

RESEARCH ARTICLES

Derivation and Significance of Valence Molecular Connectivity

LEMONT B. KIER** and LOWELL H. HALL†

Received September 8, 1980, from the *Department of Pharmaceutical Chemistry, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298, and the †Department of Chemistry, Eastern Nazarene College, Quincy, MA 02170. Accepted for publication January 7, 1981.

Abstract □ The physical basis for valence molecular connectivity was studied. The δ^v and δ values are cardinal numbers describing the electronic structure of atoms in their valence states. The value $\delta^v + \delta$ describes the volume of a bonding atom while the value $\delta^v - \delta$ describes the electronegativity. By using the principle of electronegativity equalization, bond electronegativity is defined as $(\delta_i^v \delta_j^v)^{-1/2}$, and the valence molecular connectivity index (${}^1\chi^v$) is derived as a sum of these bond descriptions. The valence chi index is interpreted in terms of the information encoded, describing both the volume and electronic characteristics of bonds in molecules. Examples of close relationships with molecular volume and electronic properties are shown. A new way of estimating valence state electronegativity is proposed from a count of exterojacent electrons divided by the quantum number squared for at least the first three quantum levels.

Keyphrases □ Molecular connectivity indexes—relationship between molecular volume and electronic composition, mathematical estimation of valence state electronegativity □ Structure—activity relationships—usefulness of molecular connectivity indexes, mathematical estimation, relationship between molecular volume and electronic composition □ Electronegativity—mathematical estimation, usefulness of molecular connectivity indexes, relationship between molecular volume and electronic composition

Molecular connectivity is a method for the quantitation of molecular structure that encodes information about size, branching, cyclization, unsaturation, and heteroatom content (1). It was developed in several stages, beginning with an alkane branching index proposed by Randic (2), a treatment of unsaturation (1, 3), a development of extended bond quantitation (4), and, finally, the most useful addition for the medicinal chemist, a rational way of quantifying heteroatom content (5). The treatment of heteroatoms, introduced by Kier and Hall (1) and called valence molecular connectivity, brought molecular connectivity into the role of a structure quantitation method useful in structure-activity analyses of drug molecules. Several investigators have employed valence molecular connectivity in structure-activity relationship analyses of drugs (Table I).

The growing use of molecular connectivity has resulted in an interest in the significance of the elements of the method and the meaning of the calculated indexes. Several misconceptions have arisen as a result of incomplete understanding of molecular connectivity or of previous bias grounded in physicochemical relationships to drug action. This study explores the fundamental basis of molecular connectivity.

DISCUSSION

Method—The simple molecular connectivity index arises from the assignment of numerical adjacency values to each atom other than hydrogen (the molecular skeleton) in a molecule. These values, called δ values, are cardinal numbers enumerating the presence of nonhydrogenic-bonded neighbors, irrespective of what the element is or what multiplicity of bonding is present. The assignment of δ values for acetic acid is shown in Fig. 1a. The molecule is dissected into bonds, each described by two δ values.

A term for each bond (C_{ij}) is calculated according to $C_k = (\delta_i \delta_j)_k^{-1/2}$. These terms are summed over the entire molecule to give the simple molecular connectivity index of the first order (${}^1\chi$):

$${}^1\chi = \sum (\delta_i \delta_j)_k^{-1/2} = \sum C_k \quad (\text{Eq. 1})$$

Applied to acetic acid in Fig. 1a, the calculated value for ${}^1\chi = 3(1 \times 3)^{-1/2} = 1.732$.

An inspection of the acetone molecule in Fig. 1b shows an assignment of δ values. This set of δ values is the same as for acetic acid. It is possible that the ${}^1\chi$ value of 1.732 for both of these molecules is sufficiently descriptive so that certain general properties, e.g., size and shape, may be predicted from a relating equation. However, it is more than likely that the obvious difference in the structures must be encoded to relate the molecules to many properties.

To accomplish the encoding of structural differences, Kier and Hall went beyond the structural fact of adjacency and considered the count of the valence electrons participating in sigma, pi, and lone-pair orbitals on each atom, exclusive of bonds to hydrogen. This count, again a cardinal number, is the valence delta (δ^v) used to calculate a valence molecular connectivity index (${}^1\chi^v$). The assignment of the δ^v value to an atom is based on the number of valence electrons (Z^v) not involved in bonds to hydrogen. Thus, $\delta_i^v = Z_i^v - h_i$, where h_i is the number of bonded hydrogens on atom i . The significance of suppressing hydrogen in this count

Table I—Use of Valence Connectivity Indexes in Structure-Activity Relationship Studies

Study	Molecules	Reference
Toxicity of anesthetics	Ethers	14
Anesthetic potency	Ethers, ketones	15
Barbiturate potency	Barbiturates	16
Partition coefficient	Hydroxyureas	17
Carminative activity	Alcohols, esters	18
Chromatographic retention	Alcohols	19
Hückel pi electron energy	Aromatic hydrocarbons	20
Taste threshold	Alcohols, carbonyls, acids, esters	21
Bitter taste	Amino acids, peptides	22
Boiling point	Alcohols, amines, halides	23
Solubilities		23
Molar refraction		23
General anesthetics	Mixed	24
Cytochrome conversion	Phenols	25
Sweet taste	Nitroanilines	25
Enzyme inhibition	Benzyl alcohols	25
Toxicity	Phenyl phosphates	25
Odor	Benzaldehydes, nitrobenzenes	26
Hallucinogenic activity	Phenylisopropyl amines	27
Mutagenicity	Nitrosamines	28
Serotonin antagonists	Phenalkylamines	29
Serotonin antagonists	Lysergic acid amides	30
Muscarinic antagonists	Amines	31
Antimicrobial activity	Phenyl propyl ethers	32
Chromatographic retention	Mixed	33
General anesthetics	Halocarbons	34
Hallucinogenic agents	Phenethylamines	35
Monoamine oxidase inhibitors	Hydrazines	36
Hormone class	Steroids	37
Reductase inhibitors	Benzohydroxamic acids	38

will be discussed later. The calculation of ${}^1\chi^v$ proceeds exactly as described for the simple molecular connectivity index.

Figures 1c and 1d show the assignments of δ^v values to acetic acid and acetone, respectively. The calculated ${}^1\chi^v$ values are clearly different, 0.928 for acetic acid and 1.204 for acetone. Structural information is encoded in these two different values and also in the common value for the simple ${}^1\chi$ index. Studies have employed ${}^1\chi$ and ${}^1\chi^v$ indexes singly and in concert as structural descriptors in structure-activity relationship analyses. Since molecular information encoded in these χ values is derived from the δ values of the atoms, insight into the significance of molecular connectivity begins with analysis of the δ values.

