

Molecular connectivity: intermolecular accessibility and encounter simulation

Lemont B. Kier^{a,*}, Lowell H. Hall^b

^a Department of Medicinal Chemistry, Virginia Commonwealth University, P.O. Box 980540, Richmond, VA 23298-0540, USA

^b Department of Chemistry, Eastern Nazarene College, Quincy, MA 02170, USA

Received 1 September 2000; received in revised form 9 January 2001; accepted 11 January 2001

Abstract

The simple molecular connectivity indices are interpreted as summations of bond accessibilities to bimolecular encounters with another, identical molecule. To transcend this model, a molecule is treated as *dissecta membra* with each bond modeled as a discrete cell in a dynamic simulation of many molecules. Each bond accessibility is transformed into a cellular automata rule. The dynamics are run for each of 38 alkanes, recording the average number of cell encounters, β . The β values show a high correlation with the boiling points. The significance of the bond accessibilities and the concept of intermolecular encounters explaining the molecular connectivity indices is supported by these findings. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Quantitation; Topology; Molecular connectivity

1. Molecular structure

Quantitative information analysis has followed two paths over the years. One path makes use of physical properties to describe a molecule in a model, relating it to a measured response [1]. This was called quantitative structure–activity relationship (QSAR), where the word “structure” was loosely used to denote measured or estimated physical properties. At the same time, an alternative paradigm was developed in the form of theoretical models of structure derived from molecular orbital theory [2]. These models attempted to relate defined structure quantitation to measured responses in an effort to create a predictive model. Since that time, innovations to the quantitation of structure have emerged that extend the practical utility of QSAR models. In particular, the introduction of graph-based topological indices, a quarter of a century ago by Randic [3] and developed by Kier and coworkers [4–6], have made possible a description of structure that is simple and demonstratively valuable in predictive power.

It is important to distinguish between physical properties and structure concepts that have often become blurred. Attempts to rectify this misunderstanding were made by Norrington et al. [7] who proposed the term *property–activity*

relationships as a distinct sub-class of structure–activity models. Hoffmann and Laszlo have written eloquently on this subject from a broad chemical perspective [8] and more recently Testa and Kier [9] have presented detailed descriptions of these attributes in a comprehensive review.

We regard molecular structure as a collective term for codes by which we describe a molecule in quantitative terms. It is a model of the form of a molecule, which produces a series of functions called properties. The functions or properties are measured attributes, which, in the case of molecules, are averaged values of responses to chemical or physical input into a system containing the molecule in question. The numerical values of the properties present a mosaic of information about the system. From this information we weave a model of what the structure may be that gave birth to our measurements. Structure is a model; it is a presumption of what is *there*, functioning in response to our physical inquiry. A representation of what is there may take the form of statements of content, probability, accessibility, topology, relationships, or complexity. In this sense, any structure representation may be incomplete yet judged adequate for a particular purpose. The function or properties of a molecule are dependent upon the form or structure. It is an immutable relationship; structure is the antecedent to properties; form precedes functions.

In the case of structure description, there is an implied obligation to look for the meaning, interpretation or information conveyed in a structure descriptor. For 25 years, this

* Corresponding author. Tel.: +1-804-786-8483; fax: +1-804-371-7625.
E-mail addresses: kier@hsc.vcu.edu (L.B. Kier), hall@enc.edu (L.H. Hall).

has not been revealed for the molecular connectivity indices. In this paper, we address this issue and offer a contribution to the interpretation of these indices.

2. Molecular connectivity

2.1. The Randic branching index

Since Randic proposed an algorithm to encode bond contributions to a molecular branching index [3], 25 years ago, it has become possible to offer quantitative statements about the extent of branching in a molecule. Randic demonstrated that alkanes could be ordered by this scheme to correlate with a physical property, the boiling point. This branching algorithm formed the basis of a structure description paradigm called molecular connectivity developed over the next decade by Hall and Kier [6], and Kier and Hall [5,10]. This is now a very widely used non-empirical structure description in quantitative structure–activity analyses and model development [11].

