Revue Roumaine de Chimie, **2006**, *51*(11), 1131–1135

VERTEX- AND EDGE-WEIGHTED MOLECULAR GRAPHS FOR AMINES

Zoița BERINDE

Nord University of Baia Mare, Department of Chemistry, Str. Victoriei 76, RO-430122 Baia Mare, Roumania,
E-mail : zoita_berinde@yahoo.com

Received December 12, 2001

In the particular case of the bonds realized by the nitrogen atom in the class of amines, the weighted electronic distances (w.e.d.) are calculated. Using the w.e.d., we introduce the weighted electronic connectivity matrix, CEP, associated to the chemical graphs of amines. Using the CEP matrix we then calculate the ZEP topological index, which will be correlated with some properties such as the boiling point (b.p.), the molar refraction (MR) and the partition coefficient (log P), for a set of 40 amines.

INTRODUCTION

A chemical compound can be represented by using the concept of chemical graph. A chemical or molecular graph can itself be represented by several topological matrices.[1-5] The most frequently matrices used are the adjacency matrix of the graph G, $A = A(G)$, and the distance matrix of the graph G, $D = D(G)$. The mathematical modelling of the chemical structures using matrices associated to the molecular graphs has developed rapidly, as these matrices provide a source for obtaining some important molecular descriptors.[6-8] Among these descriptors, the topological indices have been widely used in QSPR (Quantitative Structure-Property Relationship) and QSAR (Quantitative Structure−Activity Relationships). Both adjacency and distance matrices, in their original formulations,[3-5] are appropriate for modelling chemical compounds that do not contain multiple bonds or heteroatoms, because to structurally different molecules, as, for example, isobutane, isobutene, 2-propanol, 2-cloropropane, 2-aminopropane, there corresponds the same chemical graph, and hence, the same adjacency matrix. This poses a very difficult problem to molecular topology: how to represent the multiple bonds and heteroatoms in such a way that different chemical structures should have different representations. A first approach to represent heteroatoms and multiple bonds has been proposed by Kier and Hall,[9] and also by Barysz and collaborators.[10] Basak and collaborators,[11] have treated the presence of heteroatoms by applying the information theory. The method proposed by Balaban,[12] considers the electronegativities, the atomic number and the number of the group in Mendeleev's short form of the periodic chart. The same problem has been also tackled successfully by Randić,[13] and Estrada.[14]

In the present paper we shall treat the presence of heteroatoms, especially those in the amines class by means of the weighted electronic distance (w.e.d.), a new local invariant considered by the author in.[15] By replacing the usual topological distances from the adjacency matrix, by the values of the weighted electronic distances, we obtain a new connectivity matrix, CEP.[15] Using this matrix we introduce a new topological index, ZEP, which is then correlated to the boiling point, the molar refraction and the partition coefficient for a set of 40 amines.

THE WEIGHTED ELECTRONIC DISTANCE

In order to ensure that to different chemical structures will correspond different matrix representations, we consider the concept of weighted electronic distance (w.e.d.) as well as that of the weighted electronic connectivity matrix (CEP). The weighted electronic distance is computed by means of the following formula, obtained in,[15]

$$w.e.d.(i,j) = \frac{1}{b_r} \cdot \frac{Z_i^{'} + Z_j^{'}}{v_i \cdot v_j} \quad , \tag{1}$$

where b_r is the bond weight with values: *1*, for single bond, *2* for double bond, *3* for triple bond and 1.5 for aromatic bond,

- v_i, v_j denote the degree of vertices *i*, and *j*, respectively;
- Z_i' denotes the formal degree of vertex *i*, and it is defined by

$$Z_i' = Z_i \cdot v_i ; \tag{2}$$

- Z_i denotes the order number of atom *i* in Mendeleev's periodic system.

The formal degree Z_i' represents a local vertex invariant (LOVI) in the molecular graph.

Similar formulae which involve the order numbers as well as the multiplicity order of covalent bonds have been used but in a different context by Barysz and collaborators,[10] and by Balaban.[12] Their approaches detect the presence of multiple bonds and heteroatoms but are not able to represent the information related to the neighbourhood of bonds, while our method is very sensitive in this respect. Having in view that for nitrogen we have $Z = 7$, it results that its formal degrees will be $Z_N' = 7$ for primary amines, $Z_N' = 14$ for secondary amines and $Z_N' = 21$ for tertiary amines. The carbon atom linked to nitrogen may be itself primary, secondary or tertiary, having the formal degree Z_j' equal to 12, 18 and 24, respectively. These aspects are illustrated in Figure 1, on the molecular graphs of isopropylamine(G.1), isopropyl-methylamine(G.2) and ethyl-methyl-tert-butylamine(G.3).