Significance of Delta—Within the context of molecular connectivity, as developed by Kier and Hall, an atom in its valence state can be described by two cardinal numbers: δ , the number of bonded neighbors (excluding hydrogen), and δ^v , the number of valence electrons (excluding those bonding hydrogen). This dual description of atoms is revealed in the matrix shown in Fig. 2 for second quantum level atoms in their valence states. Superimposed on this same matrix are the hydrides of these atoms.

It is apparent that δ_i is a count of the nonhydrogen sigma-bond electrons contributed by atom i . It follows that δ_i^v is a more inclusive count of all valence electrons (not bonding to hydrogen), including those sigma electrons (σ_i) enumerated by δ_i . Thus, the relationship is:

$$\delta_i^v = \delta_i + p_i + n_i = \sigma_i + p_i + n_i - h_i \quad (\text{Eq. 2})$$

where p_i is the number of p orbital electrons and n_i is the number of lone-pair electrons on atom i . The values of δ^v and δ depend on the element represented, its valence (hybrid) state, and the number of bonded hydrogen atoms. Thus, from Fig. 2, $\delta = 2$; $\delta^v = 4, 5,$ and 6 describes an sp carbon, an sp^2 nitrogen, and an sp^3 oxygen, respectively. In contrast, δ^v

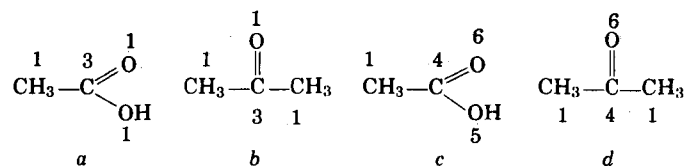


Figure 1—Assignment of δ and δ^v values to heteroatom molecules. Key: a, acetic acid with simple δ values; b, acetone with simple δ values; c, acetic acid with valence δ^v values; and d, acetone with valence δ^v values.

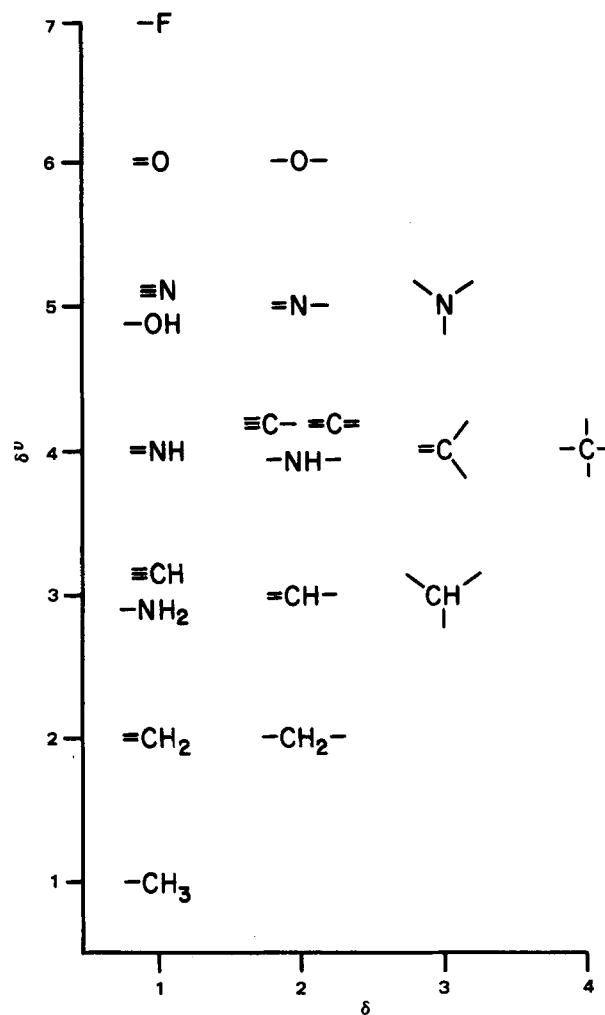


Figure 2—Atoms (and hydrides) defined by δ and δ^v values.

$= 4$; $\delta = 2, 3,$ and 4 describes an sp carbon, an sp^2 carbon, and an sp^3 carbon, respectively.

Both δ^v and δ contribute to a description of the hybrid state of an element or hydride. With these two cardinal numbers, 20 second quantum level atom valence states and hydrides are described in the matrix in Fig. 2.

It might be anticipated that the information available from δ^v and δ_i , when formulated into bond descriptions and ultimately summed to a molecular value, would be of considerable utility in ranking molecules according to physical property values and biological activities. A wealth of published experience confirms this expectation.

Two basically different kinds of information are encoded in the δ^v and δ values of atoms and groups shown in Fig. 2. First, since δ^v and δ count sigma, pi, and lone-pair orbitals, it is expected that the space occupied by these orbitals might be described. This orbital space defines the expanse of the atom in its valence state due to the number and kind of orbital domains of the atom. The measured or calculated volume of an atom (hence molecule) is a consequence of the number and kind of orbitals. The physicochemical manifestations of the volume (and closely related surface area) of a molecule are well known and are expected to be related to the volumes of the constituent atoms.

The second type of information expected from the δ^v and δ values of a covalently bound atom is of an electronic character. The numbers of sigma, pi, and lone-pair electrons must govern certain electronic properties that may be closely related to δ^v and δ of the atoms. When translated into bonds and molecules, this electronic information may bear a significant relationship to certain molecular properties.

Volume Information from δ and δ^v —A number of investigators have dissected molecules into constituent fragment volumes to estimate the molecular volume. A notable contribution in this regard is the work of Bondi (6). Superimposing the Bondi atom volumes onto the δ^v, δ matrix produces a pattern for direct comparison in Fig. 3.