An important feature of molecular connectivity is the encoding of structure in a non-empirical way. It is not a measured property nor does it derive from or directly translate into a particular physical property. This has, strangely, not been understood by some investigators employing structure–activity models. In this paper, we will examine the Randic [3] algorithm, propose an alternate structure concept, offer an interpretation of the information encoded by molecular connectivity, and demonstrate a dynamic simulation of many molecular bond encounters.

2.2. The Randic algorithm

The molecule is regarded as a sum of the bonds connecting pairs of atoms. Each atom in a molecule is encoded by a cardinal number, δ , the count of all bonded atoms other than hydrogen. In chemical graph theory terms, the molecule is a hydrogen-suppressed graph illustrated by 3-methylpentane, I. In common usage, the hydrogen atoms are collected with the attached atom as a hydride group such as CH_3 and NH_2 and are represented by a single δ value. The molecule is dissected into fragments or bonds, each retaining the δ values assigned in the original graph as shown for 3-methylpentane, II. This decomposition produces a set of fragments encoded by the two δ values of the atoms comprising each bond, shown in Table 1.

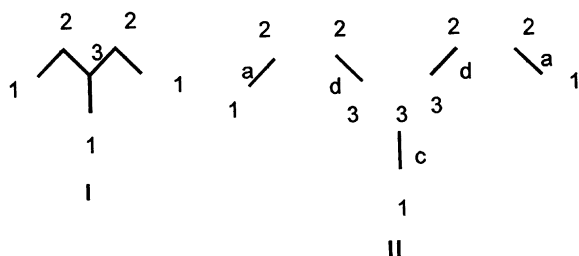


Table 1
Bond types in C sp^3 molecules

Bond type (δ_i, δ_j)	Code	$\delta_i \delta_j$	$C_{ij} = A_{ij}$
(1, 1)	Unique	1	1.000
(1, 2)	a	2	0.707
(1, 3)	c	3	0.577
(1, 4)	f	4	0.500
(2, 2)	b	4	0.500
(2, 3)	d	6	0.408
(2, 4)	g	8	0.354
(3, 3)	e	9	0.333
(3, 4)	h	12	0.289
(4, 4)	i	16	0.250

Randic combined these bond descriptors into a single number, encoding the information about the differences between the fragments. He elected to use the product of the δ values, $(\delta_i \delta_j)$, taken to the -0.5 power. Stepwise, this is the product, the square root and the reciprocal. These values are shown in Table 1. Kier and Hall [5] have designated this as the C_{ij} value

$$C_{ij} = (\delta_i \delta_j)^{-0.5} \quad (1)$$

The step leading to the creation of a molecular index was the combination of these values by summation. This index was proposed by Randic as a branching index and was shown to correlate well with alkane boiling points [3]. Kier and Hall and later generalized this as, $^1\chi$, a chi index of the first order [5]

$$^1\chi = \sum (\delta_i \delta_j)^{-0.5}, \quad \text{summed over all skeletal bonds} \quad (2)$$

Interpreting the structure information inherent in the chi indices begins with an examination of the indices characterizing the bonds. All bonds possible in alkanes are the 10 entries in Table 1. The first bond type, (1, 1), is unique for ethane among the alkanes. The others are found among all alkanes, and in the skeletons of all covalently bonded molecules. The bonds shown in Table 1 are listed according to the decreasing value of C_{ij} which corresponds to an increase in branching at one or both of the atoms comprising the bond. The C_{ij} index is thus successful in ranking this attribute.

The question of the significance of this algorithm and the structural interpretation of C_{ij} , is now addressed. Clearly C_{ij} parallels the degree of branching associated with each fragment. The algorithm to calculate this value was not derived, it was stated. The objective was to encode the influence of the branched state of each atom in the bond, on the relative contribution of that bond to the whole molecule. In other words, the characterization of some aspect of the structure of a molecule was the goal in assigning a numerical value which might correlate with some measured properties.

