

The E-State as an Extended Free Valence

Lemont B. Kier*

Department of Medicinal Chemistry, Virginia Commonwealth University, Richmond, Virginia 23298

Lowell H. Hall

Department of Chemistry, Eastern Nazarene College, Quincy, Massachusetts 02170

Received January 24, 1997[®]

The relationship between the electrotopological state index (E-state) and the molecular orbital based index of free valence is examined. The dependence of both on electronegativity and topology are developed from comparative values. Calculations of both indices for a diverse set of conjugated carbons reveal a significant correlation. Since the E-state index is calculated for heteroatoms and all σ bonded atoms, it represents an extension of the concept of free valence applied to all atoms in a molecule.

INTRODUCTION

In our continuing effort to define molecular structure with nonempirical codes, we have recently focused attention on atom and hydride groups in molecules.^{1–5} This work has led to the development of a structure index encoding both the relative electronegativity of an atom and its topological state in terms of the number and arrangement of σ bonds. We have recognized the biphasic character of this information by calling this complex structure code the electrotopological state or E-state index.

In a comprehensive review we have described in detail the development of the ingredients of the algorithm, the significance of these ingredients, the information generated, and strategies for structure–activity applications.⁵ In that review we introduced some interpretations of the origins and structural significance of the calculated E-state values. In this article we would like to extend our examination of these issues. In particular we want to study a possible parallelism between the E-state indices and the molecular orbital reactivity index called free valence.

In pursuit of this objective we will recount the highlights of the derivation of the E-state algorithm with emphasis on those features with a link to molecular orbital theory. We will then review the concept and calculation of free valence along with a brief review of its use. The heart of this study will be the presentation of arguments expressing a common basis for these two indices. We believe that the E-state index extends the information anticipated for the elusive free valence index for heteroatoms. The consequences of this conclusion will be expressed in terms of future work and possible new insights into predicting intermolecular reactivity.

Formation of an Atom-Centered Index. We represent a molecule as a hydrogen-suppressed graph (i.e., a chemical graph) in which the atoms are identified as elements with certain valence states. The σ bonds are represented by dimension-less connections between atoms. Our goal is to derive an index that encodes some electronic structure information about each atom in the chemical graph along with information about its topological environment. The ingredients in the chemical graph are (a) the presence of an

atom, (b) its valence state, hence by inference the counts of σ , π , and lone-pair electrons, and (c) the degree of adjacency. It is with these ingredients that we have created a parameter reflecting the electronic and the topological state of an atom in the chemical graph.

The topological state of an atom in a molecule can be perceived from direct examination of the graph. This attribute of an atom is a condition within a molecule which ranges between mantle atom and buried atom status. If we consider the case of neopentane, the four methyl groups are on the periphery or the mantle of the molecule. In contrast the central atom is buried within the molecule. The quantitation of this attribute may be accomplished if we use the degree of adjacency, which is equal to the count of σ electrons contributed by an atom in the chemical graph, a value which we call δ in the molecular connectivity paradigm. We can use the value of δ directly, but it may be more convenient to reciprocate the δ to give a primary atom the largest value of this attribute. Thus $1/\delta$ becomes an index of the topological state of an atom in a chemical graph.

The electronic attribute of an atom in a chemical graph is also encoded into the E-state index. We have at hand information about the number of π , σ , and lone-pair electrons associated with each atom. Since we are interested in applications involving the chemistry and biological activity of organic covalent compounds, we are primarily interested in the count of π and lone-pair electrons. These are the reactive, interactive species which give rise to intermolecular events other than covalent bond formation. The σ electrons, on the other hand, are far less responsible for intermolecular events, but they form the skeletal framework of the molecule, i.e., the topology of the molecule.

