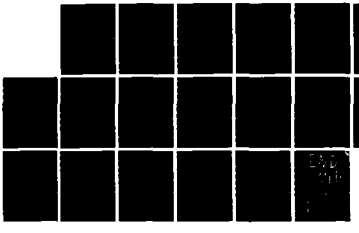
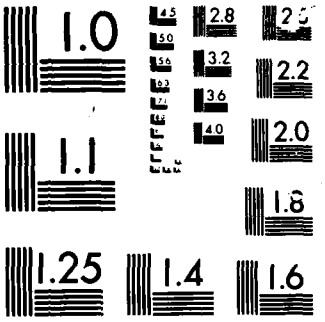
AD-A166 986 THE PREDICTION OF BIOLOGICAL ACTIVITY USING MOLECULAR CONNECTIVITY INDICES(U) GEORGIA UNIV ATHENS DEPT OF CHEMISTRY D H ROUVRAY 23 APR 86 TR-21 N00014-84-K-0365 1/1

UNCLASSIFIED F/G 6/15 NL





MICROCOPY CHART

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(12)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 21	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE PREDICTION OF BIOLOGICAL ACTIVITY USING MOLECULAR CONNECTIVITY INDICES		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) D.H. Rouvray		8. CONTRACT OR GRANT NUMBER(s) N00014-84-K-0365
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Georgia Department of Chemistry Athens, GA 30602		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 051-861
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlginton, VA 22217		12. REPORT DATE 4/23/86 13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/ DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE APR 30 1986 S D B

18. SUPPLEMENTARY NOTES

To be published in Acta Pharmaceutica Jugoslavica

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Biological Activity
Molecular Connectivity Indices
Precition of Bioactivity

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The set of topological indices widely referred to molecular connectivity indices has been used extensively for the prediction of biological activity in many different classes of molecules. Here we review the major steps taken since the first molecular connectivity index was introduced by Randić some twelve years ago. Developments to the end of 1985 are covered.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

OFFICE OF NAVAL RESEARCH

Contract N00014-84-K-0365

TECHNICAL REPORT NO. 21

The Prediction of Biological Activity Using Molecular Connectivity Indices

by

Dennis H. Rouvray

Prepared for publication in

Acta Pharmaceutica Jugoslavica

University of Georgia
Department of Chemistry
Athens, Georgia 30602

April 23, 1986

The Prediction of Biological Activity Using Molecular Connectivity Indices

D.H. Rouvray, Department of Chemistry, University of Georgia

Athens, Georgia 30602, United States of America

Abstract

The set of topological indices widely referred to molecular connectivity indices has been used extensively for the prediction of biological activity in many different classes of molecules. Here we review the major steps taken since the first molecular connectivity index was introduced by Randić some twelve years ago. Developments to the end of 1985 are covered.

The Prediction of Biological Activity Using Molecular Connectivity Indices

D.H. Rouvray, Department of Chemistry, University of Georgia
Athens, Georgia 30602, United States of America

Abstract

The set of topological indices widely referred to molecular connectivity indices has been used extensively for the prediction of biological activity in many different classes of molecules. Here we review the major steps taken since the first molecular connectivity index was introduced by Randić some twelve years ago. Developments to the end of 1985 are covered.

Dedicated to Professor Milan Randić

General Introduction

A highly influential paper by Randić entitled 'On characterization of molecular branching' appeared in November, 1975 in the Journal of the American Chemical Society.[1] In this paper a new topological index[2-4] was introduced which has since become known as the molecular connectivity index. Topological indices (TIs) are graph-theoretical invariants which are employed extensively as mathematical descriptors for a wide range of molecular species.[5-7] They have been used mainly for the purpose of correlating the properties of molecular species with their topological structure. The molecular connectivity index was originally put forward with the aim of characterizing the branching in alkane species, but more recently has been shown to have numerous applications in both the physical

and biological sciences. To date, it is the only TI to have had a whole book devoted to it.[8] In this review we shall focus on the use of the index in the more biologically oriented sciences, such as pharmacology and toxicology. In particular, we shall discuss the prediction of biological activity in chemical species by means of this index. Little mention will be made here of other applications of this index; these are reviewed comprehensively elsewhere.[8,9]

The molecular connectivity index as originally presented by Randić[1] was designed to parallel the ordering obtained for sets of isomeric alkane molecules based on their physicochemical properties. The index, which is nowadays universally represented by the symbol χ, was defined in the following way:

$$\chi = \sum_{\text{edges}} (v_i \, v_j)^{-\frac{1}{2}} , \qquad (1)$$

where v_i and v_j are the degrees of a pair of neighboring vertices joined by the edge $\{i, j\}$ in the chemical graph of the molecule under consideration. The index was soon generalized by Randić and coworkers[10] to include a variety of subgraphs of G other than edges; these subgraphs are referred to as paths, clusters, path/clusters, and chains and are illustrated in Figure 1. The generalized form of equation (1) may be written explicitly as:

$${}^{h}\chi_t = \sum_{\text{subgraphs}} (v_i \, v_j \ldots . \, v_{h+1})^{-\frac{1}{2}} \qquad (2)$$

or, more concisely, as:

$${}^{h}\chi_t = \sum_{k=1}^{n_h} \prod_{i=1}^{h+1} (v_i)_k^{-\frac{1}{2}} , \qquad (3)$$

where h is the number of edges in the subgraph of G used, t is the type of subgraph

considered (see Figure 1), $\underline{n}_{\underline{h}}$ is the number of subgraphs of type t having $\underline{h}$ edges, and the index k extends over all the $\underline{n}_{\underline{h}}$ subgraphs.