Figure 3 reveals that the volume of the atom or hydride diminishes with

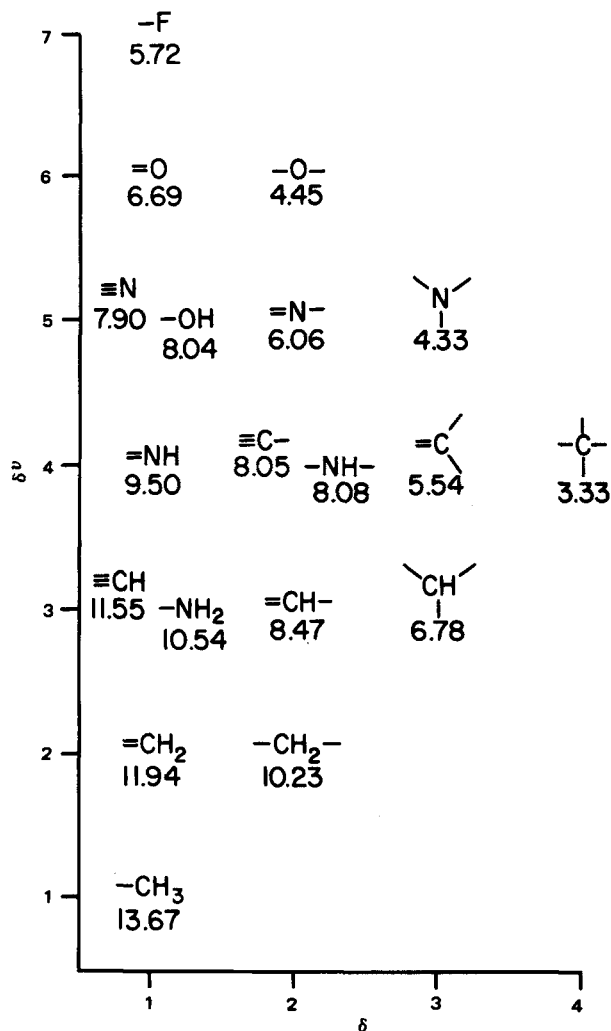


Figure 3—Bondi atom and hydride volumes and δ^ν , δ . The value for $-O-$ is averaged from alkyl aryl ethers.

an increasing δ^ν value when holding δ constant. The volume also diminishes when δ increases for a constant δ^ν value. These observations are not unexpected among cases where the number of hydrogens in the group diminishes. In cases where the fragment is a single atom in its valence state, the volume increases with an increase in the number of pi orbitals. This phenomenon occurs even though the number of valence electrons may be less. It is also apparent that the volume diminishes when the number of sigma electrons increases for a constant δ^ν value.

From these observations, it is concluded that the structural features encoded by δ^ν and δ act cooperatively to diminish the volume as numerical values increase. Such a relationship is revealed using regression analysis:

$$V_{\text{Bondi}} = 17.03 - 1.59(\delta^\nu + \delta) \quad (\text{Eq. 3})$$

$$r = 0.990 \quad s = 0.52 \quad n = 20 \quad F = 859$$

The calculated values are compared with the Bondi volumes in Table II. With Eq. 3, additional fragments and small molecules are predicted. The ability to predict fragment values where experimental data are lacking is an important consequence of this relationship.

So far, only second quantum level elements ($N = 2$) have been considered. A more general equation is possible, accounting for the greater distance of the valence electron orbitals from the nuclei in third quantum level atoms:

$$V_{\text{Bondi}} = 17.03 - 6.36 \frac{(\delta^\nu + \delta)}{N^2} \quad (\text{Eq. 4})$$

where N is the principal quantum number. For third-level atoms ($N = 3$), the equation predicts the volumes for $-S-$ and $-Cl$ of 11.38 cm^3/mole each, compared to 10.8 and 11.62 cm^3/mole from Bondi (6).

Bondi used the fragment values in a roughly additive way to estimate

Table II—Delta Values of Valence-State Atoms or Hydrides and Bondi Volumes

Atom/Hydride	δ	δ^ν	Volume, cm^3/mole	
			Bondi ^a	Calculated ^b
	4	4	3.33	4.31
	3	3	6.78	7.49
	2	2	10.23	10.67
$-\text{CH}_3$	1	1	13.67	13.85
$=\text{CH}_2$	1	2	11.94	12.26
$-\text{C}\equiv, =\text{C}=\equiv$	2	4	7.50 ^c	7.49
$\equiv\text{CH}$	1	3	11.55	10.67
$=\text{CH}-$	2	3	8.47	9.08
$-\text{O}-$	2	6	4.45 ^c	4.31
$-\text{OH}$	1	5	8.04	7.49
$-\text{NH}_2$	1	3	10.54	10.67
	2	4	8.08	7.49
	3	5	4.33	4.31
$\equiv\text{N}$	1	5	7.90	7.49
$\equiv\text{C}-$	3	4	5.54	5.90
$=\text{O}$	1	6	6.79	5.90
$=\text{NH}$	1	4	9.50	9.08
$=\text{N}-$	2	5	6.06	5.90
CH_4	0	0	17.12	17.03
NH_3	0	2	13.80	13.85

^a From Ref. 6. ^b From Eq. 3. ^c Average values.

molecular volume. Empirically based correction factors were introduced to account for branching, cyclization, and other structural effects. The formalism of molecular connectivity achieves a similar result without tabulated correction factors. The δ values are used to describe bonds that are summed to give a molecular index. The molecular index (${}^1\chi^\nu$) can correlate closely with experimental or computed volumes derived from various methods or procedures. For example, amino acid side-chain molar volumes in Table III correlate with a single-valence connectivity index, ${}^1\chi^\nu$ (7):

$$V(A)^3 = 29.46 {}^1\chi^\nu + 36.30 \quad (\text{Eq. 5})$$

$$r = 0.989 \quad s = 3.93 \quad n = 18 \quad F = 696$$

Electronic Information from δ and δ^ν —The δ^ν and δ values are

Table III—Amino Acid Side-Chain Volume and ${}^1\chi^\nu$

Amino Acid, R	${}^1\chi^\nu$	$V(A)^3$ ^a	$V(A)^3$ calc.
Glycine	0	36.3	36.3
Alanine	0.577	52.6	53.3
Valine	1.488	85.1	80.1
Leucine	1.971	102.0	94.4
Isoleucine	2.026	102.0	96.0
Phenylalanine	2.672	113.9	115.0
Proline	1.408	73.6	77.8
Tryptophan	3.417	135.4	137.0
Tyrosine	2.807	116.2	119.0
Histidine	2.079	91.1	97.6
Threonine	1.169	71.2	70.7
Serine	0.724	54.9	57.6
Asparagine	1.255	72.4	73.3
Glutamine	1.755	92.7	88.0
Aspartic acid	1.190	68.4	71.4
Glutamic acid	1.690	84.7	86.1
Lysine	2.224	105.1	101.8
Arginine	2.550	109.1	111.4

^a From Ref. 7.

