ELSEVIER

www.elsevier.nl/locate/farmac

Il Farmaco 54 (1999) 346–353

IL FARMACO

The E-state in database analysis: the PCBs as an example☆

Lemont B. Kier [a,*], Lowell H. Hall [b]

[a] *Department of Medicinal Chemistry, School of Pharmacy, Virginia Commonwealth University, Richmond, VA 23298, USA*
[b] *Department of Chemistry, Eastern Nazarene College, Quincy, MA 02170, USA*

Received 17 March 1999; accepted 25 March 1999

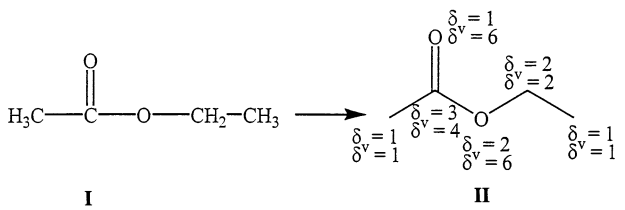
Abstract

The electrotopological state (E-state) is presented as a representation of atom and molecular fragment structure useful for structure–activity analysis and chemical database management. The E-state formalism is presented along with its extension to the atom-type E-state formalism. An approach to database analysis, using the polychlorobiphenyls (PCBs) as an example, reveals the descriptive power of the E-state paradigm. © 1999 Elsevier Science S.A. All rights reserved.

Keywords: Electrotopological state; Database management; PCBs; Parameter space; Atom indices

1. Introduction

With the development of combinatorial chemistry and high throughput screening, the availability of enormous databases has created a new and urgent need for the management of large amounts of molecular and biological information. Libraries of structure and biological information, either real or virtual, are now commonplace in the laboratories pursuing drug design. The cataloging, searching and analysis of databases involves both similarity and diversity methods. To carry out these analyses, the structures of molecules must now be encoded into coherent patterns for rapid calculation and manipulation. At this time, one of the best hopes for effective database management lies with the non-empirical structure indices encoding the topology and the electronic attributes of molecules and sub-fragments. The premiere example of this structure encryption is the electrotopological state (E-state) [1,2]. The utility of this paradigm is illustrated in this report, using the polychlorobiphenyls (PCBs) as an example.



I II

2. The electrotopological state (E-state)

We are aware through experiments and good models that fragments of molecules play significant roles in their biological function. It is important to identify these fragments, recognizing that they are a part of the whole, and at the same time playing semi-independent roles in biological encounters. This challenge has led us to consider the development of indices encoding structure information about the atoms and fragments within the context of the entire molecule. Any approach to this quantitation must be built upon the relationships and forces operating within the complex system of the molecule.

We can view each atom in a molecule as existing in a field within a molecule in which all other atoms participate [1,2]. This field participation is characteristic of any atom in a particular molecule. The methyl groups in ethyl acetate (**I**) are different from each other and from all other methyl groups in other molecules by virtue of their context, in spite of their common intrinsic states as a methyl groups. Quantifying the methyl

☆ Presented at the XIV Meeting of the Division of Medicinal Chemistry of the Italian Chemistry Society Salsomaggiore Terme, 21–25 September 1998.

* Corresponding author. Fax: +1-804-828 7625.

PII: S0014-827X(99)00039-7

group requires both an identity as a methyl group and its modification through the relationships to all other atoms in the molecule in which it resides. This modifying influence is propagated primarily through the for-mal bonding relationships within the molecule. It is a very good assumption that the covalent bonding structure or an abbreviation of it, the chemical graph, is the framework over which this influence is propagated. Consider the molecule of ethyl acetate, drawn as a chemical graph, (**II**). There are characteristics of the methyl groups and all other atoms which must be identified in order to describe them in a quantitative way, unique to their presence in ethyl acetate. We consider each of these characteristics in turn.