Since it is of critical importance in making correlations involving biological parameters to be able to deal effectively with so-called heteroatoms, i.e. atoms other than carbon, in molecular graphs, a further generalization of the index was made by Kier and Hall.[11] This generalization provides a means of taking into account situations in which molecular species contain heteroatoms; such species can be conveniently modelled using rooted graphs.[8] The differing valencies which heteroatoms will have are allowed for by assigning values according to the formula:

$$\delta = Z - N_H , \qquad (4)$$

where δ is the difference between the number of valence electrons in the heteroatom and the number of hydrogen atoms to which it attached, Z is the number of valence electrons, and N_H the number of attached hydrogen atoms. Some values of for various heteroatoms are presented in Table 1. Whenever valence corrected atoms are used in the calculation of the molecular connectivity index, the index is written with a superscript v as $^{v}\chi$ to indicate this fact.

In the succeeding sections we document in some detail the manifold correlations with biological activity which have been obtained up to the end of 1985 using several of the various molecular connectivity indices. The correlations span a very broad range of biological activities, and, for convenience, will be discussed under the four headings (i) conscious state modification; (ii) olfactory and gustatory stimulation; (iii) inhibitory and depressant activity; and (iv) mutagenic and toxic behavior. It is of interest to note that molecular connectivity indices first began to be employed for correlations involving biological parameters after it was

Chemical Group	Value of δ
-OH	5
-O-	6
=O	6
$-NH_2$	3
-NH-	4
=NH	4
-N-	5
≡N	5
-Cl	0.690
-Br	0.254
-I	0.085

Table 1. Values for δ for several different chemical groups.

discovered that the indices correlate well with certain of the physicochemical properties of organic compounds, specifically the octanol/water partition coefficient[12] and the molecular polarizability[13] (see Figures 2 and 3). These two parameters were already well known to correlate with biological responses.[14,15] As observed by Randić, connectivity indices encode in a simple manner information on the bond environments of the various atoms comprising the molecule. Moreover, since the indices relate to properties which are dependent upon both the molecular volume and surface area, they should effectively reflect both of these geometrical features of a molecule.[17] This would imply a relationship to the atom-bond (polarization) and bond-bond (dispersion) behavior of molecules. It is reasonable therefore to expect that connectivity indices are able to model biological responses which are determined by such forces.[17] Interactions based only on forces of this type have been described by Sabljić and Trinajstić as nonspecific interactions.[6]

Conscious State Modification

In this section we discuss the correlations obtained for anesthetic inhalation gases, local anesthetics, narcotic and hallucinogenic agents. Some of this material has already been covered by previous reviewers,[5-8] but is included here for the sake of completeness. Throughout this paper, correlation statistics will be given in terms of three parameters: r, the correlation coefficient; s, the standard deviation; and n, the number of data points. For a set of 28 inhalation anesthetics, including several halocarbon species, a correlation of the type:

$$\log \frac{1}{p} = 0.571\ {}^{0}\chi^{v} - 0.638$$

($r = 0.881$, $s = 0.496$, $n = 28$) was obtained[18] using the single variable $^{o}\chi^{v}$, where p is the effective anesthetic pressure (ATA) in atmospheres. By inclusion of a variable Q_H representing the charge on the polar hydrogen atoms, a signficantly improved result ($r = 0.966$, $s = 0.278$, $h = 28$) was found. A later study[19] on 45 halogenated hydrocarbons confirmed these earlier results by again showing that both $^{o}\chi^{v}$ and Q_H were necessary to achieve a good correlation ($r = 0.975$, $s = 0.27$, $n = 45$). A similar study[20] on 28 anesthetic aliphatic ethers, however, revealed that $^{1}\chi$ was the relevant variable for correlation with the AD_{50} molar concentration. Linear correlation with $^{1}\chi$ produced statistics ($r = 0.979$, $s = 0.076$, $n = 28$) which were improved when parabolic correlation, i.e. the addition of a $(^{1}\chi)^{2}$ term, was attempted ($r = 0.986$, $s = 0.063$, $n = 28$)

Barbiturates and their derivatives were the subject of investigations undertaken by Murray et al.,[21] Bonjean and Luu Duc,[22] and Basak et al.[23] The correlations were made between the logarithm of the partition coefficient and $^{1}\chi^{v}$; the correlation coefficients ranged from 0.847 to 0.995. Use was made of both linear and parabolic correlations, and it is interesting to note that $^{1}\chi^{v}$ was superior to log p in the correlation of the <u>in vivo</u> activity of barbiturates.[23] Several studies on hallucinogenic agents have demonstrated the great predictive potential of the indices in this domain. To date, studies have been carried out on the phemethylamines (mescaline analogs),[24] the phenylisopropylamines (amphetamine analogs),[25,26] and the indoleamines (lysergamide analogs).[27] The correlations of hallucinogenic activity were made in mescaline units and based on a two- or three-parameter approach. For the amphetamine analogs, the equation assumed the form:

$$\log \mu = 45.16 \left(\frac{1}{^{3}\chi_p}\right) + 1.288\ {}^{6}\chi_p - 4.298 \left(\frac{1}{^{4}\chi^{v}_{pc}}\right) - 5.592$$

($r = 0.920$, $s = 0.251$, $n = 23$), where μ is the molar activity. This equation was found to be so reliable that it was subsequently used to predict the activities of various other amphetamines as well as those of mescaline and tryptamine species.[25]