The Intrinsic State. To define the electronic attributes of an atom in a molecule, we use the electronegativity, which describes the polarity effects taking place between atoms within a molecule. The quantitation of electronegativity has been of great interest for many years. One of the best general schemes is that developed by Mulliken⁶ and extended by Hinze and Jaffe⁷ in which the electronegativity of valence state atoms is the average value of the ionization potential and electron affinity. In previous work on molecular connectivity we have demonstrated that the Mulliken valence

[®] Abstract published in *Advance ACS Abstracts*, May 1, 1997.

state electronegativity could be closely correlated with the difference between the count of valence electrons and the count of σ electrons on an atom.⁸ This count is referred to as the Kier–Hall electronegativity. The valence electron count on an atom in a chemical graph is designated by δ^v in molecular connectivity formalism. It is equal to the count of valence electrons contributed by an atom in a molecule, Z^v , minus the count of hydrogens on that atom, H . Thus $Z^v - H = \delta^v$ for second row elements. From these relationships it follows that the Kier–Hall electronegativity is $\delta^v - \delta$. Dissecting each of these terms into their electronic composition we obtain

$$\delta = \sigma \text{ electrons}$$

$$\delta^v = \sigma + \text{lone-pair electrons} + \pi \text{ electrons}$$

Therefore

$$\delta^v - \delta = \pi + \text{lone-pair electrons}$$

Thus the expression for the Kier–Hall electronegativity is identical to the count of electrons identified earlier as being related to intermolecular events among organic molecules. (A simplified statement of this electronegativity value is the column number in the periodic table minus the count of adjacent atoms). The Kier–Hall electronegativity index provides a way of enumerating the relative structure of an atom in a chemical graph in such a way that it conveys information about the potential for both intramolecular and intermolecular phenomena.

To unify the electronic attribute ($\delta^v - \delta$) and the topological attribute ($1/\delta$), we may consider the product of these two to obtain a single index:

$$(\delta^v - \delta)/\delta$$

An alternative way of interpreting this derivation is to state that it arises from a ratio of electronegativity of an atom to the σ bond count at that atom.

It is evident that in the case of the methyl, methylene, methine, and quaternary fragments, the value $\delta^v - \delta$ is zero. This expression does not distinguish among the four adjacency patterns of these groups. The expression can be scaled to distinguish among these by adding 1 to the numerator. The expression becomes

$$(\delta^v - \delta + 1)/\delta$$

We simplified this expression by adding 1. This produced a value, called I , for the intrinsic state of an atom in a chemical graph:

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

The values of the I index for second quantum level chemical graph atoms are shown in Table 1. If we reflect on this equation it is possible to detect an alternative interpretation for this index. Removing the constant, 1, leaves a simple ratio of the count of valence electrons to the adjacency of the atom in the chemical graph. We might think of δ^v/δ as a “density” of valence electrons associated with the atom and its σ bonded paths. Consider the case of a tertiary nitrogen atom in a chemical graph. There are five valence electrons and the adjacency is three. We can consider the adjacency to be equal to the number of paths

Table 1. Intrinsic State Values

atom (skeletal hydride group)	I [($\delta^v + 1$)/ δ]	atom (skeletal hydride group)	I [($\delta^v + 1$)/ δ]
>C<	1.250	—O—	3.500
>CH—	1.333	=CH, —NH ₂	4.000
—CH ₂ —	1.500	=NH	5.000
>C=	1.667	=N, —OH	6.000
—CH ₃ , =CH—, >N—	2.000	=O	7.000
≡C—, —NH—	2.500	—F	8.000
=CH ₂ , =N—	3.000		

associated with the nitrogen. This density quantitates the possible partitioning of π and lone pair electron influence along the paths emanating from the nitrogen atom. It also reflects the partitioning of influence that the nitrogen atom may have on atoms in its immediate environment. The less partitioning of the valence electron influence, the more available are the valence electrons for intermolecular interaction.

The Electrotological State Index. The intrinsic state of an atom in a chemical graph reflects its electronic and topological attributes in the absence of interaction with the rest of the molecule. The influence of all other atoms upon this atom mitigates the value of this intrinsic state. It is necessary to make some estimate of this perturbation in order to construct a model of the electronic and topological states of atoms in chemical graphs.

To estimate the perturbation of an intrinsic state value of an atom due to the influence of all other atoms, we dissect the graph into definable fragments. These are the sequence of skeletal bonds from atom i to atom j along the shortest path. Accordingly, we consider all such paths emanating from a specific atom. These paths begin with an atom under study, i , and end with the last atom in the path, j . This constitutes a fragment of two atoms separated by the remaining number of atoms in that path. The distance, r_{ij} , between the two extreme atoms is taken as the total number of atoms in the path.