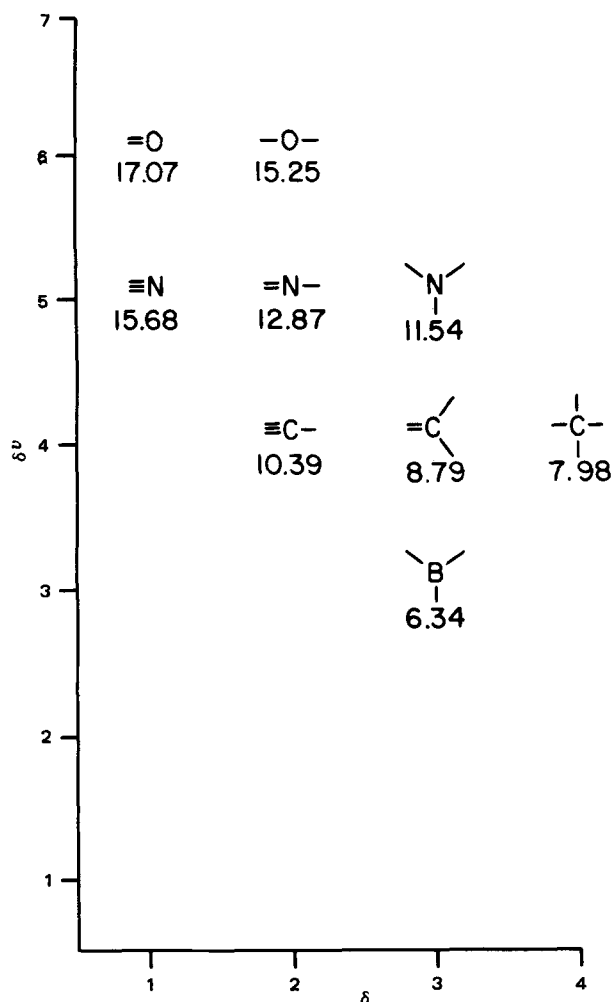
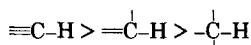


Figure 4—Mulliken electronegativities and δ^ν , δ .

counts of electrons in sigma, pi, or lone-pair orbitals. The number and distribution of electrons certainly influence the electronic characteristics associated with atoms in the valence state. An obvious electronic property of a bonded atom in a molecule is electronegativity. Pauling (8) proposed this concept and defined it as a measure of the power of a chemically bonded atom to attract electrons to itself. Several measures of this characteristic were suggested, the most satisfactory being that of Mulliken (9). He computed electronegativity (X_M) as the average value of the ionization potential and the electron affinity. Expressed in electron volts, this value was computed for atoms in their bonding valence states (10). The electronegativity scale of Pauling (8) is more familiar to chemists but does not distinguish among hybrid states of the same atom, although they certainly have different electronegativities (10).

Figure 4 shows the Mulliken electronegativity values on the δ , δ^ν matrix for some second-row atoms. Inspection reveals some obvious trends. For constant $\delta^\nu = 4$, $\delta = 2, 3$, and 4 describes sp , sp^2 , and sp^3 carbon atoms. Electronegativity diminishes in the same order. This electronic effect manifests itself in the decreasing acidity of the bonds:



As δ^ν decreases for constant δ , electronegativity decreases. The trends suggest that electronegativity is some function of $\delta^\nu - \delta$. For the nine second-row atom valence states in Fig. 4, the following relationship obtains:

$$X_M = 2.05(\delta^\nu - \delta) + 6.99 \quad (\text{Eq. 6})$$

$$r = 0.989 \quad s = 0.60 \quad n = 9 \quad F = 305$$

This correlation is excellent, with the standard deviation being less than the estimate of Hinze and Jaffe (10) due to uncertainties in the electron affinities. The intercept is close to the electronegativity of hydrogen (7.17 eV), which would have a $\delta^\nu - \delta$ value of zero.

The interpretation of Eq. 6 is clear. It states that the electronegativity of an atom in its valence state is closely related to the number of electrons in pi and lone-pair orbitals. This is apparent from the relationships:

$$\delta^\nu = \sigma + p + n - h \quad (\text{Eq. 7a})$$

$$\delta = \sigma - h \quad (\text{Eq. 7b})$$

$$\delta^\nu - \delta = p + n \quad (\text{Eq. 7c})$$

where σ is the number of sigma bonding electrons other than to hydrogen. The observation that the electronegativity of a bound atom is due to the presence of pi and lone-pair orbitals leads to the conclusion that these electrons are inferior to sigma electrons on an atom in screening the core charge. This follows from the usual model that electronegativity manifests itself through sigma bonds; thus, electrons outside of these orbitals leave the core charge ($+Z^\nu$) less screened than when all valence electrons are sigma bonding (Csp^3).

The relationship can be made more general to include atoms of other rows by realizing that the nuclear attraction for valence electrons (hence, the electronegativity) must be related to their distance from the nucleus and the screening effect of intervening electrons. This distance effect may be approximated by the square of the principal quantum number, N (11). The general expression relating the Mulliken electronegativities to the δ values for 19 atom valence states of the first, second, and third rows, which are prominent in forming covalent bonds, takes the form of:

$$X_M = 7.99 \frac{\delta^\nu - \delta}{N^2} + 7.07 \quad (\text{Eq. 8})$$

$$r = 0.988 \quad s = 0.48 \quad n = 19 \quad F = 660$$

The δ^ν value in Eq. 8 for third-row atoms is defined as $\delta^\nu = Z^\nu - h$. The usual expression $\delta^\nu = (Z^\nu - h)/(Z - Z^\nu)$ is not used since the use of N^2 corrects for the distance of the valence electrons from the nucleus and is more useful in relating the electronegativity directly to the count of exterrjacent electrons, $\delta^\nu - \delta$. The two expressions are numerically very close.

The experimental and calculated results are shown in Table IV. The fifth column in Table IV, $(\delta^\nu - \delta)/N^2$, is the molecular connectivity valence state electronegativity. This value can be compared to the Pauling estimate of electronegativity by adding 2.5 to the value, although the Pauling value is for an atom irrespective of the valence state.

To conclude this section, it is stated that the number derived from $\delta_i^\nu - \delta_i$ for atom i describes a property of a valence state atom closely approximated by the electronegativity. Also, the number derived from $\delta_i^\nu + \delta_i$ for atom i describes a property of a valence state atom closely approximated by the van der Waals volume contribution to the molecule. Mathematically, $\delta_i^\nu - \delta_i$ and $\delta_i^\nu + \delta_i$ are orthogonal; that is, they are independent of each other. This orthogonality corresponds to the observation that the Bondi volume and the Mulliken electronegativities from Figs. 3 and 4 autocorrelate only to the extent of $r^2 = 0.21$. Orthogonal descriptors $\delta^\nu - \delta$ and $\delta^\nu + \delta$, simply derived from an inspection of atomic valence state structure and relating to electronegativity and van der Waals volume, should be of great practical use.