3. A bimolecular encounter model

3.1. Accessibility interpretation of bond fragments

Another approach to the definition of molecular structure and its relationship to the value of a physical property is proposed here. This focuses on the bond indices C_{ij} , while the conventional wisdom has been directed at the numerical descriptors of whole molecules. The basic premise of this alternative is that a useful description of molecular structure represents the molecule in a milieu of other molecules. This idea was expressed by Seybold et al. [12] who wrote about the structure influence on packing of molecules in their milieu, an attribute dependent upon their shape. Consideration of a milieu of molecules leads to an average quality of intermolecular dispersion that would account for several physical properties. In fact a measurement of a physical property reflects the collective influence of encounters of each molecule with other molecules in its immediate environment. Our model places emphasis on the encounters among molecules, in particular the possibilities of fragments of one encountering fragments of another. Kier and Hall have focused on this concept by exploring the possibilities of encounters among two molecules [13]. This bimolecular encounter model reflects in a microcosm what is very likely taking place in a manifold of molecules in a bulk system. Analysis of these encounters form the basis of an interpretation of parts of the Randic algorithm.

3.2. The δ values

Our interpretation of the significance of a C_{ij} value is that it encodes the relative accessibility of a bond to encounter another bond in another molecule, leading to an intermolecular interaction. The term accessibility is defined as the topological and electronic availability of one bond to engage in some interaction with a bond in another molecule. How then can we develop de novo, numerical expressions for accessibility in this context? Consider the hydrogen-suppressed graph fragments in II. Each atom (hydride group) in each fragment is designated by a δ value, which is the count of σ -electrons contributed to bonds formed with adjacent atoms. The atoms in II with $\delta = 1$ contribute all of their non-hydrogen σ -electrons to the bond. In the fragments in II where one atom has a $\delta = 2$ value, there is encoded the contribution of 1/2 of its non-hydrogen σ -electrons to the bonds of which it is a part. By analogy, the fragments in II with atoms designated $\delta = 3$, have only 1/3 of their non-hydrogen σ -electrons contributed to the bond shown.

The δ values therefore have a dual meaning. First the δ value is the count of neighboring atoms bonded to an atom in the hydrogen-suppressed graph. This corresponds to the count of the σ -electrons contributed by that atom to bonded, non-hydrogen atoms. Secondly, the reciprocal of the δ value, $1/\delta$, is the fraction of the total number of

non-hydrogen σ -electrons contributed to each bond formed with that particular atom.

3.3. The bond terms

The $1/\delta$ terms reflect effective contributions to the relative accessibility of the bond in which they are a part. Galvez [14] has compared these δ values to molecular orbital parameters and has derived a relationship between the bond order and the $(1/\delta_i, 1/\delta_j)^{1/2}$ value. From this we interpret the term to be a relative bond accessibility value which we label A_{ij} . All possible alkane bond A_{ij} values are shown in Table 1. When this value is high for a bond, there is an expectation that the bond is relatively accessible to other bonds in the milieu. Conversely, a low value of A_{ij} infers a low accessibility. From structure II, we conclude that the (1, 2) bond has a greater accessibility to other bonds in its environment than does the (1, 3) or (2, 3) bonds in II.

3.4. Bimolecular encounter parameters

Intermolecular encounters of molecules, governed by the pattern of atoms and bonds, influence many physical property values. Kier and Hall have used the counts of the numbers of bimolecular encounters and have formulated a series of parameters encoding a relationship with some physical properties [13]. Expanding on that work we propose a model based on the encounters of bonds in two molecules using 3-methylpentane, I, as an example. The encounter of two molecules of 3-methylpentane is considered to be some function of the interaction of the bonds between molecules. A series of bimolecular bond encounter parameters is calculated from the manifold of bond accessibilities operating between two molecules of 3-methylpentane. In this model, the encounter accessibility of any two bonds on two different molecules is calculated as the product of the individual accessibilities, A_{ij} , of each. Table 2 shows all possible bimolecular bond encounter terms for any alkane.

Stated another way, the encounter of two 3-methylpentane molecules includes the possible interaction of the two fragments of the type (1, 2) shown in II. The accessibility, A_{ij} , of each of these bonds is 0.707 from Table 2. The possibility of the encounter of two such bonds on separate molecules is evaluated as the product of the accessibilities, 0.707×0.707 or 0.5. The complete set of encounter accessibilities of two 3-methylpentane molecules can be extracted from Table 2. The products of all such interactions, $(A_{ij})(A_{kl})$, for two 3-methylpentane molecules are summed to give a total bimolecular encounter possibility for 3-methylpentane. From the data in Table 2 we calculate the total bimolecular interaction possibility, P_T , for two 3-methylpentane molecules, m and n , to be

$$\sum_m (A_{ij})_m \sum_n (A_{kl})_n = 7.885 \quad (3)$$

The bimolecular encounter possibility, P_T , is the product of the attributes of each molecule, P_m and P_n , in the en-