A number of important correlations have been established between either $^1\chi$ or $^1\chi^v$ and the local anesthetic behavior or the narcotic activites of various organic molecules toward living organisms. A significant correlation, illustrated in Figure 4, was found by Kier et al.[13] between the minimum blocking concentration, MBC, of 36 local anesthetics using cells under standard conditions and $^1\chi$ as follows:

$$\log MBC = 3.55 - 0762\ ^1\chi$$

($r = 0.983$, $s = 0.39$, $n = 36$). Similar relationships were recorded[13] with the biologically relevant properties of solvent cavity surface area and molecular polarizability of these 36 molecules. Whole animal narcosis studies have related the narcotic effects of molecules on larvae or frog tadpoles and are summarized in Kier and Hall's book.[8] These studies have included correlation of the narcotic concentration of 20 organic compounds for Arenicola larvae,[8] the narcotic concentration of 15 alcohols with barnacle larvae,[12] the effective concentration of 52 organic compounds for tadpole narcosis,[8] and a repeat of the latter correlation using only 36 organic compounds.[17] In these four latter linear correlations the correlation coefficients were respectively 0.979, 0.987, 0.886 and 0.956. The differing coefficents for tadpole narcosis reflect the fact that two different mechanisms operate.[8] Similar results were reported by Murray et al.[12] who investigated the I_{100} movement factor for both 2.5-day-old and 12-day-old tadpoles at 18°C.

Olfactory and Gustatory Stimulation

A satisfactory basis for the classification of various odorant molecues in terms of their shape and size has been advanced by Amoore.[28,29] The molecules in the various categories proposed by Amoore[29] have been ranked according to their degree of odor similarity, OS. Kier et al.[30] investigated the relationship of OS to molecular connectivity indices for four categories of odorants. The results obtained were as follows:

For 15 etheral odorants:

$$OS = 7.47 - 1.84\,{}^{2}\chi + 1.34\,{}^{3}\chi_{c}$$

($r = 0.921$, $s = 0.395$, $n = 15$).

For 16 floral odorants:

$$OS = 3.12e^{-1.66({}^{0}\chi - 9.51)^{2}} + 3.43$$

($r = 0.959$, $s = 0.225$, $n = 16$).

For 14 odiferous carboxylic acids:

$$\log \Delta c_t = 3.534\,{}^{1}\chi - 0.387({}^{1}\chi)^{2} - 3.017\,{}^{5}\chi_{p} - 3.184$$

($r = 0.949$, $s = 0.343$, $n = 14$), where Δc_t is the threshold concentration of the

acid needed to overcome anosima.

For 24 substituted benzaldehydes and nitrobenzenes:

$$OS = 15.02 - 2.122\ {}^{o}\chi^{v} + 2.424\ {}^{3}\chi^{v}_{c}$$

($r = 0.937$, $s = 0.60$, $n = 24$).

The meaningful correlations obtained between quantitative odor parameters and various connectivity indices indicate that the latter reliably model structure-activity relationships for several different categories of odorants.

The relationships between sweet and bitter tastes and molecular structures have also been explored using connectivity indices. Hall and Kier[31] studied a set of substituted nitroanilines which possess a sweet taste potency of up to 4000 times that of sucrose. By expressing sweetness as the relative molar sweetness with respect to sucrose, RS, a relationship of the following form was obtained:

$$\log RS = 0.350\ {}^{1}\chi + 0.694\ {}^{1}\chi^{v} - 3.856$$

($r = 0.953$, $s = 0.222$, $n = 9$).

Another study[3] on a set of cyclohexyaldoxines attempted to assign molecules to sweet or bitter classes. The assignment was made on the basis of the value assumed by a linear discriminant function made up of ${}^{1}\chi$ and ${}^{4}\chi_{p}$:

$$y = 1.21\ {}^{1}\chi - 3.88\ {}^{4}\chi_{p}.$$

When $y > -3.27$ assignment was made to the sweet class; when $y < -3.27$

assignment was made to the bitter class. Using this criterion it was possible to assign correctly 9 out of 10 molecules to the sweet class and 8 out of 10 to the bitter class. The statistical significance, stated as the T^2 value, was 5.9. Taste thresholds were also investigated by Gardner[33,34] for various homologous series of compounds, amino acids and peptides of importance in the production of beer. Linear correlations with $^1\chi^v$ were obtained, with correlation coefficients lying between 0.701 and 0.943.

Inhibitory and Depressant Activity

A number of studies have correlated the inhibition of enzymes with different connectivity indices. Such studies have included the inhibition of succinate[17] and hydrazide monoamine[35] oxidase, thymidine phosphorilase,[17] adenosine deaminase,[17] butyrylcholesterinase,[17] and ribonucleotide reductase.[36] For succinate oxidase a model using only $^1\chi$ was presented:[17]

$$pC = 0.916\ ^1\chi - 1.582$$

($r = 0.966$, $s = 0.169$, $n = 13$)

whereas for hydrazide monoamine oxide a three-parameter model was required:

$$pI_{50} = -5.2 - 29E - 0.82\ ^2\chi + 1.8\ ^3\chi^v_p$$

($r = 0.941$, $s = 0.201$, $n = 24$),

where pC is the concentration of $I_{15\text{-}20}$ *succinate oxidase, and E is the half-wave*

potential. The correlation obtained for ribonucleotide reductase[36] was also a three-parameter equation:

$$pC = 2.36\ {}^{3}\chi_p - 3.98\ {}^{o}\chi^{v} + 0.97\ ({}^{1}\chi^{v})^2 + 9.20$$

($r = 0.943$, $s = 0.21$, $n = 28$). The study suggested that the benzohydroxamic acid inhibitors used display inhibitory potency in two ways associated with two different parts of the molecule.