3. The intrinsic state of an atom

In ethyl acetate, we recognize that each atom or group possesses basic attributes that are identified as 'intrinsic' [1,2]. The first of these is the elemental content. In the case of the methyl group this is a composite of carbon and hydrogen atoms. A single atom in a molecule is described as a particular element while a group is defined by it component elements. The second attribute is the electron distribution which we represent as the hybrid state or more simply the valence state of the atom or group. This characterization includes the counts of σ, π and lone pair electrons comprising the valence electrons. In the case of a group such as methyl, we also must encode the number of hydrogens to distinguish it from the other hydrides of carbon, with other groups containing nitrogen and oxygen. A third attribute in an intrinsic state description is the degree of adjacency or more generally the local topological state of the atom or group. This attribute is important in defining the position of the atom relative to the topology of a molecule The carbon atoms in the three isomers of pentane each have a different spatial domain or accessibility. The terminal methyl groups in pentane are mantle fragments residing on the periphery of the molecule, hence they are easily accessible to interactions with neighboring molecules. In contrast, the methylene group in isopentane is located within the structure of the molecule with somewhat diminished accessibility to an intermolecular contact. Finally, the central atom of neopentane is buried deep within the molecule with no accessibility to any intermolecular interaction.

4. The atom representation

To develop a representation of molecular structure at the atom level, it is first necessary to characterize the atoms in such a manner that the atom descriptors can be used at the higher molecular level of description. For this we draw upon structural information encoded in the δ values.

5. The δ values

The atom description is essentially a statement of topological and electronic structure and the distribution of valence electrons among various orbitals in hybrid states. These are the defining characteristics among covalently bound atoms in organic molecules, including those of interest in biological systems. Using the chemical graph we have encoded this information from counts of electrons [3–6]. We associate with each atom, two values used in our earlier work on molecular connectivity descriptions. The first is the simple δ value which is the count of adjacent atoms other than hydrogen. In **II** the δ values in the molecule of ethyl acetate are shown. Note that just the adjacency is encoded which is equivalent to describing just the σ bond skeleton structure. A second set of δ values is included for each atom in the molecule in **II**. These are based on the total number of valence electrons on the atom excluding those bonding hydrogen. These have been called valence δ^v values in earlier work. These two sets of δ values are the basic ingredients for the definition of the intrinsic state of atoms in molecules. We summarize the definitions and the attributes of these parameters.

The simple δ is defined:

$$\delta = \sigma - h$$

where σ is the count of electrons in σ orbitals and h is the count of bonded hydrogen atoms. The simple δ value encodes: (i) the count of adjacent atoms excluding hydrogen; (ii) the count of σ electrons on an atom excluding hydrogen; (iii) the count of bonds joining to an atom other than hydrogen; and (iv) the immediate topological environment of the atom in the molecule.

The valence δ value is defined as follows:

$$\delta^v = Z^v - h = \sigma + \pi + n - h$$

Table 1
δ values for carbon, nitrogen and oxygen

Atom	Hybrid state	δ^v	δ
>C<	sp^3	4	4
=C<	sp^2	4	3
≡C–	sp	4	2
>N–	sp^3	5	3
=N–	sp^2	5	2
≡N	sp	5	1
–O–	sp^3	6	2
=O	sp^2	6	1

in which π is the number of electrons in π orbitals and n is the number of electrons in lone pair orbitals. The valence δ^v value encodes: (i) the count of valence electrons on an atom other than to hydrogen; and (ii) the count of σ, π and lone pair electrons excluding bonds to hydrogen.

Table 1 summarizes the δ and δ^v values for some covalently bonding atoms in molecules. It is evident that $\delta^v - \delta = \pi + n$, the count of π and lone pair electrons on an atom in a molecule [7]. This information referred to as the Kier–Hall electronegativity provides a quantitative measure of the potential of the atom for intermolecular interaction and reaction. It has a high correlation with the Mulliken–Jaffe electronegativity of atoms in their valence states [8]. The electronegativity of an atom in a molecule is of major importance within the context of the general information field described earlier. As a consequence this simple statement of structure has a significant role in encoding the intrinsic state of the atom.

6. The intrinsic state algorithm

The derivation of an intrinsic state value, labeled I, begins with the use of the $\delta^v - \delta$ term. Of equal importance in defining an intrinsic state is the adjacency or topology of the atom in the molecule [9]. Accordingly, the intrinsic state encodes two attributes: (1) the availability of the atom or group for intermolecular interaction (its potential for electronic interaction); and (2) the manifold of bonds over which adjacent atoms may influence, and be influenced by its state.

The adjacency, encoded by the simple δ value must therefore be a companion descriptor with the electronegativity in defining the intrinsic state. One possibility is to use the reciprocal of the adjacency, $1/\delta$, as an index of topological accessibility. The larger this value, the greater the topological accessibility of an atom or group. The product of the two terms produces a provisional description of the intrinsic state value, $(\delta^v - \delta)/\delta$. In this form, the intrinsic state may be viewed as the ratio of the π and lone pair electron count to the count of avenues of intramolecular interaction, the number of σ bonds in the skeleton for this atom. That is, due to the intramolecular interaction associated with the atom, π and lone pair electron density may be redistributed across the bonding network, which is the set of σ bonds in the molecular skeleton.