The perturbing influence within each path on a terminal atom in that path may be estimated as the difference between the intrinsic states of the terminal atoms, $I_i - I_j$. We assumed that the electronic influence of atom j upon atom i in a fragment follows an inverse-square law. We may quantitate the perturbation of atom i due to each bonded fragment to be

$$(I_i - I_j)/r_{ij}^2$$

The total perturbation of I_i , expressed as ΔI_i , is therefore

$$\Delta I_i = \sum (I_i - I_j)/r_{ij}^2 \quad (2)$$

The summation is extended over the entire molecule. The state of each atom in a chemical graph due to the intrinsic state of that atom and the molecular field is called the electrotopological state, S_i . It is calculated to be the sum of these two terms.

$$S_i = I_i + \Delta I_i \quad (3)$$

Free Valence. The concept of free valence was introduced by Coulson as a molecular orbital description of potential reactivity at an atom.^{9,10} This idea goes back further to Thiele’s partial valency and Werner’s residual affinity. Free valence, F_r , is defined

Table 2. The Set of Hydrocarbons for Which Free Valence and *E*-State Values Were Calculated^a

Atom No.	Molecular Structure	Atom id	Group	Free Valence	E-State Value	Atom No.	Molecular Structure	Atom id	Group	Free Valence	E-State Value
1		a	CH2=	0.84	3.36	34		a	aCHa	0.44	2.08
2		b	=CH-	0.39	1.64	35		b	aCHa	0.40	2.03
3		a	CH2=	0.86	3.46	36		c	aCHa	0.41	2.02
4		b	=CH-	0.38	1.71	37		d	aCas	0.12	1.18
5		c	=CH-	0.46	1.83	38		e	=C<	0.14	1.08
6		a	CH2=	0.91	3.56	39		f	CH2=	0.89	4.10
7		b	=C<	0.10	0.87	40		a	aaCa	0.15	1.39
8		c	=CH-	0.42	1.66	41		a	=CH-	0.59	2.09
9		d	CH2=	0.82	3.46	42		a	CH2=	0.97	3.68
10		a	CH2=	0.88	3.52	43		a	aaCa	0.22	1.40
11		b	=CH-	0.37	1.71	44		a	aaCa	0.24	1.36
12		c	=CH-	0.52	1.85	45		b	=C<	0.62	2.12
13		d	=C<	0.15	1.00	46		a	CH2=	0.97	3.73
14		e	=CH-	0.41	1.73	47		b	=C<	0.07	1.04
15		f	CH2=	0.84	3.56	48		a	=C<	0.09	1.01
16		a	CH2=	0.83	3.52	49		a	CH2=	1.02	3.86
17		b	=CH-	0.41	1.66	50		b	=C<	0.06	1.04
18		c	=C<	0.12	0.83	51		a	=C<	0.24	1.31
19		d	CH2=	0.91	3.66	52		a	=CH-	0.62	2.28
20		a	CH2=	0.85	3.68						
21		b	=CH-	0.40	1.85						
22		c	=C<	0.19	0.96						
23		a	aCHa	0.44	2.02						
24		b	aCHa	0.39	2.00						
25		c	aCHa	0.41	1.99						
26		d	aCas	0.11	1.17						
27		e	=CH-	0.41	1.83						
28		f	CH2=	0.82	3.63						
29		a	aCHa	0.45	2.10						
30		b	aCHa	0.39	2.05						
31		c	aCHa	0.42	2.03						
32		d	aCas	0.10	1.23						
33		e	=CH-	0.48	2.12						

^a Showing the structure of the hydrocarbon and the identification of each particular atom (by letter) used in this study along with the free valence and the computed electrotopological state index (*E*-state).

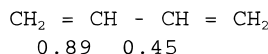
$$F_r = N_{\max} - N_r \quad (4)$$

where N_r is the sum of bond orders joining atom r , while N_{\max} is the maximum possible value. This is usually taken to be $3 + 3^{0.5}$ or 4.73 derived from the Hückel MO

calculation of N_r for the central atom in trimethylene-methane.

This N_{\max} value arises from the π bond order calculated to be $3^{0.5}$ plus an assumed value of 1.0 for each of the σ bonds.

An illustration of the results of this calculation is the molecule butadiene.