Table 2
Intermolecular encounter possibilities

Bond types	Bond types							
	a (1, 2)	c (1, 3)	b/f (2, 2)/(1, 4)	d (2, 3)	g (2, 4)	e (3, 3)	h (3, 4)	i (4, 4)
a (1, 2)	0.50	0.41	0.35	0.29	0.25	0.24	0.20	0.18
c (1, 3)		0.33	0.29	0.24	0.20	0.19	0.17	0.14
b/f (2, 2)/(1, 4)			0.25	0.20	0.18	0.17	0.15	0.13
d (2, 3)				0.17	0.14	0.13	0.12	0.10
g (2, 4)					0.13	0.12	0.10	0.09
e (3, 3)						0.11	0.10	0.08
h (3, 4)							0.08	0.07
i (4, 4)								0.06

counter. In the case of 3-methylpentane, the contribution to P_T from molecule, m (and also molecule n) is $P_m = 2.808$.

4. Comparison of the two models

4.1. The equivalence of $^1\chi$ and P_m

The algorithm of Randic computes a term for each bond type and then sums these to give a molecular connectivity index. Using 3-methylpentane as our example, we assign to each different bond type a letter code, designating a different bond type, shown in Table 1. The three different bond types for 3-methylpentane are coded a, c, and d. The appropriate summation of these indices in 3-methylpentane gives the molecular connectivity index of the first order, $^1\chi$. The calculation for 3-methylpentane is

$$^1\chi = 2a + c + 2d \quad (4)$$

In our bimolecular model, we compute each possible bond encounter as a product of bond accessibilities among two molecules of 3-methylpentane. Using the same codes for the bond types in Table 1, all possible encounters of bonds among two 3-methylpentane molecules are

$$aa, aa, aa, aa = 4a^2$$

$$dd, dd, dd, dd = 4d^2$$

$$ad, ad, ad, ad, ad, ad, ad, ad = 8ad$$

$$ac, ac, ac, ac = 4ac$$

$$cd, cd, cd, cd = 4ac$$

$$cc = c^2$$

The total interaction possibility, P_T , is the product of the two interaction possibilities for molecules m and n

$$P_T = P_m P_n = \sum (A_{ij})(A_{kl}) \\ = 4a^2 + 4ac + 8ad + 4ac + 4d^2 + c^2 \quad (5)$$

The sum of the number of bond encounters is the product of the number of bonds in each. The interaction possibility of one 3-methylpentane molecule P_m is

$$P_m = 2a + c + 2d \quad (6)$$

To test this, the sum of bimolecular interaction possibilities for 3-methylpentane is 7.885, and the contribution of each molecule, m or n , is $P_m = 2.808$. The equality exists

$$7.885 = 2.808 \times 2.808 \quad (7)$$

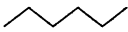

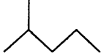


4.2. The meaning of the molecular connectivity index

We conclude that the molecular connectivity index is the contribution of one molecule to the bimolecular interactions arising from encounters of all bonds among two like molecules. The molecular connectivity index, $^1\chi = P_m$.

This result is confirmed by relating the $^1\chi$ indices of other molecules, y , to the P_y values from their bimolecular interaction values as shown in Table 3. It can be shown that the extended molecular connectivity indices are equal to the $P_{i,...,n}$ values of bimolecular interactions of higher order fragments.

In the simplest case, the graph of a molecule with only one bond type (1, 1) has an accessibility value of y , while the bimolecular accessibility value, $\sum_m (A_{ij})_m \sum_n (A_{kl})_n = y^2$. The equality, $y = y^2$ can only exist for positive values when

Table 3
Bimolecular encounter accessibilities for hexane isomers

Hexane isomer	$\sum (A_{ij})(A_{kl})$	$^1\chi = P_{ij}$
	8.491	2.914
	7.885	2.808
	7.328	2.707
	6.985	2.643
	6.559	2.561

$y = 1$. This is the only ${}^1\chi$ value allowed for a molecular graph of one bond, in the simple case, the graph for ethane.