This expression for the intrinsic state for the carbon sp^3 atom is zero since $\delta^v = \delta$ in every case. If we scale the $\delta^v - \delta$ term by one, the zero values are eliminated and there is a discrimination among the various hydrides of carbon, arising from the different values of δ. This modification leads to the expression $(\delta^v - \delta + 1)/\delta$. This achieves the objective of encoding electronic structure, topology, and a fortuitous effect, an approximation of the valence state electronegativity. A simplification of this expression can be made by adding 1 to produce the intrinsic state I, for an atom or group in a molecule.

$$I = (\delta^v + 1)/\delta \quad (1)$$

Table 2 shows the intrinsic states of second row atoms and groups in the second quantum level. For higher quantum level atoms, the equation is modified by $(2/N)\delta^v$ where N is the principle quantum number. The general equation for the intrinsic state of an atom or group becomes:

$$I = [(2/N)^2\delta^v + 1]/\delta \quad (2)$$

Table 2
Intrinsic state values of second row hydrides

Atom hydride group	$I = [(\delta^v + 1)/\delta]$
>C<	1.250
>CH–	1.333
–CH_2–	1.500
>C=	1.667
–CH_3, =CH–, >N–	2.000
≡C–, –NH–	2.500
=CH_2, =N–	3.000
–O–	3.500
≡CH, –NH_2	4.000
=NH	5.000
≡N, –OH	6.000
=O	7.000

7. Field influences on the intrinsic state

We have derived the intrinsic state of an atom or group, but this expression does not reflect its position or influence within the field of other atoms in a molecule. This influence may take the form of a perturbation of the intrinsic state using some characteristic of every other atom in the molecule. A reasonable choice is the intrinsic state of each other atom in the molecule. This approach utilizes electronegativities of other atoms to modify the state of each atom within the field of the overall molecular structure. The conduit for this influence is the network of bonds linking each atom with all others, synonymous with the chemical graph model.

A second consideration is the effect of separation of two atoms in a molecule on the influence each has on the intrinsic state of the other. Since the chemical graph is the model of the presence and connectivity of atoms within the molecule, the count of bonds or atoms in paths separating two atoms is one measure of the distance between them. This count was chosen for the unit of distance between any two atoms in a molecule. More precisely, the count of atoms in the minimum

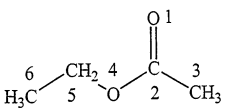
path length, r_{ij}, separating two atoms, i and j, is the distance selected to encode the influence between two atoms. Note that this count is equal to the usual graph distance, d_{ij}, plus 1 so that $r_{ij} = d_{ij} + 1$. From this model we have chosen the difference between the intrinsic states of atom i and atom j, $(I_i - I_j)$, as the perturbation on each other. This effect is assumed to diminish by some power, m, of the distance, hence, the perturbation of I_i, called ΔI_i, is expressed as:

$$\Delta I_i = (I_i - I_j)/r_{ij}{}^m \quad (3)$$

Most studies to date have employed a value of $m = 2$. The total perturbation of atom i is a consequence of the influence of all other atoms in the molecule. Accordingly, the total perturbation of the intrinsic state of atom i, ΔI_i, should be a sum of these individual perturbations, $\Sigma_j \Delta I_{ij}$, a sum of all terms expressed by Eq. (3). The actual state of atom i in a molecule is the intrinsic state, I_i, plus the sum of all perturbations included in ΔI_{ij}. This bonded state of atom i is called the electrotopological state, S_i, and is expressed as:

$$S_i = I_i + \Sigma_j \Delta I_{ij} \quad (4)$$

For brevity, the S_i term is called the E-state for atom i. The E-state index was introduced in a series of articles [1,2,10–13]. The E-state values for ethyl acetate are shown in **III**.