Bond Formation—The δ^ν and δ values for atoms in their valence states encode information about the electronegativity of atoms. The electronegativity of atoms in bonds influences several properties including ionicity, bond dipole, partial atomic charge, and bond strength, all resulting from electronegativity equalization (12). The calculation of molecular connectivity indexes as a sum of bond indexes thus incorporates some information about electronic structure, encoded in the δ values.

The δ^ν and δ values contain information about electron domains or orbitals associated with atoms. Thus, a volume characteristic is encoded in the δ values. As a result, molecular connectivity indexes contain volume information. With the structural and physical significance of δ^ν and δ revealed, it is possible to derive the bond index encoding information about structure and properties. A series of concise statements develops this index.

1. The δ^ν value is the number of nonhydrogenic valence electrons on an atom. The δ value is the number of δ^ν electrons participating in nonhydrogenic σ bonds from an atom.

2. The properties of electronegativity and volume encoded in δ_A^ν for an atom (A) are considered to be contributed equally from all valence electrons or their orbitals. Thus, one electron from atom A forming one σ bond will possess the fraction $1/\delta_A^\nu$ of the properties encoded in δ_A^ν .

3. Atom A brings to the bond $A-B$ the fraction $1/\delta_A^\nu$ of the properties encoded by δ_A^ν . Atom B brings to the same bond the fraction $1/\delta_B^\nu$ of the properties encoded by δ_B^ν .

Table IV—Delta Values of Valence State Atoms and Mulliken Electronegativities

Atom	Hybrid State	δ^v	δ	$\frac{\delta^v - \delta}{N^2}$	X_M^a, ev	$X \text{ calc.}^b$
H	s	1	1	0	7.17	7.05
B	trtrtr	3	3	0	6.34	7.05
C	didi $\pi\pi$	4	2	0.50	10.39	11.04
	trtrtr π	4	3	0.25	8.79	9.05
N	tetetete	4	4	0	7.98	7.05
	di ² di $\pi\pi$	5	1	1.00	15.68	15.04
	tr ² trtr π	5	2	0.75	12.87	13.04
O	te ² tetete	5	3	0.50	11.54	11.04
	tr ² tr ² tr π	6	1	1.25	17.07	17.04
	te ² te ² tete	6	2	1.00	15.25	15.04
Si	didi $\pi\pi$	4	2	0.22	9.06	8.82
	trtrtr π	4	3	0.11	7.90	7.93
	tetetete	4	4	0	7.30	7.05
	di ² di $\pi\pi$	5	1	0.44	11.25	10.60
P	tr ² trtr π	5	2	0.33	9.67	9.71
	tetetete	5	3	0.22	8.90	8.82
	tr ² tr ² tr π	6	1	0.55	10.88	11.49
S	te ² te ² tete	6	2	0.44	10.14	10.60
	te ² te ² te ² te	7	1	0.66	11.84	12.37

^a From Ref. 9. ^b From Eq. 8.

4. If the property under consideration is the electronegativity, then orbitals from *A* and *B* contributing the fractions $1/\delta_A^v$ and $1/\delta_B^v$, respectively, will become adjusted toward an equal intermediate electronegativity, which may be ascribed to the bond. This principle of electronegativity equalization was set forth by Sanderson (12, 13), who proposed that the geometric mean of the atom electronegativities leads to the best description of the bond electronegativity.

5. From Sanderson (12, 13), $[(1/\delta_A^v)(1/\delta_B^v)]^{1/2}$ is the algorithm appropriate for two bonded atoms. This algorithm is identical to the molecular connectivity description of a bond (1) as shown in Eq. 1.

6. The volume contributed by atom *A* to a bond *A-B* is the fraction $1/\delta_A^v$ of the total. The fraction $1/\delta_B^v$ is the volume contributed from atom *B*. The bond has a "volume" due to orbital overlap that can be approximated by $2[(1/\delta_A^v)(1/\delta_B^v)]^{1/2}$. With a constant of 2, $[(1/\delta_A^v)(1/\delta_B^v)]^{1/2}$ reflects the relative volume of the bond *A-B*. This is the same expression describing the electronegativity of the *A-B* bond from statement 5.

7. The orthogonal expressions $\delta_A^v - \delta_A$ and $\delta_A^v + \delta_A$ reflect the electronegativity and volume of atom *A*, respectively. What is the role of δ_A in the calculation of ${}^1\chi^v$? Each δ_A^v appears δ_A times in the summation of $[(1/\delta_A^v)(1/\delta_B^v)]^{1/2}$ to give ${}^1\chi^v$. Thus, ${}^1\chi^v$ can be expected to correlate with some properties related to molecular volume and electronic character.

Hydrogen Suppression—It has been stated that hydrogen atoms are "suppressed" in the count of valence electrons and bonding neighbors to an atom (1). Thus:

$$\delta^v = Z^v - h \quad (\text{Eq. 9a})$$

$$\delta = \sigma - h \quad (\text{Eq. 9b})$$

where Z^v is the number of valence electrons, h is the number of hydrogens, and σ is the number of bonding sigma electrons. The hydrogen influence in reality is subsumed into modified δ^v and δ values reflecting the number of hydrogen atoms on the atom in question. The δ^v value in combination with the δ value thus is a description of an atom possessing volume and electronegativity characteristics associated with the hydride

group, which is treated as a single atom. The δ^v and δ values accurately reflect the volume of the hydrides as seen in Table II. The δ^v and δ values for hydride groups also correctly order the group electronegativities of OH, NH₂, SH, and CH₃, based on various estimates of these values. Thus, from Eq. 8, the predicted group electronegativities are $X_M(\text{OH}) = 15.07$, $X_M(\text{NH}_2) = 11.07$, $X_M(\text{SH}) = 10.63$, and $X_M(\text{CH}_3) = 9.07$, all on the Mulliken scale.

The uniform exclusion of hydrogen makes the electronegativity and the volume of this atom a reference point. Note that in Eq. 8 relating to Mulliken electronegativity, the y intercept (when $\delta^v - \delta = 0$) is the electronegativity of hydrogen. Similarly in Eq. 4 relating to the Bondi fragment volume, the y intercept (when $\delta^v = \delta = 0$) for methane is the value of the sp^3 carbon fragment plus all four hydrogen atoms.

