Journal of Mathematical Chemistry 11(1992)179–186 

# MOLECULAR TOPOLOGICAL INDEX: AN APPLICATION IN THE QSAR STUDY OF TOXICITY OF ALCOHOLS

Albin JURIĆ and Mirko GAGRO
*Chemistry Department, Agricultural Institute, 43260 Križevci, The Republic of Croatia*

Sonja NIKOLIĆ* and Nenad TRINAJSTIĆ
*The Rugjer Bošković Institute, P.O. Box 1016, 41001 Zagreb, The Republic of Croatia*

## Abstract

An application of the molecular topological index, adapted for heterosystems, in the QSAR study of toxicity of alkyl alcohols on fathead minnows (*Pimephales promelas*) is examined. The obtained QSAR model is comparable to models based on the Wiener number and the connectivity index.

Recently, a novel graph-theoretical index, named the molecular topological index (MTI) or the Schultz index after its originator [1], has been introduced [2]. It has also been extended to heterosystems [3]. Since the MTI index is a molecular descriptor with a number of attractive features [2], we decided to test its applicability in quantitative structure–activity relationships (QSAR) studies. The toxicity of alcohols on fish was selected for this purpose because the problem has already been treated with the connectivity index of Randić [4] with some success [5]. The connectivity index and its variants are the most successful graph-theoretical descriptors to date to be used in QSAR work [6,7] and thus we will be able to investigate how the MTI descriptor performs in comparison with the connectivity index on the same sample. We will also compare the QSAR model based on MTI with the model based on the Wiener number. The Wiener number [8] is another topological index which is used in structure–property–activity studies with considerable success [6,9].

The toxicity of alcohols on fathead minnows (*Pimephales promelas*) are taken from the data base of toxicity of commercial chemicals on fish published by the Center for Lake Superior Environmental Studies [10]. Toxicity tests on fish were conducted under controlled and uniform conditions. Mortality was recorded at 96 h and their $LC_{50}$ data were calculated and reported as g/ℓ or mg/ℓ. Those data were recalculated in molar units (mmol/ℓ) and their logarithms are used as toxicity variables in this work.

*To whom correspondence should be addressed.

The MTI is defined as [11]

$$\mathrm{MTI} = \sum_{i=1}^{N} e_i, \tag{1}$$

where $e_i$ are elements of the row marix

$$\upsilon[\mathbf{A} + \mathbf{D}] = [e_1\, e_2, \ldots, e_N]. \tag{2}$$

**A** is the adjacency matrix, **D** is the distance matrix, and $\upsilon$ is the valency row matrix.

The MTI index is extended to heterosystems [3] by replacing the elements of the adjacency matrix and the distance matrix corresponding to the heteroatoms and heterobonds with the values representing the corrections introduced for the change from the carbon atom to the heteroatom. This was done by considering their atomic numbers.

The corrections related to heteroatoms and heterobonds, denoted as the atomic parameters and the bond parameters, are

$$\text{atomic parameter} = 1 - \frac{Z_C}{Z_i}, \tag{3}$$

$$\text{bond parameter} = \frac{1}{b_r}\,\frac{Z_C^2}{Z_i Z_j}, \tag{4}$$

where $Z_C$ is to atomic number of the carbon and $Z_i$ the atomic number of atom $i$. The values of $b_r$ are 1, 2 and 3 for a single bond, double bond and triple bond, respectively. In the case of the adjacency matrix, the diagonal elements corresponding to heteroatoms are identical to (3) and the off-diagonal elements corresponding to heterobonds to (4). In the case of the distance matrix, the diagonal elements corresponding to heteroatoms are also identical to (3). Since the off-diagonal elements in the distance matrix represent the length of a shortest path between sites $i$ and $j$, eq. (4) represents a weight of a single bond which contributes its weight to the path count between $i$ and $j$.

Tables of the values corresponding to (3) and (4) are available [12] or can be easily computed.

In table 1, we give as an illustrative example the computation of the MTI for 2-methyl-1-propanol.