E-state values

S(1) = 9.82

S(2) = −0.20

S(3) = 1.38

S(4) = 4.41

S(5) = 0.48

S(6) = 1.78

III

8. Atom-type E-state indices

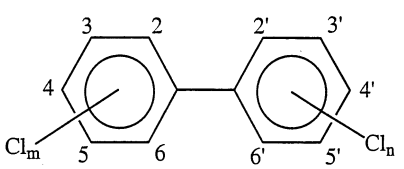
In an extension of the E-state method, an atom-type E-state value is defined for each type of atom (or hydride group such as –Cl, –OH, $-CH_2-$) in a molecule. Each atom is classified according to its valence state, number of bonded hydrogens and aromaticity [14]. For an atom-type E-state index, the E-state values are summed for all atoms of the same type in the molecule. The symbol for an atom-type E-state index is $S^T(X)$ where X denotes the atom (hydride group) such as $S^T(-Cl)$, $S^T(-OH)$ and $S^T(\cdots CH\cdots)$ for an aromatic carbon [1,2]. MOLCONN-Z currently recognizes 80 atom-types [15]. Atom-type E-state indices encode three distinct types of chemical structure information: (1) electron accessibility for the atom-type; (2) presence/absence of the atom-type; and (3) (approximate) count of atom-type present in the molecule.

The atom-type E-state indices are used for heterogeneous data sets for both QSAR/QSPR data sets [16–19] as well as for database analysis [20]. The atom level E-state indices defined earlier are only used in the cases in which there is a common molecular skeleton, as small as an amide group [17] or much larger, such as steroids [21].

9. Organization of database subsets

The individual E-state values for an atom or group may be thought of as numerical components of a space or basis vectors in a manifold containing all possible atoms or groups. Each dimension is a parameter calculated for a particular atom or group. For example, $-CH_3$, $-CH_2-$, –OH, etc. constitute a three-parameter description of the data set seen in Table 3. These structures can be displayed in a structure realm as shown in Fig. 1.

Within database subsets of molecules there are patterns of structure variation that are of interest in compound design. These patterns characterize the relative similarity or diversity within the subset. The E-state indices possess the ability to organize the subset in a manner which facilitates the design of molecular modifications, selection of diverse structures for testing, cluster analysis based on structure, and structure–activity analyses. To illustrate this organization and how the E-state indices can accomplish this, we look at the notorious PCBs (**IV**).



IV

Three atom-types are present in this series, each designated by its atom-type E-state code. These atom-types are the chlorine atom –Cl, the aromatic ···CH··· group and the substituted aromatic carbon atom. The atom-type E-state symbols for the first two are $S^T(-Cl)$ and $S^T(\cdots CH\cdots)$. Using the sum of E-states for each atom-type, a parameter space is created. To illustrate the structure organization possible, the two parameters $S^T(-Cl)$ and $S^T(\cdots CH\cdots)$ are used to create a two-dimensional space. A view of this space over a relatively wide range reveals many of the possible polychlorobiphenyls from the unsubstituted through the tetrasubstituted derivatives (see Fig. 2). The major parameter governing the position in this space is the count of the

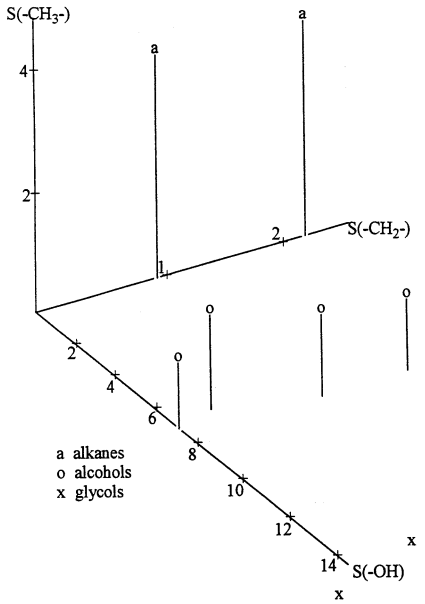


Fig. 1. Three atom-types used as coordinates in a structure space.

Table 3
Sum atom-type E-state values for a set of alcohols and glycols

	Molecule	$S^T(–CH_3)$ [a]	$S^T(–CH_2–)$ [b]	$S^T(–OH)$ [c]
1	$CH_3–OH$	1.00	0.00	7.00
2	$CH_3–CH–OH$	1.68	−0.25	7.57
3	$CH_3–CH_2–CH_2–OH$	1.93	1.19	7.88
4	$CH_3–CH_2–CH_2–CH_2–OH$	2.05	2.38	8.07
5	$HO–CH–CH_2–OH$	0.00	−0.25	15.25
6	HO–CH–CH–CH–OH	0.00	0.69	15.81
7	$CH_3–CH_2–CH_3$	4.25	1.25	0.00
8	$CH_3–CH_2–CH_2–CH_3$	4.36	2.64	0.00

[a] Symbol for the sum of E-state values for $–CH_3$ groups in the molecule.