Index for a Molecule—The index for the molecule, ${}^1\chi^v$, is written as a summation of the bond terms:

$${}^1\chi^v = \sum(\delta_i^v \delta_j^v)^{-1/2} \quad (\text{Eq. 10})$$

This is the valence molecular connectivity index of first order since it is summed over all bonds that are single-path fragments of the molecule.

Valence molecular connectivity is built on the sigma electron or interjacent electron framework of the molecule. The presence of an atom, *i*, with nonsigma valence electrons, pi, or lone-pair electrons, which can be called exterjacent electrons, results in the fraction $1/\delta_i^v$ being smaller than $1/\delta_i$. The value of ${}^1\chi^v$ for the molecule containing atom *i* with exterjacent electrons reflects the total number of bonds and their adjacency, but it also reflects the diminished value, $1/\delta_i^v$, due to the presence of exterjacent electrons on atom *i*.

With the data in Table V, a series of comparisons reveals the encoded structural information in ${}^1\chi^v$ values. Molecule 1, butane, is the reference structure where all $\delta_i^v = \delta_i$; hence, ${}^1\chi^v = {}^1\chi = 1.914$. There are no exterjacent electrons so ${}^1\chi - {}^1\chi^v = 0$.

Molecules 2 and 3 in Table V are four-atom amines (excluding hydrogens), each with two exterjacent electrons on the nitrogen. Since the positions of the nitrogen atoms in the chains are different in each case, the fractional contributions of valence electrons forming nonhydrogenic

Table V—Information Encoded in ${}^1\chi^v$ and δ^v Values for Four-Atom Chains with δ Values of $A^1-B^2-C^2-D^1$

Number	Molecule	A	B	C	D	${}^1\chi^v$	Exterjacent Electrons	${}^1\chi - {}^1\chi^v$
1	CH ₃ CH ₂ CH ₂ CH ₃	1	2	2	1	1.914	0	0
2	CH ₃ CH ₂ CH ₂ NH ₂	1	2	2	3	1.615	2	0.299
3	CH ₃ CH ₂ NHCH ₃	1	2	4	1	1.561	2	0.353
4	CH ₃ CH ₂ CH ₂ OH	1	2	2	5	1.523	4	0.391
5	CH ₃ CH ₂ -O-CH ₃	1	2	6	1	1.404	4	0.510
6	CH ₃ CH ₂ CH=CH ₂	1	2	3	2	1.524	2	0.390
7	CH ₃ CH=CHCH ₃	1	3	3	1	1.488	2	0.426
8	CH ₃ CH ₂ CH=O	1	2	3	6	1.351	6	0.563
9	CH ₃ CH ₂ C≡CH	1	2	4	3	1.349	4	0.565
10	CH ₃ C≡CCH ₃	1	4	4	1	1.250	4	0.664
11	HO-CH ₂ CH ₂ -OH	5	2	2	5	1.132	8	0.782
12	CH ₃ CH=N-OH	1	3	5	5	1.036	8	0.878
13	O=CH-CH=O	6	3	3	6	0.805	12	1.109

Table VI—Mixotropic Series Solvents Ranked by Polarity and ${}^1\chi^v$

Solvent	${}^1\chi^v$	Dielectric Constant
Hexane	2.914	1.89
Carbon tetrachloride	2.390	2.24
<i>p</i> -Dioxane	2.155	2.21
Benzene	2.000	2.30
Ethyl ether	1.992	4.34
Ethyl acetate	1.904	6.02
Methyl ethyl ketone	1.765	18.50
Propanol	1.523	20.10
Acetone	1.204	20.70
Ethanol	1.023	24.3
Methanol	0.447	32.80
Water	0	78.0

bonds are different. The nitrogen in Molecule 2 contributes one-third $[(1/\delta_N) = 1/3]$ of its nonhydrogenic valence electrons to form one sigma bond to carbon. The nitrogen in Molecule 3, contributes one-fourth $[(1/\delta_N) = 1/4]$ of its nonhydrogenic valence electrons to each of two sigma bonds with adjacent carbons. The indexes for bonds *B-C* and *C-D* thus are different for these cases. As a consequence, the ${}^1\chi^v$ indexes are different for Molecules 2 and 3, and each, in turn, is different and less than the value for Molecule 1.

Viewed in terms of valence electron structure information, it is obvious that in Molecule 2 the nitrogen bonds only once to a carbon and, therefore, participates only once as an amine group in an electronegativity equalization event (12). In Molecule 3, the nitrogen is located midchain and, therefore, participates in two bonds to carbons, with two opportunities to participate in electronegativity equalization. Clearly, electronically governed properties differ and are potentially correlated with the ${}^1\chi^v$ values. The formation of two sigma bonds by nitrogen in Molecule 3 leads to two electronegativity equalization events, resulting in a greater buildup of charge on the nitrogen in Molecule 3. Thus, Molecule 3 is more basic than Molecule 2. Within closely related amines, the ${}^1\chi - {}^1\chi^v$ difference appears to parallel this effect. It follows that ${}^1\chi - {}^1\chi^v$ for trimethylamine is 0.390, suggesting that it is more basic than Molecule 3. This is true in the gas phase.

The two amines, Molecules 2 and 3, have the same number and kind of atoms. They differ from the reference structure, butane, by having two exterrajacent electrons on one atom. The calculated ${}^1\chi^v$ value reflects both the presence of the exterrajacent electrons and the adjacency relationship of that atom (nitrogen) to other atoms in the molecule. The diminished value of ${}^1\chi^v$ for Molecule 3, relative to butane and to its positional isomeric amine, parallels structural differences that can be interpreted in electronic terms such as relative charge and basicity.

As a companion to the electronic interpretation of the information encoded by the different ${}^1\chi^v$ values, the volumes are expected to be different. Thus, properties governed by volume are expected to, and have been shown to, parallel the calculated ${}^1\chi^v$ values for a series of amines (1).

In Molecules 4 and 5 (Table V), oxygen is the atom contributing exterrajacent electrons to the molecules. The same pattern of results is found as for the amines. The terminal oxygen in Molecule 4 with six valence electrons has a bonded hydrogen; hence, $\delta^v = Z^v - h = 6 - 1 = 5$. Thus,

Table VII—Relative Benzene-Likeness of Hydrocarbons and Heterocycles

Molecule	${}^1\chi^v$	${}^1\chi^v/\text{Bonds}$	B^a
Benzene	2.000	0.333	1.00
Naphthalene	3.405	0.310	0.93
Phenanthrene	4.815	0.301	0.90
Anthracene	4.808	0.300	0.90
Pyrrole	1.577	0.315	0.95
Pyridine	1.850	0.308	0.93
Furan	1.471	0.294	0.88
Pyrazine	1.716	0.286	0.86
Imidazole	1.427	0.285	0.86
Pyrimidine	1.699	0.283	0.85
Isoxazole	1.343	0.269	0.81
Oxazole	1.321	0.264	0.79
Sydnone	1.338	0.223	0.67

^a Benzene-likeness from $[{}^1\chi^v/\text{bonds (for molecule)}]/[{}^1\chi^v/\text{bonds (for benzene)}]$.