5. Simulation of bond encounters

5.1. A general model

Our interpretation of the meaning of molecular connectivity is based on bond–bond encounters and so is a reasonable model for alkanes. More complex molecules have intermolecular encounters based on these and other molecular features such as π - and lone-pair electrons. The molecular connectivity indices have been expanded in their ability to encode some of this information, which makes them of more general value in studying structure–activity relationships [6,10]. The E-state indices have been developed to encode atom contributions to intermolecular non-covalent interactions [13].

The relationship between ${}^1\chi$ and the sum of bond encounter possibilities among two molecules has been demonstrated. As a model, the bimolecular encounter is a limiting case. In the liquid state, the alkanes engage in a host of intermolecular encounters among many molecules. By modeling the bond encounters among a large number of molecules, we might simulate bulk conditions and transcend the information resident in just the bimolecular ${}^1\chi$ index model. One approach, is to regard the bonds in a molecule as scattered parts or fragments, a system that could be called *disjecta membra*. Each bond of a molecule is modeled as a disconnected entity, endowed with a state that corresponds to one of the types of bonds found in Table 1. These bonds move through a system containing other bonds in other states from other molecules of the same structure. This model can be simulated with cellular automata dynamics. That is the approach that we have pursued in this study.

5.2. Cellular automata

5.2.1. The model

Cellular automata are computational systems that are discrete in space, time and state whose dynamic behavior is specified by rules governing local relationships. They are an attempt to simplify the often numerically intractable dynamic simulations into a set of simple rules that mirror intuition and that are easy to compute. As an approach to the modeling of emergent properties of complex systems it has a great benefit in being visually informative of the progress of dynamic events. From the early development by von Neumann [15] a variety of applications ranging from gas phenomena to biological applications have been reported [16–19].

The model in our study is composed of a grid of spaces called cells on the surface of a torus to remove boundary conditions. Each cell i has four tessellated neighbors, j , and

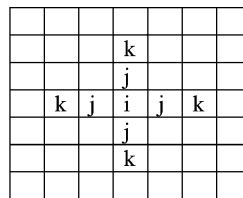


Fig. 1. The von Neumann neighborhood focused on cell i .

four extended neighbors, k , in what is called an extended von Neumann neighborhood [15], Fig. 1. An occupant of a cell has a state governing its characteristics. The occupants of a cell moves, joins with another occupied cell or breaks from a tessellated relationship according to probabilistic rules. The illusion of movement is achieved by the change of states in adjacent cells, an unoccupied cell assuming a zero value while the adjacent occupied cell assumes a non-zero value. These rules are established at the beginning of each simulation. The rules are applied one after another to each cell at random, the complete application of the rules to all cells constituting one iteration. The rules are applied uniformly to each cell type and are local, thus there is no action at a distance. Our cellular automata model is thus kinematic, asynchronous and stochastic. The initial conditions are random hence they do not determine the ultimate state of the cells, called the configuration. The same initial conditions do not yield the same set of configurations after a certain number of iterations except in some average sense. The configurations achieved after many iterations reach a collective organization that possesses a relative constancy in appearance and in reportable counts of attributes. What we observe and record from the cellular automata simulations are emergent attributes of a complex system.

5.2.2. The rules

The parameters adopted in our model govern the probabilities of a cell to encounter other cells in the grid. The breaking probability, P_B , is the probability for a cell at i to break away from another at j when there is exactly one occupied j cell (see Fig. 1). The value for P_B lies in the closed unit interval. The second parameter, J , describes the movement of a cell at i toward or away from the cell at k in the extended von Neumann [15] neighborhood when the intermediate j cell is vacant. J is a positive real number. When $J = 1$, it indicates that the particle, i , has the same probability of movement toward or away as for the case when k is empty. When $J > 1$, it indicates that i has a greater probability of movement toward an occupied cell, k than when k is empty. When $J < 1$, it indicates that i has a lower probability of such movement. Our studies on water [19] and other simulated liquids have supported the use of the standard relationship between P_B and J expressed as

$$\log J(W) = -1.50P_B(W) + 0.60 \quad (8)$$

which we adopt in this study.