A congeneric series of quarternary ammonium salts was employed[21] for correlations of minimum inhibitory concentrations, minimum killing concentrations and minimum concentrations for 50% hemolysis against various microorganisms, including Staphylococcus typhosa and Staphylococcus aureus. The correlations all assumed a parabolic form using ${}^{1}\chi$ and $({}^{1}\chi)^2$; correlation coefficients ranged from 0.872 to 0.982. It was conjectured that the correlations were due to the fact that ${}^{1}\chi$ mirrors well the size and shape of the molecules involved. Similar results were reported[21] for a set of aromatic and heterocyclic amines using the minimum inhibitory concentration on Mycobacterium tuberculosis. The inhibition of the microsomal p-hydroxylation of aniline by alkanols was investigated by Sabljić and Protić-Sabljić.[37] They found the best two-parameter equation to be the following:

$$pIC_{50} = -6.88\ \left(\frac{1}{{}^{o}\chi^{v}}\right) - 1.14\ {}^{4}\chi_{pc} + 1.85$$

($r = 0.983$, $s = 0.156$, $n = 20$), and concluded that the alcohols bind at the enzyme-active site, with the inhibiting effect being the result of a fine balance between the size and degree of branching in the alkyl chain.

The correlating equation between the muscarinic receptor affinity of 104

optically inactive acetylcholine antagonists and three different connectivity indices was found by Kier and Hall[38] to be:

$$\log K = 0.749\ {}^{4}\chi_{pc} + 0.258\ {}^{1}\chi^{v} + 1.340\ {}^{3}\chi_{c}^{v} + 0.827$$

($r = 0.962$, $s = 0.331$, $n = 104$), where log K expresses the receptor affinity. A check on the reliability of the correlation was performed by substituting random numbers for all of the 22 different indices considered and repeating the procedure 100 times. The highest correlation coefficient obtained in this way was only 0.45; the relation thus reflects the systematic variation of log K with molecular structure. Moreover, the analysis revealed that the separate ends of the molecule are of greatest importance in determining the observed activity, namely the onium group and the side chain. The contribution to log K from the increasing number of atoms in the molecules (reflected in ${}^{1}\chi^{v}$) was offset by a decline in ${}^{3}\chi_{c}^{v}$ together with a levelling of the ${}^{4}\chi_{pc}$ values in a set of eight test onium compounds. When the relationship was used to predict both agonist and antagonist affinities, very satisfactory results were obtained.[38]

The behavior of isatin derivatives in regard to their mitodepressant activity and binding to human serum albumin has been investigated by Sablić et al.[39] using both ${}^{1}\chi^{v}$ and the Hückel energy of the highest occupied molecular orbital, E_{HOMO}. The best multiple correlation obtained for protein binding of 16 isatin derivatives was:

$$\text{P.B.} = 1.672\ ({}^{1}\chi^{v})^{2} - 16.924\ {}^{1}\chi^{v} - 131.538\ (E_{HOMO})^{2} + 129.285\ E_{HOMO} + 42.155$$

with a correlation coefficient of 0.892. The equation obtained for mitodepressant

activity was the parabolic one with $^1\chi^v$:

$$\log \frac{1}{c} = -0.060\,(^1\chi^v)^2 + 0.884\,(^1\chi^v) - 1.710$$

($r = 0.820$, $n = 26$), though the best multiple correlation equation assumed the form:

$$\log \frac{1}{c} = -4.624\,(E_{HOMO})^2 + 4.151\,E_{HOMO} + 0.126\,^1\chi^v - -0.253$$

($r = 0.841$, $n = 26$). The comparatively low values of the correlation coefficients were ascribed to the low accuracy of the biological data used.

The antimicrobial activity of phenyl propyl ethers against skin fungi and the antiviral activity of benzimidazoles against the Lee strain of B influenza virus were studied using connectivity indices by Hall and Kier.[40] After showing that these two types of activity correlated well (correlation coefficients of 0.955 and 0.966 for a three-parameter and a two-parameter equation, respectively), they went on to partition the significant regression variables into subgraph terms reflecting various structural features of the molecules in question. This analysis led to several instructive conclusions about the nature of the principal interactions involved. For instance, it was found that in the phenyl propyl ethers the principal interactions focus on the para-position of the phenyl ring; and in the benzimidazoles that substitution of branched or cyclic alkyl groups on the 2-position (five-membered ring) is important for high activity. Such results are clearly superior to earlier studies based on electronic factors. Two other studies which set out to compare the molecular connectivity approach with other approaches also revealed the power of the former approach.[41,42] The activity

chosen in each case was the antimicrobial activity of an identical set of 50 chloro-, bromo- and alkylphenols against Staphylococcus aureus. Both studies established the substantial advantages inherent in the use of purely topological methods to carry out structure-property correlations, including the extreme simplicity of the method and the high reliability of the correlations obtained. A further confirmation of the value of molecular connectivity indices was presented by Samanta et al.[43] who investigated the anitfungal activity of 12 phenolic compounds and were able to show that such activity was modelled better by use of $^{2}\chi^{v}$ than by the logarithm of the octanol/water partition coefficient.