$(1/\delta^v) = 1/5$ describes the sigma electron fraction forming a hydroxyl oxygen bond to a carbon. Only one sigma bond involves the hydroxyl oxygen group. In Molecule 5, the δ^v for oxygen is 6; hence, $(1/\delta^v) = 1/6$ describes each of two sigma electrons bonding to two adjacent carbons. Electronegativity equalization and atom volume overlapping are possible through two bonds in Molecule 5 but only one in Molecule 4. The greater ${}^1\chi - {}^1\chi^v$ calculated for Molecule 5 reflects the consequences of this greater degree of adjacency. Note also that the ${}^1\chi^v$ for Molecule 4 is less than that of Molecule 2 (and 5 is less than that of 3), reflecting more exterrajacent electrons on the oxygen-containing molecules. A physical manifestation mirroring this result is the lower ionization potentials of ethers relative to alcohols of the same size. A lone-pair electron is the one lost in this process.

Molecules 6 and 7 in Table V describe a different situation. In these isomeric alkenes, there are two exterrajacent electrons but they are not paired on the same atom, as with the amines; they are located on adjacent atoms. Thus, the δ^v values for Molecule 6 are different from the reference butane structure in the case of two atoms. This situation is typical of an unsaturated bond. As a consequence, of the three bond indexes calculated for each alkene molecule, two differ from the reference molecule for Molecule 6 and all three differ for Molecule 7. Thus, the ${}^1\chi^v$ for Molecule 7 is lower than that for Molecule 6. The lower ${}^1\chi^v$ value for Molecule 7 relative to Molecule 6 presages a greater availability of the pi electrons. Thus, Molecule 7 reacts faster with electrophiles and has a lower heat of hydrogenation and a lower ionization potential.

By extending this list of four-atom molecules in Table V, it can be seen that with an increasing number of exterrajacent electrons, ${}^1\chi^v$ decreases. The series might be called an increasing "exterrajacency series." There is a parallel with physical experience relating to stability, polarity as a solvent, and chemical reactivity. Table VI ranks common solvents by polarity, expressed as the dielectric constant correlating with ${}^1\chi^v$. The relative exterrajacency, encoded in the ${}^1\chi^v$ value, manifests itself as a structural description in another way. For benzene, for example, the simple δ values for the ring atoms are uniformly 2. The ${}^1\chi$ value is 3.000. In contrast, the δ^v values are uniformly 3, leading to a ${}^1\chi^v$ value of 2.000. The specific exterrajacency (per bond) $({}^1\chi - {}^1\chi^v)/6$ is 0.333. It is now possible to rank cyclic heterocycles relative to benzene to compare their likeness to it (Table VII).

The molecular volume information encoded in ${}^1\chi^v$ is easy to envision. A glance down the molecules in Table V reveals that ${}^1\chi^v$ decreases with decreasing bond lengths. Thus, the ranking closely parallels the molecular volume.

SUMMARY

The valence molecular connectivity index (${}^1\chi^v$) is derived from a cardinal number (δ^v) enumerating the valence electrons on an atom not bonding to hydrogen. The δ^v number is used δ times in calculating ${}^1\chi^v$, where δ is the number of sigma electrons on an atom not bonding to hydrogen. These two cardinal numbers specify the hybrid states or valence states of any atom within a quantum level.

In this study, it was shown that the atomic fragment volume approximations used in assessing molecular volume are highly correlated with the value of $\delta^v + \delta$ for an atom in its valence state. It also was shown that the valence state electronegativity estimates of Mulliken correlate well with the value of $\delta^v - \delta$ for an atom in its valence state. These two relationships are general for atoms in at least the first three quantum levels.

The δ^v value, when formulated into the ${}^1\chi^v$ value for a molecule, carries with it information relating to both volume and electronic character. A simple analogy is the observation that the number 12 carries information about the space occupied by a dozen freshly laid eggs and the number of embryonic heartbeats. Thus, an sp^3 nitrogen atom has a $\delta^v = 5$. The cardinal number, 5, counts orbitals contributing to a spatial property, volume. The number 5 concurrently counts the number of valence electrons, which govern electronic characteristics. This duality in the information encoded in the cardinal number (δ^v) leads to a duality of structural information encoded in the derived ${}^1\chi^v$ value.

The volume information encoded in ${}^1\chi^v$ values has been demonstrated on numerous occasions in the form of close correlations with volume-related physical properties (39). Other studies showed that properties largely governed by electronic composition also correlate with ${}^1\chi^v$ indexes. Inspection of the many studies using valence connectivity indexes (Table I) show both aspects of structure relating to these indexes.

In this study, for the first time, a derivation of ${}^1\chi^v$ from δ^v has been shown. It is based on the assignment of fractional parts of properties ascribed to the valence electrons of a bonding atom. The fractions $1/\delta^v$

for bonding atoms contribute to the bond so that the geometric mean of the fraction products characterizes the new physical environment. Thus, a bond index is born. The bond indexes $[(1/\delta_A)(1/\delta_B)]^{1/2}$ are summed to form the molecular connectivity index, ${}^1\chi^v$.