5.3. A model of molecular bonds

Each bond type was simulated by a cell with a particular state, derived from the classifications in Table 1. The rules for joining and breaking of each type of cell were derived from the encounter possibilities, $(A_{ij})(A_{kl})$, among two molecules as shown in Table 2. To scale these possibilities to the trajectory rules for our cellular automata model we have adopted the relationship

$$J : (A_{ij})(A_{kl}) = 4(A_{ij})(A_{kl}) \quad (9)$$

The breaking probabilities, $P_B(A_{ij})(A_{kl})$, were derived from Eq. (8). A complete set of joining and breaking rules

Table 4
Breaking and joining rules for all bond types

Interacting pair	P_B	J
aa	0.20	2.00
ab	0.30	1.40
ac	0.26	1.64
ad	0.36	1.16
ae	0.41	0.96
af	0.30	1.40
ag	0.40	1.00
ah	0.47	0.80
ai	0.50	0.72
bb	0.40	1.00
bc	0.36	1.16
bd	0.47	0.80
be	0.51	0.68
bf	0.40	1.00
bg	0.50	0.72
bh	0.55	0.60
bi	0.59	0.52
cc	0.32	1.33
cd	0.41	0.96
ce	0.48	0.76
cf	0.36	1.16
cg	0.47	0.80
ch	0.51	0.68
ci	0.57	0.56
dd	0.51	0.68
de	0.59	0.52
df	0.47	0.80
dg	0.57	0.56
dh	0.61	0.48
di	0.67	0.40
ee	0.64	0.44
ef	0.51	0.68
eg	0.61	0.48
eh	0.67	0.40
ei	0.73	0.32
ff	0.40	1.00
fg	0.50	0.72
fh	0.55	0.60
fi	0.58	0.52
gg	0.59	0.52
gh	0.67	0.40
gi	0.70	0.36
hh	0.73	0.32
hi	0.77	0.28
ii	0.81	0.24

for the nine bond types is shown in Table 4 using the letter codes from Table 1 to identify the bond encounter pairs. Each bond in a molecule was represented by a particular state of a cell. One hundred molecules were modeled in a grid of 3025 cells. For example, the *dissecta membra* model of 3-methylpentane, II, uses 200 cells with a state corresponding to bond type a, 100 cells with a state corresponding to bond type c, and 200 cells with a state corresponding to bond type d. Each cell moved randomly during one iteration, joining another cell, breaking from another cell or moving freely in unoccupied grid space. The dynamics were run for 990 iterations and then during the next 10 iterations, the count of the number of joined cells was recorded. This process was repeated for 25 runs and the count of joined cells was averaged from this data. The count of joined cells was called the, β , value. Thirty eight alkanes including all of the pentanes, hexanes, heptanes and octanes and three cycloalkanes were modeled and the β values recorded in Table 5.

5.4. Results

The calculated β values for the 38 alkanes are shown in column 2 of Table 5. The interpretation of this value is a possible relationship with a physical property that is highly dependent upon intermolecular interactions. Such a property is the boiling point, entered for each alkane in column 4 of Table 5. In comparison, the molecular connectivity $^1\chi$ values for the alkanes are recorded in column 3 of the same table. The β values, expressed in a quadratic equation, relates to the boiling point ($^{\circ}\text{C}$) with the statistics shown in Eq. (10):

$$\text{bp} = 0.584\beta - 0.0004\beta^2 - 71.517 \quad (10)$$

$$r^2 = 0.991, \quad s = 2.996, \quad n = 38, \quad F = 2027$$

In comparison the $^1\chi$ indices for the same molecules produced a quadratic equation

$$\text{bp} = 115.50^1\chi - 9.33(^1\chi)^2 - 184.18 \quad (11)$$

$$r^2 = 0.984, \quad s = 4.158, \quad n = 38, \quad F = 1044$$

6. Discussion

The count of cell encounters averaged over time (iterations), encoded by the β value is very closely related to the boiling point of the 38 alkanes. The standard deviation of only 3.00° is better than any one-variable quadratic analysis we have found reported. It is better than the correlation with the molecular connectivity index, even though both $^1\chi$ and β are derived from identical accessibility values. The cellular automata dynamics improves the simulation of bond encounters compared to the static count of all possible bimolecular encounters. The quality of the relationships revealed here supports the model that we have developed

