Tetrahedron Vol. 33, pp. 1953 to 1957. Pergamon Press 1977. Printed in Great Britain

A MOLECULAR CONNECTIVITY STUDY OF ELECTRON DENSITY IN ALKANES

LOWELL H. HALL
Eastern Nazarene College, Quincy, Massachusetts, U.S.A.

and

LEMONT B. KIER
Department of Pharmaceutical Chemistry, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

(*Received in USA 30 December 1976; Received in UK for publication 17 March 1977*)

Abstract—Molecular connectivity indices have been developed which characterize contributions of neighboring atoms to the CNDO/2 calculated electronic charge of a carbon atom. An analysis of these indices reveals their ability to predict this charge to 0.001 electron.

INTRODUCTION

It is known that alkanes exhibit small but real dipole moments. Since each C atom has a comparable core charge and each C–C bond is derived from very similar hybrid states, the "electron densities" at each C must be largely dependent upon the molecular topology.

The topology of a molecule includes the number and type of constituent atoms along with the pattern of branching in the molecule. Thus, in an alkane, the charge at any C is a result of its degree of branching and the branching degree of the attached atoms. Because of the valence requirements of alkane C atoms, the number of H atoms attached to a C is implied by specifying the degree of branching as a C.

It follows from this, that if the topology of alkanes could be quantified so that each C atom bore a numerical description, then such a description should bear some close relationship to the charge on the C atoms. Evidence from a number of recent studies[1-6] suggests that the method known as molecular connectivity is capable of providing such a topological description. In this paper, we examine the topological influence on alkane C atom charges.

Molecular connectivity indices

Molecular connectivity is a general, non-empirical approach to quantification of branching in molecular skeletons as well as structural features such as cyclization, unsaturation, heteroatom type and position.[1] The method has its origin in a branching index proposed by Randić.[2] Molecular connectivity has been utilized extensively to describe molecular structural features which govern physiochemical properties as well as activity of drug molecules.[1]

The calculation of connectivity indices ${}^m\chi$ is based upon the hydrogen-suppressed graph, the carbon skeleton of the molecule. An atom connectivity value is assigned to each atom i (graph vertex) as the number of attached C atoms. Thus, there are five types of alkane C atoms: $\delta = 0, 1, 2, 3, 4$. The first order term, ${}^1\chi$, is computed as a sum of edge terms, ${}^1c_i$:

$$ {}^1\chi = \sum_i (\delta_i\delta_j)_i^{-1/2} = \sum {}^1c_i $$

where the sum is over all connections or bonds in the graph. Atoms i and j are formally bonded.

An extended chi term of second order, ${}^2\chi$, is computed:

$$ {}^2\chi = \sum_i (\delta_i\delta_j\delta_k)_i^{-1/2} = \sum {}^2c_i $$

where i, j, and k represent atoms bonded in sequence or in a path, and the sum is overall distinct sets of two-edge paths. The zeroth order term, derived from single atom considerations, is similarly defined:

$$ {}^0\chi = \sum_{i=1}^{n} (\delta_i)^{-1/2}. $$

A general expression may be written for chi terms based on sets of m adjacent bonds, ${}^m\chi_p$:

$$ {}^m\chi_p = \sum_{j=1}^{{}^mN_j} \left[\prod_{i=1}^{m+1} \delta_i \right]_j^{-1/2} = \sum_{j=1}^{{}^mN_j} {}^mc_j. $$

The symbol ${}^mN_j$ is the number of paths with m bonds. The connectivity indices describe the topology of the graph which represents the whole molecule.

Application of molecular connectivity to charge

A molecular connectivity index computed for individual atoms in a graph has been developed by partitioning the ${}^mc_i$ terms between the atoms. For the first order term, the procedure is as follows:

1. The c_i value for each edge is partitioned equally to each vertex so that the edge contribution to a vertex is $\frac{1}{2}c_i$.

2. At each vertex, the $\frac{1}{2}c_i$ values are summed to give a vertex connectivity index, ${}^1\bar{\chi}$, so that

$$ {}^1\bar{\chi} = \frac{1}{2}\sum_{i=1}^{\delta} {}^1c_i. $$

3. In a similar fashion, each ${}^2c_i$ value for paths of length two are partitioned among the three contributing atoms so that each atom receives a value of $\frac{1}{3}{}^2c_i$.