The π bond orders calculated using simple Hückel molecular orbital theory reveal a greater bond order for the 1,2 and 3,4 bonds. The total bond order values are estimated assuming that all C-H and C-C σ bonds have bond orders of 1.0. Summing the bond orders at each atom then gives 3.89 for the 1 and 4 position atoms and 4.34 for the 2 and 3 position atoms. The free valence values are then calculated to be 0.84 for the 1 and 4 positions and 0.39 for the 2 and 3 positions.

The interpretation is that this number encodes the relative amount of leftover or residual π bonding capability available for intermolecular interactions. This manifests itself in the potential reactivity at a particular atom. In the case of butadiene, the free valence correctly predicts the atom that is the most reactive to neutral reagents. A number of examples of the use of free valence to predict reactivity in aromatic hydrocarbons have been reviewed.¹¹

Gutman has analyzed the free valence algorithm in the light of the topology of the atoms in a conjugated system. He has shown a clear topological relationship between N_r in eq 4 and the summation of bonds at a graph vertex.

The general view is held that free valency is related to the forces starting a reaction which are largely due to residual affinity for approaching atoms rather than to strong interactions leading to transition state perturbations. These affinities permit the formation of incipient bonds without any severe deformation of the reagents. It is clear that in modern terms, we are looking at an index which mirrors the phenomena associated with receptor or protein, noncovalent bonding.

As more sophisticated all-valences electron MO methods emerged in the 1960s the concept of free valence was largely shelved in favor of population analyses, frontier orbitals, and electrostatic potentials to predict the relative reactivity of parts of molecules. The shortcoming of the classical free valence formulation is that it was derived for conjugated hydrocarbon systems and its extension to atoms with lone pair electrons was never made. If we can find some relationship between E-state and free valence of conjugated molecules, we may conclude that the E-state has an interpretation, comparable to an extended free valence index which includes sp^3 atoms and heteroatoms. Accordingly we have examined several polycyclic aromatic hydrocarbons with computed free valences in an effort to find possible correlations.

Comparison of Free Valence and E-State Values. To determine whether there is a significant relationship between the free valence and the corresponding E-state indices, we have extracted free valence values from a compilation of Hückel molecular orbital calculations by Coulson and Streitweiser¹³ of a wide variety of conjugated carbon atoms (Table 2). Examination of the atoms and molecules listed in Table 2 indicates that there is a limited number of possible types of carbon atoms. For these hydrocarbons, carbon atoms in a conjugated system are all designated as sp^2 in valence bond notation. In Table 2 there are six distinguishable bonding environments, as follows: $=CH_2$, $=CH-$, aaCa, aCHa, $=C<$, and aaC-, where an aromatic bond is

indicated by the letter "a". These atom types are distinct, that is, there are no intermediate states of bonding.

It is clear that the free valence values reflect this discrete character of atoms types, and hence the free valence values (Table 2) are somewhat clustered. Furthermore, the E-state index values preserve this chemically meaningful arrangement of atom types. The plot of E-state versus free valence values is shown in Figure 1. In the figure we have plotted, for each atom type, the mean values for E-state index and free valence. The "C" symbol in each atom type is located on the mean value point. The correlation for this six plot is $r = 0.98$, suggesting a very significant parallel between these values for E-state and free value.

Free Valence Indices for Unspecified Atoms. The free valence has been computed for a series of unspecified atoms attached to a general conjugated system $C=C-X$ where X may be a heteroatom contributing π or lone-pair orbital electrons to the system. The inclusion of atoms or groups such as this have been treated by assuming that their influence on the conjugated system is due to their variable electronegativity. This has been encoded into the Hückel MO calculation as parameters modifying the Coulomb integral, α , and the resonance integral, β .^{11,14}

In Hückel MO calculations, the Coulomb integral is an approximation of the attraction between a π electron and the atom core. It is expressed in this case by employing a parameter, α , encoding the relative electronegativity of the unspecified atom X. The parameters are h , modifying the Coulomb integral, and k , modifying the resonance integral. The modifiers take the form

$$\alpha_x = \alpha_c + h\beta \quad \beta_{cx} = k\beta_{cc}$$

The calculated free valence values for a variety of X groups is shown in Table 3. As the electronegativity increases with an increase in h , the free valence index also increases. This clearly establishes a relationship between this electronic attribute for an atom and the free valence index.