Table 5
Boiling points of alkanes vs. β values

Alkane	β	$^1\chi$	Boiling point	Boiling point calculated ^a	Response
Pentane	219.00	2.41	36.10	37.17	−1.07
2-Methylbutane	202.00	2.27	27.90	30.10	−2.20
2,2-Di-methylpropane	160.00	2.00	9.50	11.66	−2.16
Cyclobutane	160.00	2.00	13.00	11.66	1.34
Hexane	303.00	2.91	68.70	68.69	0.01
2-Methylpentane	281.00	2.77	60.10	60.98	−0.88
3-Methylpentane	291.00	2.81	63.50	64.53	−1.03
2,3-Di-methylbutane	258.00	2.64	57.90	52.50	5.40
2,2-Di-methylbutane	251.00	2.56	49.70	49.84	−0.14
Cyclopentane	243.00	2.50	49.30	46.75	2.55
2,4-Di-methylpentane	355.00	3.13	80.50	85.37	−4.87
2,2,3-Tri-methylbutane	330.00	2.94	81.00	77.62	3.38
Heptane	404.00	3.41	99.30	99.11	0.19
2-Methylhexane	372.00	3.27	90.00	90.36	−0.36
3-Methylhexane	386.00	3.31	93.40	94.29	−0.89
2,3-Di-methylpentane	367.00	3.18	89.70	88.91	0.79
2,2-Di-methylpentane	346.00	3.06	79.20	82.64	−3.44
3,3-Di-methylpentane	356.00	3.12	86.00	85.67	0.33
3-Ethylpentane	396.00	3.35	93.40	97.00	−3.60
Cyclohexane	332.00	3.00	80.70	78.26	2.44
Octane	506.00	3.91	125.70	121.57	4.13
3-Methylheptane	489.00	3.81	119.00	118.41	0.59
3-Ethylhexane	496.00	3.85	118.60	119.74	−1.14
3-Methylpentane, 3-ethylpentane	480.00	3.68	118.3	116.64	1.66
3,4-Di-methylhexane	483.00	3.72	117.70	117.23	0.47
2-Methylheptane	482.00	3.77	117.70	117.04	0.66
4-Methylheptane	493.00	3.81	117.70	119.17	−1.47
2-Methylpentane, 3-ethylpentane	483.00	3.72	115.70	117.23	−1.53
2,3-Di-methylhexane	467.00	3.68	115.60	113.97	1.63
2,3,3-Tri-methylpentane	444.00	3.50	114.80	108.91	5.89
2,3,4-Tri-methylpentane	445.00	3.55	113.50	109.14	4.36
3,3-Di-methylhexane	461.00	3.62	112.00	112.69	−0.69
2,2,3-Tri-methylpentane	444.00	3.48	109.90	108.91	0.99
2,4-Di-methylhexane	468.00	3.66	109.50	114.18	4.68
2,5-Di-methylhexane	461.00	3.63	109.10	112.69	3.59
2,2-Di-methylhexane	450.00	3.56	106.90	110.27	3.37
2,2,3,3-Tetra-methylbutane	409.00	3.25	106.30	100.41	5.89
2,2,4-Tri-methylpentane	427.00	3.42	99.30	104.90	5.60

^a Calculated using the β equation.

centering on the accessibility of bonds, at least in alkanes, to intermolecular encounters.

The results of this study lead to a detailed structural interpretation of the molecular connectivity indices. After a quarter of a century of productive use, we can support our claim to the concept that molecular connectivity indices are indeed non-empirical structure descriptions that are rich in information. This study allows us to propose a meaning of the chi indices, that is the encoding of bimolecular accessibility. It is now clear why these indices are so useful in quantifying the relationship between structure and physical properties, particularly those that arise from intermolecular interactions. This realization should finally, correctly categorize these indices in their rightful place in the pantheon of quantitative descriptions of molecular structure.