4. At each vertex, the $\frac{1}{3}{}^2c_i$ values are summed so that

$$ {}^2\bar{\chi} = \frac{1}{3}\sum {}^2c_i. $$

5. The ${}^0\bar{\chi}$ value for an atom is simply the $\delta_i^{-1/2}$ value:

$$ {}^0\bar{\chi} = \delta_i^{-1/2}. $$

It should be noted that the sum of ${}^m\bar{\chi}$ values over all of the non-hydrogenic atoms in the molecular yields the overall connectivity index ${}^m\chi$:

$$ {}^m\chi = \sum_{i=1}^{n} {}^m\bar{\chi} $$

where n is the number of non-hydrogenic atoms in the molecule (the number of vertices in the graph).

Relationship of ${}^m\bar{\chi}$ values to MO calculated alkane charges

To study the relationship between the ${}^m\bar{\chi}$ values and computed atom charges, we have used the CNDO/2 method to evaluate "electron densities".[8] The calculations were performed using standard bond lengths and angles[9] and in staggered, extended conformations. For acyclic alkanes there are five types of C atoms, with respect to topology, or the number of carbon attachments corresponding to 4, 3, 2, 1, 0. When considering only the number and kinds of C atoms which can be attached to a single atom, there are 69 possibilities. When considering atoms beyond the alpha position, the number rises dramatically. For groups of three bonded carbons, the number is 6435. These numbers correspond to the number of possible values for ${}^0\bar{\chi}$, ${}^1\bar{\chi}$ and ${}^2\bar{\chi}$, respectively, in alkanes.

We have included all the alkanes, propane through the isomeric octanes, in our study. Although all of the above structural possibilities are not present, a reasonably representative selection may nonetheless be studied. This present set includes 263 C atoms. In the regression analysis we have limited ourselves to all symmetry independent atoms as indicated in Table 1. Thus, the statistical studies deal with 193 atoms.

RESULTS

We have performed a regression analysis with several terms of ${}^m\bar{\chi}$ against the computed CNDO/2 charge (expressed as net charge). The best single variable equation is:

$$ q_C = 0.1017 - 0.1152\,{}^0\bar{\chi} $$

$$ r = 0.979, \quad s = 0.0045, \quad F = 4397.8 \qquad (1) $$

for two terms:

$$ q_C = 0.0438 - 0.0710\,{}^0\bar{\chi} + 0.0462\,{}^1\bar{\chi} $$

$$ r = 0.990, \quad s = 0.0032, \quad F = 4219.5 \qquad (2) $$

and for three terms:

$$ q_C = 0.0162 - 0.0493\,{}^0\bar{\chi} + 0.0493\,{}^1\bar{\chi} + 0.0217\,{}^2\bar{\chi} $$

$$ r = 0.998, \quad s = 0.0014, \quad F = 15553. \qquad (3) $$

Based on the F test for the addition of a variable, the results in each case exceed the 99.99% confidence level. Regression analysis with more than three terms does not significantly improve the result.

DISCUSSION

Our regression analysis indicates that three variables are required to achieve an excellent correlation. This

Table 1

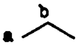
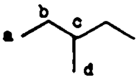
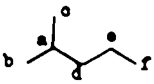
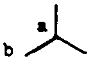
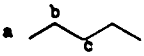
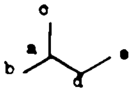
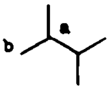
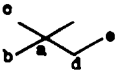
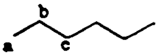
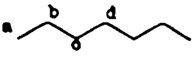
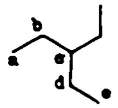
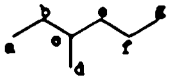
		Carbon Charge e/1000				Carbon Charge e/1000	
Alkane Graph	Atom	CNDO/2	This Work	Alkane Graph	Atom	CNDO/2	This Work
	a	- 9	-11		a	-13	-13
	b	24	22		b	17	17
	a	-12	-12		c	39	36
					d	-12	-13
	b	20	18		a	40	38
	a	46	43		b	-11	-12
	b	- 9	-11		c	-11	-12
	a	-11	-12		d	17	15
	b	17	17		e	16	17
	c	17	16		f	-11	-12
	a	43	39		a	41	38
	b	-12	-12		b	-13	-12
	c	-11	-12		a	58	58
	d	20	18		b	-11	-11
	e	-13	-13		c	-10	-11
	a	62	63		d	20	18
	b	- 9	-10		e	-15	-13
	a	-11	-12		a	-11	-12
	b	18	17		b	18	17
	c	14	15		c	14	15
					d	11	14