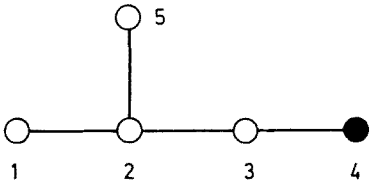
Protić and Sabljić [5], in their QSAR study of toxicity of commercial chemicals, used the zero-valence connectivity index ($^0\chi^v$). This index is defined as [7,13]

$$^0\chi^v = \sum_{i=1}^{N} \left(\delta_i^v\right)^{-0.5}, \tag{5}$$

Table 1

The computation of the MTI for 2-methyl-1-propanol

(1) Molecular graph for 2-methyl-1-propanol. The black dot denotes the position of the oxygen atom in the graph




(2) Its valency matrix

$$\upsilon = [1\ 3\ 2\ 1\ 1]$$

(3) Its adjacency matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0.75 & 0 \\ 0 & 0 & 0.75 & 0.25 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(4) Its distance matrix

$$\mathbf{D} = \begin{bmatrix} 0 & 1 & 2 & 2.75 & 2 \\ 1 & 0 & 1 & 1.75 & 1 \\ 2 & 1 & 0 & 0.75 & 2 \\ 2.75 & 1.75 & 0.75 & 0.25 & 2.75 \\ 2 & 1 & 2 & 2.75 & 0 \end{bmatrix}$$

(5) Its adjacency-plus-distance matrix

$$\mathbf{A} + \mathbf{D} = \begin{bmatrix} 0 & 2 & 2 & 2.75 & 2 \\ 2 & 0 & 2 & 1.75 & 2 \\ 2 & 2 & 0 & 1.50 & 2 \\ 2.75 & 1.75 & 1.50 & 0.50 & 2.75 \\ 2 & 2 & 2 & 2.75 & 0 \end{bmatrix}$$

(6) Its $\upsilon[\mathbf{A} + \mathbf{D}]$ matrix

$$\upsilon[\mathbf{A} + \mathbf{D}] = [14.75\ \ 9.75\ \ 11.5\ \ 14.25\ \ 14.75]$$

(7) Its MTI index

$$\text{MTI} = 65.00$$

where $\delta_i^v$ is the valence delta value of atom $i$. This quantity may be computed by means of the following expression:

$$\delta_i^v = \frac{Z_i^v - H_i}{Z_i - Z_i^v - 1}, \tag{6}$$

where $Z_i$ is the atomic number and $Z_i^v$ the number of valence electrons of atom $i$. $H_i$ in (6) is the number of hydrogen atoms attached to atom $i$.

The Wiener number $W$ is given as a half-sum of the elements of the distance matrix [14]

$$W = \tfrac{1}{2} \sum_i \sum_j (\mathbf{D})_{ij} . \tag{7}$$

In the case of the heterosystems, the distance matrix is constructed as described above. In fact, for a given (hetero)system, the same distance matrix is used to compute MTI and $W$.

The toxicities (in terms of log $LC_{50}$) of 12 alkanols (alkyl alcohols) together with their MTI values, Wiener numbers and zero-order valence connectivity indices are given in table 2.

Table 2

Toxicities of alkanols on fathead minnows and the corresponding MTI values, Wiener numbers and zero-order valence connectivity indices.

| Alkanol | log $LC_{50}$ [a] | MTI | $W$ | $^0\chi^v$ |
|---|---|---|---|---|
| methanol | 2.96 | 3.50 | 0.875 | 1.447 |
| ethanol | 2.50 | 14.50 | 3.625 | 2.154 |
| 1-propanol | 1.88 | 35.75 | 9.375 | 2.861 |
| 1-butanol | 1.37 | 71.00 | 19.125 | 3.569 |
| 2-propanol | 2.16 | 33.50 | 8.375 | 3.025 |
| 2-methyl-1-propanol | 1.28 | 65.00 | 17.125 | 3.732 |
| 1-hexanol | – 0.02 | 199.50 | 54.625 | 4.983 |
| 2-ethyl-1-hexanol | – 0.66 | 370.00 | 102.125 | 6.560 |
| 1-octanol | – 0.97 | 432.00 | 118.125 | 6.397 |
| 1-nonanol | – 1.40 | 597.25 | 162.875 | 7.104 |
| 1-decanol | – 1.82 | 800.50 | 217.625 | 7.811 |
| 1-undecanol | – 2.22 [b] | 1045.75 | 283.375 | 8.518 |

[a] Experimental values are taken from ref. [10].
[b] This value is taken from ref. [15].