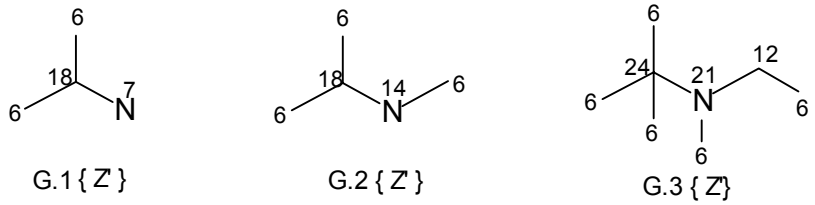


Fig. 1 – The graphs G.1, G.2 , G.3 and formal degrees (Z').

By replacing (2) in (1) and taking into account that $b_r = 1$, because the nitrogen atom establishes only simple bonds in the class of amines, we obtain

$$w.e.d.(i,j) = \frac{Z_i}{v_j} + \frac{Z_j}{v_i} \cdot \tag{3}$$

The weighted electronic distances computed by means of equation (3) for the covalent bonds carbon−carbon and nitrogen−carbon represent local edge invariants (LOEI).
On the edges of each graph are written the w.e.d., computed by (3).

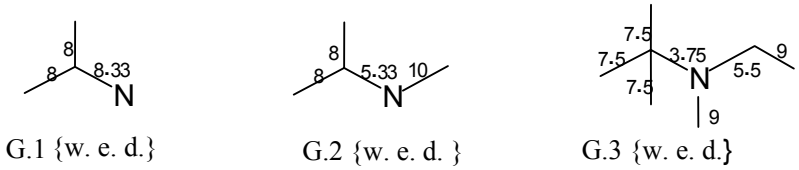


Figure 2. The graphs G.1, G.2, G.3 and w.e.d.

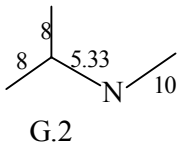
The weighted electronic distances computed by (3) were used to build the weighted electronic connectivity matrix of amines, CEP. The CEP matrix is a symmetric and quadratic $N \times N$ matrix whose entries are given by

$$CEP(G) = \{[CEP]_{ij}\,; i, j \in V(G)\} \tag{4}$$

where $V(G)$ is the set of vertices of the molecular graph *G*, and $E(G)$ is the set of edges, and

$$CEP_{ij}=w.e.d(i, j),\ if\ i\neq j \text{ and } (i,j) \in E(G) \text{ and } CEP_{ij}=0, \text{ otherwise.} \tag{5}$$

Figure 3 shows the CEP matrix for isopropyl-methylamine whose graph, G.2, is depicted in the same figure, with the values of w.e.d. computed by formula (3). SEP_i



G.2

$$\text{CEP(G.2)} = \begin{bmatrix} 0 & 8 & 0 & 0 & 0 & 8 \\ 8 & 0 & 5.33 & 0 & 8 & 21.33 \\ 0 & 5.33 & 0 & 10 & 0 & 15.33 \\ 0 & 0 & 10 & 0 & 0 & 10 \\ 0 & 8 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 3 – The CEP matrix for isopropyl-methylamine.

The CEP matrix of connectivity is a very important source of topological indices that could be used in QSPR studies.[17] Such a topological index has been introduced in:[15]

$$\text{ZEP} = \sum_{i=1}^{n} (\text{SEP}_i)^{1/2} \tag{6}$$

where

$$\text{SEP}_i = \sum_{j=1}^{n} [\text{CEP}]_{ij}\,,\ \ i = 1, 2, \ldots, n\,. \tag{7}$$

For the *G.2* graph the sums on the rows of *CEP*, denoted by SEP_i, are also given in Figure 3.

RESULTS AND DISCUSSIONS

The statistical study regarding the QSPR equations for the amines class using the topological index ZEP given by (6), have been done over a set of 40 amines by considering three physicochemical properties: the boiling point (*b.p.*), the molar refraction (*MR)* and the partition coefficient (*log P*), whose values have been taken from.[14, 18] The names, the ZEP values computed by (6) as well as the theoretical values of *b.p., MR* and *log P* are given in Table 1.