Mutagenic and Toxic Behavior

To date, very little has been attempted in correlating mutagenic behavior with molecular connectivity indices whereas toxicity has been fairly widely correlated. One reason for this may be the comparative paucity of reliable data on the mutagencity of molecules. A study carried out by Kier et al.[44] on 15 nitrosamines known to occur in cigarette smoke, nitrate-pickled meat, and smoked fish correlated their mutagenicity as revealed by the Ames test against a variety of connectivity indices. Two different equations were obtained with closely similar statistics, namely:

$$\log R = 2.398\,{}^{0}\chi - 4.095\,{}^{1}\chi^{v} - 5.590$$

($r = 0.964$, $s = 1.09$, $n = 15$),

and

$$\log R = 2.946\ {}^{2}\chi - 0.090\ {}^{4}\chi_{p}^{v} - 4.662$$

($r = 0.967$, $s = 1.05$, $n = 15$), where R is the number of revertants per nanomole. The equations were felt[44] to be good enough to serve as screening devices for untested nitrosamines. The closely related property of carcinogenicity was investigated by Jurs et al.[45] using a large, heterogeneous data set of organic compounds. The objective was to separate the set into carcinogens and noncarcinogens using a variety of descriptors, including several connectivity indices. Pattern recognition analysis demonstrated that no descriptors were completely effective here, the best achieving only a 90% correct separation. However, it was concluded[45] that connectivity indices are appropriate for the broad classification of compounds into carcinogens or noncarcinogens.

A correlation coefficient of 0.9972 was obtained[46] in a one-parameter correlation of the toxicity of polycyclic aromatic hydrocarbons in Daphnia Pulex with ${}^{o}\chi^{v}$:

$$-\log LC_{50} = 0.5346\ {}^{o}\chi^{v} - 7.004$$

($r = 0.9972$, $n = 5$), where LC_{50} is the lethal concentration at 96 hours. Although only five values were available for correlation, up to 26 values were employed for correlations of various χ^{v} terms with other hydrocarbon data. It was shown that chromatographic, partitioning, bioconcentration, and toxification processes are closely intercorrelated; correlation coefficients for correlations between the various data sets ranged from 0.8967 to 0.9989. Two studies, based on aliphatic hydrocarbons, ethers and ketones,[47] and substituted arylamines,[48] intercorrelated the anesthetic and toxic activities of these compounds. The equations obtained were two-parameter equations based on ${}^{o}\chi^{v}$, ${}^{3}\chi_{p}^{v}$ or ${}^{6}\chi_{p}^{v}$, and, once again, very

good correlations were established between anesthetic and toxic activities. Basak et al.[49] used 15 industrially important esters in a correlation of the LC_{50} values in Pimephales Promelas with $^1\chi$ and $^1\chi^v$. The best single parameter for such prediction was $^1\chi^v$, though a multiparametric relationship based on $^1\chi^v$, the octanol/water partition coefficient and a steric parameter proved to yield the most significant correlation.

A number of studies have been undertaken of the toxicity of compounds containing either chlorine or nitrogen heteroatoms. Sabljić[50] investigated the toxicity of 19 chlorinated aromatic and aliphatic hydrocarbons and phenols toward the sheepshead minnow Cyprinodon variegatus. The best single-parameter correlation was obtained for $^0\chi$, though $^1\chi$ and $^2\chi$ also gave reasonable correlations. The derived equations account for no more than 84-88% of the variation in the toxicity data; this is hardly surprising in view of the low accuracy of these data. A set of chlorinated phenols and a miscellaneous set of organic compounds containing halogenated and nitrated products were used by Koch[51] for correlations of their toxicity to a variety or organisms, including rats and guppies, with connectivity indices. The best result was from the correlation of the LC_{50} concentration of chlorophenols using guppies:

$$\log LC_{50} = 3.257 - 0.6719\ ^1\chi^v$$

($r = 0.982$, $n = 10$). The general conclusion reached was that the correlation coefficients depend on the organism used and the similarity of the structure of the chemical compounds involved, with closely similar compounds yielding the best resuits. Various nitrogen-containing heterocycles were investigated by Schultz et al.[52] for correlations of their toxicity toward axenic cultures of the freshwater ciliate Tetrahymena pyriformis. A better correlation with $^1\chi^v$

was obtained for six-membered heterocycles than that for five-membered heterocycles (r = 0.988 as against 0.966). Elucidation of the effect of structure on toxicity appears to be possible from an analysis of the correlations found.

The molecular structure of hydroxylated compounds, namely alcohols and phenols, as reflected by connectivity indices, also seems to be closely related to their toxicity. Kier and Hall[53] used the $^1\chi$ indices of a set of 20 mixed alcohols to correlate with the haemolysis of erythrocytes, the inhibition of the Tubifex worm movement, fish narcosis, and the LD_{50} dose in mice. One-parameter equations gave a fair account of the variation of potency with molecular structure in all cases, suggesting that the heteroatom plays a more or less constant role in these biological processes. A set of 25 substituted phenols was employed by Hall and Kier[54] for a toxicity correlation using the fathead minnow Pimephales promelas. The best single-parameter equation was found to be:

$$phC_{50} = 1.079\ {}^3\chi_p^v + 2.528$$

(r = 0.903, s = 0.347, n = 25); the best two-parameter equation (using $^1\chi$ and $^3\chi_p^v$) improved r to 0.934. The equations were considered good enough to predict unknown phenol toxicities.[54]

Organotin compounds have been the subject of two recent studies by Laughlin et al.[55] and Vighi and Calamari.[56] In the former study the LC_{50} values of eight triorganotin compounds for the crab zoeae Rhithropanopeus harrisii were correlated with a variety of different parameters. Although connectivity indices were not employed, it was evident that the best correlations were obtained from parameters, such as molar surface areas, which can be reliably modelled using topological indices. In the latter study explicit use was made of $^1\chi$ and $^1\chi^v$ in characterizing the toxicities of 15 organotin compounds against Daphnia magna.