We have examined several linear and non-linear correlations between log $LC_{50}$ and three topological indices (TI) used in this work. Here, we report explicitly the linear models and quadratic models (with log TI):

(a) *Linear models*

$$\log LC_{50} = 1.890(\pm 0.280) - 0.005(\pm 0.001)\ \mathrm{MTI}, \tag{8}$$
$$n = 12,\ r = 0.926,\ s = 0.716,\ F = 60.6;$$

$$\log LC_{50} = 1.892(\pm 0.276) - 0.018(\pm 0.002)\ W, \tag{9}$$
$$n = 12,\ r = 0.928,\ s = 0.707,\ F = 62.4;$$

$$\log LC_{50} = 4.115(\pm 0.132) - 0.762(\pm 0.025)\ {}^{0}\chi^{v}, \tag{10}$$
$$n = 12,\ r = 0.995,\ s = 0.193,\ F = 959.3.$$

(b) *Quadratic models (with log TI)*

$$\log LC_{50} = 3.004(\pm 0.192) + 0.393(\pm 0.218) \log (\mathrm{MTI}) - 0.717(\pm 0.057) (\log \mathrm{MTI})^2, \tag{11}$$
$$n = 12,\ r = 0.9986,\ s = 0.106,\ F = 1596.0;$$

$$\log LC_{50} = 2.971(\pm 0.081) - 0.422(\pm 0.134) \log W - 0.700(\pm 0.049) (\log W)^2, \tag{12}$$
$$n = 12,\ r = 0.9989,\ s = 0.095,\ F = 1998.2;$$

$$\log LC_{50} = 2.975(\pm 0.310) + 1.169(\pm 1.159) \log {}^{0}\chi^{v} - 7.309(\pm 0.987) (\log {}^{0}\chi^{v})^2, \tag{13}$$
$$n = 12,\ r = 0.996,\ s = 0.178,\ F = 563.7.$$

From the above results, we learn that the linear model with ${}^{0}\chi^{v}$ is superior to the related models with either MTI or $W$. The quadratic model with $\log {}^{0}\chi^{v}$ represents an insignificant improvement over the linear model. This improvement is not worth the labor required to produce the quadratic model with the $\log {}^{0}\chi^{v}$ value. Therefore, we have confirmed the known results that the linear QSAR model with ${}^{0}\chi^{v}$ represents "an accurate tool for assessing the level of toxicity" [5] of alcohols on fish. A plot of $\log LC_{50}$ versus ${}^{0}\chi^{v}$ is shown in fig. 1.

In the case of models with MTI and $W$ indices, the improvement on going from the linear model to the quadratic model is considerable. The most accurate QSAR model appears to be the quadratic model with $\log W$. However, a pleasant surprise is to notice how good is a quadratic QSAR model with log MTI. Its performance