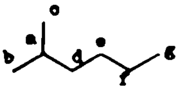
Table 1. (*Contd.*)



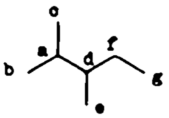
a	-13	-13
b	18	16
c	37	33
d	16	16
e	-13	-13



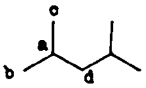
a	-13	-13
b	17	17
c	36	35
d	-12	-13
e	14	14
f	16	17
g	-11	-12



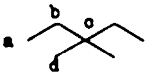
a	40	38
b	-11	-11
c	-11	-12
d	14	14
e	13	14
f	18	17
g	-11	-12



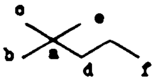
a	38	38
b	-13	-12
c	-13	-12
d	37	35
e	-15	-14
f	16	16
g	-13	-13



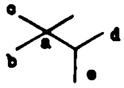
a	38	38
b	-11	-12
c	-11	-12
d	16	15



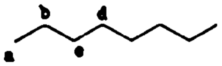
a	-15	-13
b	17	17
c	54	54
d	-12	-12



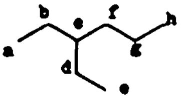
a	55	58
b	-11	-11
c	-11	-11
d	16	15
e	14	17
f	-11	-12



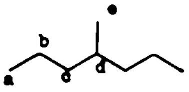
a	55	57
b	-12	-12
c	-12	-12
d	40	38
e	-15	-13



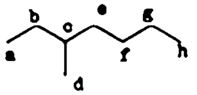
a	-11	-12
b	18	17
c	14	15
d	12	14



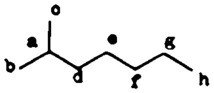
a	-13	-13
b	16	16
c	34	32
d	18	16
e	-13	-13
f	12	14
g	16	17
h	-11	-12



a	-11	-12
b	16	17
c	14	14
d	34	34
e	-12	-13

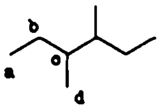
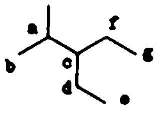
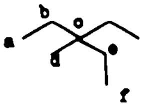
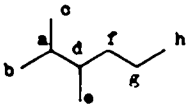
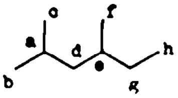
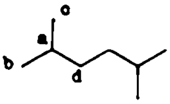
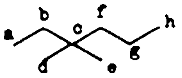
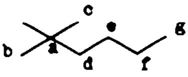
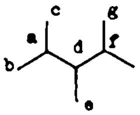
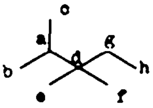
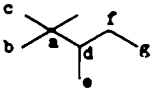


a	-13	-13
b	17	17
c	37	36
d	-12	-13
e	11	13
f	13	14
g	17	17
h	-11	-12



a	40	38
b	-11	-12
c	-11	-12
d	14	14
e	10	13
f	15	15
g	18	17
h	-11	-12

Table 1. (*Contd.*)