DISCUSSION

In the study in which we compare E-state values and free valence, we find a strong relationship. From this we surmise that at a level of structure in which free valence can be calculated, there is a strong parallel with E-state values. The free valence calculations were introduced to account for π electron influences in conjugated systems. At this level, the two indices portray a common picture when the numerical values are compared for the same molecules.

From a theoretical consideration, the perturbed intrinsic state in eq 2 and the summed bond orders impinging on an atom in eq 4 convey essentially the same information. Both are derived from relative electronegativities of bonded atoms, manifesting a polarity between the two. The cumulative effect of these polarity differences result in a certain amount of unbonded electron content at each atom. In the E-state formalism, all valence electrons are included in the electronegativity calculation hence the derived intrinsic state value contains the unbonded part of the electron component.

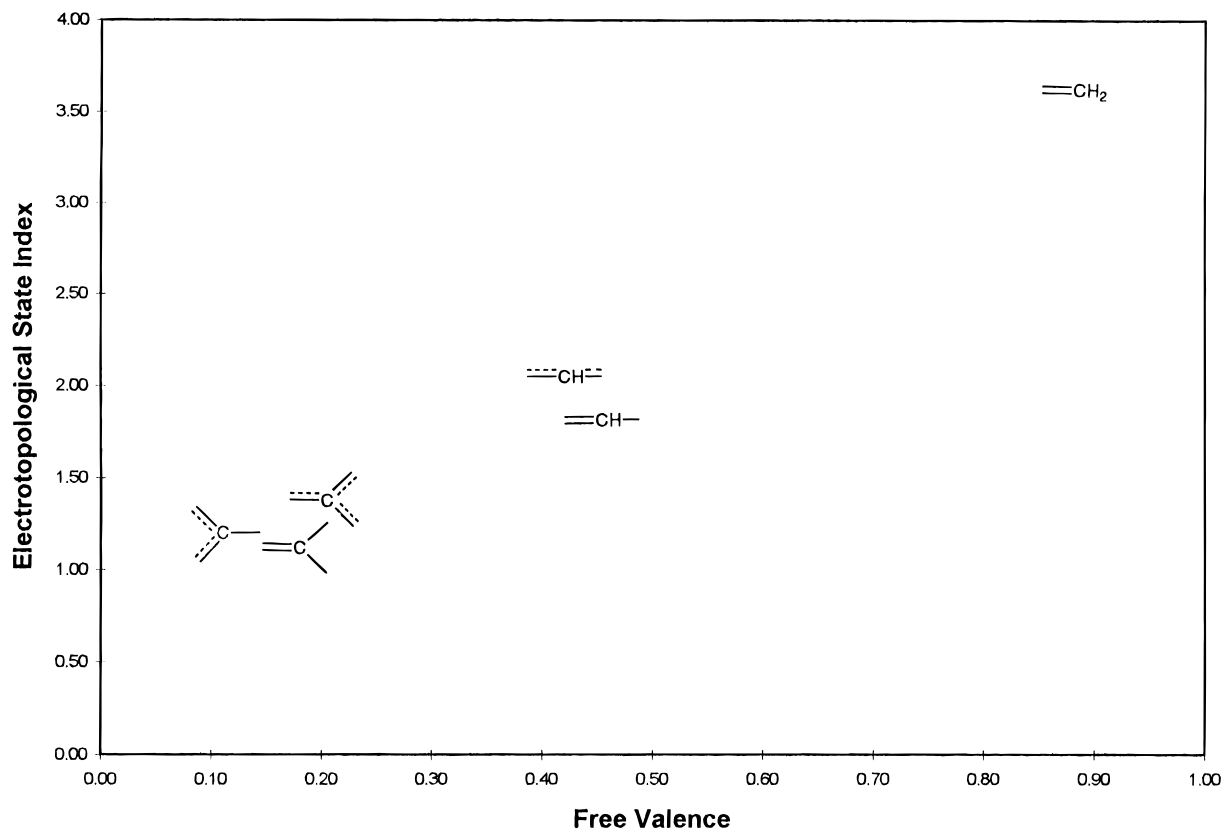


Figure 1. The relationship between the free valence for various conjugated carbon atoms and the calculated E-state indices.¹⁶ The molecules and the positions calculated are shown in Table 2.