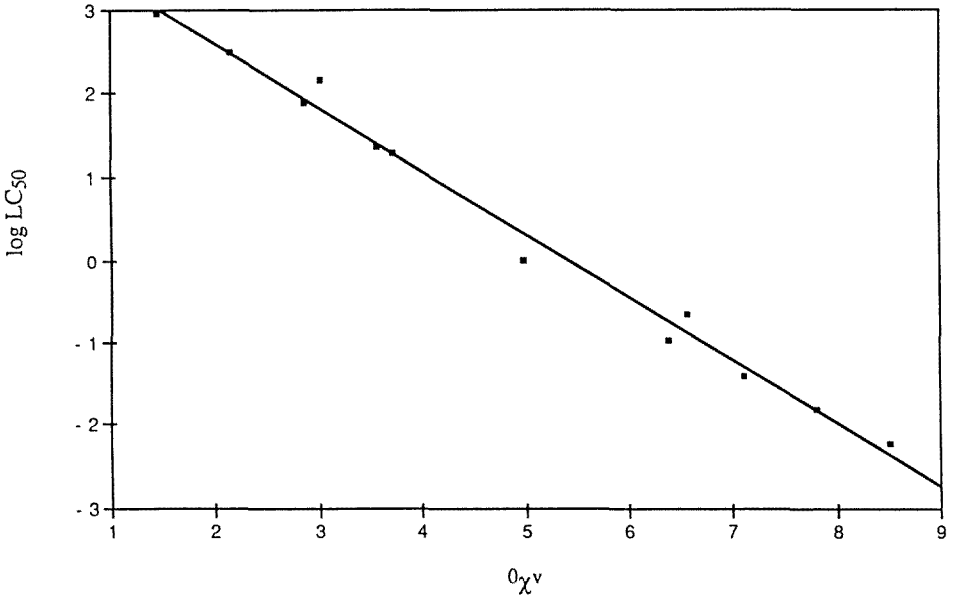



Fig. 1. A plot of log $LC_{50}$ versus ${}^{0}\chi^{v}$.

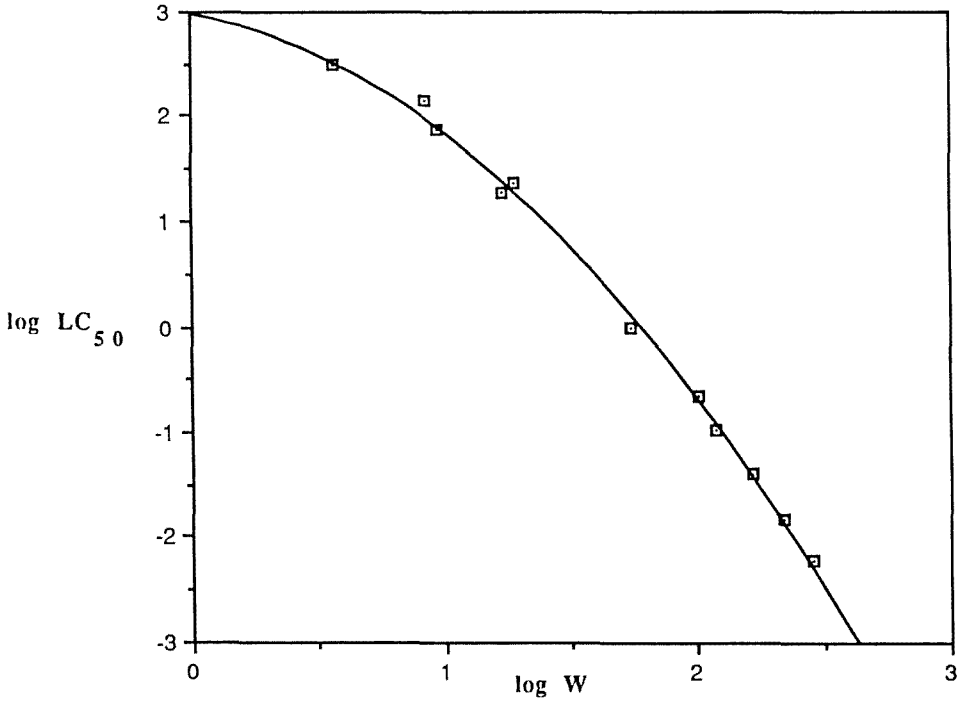



Fig. 2. A plot of log $LC_{50}$ versus log $W$.