This study confirmed that molecular topology can be employed to model the behavior of organotin compounds. With a single-parameter equation the best correlation obtained was:

$$\log\left(\frac{1}{EC_{50}}\right) = 0.749\ {}^{1}\chi^{v} - 5.63$$

($r = 0.925$, $n = 12$), where EC_{50} is the effective concentration of organotin which is lethal to 50% of the organisms. The best overall correlation was obtained using a triparametric approach based on lipophilic (log P), electronic (pK_a) and steric (${}^{1}\chi$) characteristics; this raised the correlation coefficient to 0.989.

Some Miscellaneous Applications

In this section brief mention is made of a number of publications which have some bearing on the prediction of biological behavior in chemical species. The way in which connectivity indices are related to molecular volume was investigated by Hall and Kier,[57] and the capacity of the indices to model lipophilicity was studied by Basak et al.[58] The use of connectivity indices in the evaluation of environmental pollutants has been discussed by Koch.[59] Connectivity indices have also found application in both discriminant analysis[60] and principal component analysis.[61,62] The structural information which molecular connectivities indices encode has been the subject of several papers.[63-65] The extent to which the indices can model steric phenomena has also been explored by several authors.[66-68] The possiblity of obtaining good correlations in multiple regression analyses by chance has been ascertained by Topliss and Costello,[69] and shown to be negligible for the great majority of the studies described herein.

Acknowledgment

Partial support of this project by the U.S. Office of Naval Research is gratefully acknowledged.

References

1. M. Randić, J. Am Chem. Soc. 97 (1975) 6609.
2. A.T. Balaban, I. Motoc, D. Bonchev and O. Mekenyan, Topics Curr. Chem. 114 (1983) 21.
3. D. Bonchev, Information-Theoretic Indices for Characterization of Chemical Structures, Research Studies Press, Chichester, U.K., 1983.
4. A.T. Balaban and I. Motoc, Handbook of Topological Indices, CRC Press, Boca Raton, Florida, 1986.
5. A.T. Balaban, A. Chiriac, I. Motoc and Z. Simon, Lect. Notes in Chem. 15 (1980) pp. 22-51.
6. A. Sabljić and N. Trinajstić, Acta Pharm. Jugosl. 31 (1981) 189.
7. N. Trinajstić, Chemical Graph Theory, Vol. II, CRC Press, Boca Raton, Florida, 1983, chap. 4, pp. 105-140.
8. L.B. Kier and L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
9. D.H. Rouvray, J. Math. Chem., 1 (1987) 1.
10. L.B. Kier, W.J. Murray, M. Randić and L.H. Hall, J. Pharm. Scis 65 (1976) 1226.
11. L.B. Kier and L.H. Hall, J. Pharm. Scis 65 (1976) 1806.
12. W.J. Murray, L.H. Hall and L.B. Kier, J. Pharm. Scis 64 (1975) 1978.
13. L.B. Kier, L.H. Hall, W.J. Murray and M. Randić, J. Pharm. Scis 64 (1975) 1971.
14. C. Hansch and T. Fujita, J. Am. Chem. Soc. 86 (1964) 1616.
15. D. Agin, L. Hersch and D. Holtzman, Proc. Natl. Acad. Sci. U.S. 53 (1965) 952.
16. M. Randić, J. Chromatogr. 161 (1978) 1.
17. L.B. Kier, W.J. Murray and L.H. Hall, J. Med. Chem. 18 (1975) 1272.

