ij} of the distance matrix by assigning the actual topological distance between atoms i and j calculated by means of the Hyperchem software[11]) The elements d_{ii} in the matrix $[B^A]_{(nxn)}$ is 0

Definition of C(u) index

This index is defined by the following expressions.

$$V([A_c]_{(nxn)} + [D]_{(nxn)}) = [C_1\ C_2,....,C_n]$$

$$C(u) = \sum_{i=1}^{n} C_i$$

Definition of $C^A(u)$ index

This index is defined by the following expressions

$$V^*([A_c]_{(nxn)} + [B^A]_{(nxn)}) = [A_1\ A_2 \ldots\ A_n]$$

$$C^A(u) = \sum_{i=1}^{n} C_i$$

RESULTS

Comparison of C(u) and $C^A(u)$ topological indices vs. $\chi 1^V$ L. B. Kier chemical connectivity index

Comparisons were made for subset of fourteen SU derivatives (1) identified by the numbers: 1,3,4,5,6,8,10,11,12,13,14,15,16 and 18 in Table 2. Compounds 2, 7, 9, 17 and 19 did not fit with the χ_1^v L. B. Kier index[9,12]), which is one of the most successfully applied indices to QSPR discipline.

Linear regressions of the type Y=A+B*X were performed for Ell_{50} f (C(u)),

Ell_{50} f(C^A(u)), Ell_{50} f(χ_1^v) [Note: f means function]

The three resulting equations were

$Ell_{50} = -21.9094 + 0.0045C^A(u)$ eq. 1

r = 0.5605
F =5.4968
n = 14

$Ell_{50} = -22.4166 + 0.0068C(u)$ eq. 2

r= 0.5722
F = 5.8414
n = 14

$Ell_{50} = -50.9718 + 10.8037c1V$ eq. 3

r = 0.4653
F= 3.3162
n = 14

From the correlation index, r, and the Fisher statistical parameter, F, we concluded that the topological indices C(u) and $C^A(u)$ are equivalent and superior than χ_1^v .

After analyzing the nineteen substances in a similar manner, i.e., correlating Ell_{50} with the χ_1^v index, and the topological index having the best r value, the following correlation indices 0.2279 and 0.3372 are obtained. Since the correlation values for C(u) and $C^A(u)$ indices are similar, C(u) is preferred because its numerical calculation is easier to obtain.

Evaluation of the C(u) index

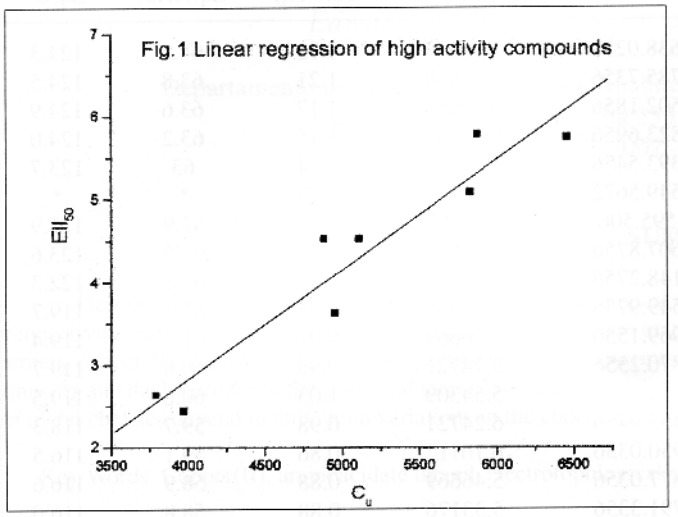
Linear regression between Ell_{50} and C(u) for the nineteen SU derivatives shows clearly two subsets of selected ordered pairs that correspond to SU derivatives with higher relative biological activity (SU-1, compounds 4,11,12,14,15,16,17,18) and those with lower biological activity (SU-2, compounds 1,3,5,6,7,8,9) (Table 3)

Although the linear regression obtained in eq. 4 is significant, the regression in eq 5 only shows a good linear tendency on r and a high error percentage in the independent variable. Four substances do not belong to any subset. The reason for this behavior is not clear yet.