[b] Symbol for the sum of E-state values for $–CH_2–$ groups in the molecule.

[c] Symbol for the sum of E-state values for –OH groups in the molecule.

number of chlorine atoms which is encoded indirectly in $S^T(–Cl)$. To extract more useful information, we examine subsets with common numbers of chlorine atoms by viewing a restricted range of parameter values.

The dichlorobiphenyls in atom-type E-state space for Cl and aromatic ···CH··· coded as $S^T(–Cl)$, $S^T(···CH···)$ space are shown in Fig. 3. The structures with their letter codes are shown in Table 4. The upper quadrant contains molecules *h* and *i* which are substituted on the rings at the 2- or 2′-positions in **IV**. The lower quadrant contains molecules with chlorine atoms far removed from the juncture of the two rings. In between these

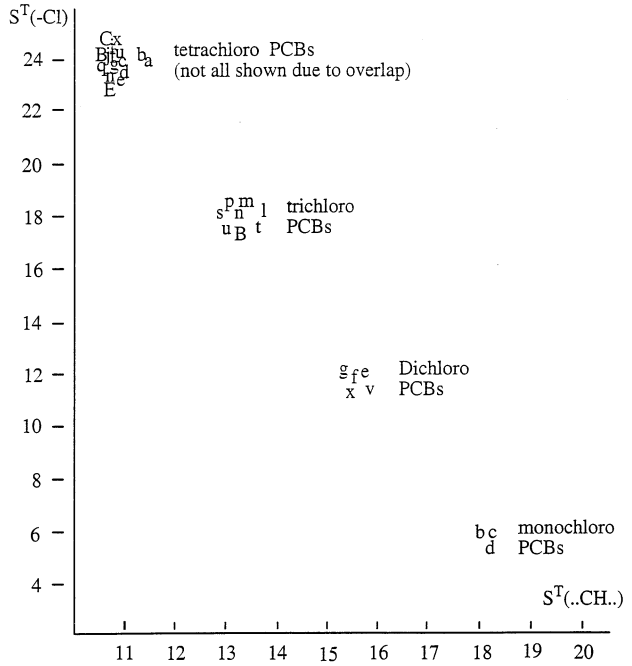


Fig. 2. A plot of PCBs in a space defined by two atom-type E-state indices mono- through tetrachloro-PCBs.

two subsets lie biphenyls substituted at intermediate positions. This set of PCBs is clearly organized in a meaningful way in terms of molecular structure and is indicated in Fig. 3 with the circles/ellipses of enclosure.

The organization of these molecules in this parameter space is more evident in the case of trichlorobiphenyls shown in Fig. 4. An examination of Fig. 4 shows a cluster of derivatives with all three chlorine atoms on the same ring. These are labeled t, s, r, and q corresponding to the derivatives 2,3,6; 2,4,6; 2,3,5; and 2,3,4. Within this cluster the molecules with chlorine atoms

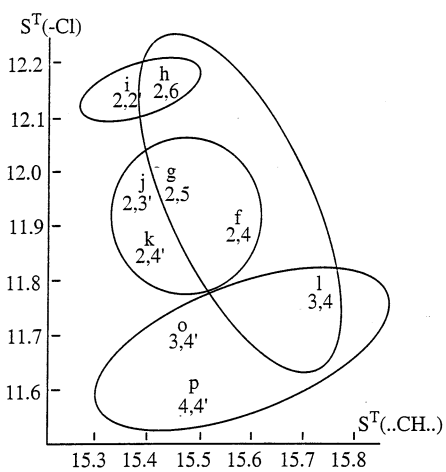


Fig. 3. A plot of dichloro-PCBs in a space defined by two atom-type E-state indices. The ellipses of enclosure are defined in the text.