18. T. Di Paolo, L.B. Kier and L.H. Hall, Mol. Pharm. 13 (1977) 31.

19. T. Di Paolo, L.B. Kier and L.H. Hall, J. Pharm. Scis 68 (1979) 39.

20. T. Di Paolo, J. Pharm. Scis 67 (1978) 564.

21. W.J. Murray, L.B. Kier and L.H. Hall, J. Med. Chem. 19 (1976) 573.

22. M.-C. Bonjean and L. Luu Duc, Eur. J. Med. Chem. 13 (1978) 73.

23. S.C. Basak, D.P. Gieschen, V.R. Magnuson and D.K. Harriss, IRCS Med. Sci. 10 (1982) 619.

24. R.A. Glennon, L.B. Kier and A.T. Shulgin, J. Pharm. Scis 68 (1979) 906.

25. L.B. Kier and L.H. Hall, J. Med. Chem. 20 (1977) 1631.

26. L.B. Kier and R.A. Glennon, Life Sciences 22 (1978) 1589.

27. R.A. Glennon and L.B. Kier, Eur. J. Med. Chem. 13 (1978) 219.

28. J.E. Amoore, Molecular Basis of Odor, C.C. Thomas, Springfield, Illinois, 1970.

29. J.E. Amoore, Nature (Lond.) 233 (1971) 270.

30. L.B. Kier, T. Di Paolo and L.H. Hall, J. Theor. Biol. 67 (1977) 585.

31. L.H. Hall and L.B. Kier, J. Pharm. Scis 66 (1977) 642.

32. L.B. Kier, J. Pharm. Scis 69 (1980) 416.

33. R.J. Gardner, Tech. Q. Master Brew. Assoc. Am. 16 (1979) 204.

34. R.J. Gardner, J. Sci. Food Agric. 31 (1980) 23.

35. A.J. Richard and L.B. Kier, J. Pharm. Scis 69 (1980) 124.

36. B. van't Riet, L.B. Kier and H.L. Elford, J. Pharm. Scis 69 (1980) 856.

37. A. Sabljić and M. Protić-Sabljić, Mol. Pharm. 23 (1983) 213.

38. L.B. Kier and L.H. Hall, J. Pharm. Scis 67 (1978) 1408.

39. A. Sabljić, N. Trinajstić and D. Maysinger, Acta Pharm. Jugosl. 31 (1981) 71.

40. L.H. Hall and L.B. Kier, J. Pharm. Scis 67 (1978) 1743.

41. L.H. Hall and L.B. Kier, Eur. J. Med. Chem. 13 (1978) 89.

42. C. Mercier and J.-E. Dubois, Eur. J. Med. Chem. 14 (1979) 415.

43. A.K. Samanta, S.K. Ray, S.C. Basak and S.K. Bose, Arzneim. Forsch. 32 (1982) 1515.

44. L.B. Kier, R.J. Simons and L.H. Hall, J. Pharm. Scis 67 (1978) 725.

45. P.C. Jurs, J.T. Chou and M. Yuan, J. Med. Chem. 22 (1979) 476.

46. H. Govers, C. Ruepert and H. Aiking, Chemosphere 13 (1984) 227.

47. T. Di Paolo, J. Pharm. Scis 67 (1978) 566.

48. F. Heymans, L. LeThérizien and J.-J. Godfroid, J. Med. Chem. 23 (1980) 184.

49. S.C. Basak, D.P. Gieschen and V.R. Magnuson, Environm. Tox. Chem. 3 (1984) 191.

50. A. Sabljić, Bull. Environm. Contam. Toxicol. 30 (1983) 80.

51. R. Koch, Chemosphere 11 (1982) 925.

52. T.W. Schultz, L.B. Kier and L.H. Hall, Bull. Environm. Contam. Toxicol. 28 (1982) 373.

53. L.B. Kier and L.H. Hall, Bull. Environm. Contam. Toxicol. 29 (1982) 121.

54. L.H. Hall and L.B. Kier, Bull. Environm. Contam. Toxicol. 32 (1984) 354.

55. R.B. Laughlin, W. French, R.B. Johannesen, H.E. Guard and F.E. Brinckman, Chemosphere 13 (1984) 575.

56. M. Vighi and D. Calamari, Chemosphere 14 (1985) 1925.

57. L.H. Hall and L.B. Kier, Eur. J. Med. Chem. 16 (1981) 399.

58. S.C. Basak, D.K. Harriss and V.R. Magnuson, J. Pharm. Scis 73 (1984) 429.

59. R. Koch, Toxicol. Environm. Chem. 7 (1984) 331.

60. D.R. Henry and J.H. Block, J. Med. Chem. 22 (1979) 465.

61. L.B. Burkhard, A.W. Andren and D.E. Armstrong, Chemosphere 12 (1983) 935.

62. M.P. Moulton and T.W. Schultz, Chemosphere 15 (1986) 59.

63. L.B. Kier and L.H. Hall, Eur. J. Med. Chem. 12 (1977) 307.

64. L.B. Kier, J. Pharm. Scis 69 (1980) 1034.

65. J.T. Edward, Can. J. Chem. 60 (1982) 480.

66. G.R. Parker, J. Pharm. Scis 67 (1978) 513.

67. B.K. Evans, K.C. James and D.K. Luscombe, J. Pharm Scis 68 (1979) 370.

68. S.C. Basak, D.P. Gieschen, D.K. Harriss and V.R. Magnuson, J. Pharm. Scis 72 (1983) 934.

69. J.G. Topliss and R.J. Costello, J. Med. Chem. 15 (1972) 1066.

Captions to Figures

Figure 1. Examples of the four different types of subgraphs used in the calculation of molecular connectivity indices.

Figure 2. The correlation of logarithm of the octanol/water partition coefficient against $^1\chi$ for a mixture of 138 esters, carboxylic acids, alcohols, amines, ketones and ethers. [Based on W.J. Murray, L.H. Hall and L.B. Kier, J. Pharm. Scis 64 (1975) 1978.]

Figure 3. The correlation of molecular polarizability against $^1\chi$ for 36 different compounds used in a nonspecific anesthesia study. [Based on L.B. Kier, L.H. Hall, W.J. Murray and M. Randić, J. Pharm. Scis 64 (1975) 1971.]

Figure 4. The correlation of logarithm of the minimum blocking concentration against $^1\chi$ for 36 different compounds used in a nonspecific anesthesia study. [Based on L.B. Kier, L.H. Hall, W.J. Murray and M. Randić, J. Pharm. Scis 64 (1975) 1971.]