This fact permits to observe the discriminatory power of the index proposed over the set of substances. The corresponding linear regressions for both cases are

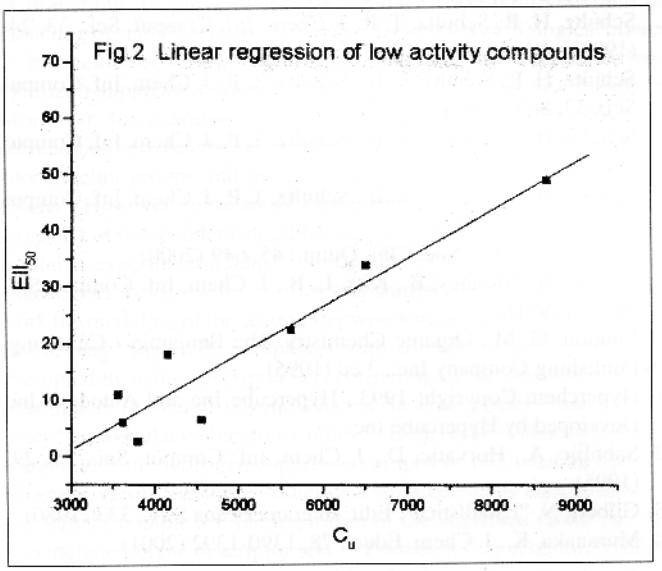
$Ell_{50} = -2.48661\ (\pm 0.88) + 0.00133\ (\pm 1.7 * 10^{-4}) * C(u)$ eq. 4
r = 0.9541
F = 60.9907
n = 8

The plot for this function is represented in Figure 1



and
EII_{50} = -6.51971 (± 10.2) + 0.00653 (± 0.00181) * C(u)
eq. 5
r = 0.84959
F= 12.9729
n = 7

The plot for the second function is represented in figure 2.



CONCLUSIONS

Both equations exhibit high correlation indices even though the dependent variable EII_{50} measures biological activity.

For equations 4 and 5, Fisher values (F) were: 60.9907 and 12.9729 respectively with corresponding critical values of 5.99 and 6.61(α = 0.05; 95% confidence). These values were calculated according to N.Gilbert[13]). The F values obtained by regression analysis were higher than the critical values, which means that the dependent, EI_{50}, and the independent, C(u) variables, exhibit high correlation (with only 5% uncertainty). The higher F value obtained makes equation 4 of greater predictable value[14]) than equation 5.

If $\Delta\delta$ (NH_2), Δ_q (NH_2), Δ_q (O) concepts (1) and χ_1^v (9) are applied as independent variables to the SU derivative subsets with higher and lower biological activity and the corresponding linear regressions are calculated, statistical parameters r, F are obtained. (Table 3). By using the same procedure, but now applying the proposed C(u) index (Eqs. 4 and 5), better values for these statistical parameters are obtained.

Collected information from this work shows that both subsets might exhibit quite different mechanisms of biological activity, which suggests dissimilar stereochemical effects in the inhibition of DHPS enzyme.

Table 1. Relative electronegativity in relation to carbon and sulfur atoms

Connection	Relative
C-C	1.00
C-N	1.19
C-O	1.35
C-Br	1.16
C-S	1.01
S-O	1.33
C-Cl	1.18
C-F	1.57

The connection between atoms means adjacency, not bond

Table 2. Topological and physicochemical parameters for the activity of diphenyl sulfone derivatives