Table 4
E-state indices for polychlorobiphenyls

Obs.	S [a]	PCB	Chlorine count	S^T(–Cl)	S^T(···CH···)	S^T(···C···)
1	a	biphenyl	0	0.000	20.781	0.0000
2	b	2-Cl	1	6.059	18.005	0.7997
3	c	3-Cl	1	5.895	18.078	0.7790
4	d	4-Cl	1	5.799	18.114	0.7775
5	e	2.3-diCl	2	12.028	15.604	1.2022
6	f	2,4-diCl	2	11.900	15.527	1.3457
7	g	2,5-diCl	2	11.980	15.429	1.4249
8	h	2,6-diCl	2	12.159	15.418	1.3680
9	i	2,2′-diCl	2	12.144	15.355	1.4456
10	j	2,3′-diCl	2	11.972	15.392	1.4696
11	k	2,4′-diCl	2	11.871	15.405	1.4959
12	l	3,4-diCl	2	11.768	15.713	1.1800
13	m	3,5-diCl	2	11.831	15.565	1.3266
14	n	3,3′-diCl	2	11.803	15.442	1.4768
15	0	3,4′-diCl	2	11.704	15.463	1.4937
16	p	4,4′-diCl	2	11.607	15.488	1.5050
17	q	2,3,4-triCl	3	17.942	13.429	1.3717
18	r	2,3,5-triCl	3	17.991	13.217	1.5959
19	s	2,3,6-triCl	3	18.155	13.144	1.6166
20	t	2,4,6-triCl	3	18.042	13.130	1.6825
21	u	2,3,2′-triCl	3	18.131	13.045	1.7389
22	v	2,3,3′-triCl	3	17.954	13.059	1.7908
23	w	2,3,4′-triCl	3	17.851	13.057	1.8357
24	x	2,4,2′-triCl	3	17.999	12.945	1.9103
25	y	2,4,3′-triCl	3	17.823	12.967	1.9528
26	z	2,4,4′-triCl	3	17.721	12.969	1.9920
27	A	3,4,2′-triCl	3	17.859	13.095	1.7893
28	B	3,4,3′-triCl	3	17.687	13.130	1.8150
29	C	3,4,4′-triCl	3	17.586	13.140	1.8447
30	E	3,5,2′-triCl	3	17.927	12.969	1.9080
31	F	3,5,3′-triCl	3	17.753	12.996	1.9431
32	G	3,5,4′-triCl	3	17.651	13.002	1.9785
33	a	2,3,4,5-tetraCl	4	23.980	11.345	1.3888
34	b	2,3,4,6-tetraCl	4	24.111	11.159	1.554
35	c	2,3,4,2′-tetraCl	4	24.060	10.938	1.8271
36	d	2,3,4,3′-tetraCl	4	23.880	10.936	1.8975
37	e	2,3,4,4′-tetraCl	4	23.774	10.923	1.9552
38	f	2,3,5,2′-tetraCl	4	24.113	10.749	2.0234
39	g	2,3,5,3′-tetraCl	4	23.931	10.740	2.1033
40	h	2,3,5,4′-tetraCl	4	23.824	10.723	2.166
41	i	2,3,6,2′-tetraCl	4	24.285	10.713	1.9994
42	j	2,3,6,3′-tetraCl	4	24.100	10.690	2.0961
43	k	2,3,6,4′-tetraCl	4	23.991	10.664	2.1688
44	l	2,4,5,2′-tetraCl	4	24.013	10.762	2.0498
45	m	2,4,5,3′-tetraCl	4	23.832	10.761	2.1202
46	n	2,4,5,4′-tetraCl	4	23.726	10.748	2.1778
47	o	2,4,6,2′-tetraCl	4	24.167	10.675	2.0932
48	p	2,4,6,3′-tetraCl	4	23.984	10.660	2.1805
49	q	2,4,6,4′-tetraCl	4	23.876	10.640	2.2475
50	r	3,4,5,2′-tetraCl	4	23.888	10.974	1.8512
51	s	3,4,5,3′-tetraCl	4	23.711	10.987	1.9047
52	t	3,4,5,4′-tetraCl	4	23.607	10.981	1.9530
53	u	2,3,2′,3′-tetraCl	4	24.132	10.802	1.9510
54	v	2,3,2′,4′-tetraCl	4	23.996	10.687	2.1409
55	w	2,3,2′,5′-tetraCl	4	24.085	10.627	2.1737
56	y	2,3,2′,6′-tetraCl	4	24.277	10.676	2.0441
57	z	2,4,2′,4′-tetraCl	4	23.861	10.577	2.3251
58	A	2,4,2′,5′-tetraCl	4	23.949	10.512	2.3635
59	B	2,4,2′,6′-tetraCl	4	24.139	10.553	2.2434
60	C	2,5,2′,5′-tetraCl	4	24.038	10.452	2.3963
61	D	2,5,2′,6′-tetraCl	4	24.230	10.501	2.2668
62	E	2,6,2′,6′-tetraCl	4	24.425	10.563	2.1204
63	F	3,4,3′,4′-tetraCl	4	23.579	10.859	2.1032
64	G	3,4,3′,5′-tetraCl	4	23.647	10.736	2.2185
65	H	3,5,3′,5′-tetraCl	4	23.717	10.618	2.3281

[a] Symbols as used in Figs. 2–5.