	a	-13	-13
	b	16	16
	c	35	34
	d	-15	-14
	a	36	37
	b	-13	-12
	c	35	32
	d	14	16
	e	-13	-13
	f	16	16
	g	-13	-13
	a	-15	-13
	b	16	16
	c	50	51
	d	-14	-14
	e	17	16
	f	-16	-13
	a	38	38
	b	-13	-13
	c	-13	-12
	d	34	34
	e	-15	-14
	f	12	14
	g	16	17
	h	-11	-12
	a	38	38
	b	-11	-12
	c	-11	-12
	d	13	14
	e	35	35
	f	-12	-13
	g	18	16
	h	-13	-13
	a	40	38
	b	-11	-12
	c	-11	-12
	d	12	14
	a	-15	-13
	b	18	17
	c	51	54
	d	-12	-12
	e	14	15
	f	14	17
	g	-11	-12
	a	55	58
	b	-11	-11
	c	-11	-11
	d	36	34
	e	11	14
	f	18	17
	g	-11	-12
	a	36	37
	b	-13	-12
	c	-13	-12
	d	36	34
	e	-16	-14
	f	37	37
	g	-13	-12
	a	38	37
	b	-15	-13
	c	-15	-13
	d	51	53
	e	-14	-13
	f	-14	-13
	g	16	17
	h	-15	-13
	a	52	56
	b	-12	-12
	c	-12	-12
	d	37	35
	e	-16	-14
	f	14	16
	g	-13	-13
	a	53	56
	b	-14	-12

suggests that three topological indices make a major contribution to the description of the structural features influencing CNDO/2 calculated charges. These factors appear to be the type of carbon atom, the degree of branching of alpha carbons, and the influence of beta carbons on the charge.

These factors are encoded in the three variables in eqn (3). The type of C atom, that is, the degree of branching is characterized by the numerical value of ${}^0\bar{\chi}$. The branching or the connectivity of atoms alpha to the carbon is encoded in the ${}^1\bar{\chi}$ terms. These two terms, ${}^0\chi$ and ${}^1\bar{\chi}$, are of nearly equal importance in eqn (3). Finally the connectivity of atoms beta to the carbon is encoded in the ${}^2\bar{\chi}$ terms; the weight of this term is about half the other two. However, the improvement in the overall calculation is most notably influenced by the introduction of the ${}^2\bar{\chi}$ term. This observation indicates that, although the numerical contribution associated with the ${}^2\chi$ term is not the largest contribution, very important effects are associated with the nature of bonds involving atoms beta to a given carbon.

It should be noted that a few graphs appear more symmetrical than the alkane molecular structures they represent. Hence, atoms *b* and *d* in 3-ethylpentane show different values for CNDO/2 calculated charge because of such conformational differences.

Molecular connectivity can be shown to bear a relationship to the topological matrix for a molecule.[7] Lykos[10] and Ruedenberg[11] have shown that since the topological matrix is related to the overlap matrix, one might expect some significant relationship between such quantities as energy and charge density with parameters derived from a topological matrix. This question is a subject of our continued investigation.

It is interesting to note that ${}^3\bar{\chi}_P$ terms, which describe the topological influence of gamma atoms on the carbon charge, do not materially improve the correlation. This leads us to suspect that the gamma atoms in alkanes play a very minor role in influencing the charge on an atom.

When other MO methods including ab initio calculations produce long lists of alkane charges, we anticipate that the connectivity calculations described here will correlate with these values.

REFERENCES

[1] L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York (1976).

[2] L. B. Kier, L. H. Hall, W. J. Murray and M. Randic, *J. Pharm. Sci.* **64**, 1971 (1975).

[3] L. H. Hall, L. B. Kier and W. J. Murray, *Ibid.* **64**, 1974 (1975).

[4] W. J. Murray, L. H. Hall and L. B. Kier, *Ibid.* **64**, 1978 (1975).

[5] L. B. Kier, L. H. Hall, M. Randic and W. H. Murray, *Ibid.* **65**, 1226 (1976).

[6] L. B. Kier and L. H. Hall, *Ibid.* **65**, 1806 (1976).

[7] M. Randic, *J. Am. Chem. Soc.* **97**, 6609 (1975).

[8] J. Pople and G. A. Segal, *J. Chem. Phys.* **44**, 3289 (1966).

[9] J. A. Pople and M. Gordon, *J. Am. Chem. Soc.* **89**, 4253 (1967).

[10] P. G. Lykos and H. N. Schmeising, *J. Chem. Phys.* **35**, 288 (1961).

[11] K. Ruedenberg, *Ibid.* **22**, 1878 (1954).