We propose that treating the bonds of a molecule as *dissecta membra* is a model with a limited objective and a lim-

ited relationship to reality. It is an example of the analysis of a complex system using reduction to isolate relevant parts followed by synthesis using cellular automata dynamics to create a model that reveals some information about emergent properties and the role that the ingredients contribute to the whole. In this case, we considered just the bonds in alkanes, endowed with numerical values reflecting their accessibility to other bonds in other molecules. Our dynamic model is an alternative to the static model of the bimolecular encounter that we have offered to explain the molecular connectivity indices [20]. It is a dynamic model that, in a limited way, simulates the conditions of a molecule in its milieu. Objectives of the study were to reinforce the central idea in our bimolecular interpretation, confirm the reasonableness of our relative bond accessibility values and to extend the concept of the importance of bonds and their attributes in relating structure to properties. These studies contribute to these objectives.

The use of cellular automata to model the dynamics of fragments of a molecule, treated as isolated entities is a new idea that may have some merit in studies of ensembles of molecules. Novel possibilities are suggested by these results.

Acknowledgements

The author acknowledges the computer programming assistance of Professor C.-K. Cheng, the statistical assistance from Professor R. Young and helpful discussions with Professor U. Desai.

References

- [1] C. Hansch, A quantitative approach to biochemical structure–activity relationships, *Acc. Chem. Res.* 2 (1969) 232–239.
- [2] L.B. Kier, *Molecular Orbital Theory in Chemistry and Drug Research*, Academic Press, New York, 1971.
- [3] M. Randic, On characterization of molecular branching, *J. Am. Chem. Soc.* 97 (1975) 6606–6615.
- [4] L.B. Kier, L.H. Hall, W.J. Murray, M. Randic, Molecular connectivity. I. Relationship to non-specific local anesthetic activity, *J. Pharm. Sci.* 64 (1975) 1971–1976.
- [5] L.B. Kier, L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
- [6] L.H. Hall, L.B. Kier, The molecular connectivity chi indexes and kappa shape indexes in structure–property modeling, in: K.B. Lipkowitz, D.B. Boyd (Eds.), *Reviews of Computational Chemistry*, VCH, New York, 1991.
- [7] F.E. Norrington, R.M. Hyde, S.G. Williams, R. Wooten, Physicochemical activity relations in practice. 1. A rational self-consistent data base, *J. Med. Chem.* 18 (1975) 604–608.
- [8] R. Hoffmann, P. Laszlo, Representation in chemistry, *Angew. Chem. Engl. Ed.* 30 (1991) 1–112.
- [9] B. Testa, L.B. Kier, The concept of molecular structure in structure–activity relationships and drug design, *Med. Res. Rev.* 11 (1991) 35–47.
- [10] L.B. Kier, L.H. Hall, *Molecular Connectivity in Structure Activity Analysis*, Wiley, New York, 1986.
- [11] Molconn-Z, Hall Associates Consulting, Version 3.15, Quincy, MA.
- [12] P.G. Seybold, M. May, U.A. Bagal, Molecular structure–property relationships, *J. Chem. Ed.* 64 (1987) 575–581.
- [13] L.B. Kier, L.H. Hall, *Molecular Structure Description: The Electropotential State*, Academic Press, San Diego, CA, 1999, pp. 175–178.
- [14] J. Galvez, On the topological interpretation of electronic and vibrational molecular energies, *J. Mol., Struct. (Theochem.)* 429 (1998) 255–264.
- [15] J. von Neumann, in: A. Burks (Ed.), *Theory of Self-Reproducing Automata*, University of Illinois, Urbana, IL, 1966.
- [16] S. Wolfram, *Theory and Applications of Cellular Automata*, World Scientific, Singapore, 1986.
- [17] F.B. Ermentrout, L. Edelstein-Keshet, Cellular automata approaches to biological modeling, *J. Theor. Biol.* 160 (1993) 97–133.
- [18] B. Chopard, M. Droz, *Cellular Automata Modeling of Physical Systems*, Cambridge University Press, Cambridge, UK, 1998.
- [19] L. Kier, C.-K. Cheng, B. Testa, Cellular automata models of biochemical phenomena, *Future Generation Comput. Syst.* 16 (1999) 273–289.
- [20] L.B. Kier, L.H. Hall, Intermolecular accessibility: the meaning of molecular connectivity, *J. Chem. Inform. Comput. Sci.* 40 (2000) 792–795.