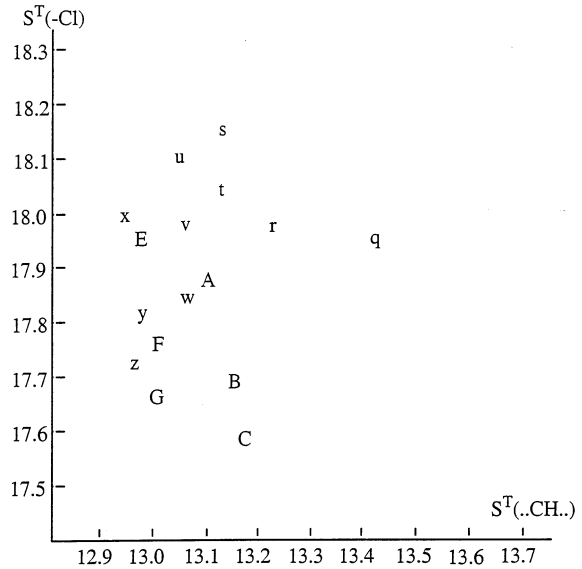


Fig. 4. A plot of trichloro-PCBs in a space defined by two atom-type E-state indices. The letter codes are defined in Table 4.

closer to the ring juncture in **IV**, are higher in the grid, ranking in this order. As can be seen here, the pattern of organization for trichloro compounds is similar to that found for the dichloro compounds.

A second pattern in the grid of trisubstituted PCBs is revealed in examination of the location of compounds substituted in the positions near the ring junctures. This includes the 2,3,6 and the 2,4,6 just discussed and also the 2,3,2′ and 2,4,2′ derivatives. At the other extreme, if substituent locations are far from the ring junctures as in the 3,4,4′; 3,5,4′; and the 3,4,3′ derivatives, these molecules are located the lower region of the grid. A central cluster of molecules in the grid are the derivatives of both rings including a 2,3 substitution pattern. Below these are the mixed ring derivatives with 3,3′ substituents always present.

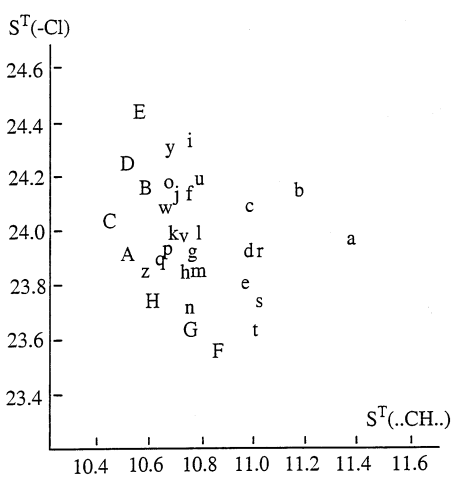


Fig. 5. A plot of PCBs in a space defined by two atom-type E-state indices. The letter codes are defined in Table 4.

The tetrasubstituted derivatives of PCB are shown in Fig. 5. The same general organization of substitution patterns is observed here. In addition, the derivatives that have two chlorines on each ring are found in a segment of the plot on the left margin. As expected, the 2,6,2′,6′ is near the top of the plot while the 3,4,3′,4′ is at the bottom of a curved area containing all evenly distributed derivatives.

The potential for organizing subsets of molecules from a database is made evident with this example. Clearly more complex structures with greater variation in their structures may show patterns which are more complex and less simply arranged but no less meaningful. The illustration here reveals the potential for two- or three- or multi-dimensional plotting of molecules to reveal similarity or diversity.