Table 3. Free Valence and Hückel MO Coulomb Integral

$$\alpha_x = \alpha_c + h_x \beta$$

$$\beta_{cx} = kc - c$$

CC X

h_x	$kc-x$	F_r	h_x	$kc-x$	F_r
0.00	0.50	1.28	1.00	1.00	1.25
0.50	0.50	1.41	0.00	1.50	0.90
1.00	0.50	1.48	0.50	1.50	0.98
0.00	1.00	1.02	1.00	1.50	1.07
0.50	1.00	1.15			

In the case of the free valence, the calculation is built around the π electrons resident in a conjugated system. The other valence electrons, the lone pairs, are only implicit in the modification of the carbon Coulomb integral used to characterize a heteroatom. In the free valence calculations, the σ bonds are counted as contributing a value of one in the summation at each atom.

We say that all of the valence electrons in the E-state calculation are integrated into the formalism in such a way that separate terms are not necessary for the estimation of their contributions. As a result, the relative availability of the electrons accessible to intermolecular interactions are quantitated more explicitly with the E-state index than has been possible to achieve with the free valence as originally formulated. We have used the E-state index to successfully encode the relative intermolecular binding of barbiturates to cyclodextrin,¹ the binding of β -carboline to the benzodiazepine receptor,³ the binding of indolalkylamines to 5-HT₂ receptors,⁴ and the creation of a 3-D CoMFA field.¹⁵ The results of this study lend a greater credibility to the E-state

when its relationship to the molecular orbital free valence index is appreciated.

ACKNOWLEDGMENT

L.H.H. acknowledges the DuPont Company for partial support of this work.

REFERENCES AND NOTES

- (1) Kier, L. B.; Hall, L. H. An Electropotential State for Atoms. *Pharm. Res.* **1990**, *7*, 801–807.
- (2) Hall, L. H.; Mohny, B.; Kier, L. B. Structure Information at the Atomic Level for Molecular Graphs. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 76–82.
- (3) Hall, L. H.; Mohny, B.; Kier, L. B. The Electropotential State: An Index for QSAR. *Quant. Struct.-Act Relat.* **1991**, *10*, 43–51.
- (4) Kier, L. B.; Hall, L. H.; Frazer, J. W. An Index of Electropotential State for Atoms in Molecules. *J. Math. Chem.* **1992**, *7*, 229–235.
- (5) Kier, L. B.; Hall, L. H. Atom Description in QSAR Models: Development and Use of an Atom Level Index. *Adv. Drug Res.* **1992**, *22*, 1–37.
- (6) Mulliken, R. S. A New Electroaffinity Scale. *J. Chem. Phys.* **1934**, *2*, 782–793.
- (7) Hinze, J.; Jaffe, H. H. Electronegativity I. Orbital Electronegativity of Neutral Atoms. *J. Am. Chem. Soc.* **1962**, *84*, 540–549.
- (8) Kier, L. B.; Hall, L. H. Derivation and Significance of Valence Molecular Connectivity. *J. Pharm. Sci.* **1981**, *70*, 583–587.
- (9) Coulson, C. A. *Valence*; Oxford University Press: 1948; p 271.
- (10) Coulson, C. A. The Theory of the Structure of Free Radicals. *Discuss. Faraday Soc.* **1947**, *2*, 9–18.
- (11) Streitwieser, A. *Molecular Orbital Theory for Organic Chemists*; John Wiley: New York, pp 329–341.
- (12) Gutman, I. Topological Formulas for Free-Valence Index. *Croat. Chem. Acta* **1978**, *51*, 29–33.
- (13) Coulson, C. A.; Streitwieser, A. *Dictionary of Pi-Electron Calculations*; W. H. Freeman Co.: San Francisco, CA.
- (14) Kier, L. B. *Molecular Orbital Theory in Drug Research*; Academic Press: New York, 1971.
- (15) Kellogg, G.; Kier, L.; Gaillard, P.; Hall, L. E-State Fields: Applications to 3-D QSAR. *J. Comp.-Aided Molec. Design* **1997**, *10*, 513–520.
- (16) Computation of E-state values done using Molconn-Z.

CI970002B