The correlations of ZEP with these properties, by linear or nonlinear regression curves, give good correlation for *b.p.*, very good results for *log P* and excellent results for *MR*. We illustrate these results by selecting one of the QSPR equations obtained, indicating for each equation its statistical parameters: *r* (correlation coefficient) and *s* (standard deviation).

$$\text{b.p.} = -\,48.5076 + 6.43591 \cdot \text{ZEP};\quad r = 0.94;\quad s = 17.0;\ (N = 31)\quad F = 225.905 \tag{8}$$

$$\text{RM} = 1.25315 + 1.33197 \cdot \text{ZEP};\quad r = 0.998;\quad s = 0.3;\ (N = 19)\quad F = 406.6 \tag{9}$$

$$\log \text{P} = -\,1.62977 + 0.13747 \cdot \text{ZEP};\quad r = 0.97;\quad s = 0.1;\ (N = 20)\quad F = 302.4 \tag{10}$$

Due to the fact that not all the values of *b.p.*, *MR* and *log* P for the set of amines considered in the study were available, each statistical equation has an indication on the number N of amines taken under study. The obtained QSPR equations with good statistical parameters could be now used to predict the same properties

for other chemical compounds whose molecular structure is known. For example, by means of equation (8), we computed the values of *MR* for 19 amines and we compared them to the observed values, given in Table 2.

Table 1

The index ZEP and the experimental values of *b.p., MR* and *log P* for a set of amines

No.	Compound	ZEP	b.p.	MR	log P
1	Propyl amine	13.892	49.0	19.401	0.48
2	2-amino-2-methyl propane	16.449	44.4	24.257	–
3	Isopropylmethyl amine	17.352	50.0	–	–
4	Diethyl amine	17.479	56.0	–	–
5	1-Aminobutane	17.356	77.8	24.079	0.75
6	2-Amino-2-methylbutane	20.155	78.0	–	–
7	1-amino-3-methyl butane	20.702	95.0	28.672	–
8	n-Propyl-n-buthylamine	27.886	–	–	2.12
9	Methyl-n-butyl amine	21.096	–	–	1.33
10	3-aminopentane	20.650	91.0	28.617	–
11	1-aminopentane	20.820	104.4	28.728	1.49
12	3-amino-2,2-dimethyl butane	23.482	102.0	25.938	–
13	Di-isopropyl amine	23.815	84.0	33.641	–
14	Butyldimethylamine	24.575	95.0	33.816	1.70
15	Triethylamine	24.485	89.0	33.794	1.44
16	Hexilamine	24.284	130.0	33.290	1.98
17	Dimethylpentilamine	28.040	123.0	38.281	–
18	2-aminoheptane	27.450	142.0	38.038	–
19	1-aminoheptane	27.748	156.9	38.003	2.57
20	Di-isobutilamine	30.866	139.0	42.920	–
21	Dibutylamine	31.350	159.0	–	2.68
22	Tripropylamine	34.854	156.0	47.784	–
23	n-octilamine	31.212	179.8	–	2.82
24	n-nonylamine	34.676	202.4	47.277	–
25	n-decilamine	38.140	220.7	–	–
26	etilamine	10.383	16.4	–	–
27	Trimethylamine	14.196	–	–	0.27
28	Isobuthylamine	17.129	67.9	–	0.73
29	Di-n-propylamine	24.422	109.0	33.515	1.67
30	Tri-n-butylamine	45.246	214.2	–	–
31	2-butylamine	17.074	62.9	–	0.74
32	Diethylmethylamine	21.087	64.2	–	–
33	Ethylbuthylamine	24.415	108.2	33.452	–
34	Terțbuthylisopropylamine	26.886	98.2	–	–
35	1-amino-2,2-dimethylpropane	20.235	–	28.471	–
36	Dimethylisobuthylamine	24327	–	33.852	–
37	Ethyl-isopropylamine	20.654	–	–	0.93
38	n-propyl-sec-butylamine	27.699	–	–	1.91
39	n-propyl-isobutylamine	27.644	–	–	2.07
40	Isopropylamine	13.475	–	–	0.26

The obtained results are given in Table 2.

Table 2

Observed and calculated values of *MR* for a class of amines

Compound	ZEP	Obs. MR	Calc. MR	Res.
Propylamine	13.89	19.40	19.75	–0.35
2-amino-2-methylpropane	16.49	24.25	23.22	1.02
Isopropyl-methylamine	17.35	24.07	24.37	–0.29
1-amino-2,2-dimethylpropane	20.23	28.47	28.20	0.26

Table 2 (continues)

Table 2 (continued)