SU	4' Substitution	EII_{50}	$C^A(u)$	C(u)	$\chi 1^v$	$\Delta\delta(NH_2)$	Δq(NH2)	Δq(O)
1	$CON(C_2H_5)_2$	49.06 (7.54)	12855.9447	8638.0256	7.76749	1.12	64.3	124.3
2	CN	67.30 (14.16)	*	4785.7356	5.51669	1.21	63.8	124.5
3	$COOCH_3$	33.96 (11.32)	9808.8407	6502.1856	6.10958	1.17	63.6	124.9
4	COOH	5.09 (1.51)	8608.6746	5823.6956	5.72082	1.15	63.2	124.0
5	$COCH_3$	43.40 (7.54)	7696.7200	5393.5456	5.99721	1.14	63	123.7
6	B_r	11.07 (1.60)	5325.0693	3549.5672	6.02367	1.08	*	*
7	Cl	20.75 (5.66)	*	3595.506	6.40658	1.07	62.9	122.9
8	$CONH_2$	22.60 (3.78)	8320.8125	5607.8756	5.78589	1.12	62.5	123.6
9	F	18.11 (4.53)	*	4148.2756	5.23207	1.04	62.1	122.3
10	OCH_3	6.60 (1.15)	7087.8551	4549.9756	5.65546	0.95	61.1	119.7
11	OH	2.45 (0.32)	5890.9794	3969.1556	5.26669	0.91	61.1	119.4
12	NHCHO	5.79 (0.85)	8581.6834	5870.2556	5.74721	0.98	60.8	119.7
13	CH_3	6.04 (0.94)	5397.3280	3614.0056	5.54309	1.03	60.6	119.5
14	$NHCOCH_3$	5.75 (2.27)	*	6453.9456	6.24721	0.98	59.7	118.3
15	$N(CH_3)_2$	3.63 (1.33)	9506.4191	4950.0356	6.16112	0.86	59.1	116.5
16	NHOH	4.53 (1.36)	7362.8459	4877.0356	5.46669	0.88	58.9	116.6
17	NH_2	2.64 (0.76)	*	3791.3356	5.33176	0.88	58.8	116.0
18	$N(C_2H_5)_2$	2.64 (0.50)	9399.2724	6307.7956	7.31366	0.88	58.8	115.9
19	$NHCH_2CH_3$	4.53 (0.76)	*	5101.6656	6.35375	0.85	58.8	116.0

* mean values not reported in literature (1) or not calculated in this work

Table3. Statistical parameters for the correlation of the two diphenyl sulfone subsets

SU GROUP	**Δδ (NH2)**	**Δq (NH2)**	**Δq (O)**	**χ_1^v**
SU-1 $EII_{50} < 10$				
	n = 8	n = 8	n = 8	n = 8
r	0.50701	0.2 3488	0.48529	0.54233
F	2.07600	0.35035	1.84831	2.49990
SU-2 $EII_{50} > 10$	n = 7	n = 7	n = 7	n = 7
r	0.65927	0.82806	0.69143	0.64668
F	3.84389	8.72601	3.66401	3.59401

REFERENCES

1. De Benedetti, P. G., Iarossi, D., Menziani, C., Caiolfa, U., Frassinetti, C., Cemnano, C., J. Med. Chem., 30, 459 (1987).
2. Schultz, H. P., J. Chem. Inf. Comput. Sci., 29, 227 (1989).
3. Schultz, H. P., Schultz, E. B., Schultz, T. P., J. Chem. Inf. Comput. Sci., 30, 27 (1990).
4. Schultz, H. P., Schultz, T. P., J. Chem. Inf. Comput. Sci., 33, 240 (1993).
5. Schultz, H. P., Schultz, E. B., Schultz, T. P., J. Chem. Inf. Comput. Sci., 33, 863 (1993).
6. Schultz, H. P., Schultz, E. B., Schultz, T. P., J. Chem. Inf. Comput. Sci., 34, 1151 (1994).
7. Schultz, H. P., Schultz, E. B., Schultz, T. P., J. Chem. Inf. Comput. Sci., 35, 864 (1995).
8. Cornwell, E., Bol.Soc. Chil. Quim.; 45, 649 (2000).
9. Hall, L. H., Mohney, B., Kier, L. B., J. Chem. Inf. Comput. Sci., 31, 76 (1991).
10. Loudon, G. M., Organic Chemistry. The Benjamin / Cummings Publishing Company Inc., 3 ed (1995).
11. Hyperchem Copyright 1993., Hypercube Inc and Autodesk Inc. Developed by Hypercube Inc.
12. Sabbljic, A., Horvatic, D., J. Chem. Inf. Comput. Sci., 33, 292 (1995).
13. Gilbert, N. "Estadística". Edit. Interamericana S.A., 333 (1980).
14. Muranaka, K., J. Chem. Educ., 78, 1390-1392 (2001).