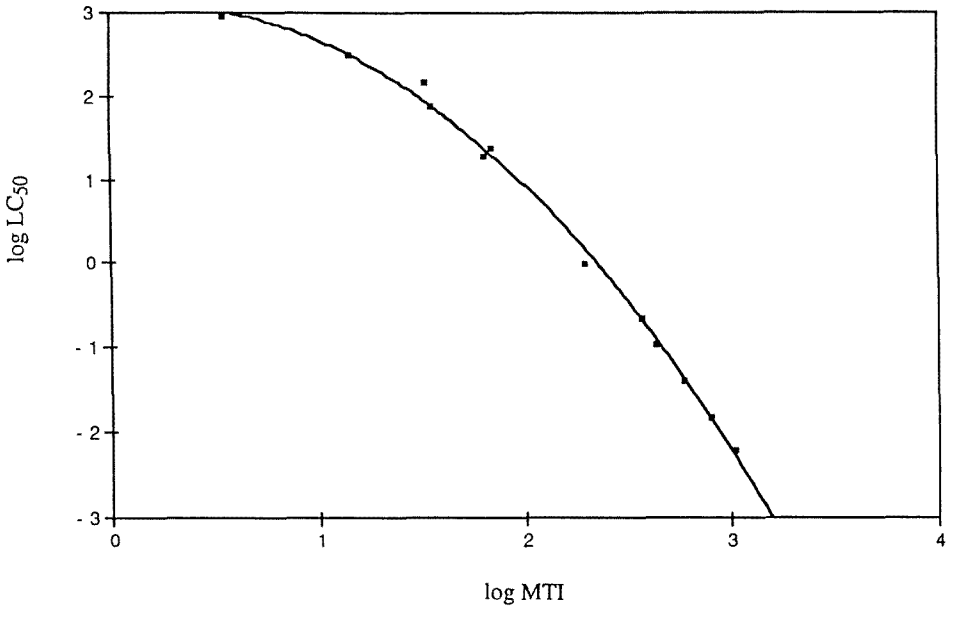



Fig. 3. A plot of log $LC_{50}$ versus log MTI.

is comparable to the achievement by the Wiener number. Plots of log $LC_{50}$ versus log $W$ and log $LC_{50}$ versus log MTI are given in figs. 2 and 3, respectively.

In a way, the above result is not so unexpected because the MTI and $W$ indices are highly intercorrelated descriptors for alkanes [16]. To conclude, we point out that the MTI descriptor shows potential for use in structure–property–activity work. However, more work is needed before the range of its applicability is established. Some research in this direction is already in progress [17].

## Acknowledgements

We thank Dr. T.-S. Balaban (Bucharest) for helpful suggestions. We also thank the referees for their comments. This work was supported by the Croatian Science Fund (Grant No. 1-07-159).

## References

[1] N. Trinajstić, *Chemical Graph Theory*, 2nd revised Ed. (CRC, Boca Raton, 1992), ch. 10.
[2] H.P. Schultz, J. Chem. Inf. Comput. Sci. 29(1989)227.
[3] S. Nikolić, N. Trinajstić and Z. Mihalić, J. Math. Chem., in press.
[4] M. Randić, J. Amer. Chem. Soc. 97(1975)6609.
[5] M. Protić and A. Sabljić, Aquatic Toxicol. 14(1989)47.
[6] D.H. Rouvray, Sci. Amer. 254(1986)40.

[7] L.B. Kier and L.H. Hall, *Molecular Connectivity in Structure–Activity Analysis* (Research Studies Press, Letchworth, 1986).
[8] H. Wiener, J. Amer. Chem. Soc. 69(1947)17.
[9] D.H. Rouvray, in: *Mathematics and Computational Concepts in Chemistry*, ed. N. Trinajstić (Horwood, Chichester, 1986), p. 295.
[10] D.L. Geiger, C.E. Northcott, D.J. Call and L.T. Brooke (eds.), *Acute Toxicities of Organic Chemicals to Fathead Minnows (Pimephales promelas)*, Vols. 1, 2 (Center for Lake Superior Environmental Studies, University of Wisconsin-Superior, Superior, Wisconsin, 1984/1985).
[11] W.R. Müller, K. Szymanski, J.V. Knop and N. Trinajstić, J. Chem. Inf. Comput. Sci. 30(1990)160.
[12] M. Barysz, G. Jashari, R.S. Lall, V.K. Srivastava and N. Trinajstić, in: *Chemical Applications of Topology and Graph Theory*, ed. R.B. King (Elsevier, Amsterdam, 1983), p. 222.
[13] L.B. Kier and L.H. Hall, J. Pharm. Sci. 65(1976)1806.
[14] H. Hosoya, Bull. Chem. Soc. Japan 44(1971)2332.
[15] G.D. Veith, D.J. Call and L.T. Brooke, Can. J. Fish. Aquat. Sci. 40(1983)743.
[16] Z. Mihalić, S. Nikolić and N. Trinajstić, J. Chem. Inf. Comput. Sci. 32(1992)28.
[17] S. Nikolić, work in progress.