10. Conclusions

The E-state methodology, especially in its atom-type form, is shown to be a sound basis for the creation of an *n*-dimensional space in which to represent chemical structures. Molecules in this space are arranged in a coherent fashion in terms of important aspects of molecular structure. Structures are grouped according to significant molecular structure features such as aromatic rings, halogens, functional groups without explicit specification of such groupings. Further, inherent in the E-state formalism is a measure of electron accessibility which is an important basis for understanding noncovalent intermolecular interactions. This space reveals organization arrangements which mirror chemical structure intuition and in addition information relative to intermolecular interaction.

Acknowledgements

L.K. thanks Professor Pier Vincenzo Plazzi and the organizers of the Convegno for the invitation to present this work.

References

[1] L.B. Kier, L.H. Hall, Molecular Structure Description: The Electrotopological State, Academic Press, San Diego, CA, 1999.
[2] L.B. Kier, L.H. Hall, The electrotopological state: structure modeling for QSAR and database analysis, in: J. Devillers (Ed.), Topological Indices and Related Descriptors in QSAR and QSPR, Gordon and Breach, London, 1999.
[3] L.B. Kier, L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
[4] L.H. Hall, L.B. Kier, Molecular connectivity and substructure analysis, J. Pharm. Sci. 67 (1978) 1743–1747.
[5] L.B. Kier, L.H. Hall, Molecular Connectivity in Structure–Activity Analysis, Wiley, London, 1986.

[6] L.H. Hall, L.B. Kier, A molecular connectivity study of electron density in alkanes, Tetrahedron 33 (1977) 1953–1957.
[7] L.B. Kier, L.H. Hall, Derivation and significance of valence molecular connectivity, J. Pharm. Sci. 70 (1981) 583–587.
[8] (a) R.S. Mulliken, A new electroaffinity scale, J. Chem. Phys. 2 (1934) 782–793. (b) J. Hinze, H.H. Jaffe, Electronegativity I. Orbital electronegativity of neutral atoms, J. Am. Chem. Soc. 84 (1962) 540–549.
[9] L.H. Hall, L.B. Kier, Determination of topological equivalence in molecular graphs from the topological state, Quant. Struct. Act. Relat. 9 (1990) 115–131.
[10] L.B. Kier, L.H. Hall, An electrotopological state index for atoms in molecules, Pharm. Res. 7 (1990) 801–807.
[11] L.H. Hall, L.B. Kier, The electrotopological state: structure information at the atomic level for molecular graphs, J. Chem. Inf. Comput. Sci. 31 (1991) 76–83.
[12] L.H. Hall, L.B. Kier, The electrotopological state: an atomic index for QSAR, Quant. Struct. Act. Relat. 10 (1991) 43–48.
[13] L.B. Kier, L.H. Hall, J.W. Frazer, An index of electrotopological state for atoms in molecules, J. Math. Chem. 7 (1992) 229–237.
[14] L.H. Hall, L.B. Kier, Electrotopological state indices for atom-types: a novel combination of electronic, topological and valence state information, J. Chem. Inf. Comput. Sci. 35 (1995) 1039–1045.
[15] MOLCONN-Z (Version 3.15) may be obtained from Hall Associates Consulting, 2 Davis Street, Quincy, MA; SciVision Inc., 128 Spring Street, Lexington, MA 02173; Edusoft, LC, PO Box 1811, Ashland, VA 23005; and Tripos, Inc., 1699 South Hanley Road, St. Louis, MO 63144.
[16] L.H. Hall, T.A. Vaughn, QSAR of phenol toxicity using electrotopological state and κ shape indices, Med. Chem. Res. 7 (1997) 407–416.
[17] J. Gough, L.H. Hall, Modeling the toxicity of amide herbicides using the electrotopological state, Environm. Tox. Chem. 18 (1999) 1069–1075.
[18] J. Gough, L.H. Hall, Modeling antileukemic activity of carboquiinones with electrotopological state and chi indices, J. Chem. Inf. Comput. Sci. 38 (1999) 356–361.
[19] L.H. Hall, C.T. Story, Boiling point and critical temperature of a heterogeneous data set: QSAR with atom-type electrotopological state indices using artificial neural networks, J. Chem. Inf. Comput. Sci. 36 (1996) 1004–1014.
[20] C. de Gregorio, L.B. Kier, L.H. Hall, QSAR modeling with the electrotopological state indices: corticosteroids, J. Comp. Aid. Molec. Des. 12 (1998) 557–561.
[21] L.H. Hall, L.B. Kier, B.B. Brown, Molecular similarity based on novel atom-type electrotopological state indices, J. Chem. Inf. Comput. Sci. 35 (1995) 1074–1080.