REFERENCES

- (1) L. B. Kier and L. H. Hall, "Molecular Connectivity in Chemistry and Drug Research," Academic, New York, N.Y., 1976.
- (2) M. Randic, *J. Am. Chem. Soc.*, **97**, 6609 (1975).
- (3) W. J. Murray, L. H. Hall, and L. B. Kier, *J. Pharm. Sci.*, **64**, 1978 (1975).
- (4) L. B. Kier, W. J. Murray, M. Randic, and L. H. Hall, *ibid.*, **65**, 1226 (1976).
- (5) L. B. Kier and L. H. Hall, *ibid.*, **65**, 1806 (1976).
- (6) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).
- (7) C. C. Bigelow, *J. Theoret. Biol.*, **16**, 187 (1967).
- (8) L. Pauling, "The Nature of the Chemical Bond," 3rd ed., Cornell University Press, Ithaca, N.Y., 1960.
- (9) R. S. Mulliken, *J. Chem. Phys.*, **2**, 782 (1934).
- (10) J. Hinze and H. H. Jaffe, *J. Am. Chem. Soc.*, **84**, 540 (1962).
- (11) C. E. Sun, *J. Chin. Chem. Soc.*, **10**, 77 (1943).
- (12) R. T. Sanderson, *J. Chem. Educ.*, **31**, 2 (1954).
- (13) R. T. Sanderson, "Chemical Bonds and Bond Energy," Academic, New York, N.Y., 1976.
- (14) T. DiPaolo, *J. Pharm. Sci.*, **67**, 564 (1978).
- (15) *Ibid.*, **67**, 566 (1978).
- (16) M. Bonjean and C. L. Duk, *Eur. J. Med. Chem.*, **13**, 73 (1978).
- (17) G. R. Parker, *J. Pharm. Sci.*, **67**, 513 (1978).
- (18) B. K. Evans, K. C. James, and D. K. Luscombe, *ibid.*, **68**, 370 (1979).
- (19) J. S. Millership and A. D. Woolfson, *J. Pharm. Pharmacol.*, **30**, 483 (1978).
- (20) S. P. Gupta and P. Singh, *Bull. Chem. Soc. Jpn.*, **52**, 2745 (1979).
- (21) R. J. Gardner, *Tech. Q. Master Brew. Assoc. Am.*, **16**, 204 (1979).
- (22) R. J. Gardner, *J. Sci. Food Agric.*, **31**, 23 (1980).
- (23) L. B. Kier and L. H. Hall, *J. Pharm. Sci.*, **65**, 1806 (1976).
- (24) T. DiPaolo, L. B. Kier, and L. H. Hall, *Mol. Pharmacol.*, **13**, 31 (1977).
- (25) L. H. Hall and L. B. Kier, *J. Pharm. Sci.*, **66**, 642 (1977).
- (26) L. B. Kier, T. DiPaolo, and L. H. Hall, *J. Theoret. Biol.*, **67**, 585 (1977).
- (27) L. B. Kier and L. H. Hall, *J. Med. Chem.*, **20**, 1631 (1977).
- (28) L. B. Kier, R. J. Simons, and L. H. Hall, *J. Pharm. Sci.*, **67**, 725 (1978).
- (29) L. B. Kier and R. A. Glennon, *Life Sci.*, **22**, 1589 (1978).
- (30) R. A. Glennon and L. B. Kier, *Eur. J. Med. Chem.*, **13**, 219 (1978).
- (31) L. B. Kier and L. H. Hall, *J. Pharm. Sci.*, **67**, 1408 (1978).
- (32) L. H. Hall and L. B. Kier, *ibid.*, **67**, 1743 (1978).
- (33) L. B. Kier and L. H. Hall, *ibid.*, **68**, 120 (1979).
- (34) T. DiPaolo, L. B. Kier, and L. H. Hall, *ibid.*, **68**, 39 (1979).
- (35) R. A. Glennon, L. B. Kier, and A. T. Shulgin, *ibid.*, **68**, 906 (1979).
- (36) A. J. Richard and L. B. Kier, *ibid.*, **69**, 124 (1980).
- (37) D. R. Henry and J. H. Block, *Eur. J. Med. Chem.*, **15**, 133 (1980).
- (38) B. Vant Riet, L. B. Kier, and H. L. Elford, *J. Pharm. Sci.*, **69**, 856 (1980).
- (39) L. H. Hall and L. B. Kier, *Eur. J. Med. Chem.*, in press.

Analysis of Drug Contamination from Parabens in Theophylline Olamine

E. C. JUENGE^{*}, D. F. GURKA, and M. A. KREIENBAUM

Received July 1, 1980, from the National Center for Drug Analysis, Food and Drug Administration, 1114 Market Street, St. Louis, MO 63101. Accepted for publication September 29, 1980.

Abstract □ Contaminants in a commercial enema sample of theophylline olamine were found to be derived from parabens and ethanolamine. These contaminants, whose presence was characterized by the loss of preservatives and solubilizing agent, were isolated directly from the drug sample and identified. A high-performance liquid chromatographic (HPLC) system was developed to separate completely and to measure quantitatively theophylline, the two impurities, 4-hydroxybenzoic acid and *N*-(2-hydroxyethyl)-4-hydroxybenzamide, and the remaining parabens. The material balance obtained from the results of quantitative HPLC indicated the formation of these impurities at the expense of parabens. TLC, IR, and UV spectrophotometry and NMR and mass spectrometric analyses were used for identification or for comparisons of the new compound and the known *N*-(2-hydroxyethyl)benzamide.

Keyphrases □ Theophylline olamine—contaminants, identification and quantitation, high-performance liquid chromatographic analysis □ Contaminants—identification and quantitation, theophylline olamine enema solutions, high-performance liquid chromatographic analysis □ Paraben preservatives—contamination in theophylline olamine enema solutions, high-performance liquid chromatographic analysis, identification and quantitation □ High-performance liquid chromatography—contamination from paraben preservatives in theophylline olamine enema solutions

During a survey of xanthine derivative drugs conducted by the National Center for Drug Analysis, assay of some enema solutions containing theophylline olamine was attempted with a high-performance liquid chromatographic

(HPLC) method that had been used successfully for theophylline elixirs. The analysis showed that, in addition to theophylline and parabens, the sample contained two impurities. This paper discusses the identification and quantitation of these two substances.

BACKGROUND

The approach followed in this laboratory study was dictated primarily by the small amounts of sample initially available. It involved, first, the original detection of the impurities 4-hydroxybenzoic acid (I) and *N*-(2-hydroxyethyl)-4-hydroxybenzamide (II) by HPLC and TLC, with characterization of their aromatic nature by UV spectrophotometry; and, second, identification by isolation of the impurities from the drug sample and direct comparison with standards by mixed melting-point determinations and IR and NMR analyses. Identities were confirmed by HPLC and TLC comparisons of the original drug solution with commercial and synthesized standards and by a wide collection of spectral data. Finally, the theophylline peak was displaced from the narrow region between the impurity peaks by adjusting the mobile phase to enable quantitative HPLC analysis.

The question of isomerism in the conventional syntheses used to prepare the amides arose because two materials whose elemental analyses were consistent for *N*-(2-hydroxyethyl)benzamide (III) showed different properties. The problem was resolved by purification that showed the materials to be identical; moreover, the mass fragmentations were consistent with the structures given for II and its structural analog, III.

This work demonstrates that parabens can interact chemically with amines used as complexing agents in the compounding of theophylline