3-amino-pentane	20.65	28.61	28.75	–0.14
1-amino-2-methylbutane	20.70	28.67	28.82	–0.15
1-aminopentane	20.82	28.72	28.98	–0.25
Diisopropylamine	23.81	33.64	32.97	0.66
Hexilamine	24.28	33.29	33.59	–0.30
Dimethyl isobuthylamine	24.32	33.85	33.65	0.19
Ethyl-butylamine	24.41	33.45	33.77	–0.32
Dipropylamine	24.42	33.51	33.78	–0.26
Triethylamine	24.48	33.79	33.86	–0.07
Butyl-dimethylamine	24.57	33.81	33.98	0.17
2-aminoheptane	27.45	38.03	37.81	0.22
1-aminoheptane	27.74	38.00	38.21	–0.20
Dimethyl-pentylamine	28.04	38.28	38.60	–0.32
Di-isobutylamine	30.86	42.92	42.36	0.55
n-nonylilamine	34.67	47.27	47.44	–0.16

It is important to stress that the statistical parameters of the QSPR equations could be improved by using multidimensional correlations involving also other topological descriptors. As the aim of this paper is mainly to illustrate the modelling of amines by chemical graphs using the weighted electronic connectivity matrix CEP, we leave the task of searching better QSPR equations for future works.

CONCLUSIONS

The model that we proposed and studied here for the set of 40 amines is intended to accurately detect the presence of heteroatom in the molecular structure. This was expressed by means of the topological index ZEP, extracted from the weighted electronic matrix CEP. The QSPR equations obtained for the three physicochemical properties under study: boiling point, molar refraction and partition coefficient allow us to conclude that the model of the chemical structure obtained by replacing the usual topological distances by weighted electronic distances, shows good, very good and excellent correlation attributes.

The w.e.d. has the merit that it is able to differentiate not only the covalent bonds carbon-nitrogen from the ones of carbon–carbon, but is also able to discriminate the covalent bonds of nitrogen to a primary carbon from that to secondary or tertiary carbons. To our best knowledge, no other models have similar discrimination power for the covalent bonds with regard to their structural neighbourhood.

REFERENCES

1. C. Berge, „Teoria grafurilor şi aplicaţiile ei”, Editura Tehnică, Bucureşti, 1969.
2. M. Randić, W.L. Woodworth and A. Graovac**,** *Int. J. Quantum. Chem*., **1983**, *24*, 435-452.
3. A.T. Balaban**,** *Rev. Chim*., **1988**, *39*, 1026-1031.
4. M. Diudea and O. Ivanciuc**,** “*Topologie moleculară”*, Comprex, Cluj-Napoca, 1995.
5. M. Randić**,** *J. Math. Chem*., **1991**, *7*, 155-168.
6. Z. Mihalić, S. Nikolić and N. Trinajstić**,** *Inf. Comput. Sci*., **1992**, *32*, 28-37.
7. O. Ivanciuc, *Rev. Roum. Chim*., **2000**, *45*, 1037-1054.
8. M. Randić, P.J. Hansen and P.C. Jurs**,** *J. Chem. Inf. Comput. Sci*., **1988**, *28*, 60-68.
9. L.B. Kier and L.H. Hall**,** “Molecular Connectivity in Structure-Activity Analysis”, John Wiley & Sons, New York, 1986.
10. M. Barysz, G. Jashari, R.S. Lall, V.K. Srivastaya and N. Trinajstić**,** in “Chemical Applications of Graph Theory and Topology”, King, R.B. (Ed), Elsevier, Amsterdam, 1983, p. 222-230.
11. S.C. Basak, V.R. Magnuson, G.J. Nieni, R.R. Regal and G.D. Veith**,** *Math. Modelling*, **1987**, *8*, 300-305.
12. A.T. Balaban**,** *J. Chem. Inf. Comput. Sci*., **1985**, *25*, 334-343.
13. M. Randić**,** *Chemometrics Intel. Lab. Syst*., **1991**, *10*, 213-227.
14. E. Estrada**,** *J. Chem. Inf. Comput. Sci*., **1995**, *35*, 701-707.
15. Z. Berinde**,** “Applications of molecular topology in the study of physico-chemical properties of organic compounds” (in Romanian), Cub Press 22, Baia Mare, 2001.
16. Z. Berinde**,** *Bul. Appl. Math.*, **1998**, *86*, 77-86.
17. Z. Berinde**,** Topological Indices Obtained from CEP Matrix of Connectivity, Proceed. "*First Southern Symposium on Computation*", Univ. of Southern Mississippi, U.S.A., 4-7 Dec. 1998 (Electronic), 1999.
18. L.B. Kier and L.H. Hall**,** “Molecular Connectivity in Chemistry and Drug Research”, Academic Press, New York, 1976.