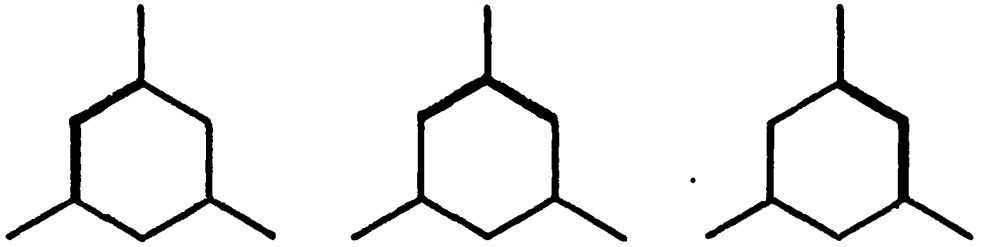
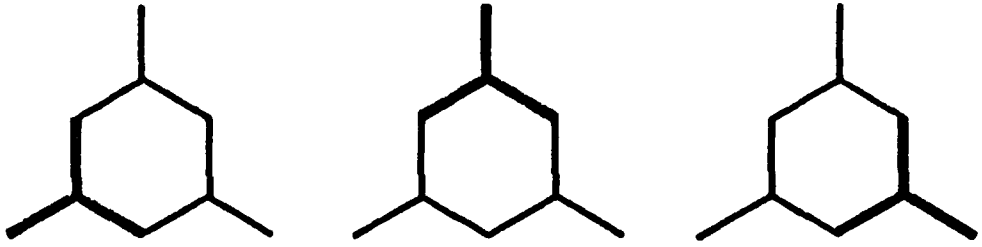
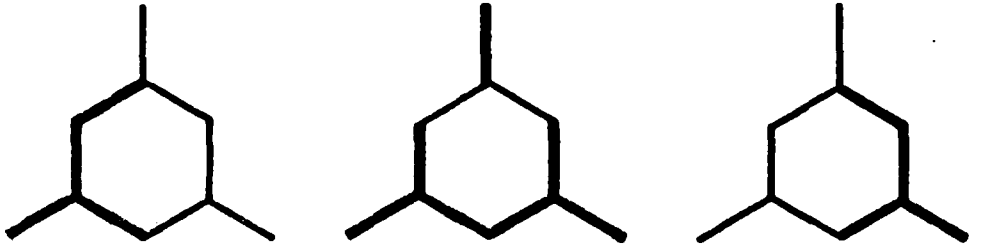
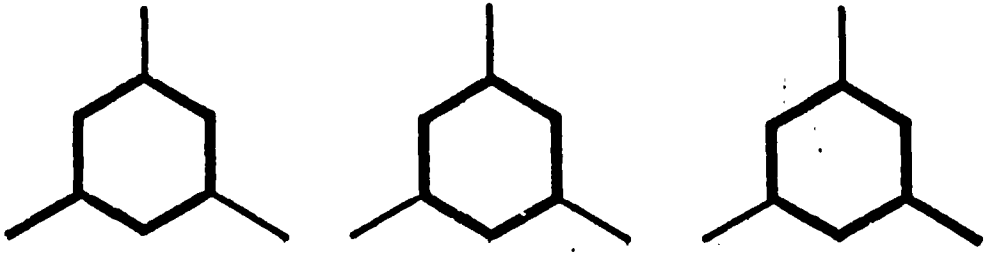
SUBGRAPH	EXAMPLES
Path	
Cluster	
Path/Cluster	
Chain	

Figure 1

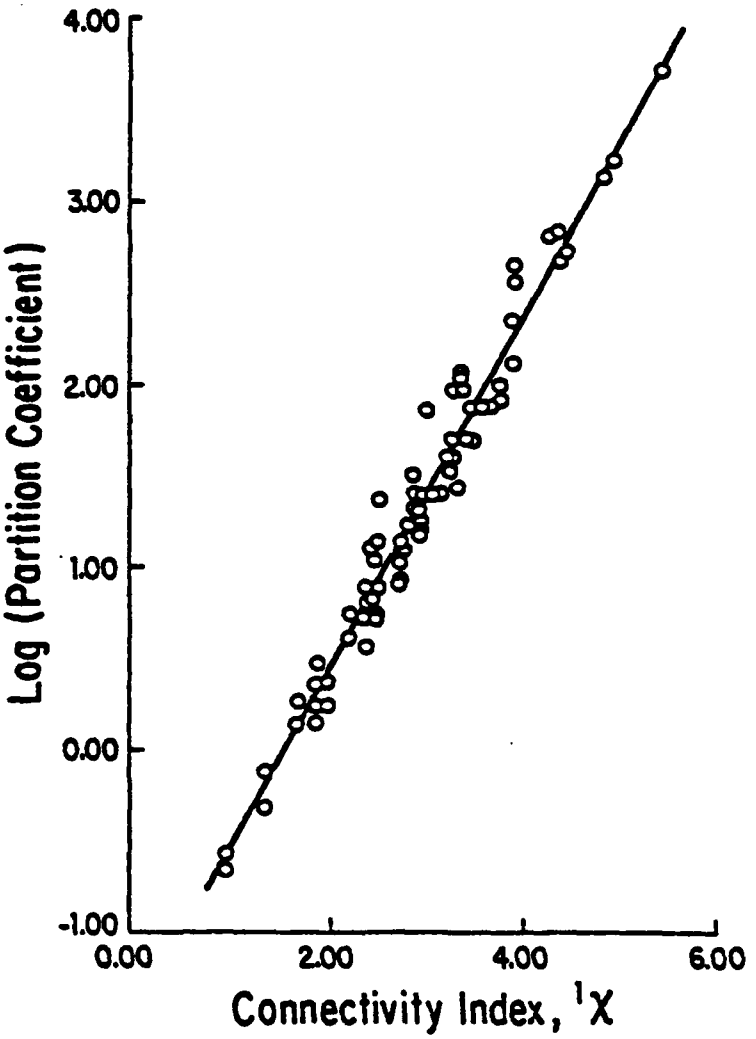

4.00
3.00
2.00
1.00
0.00
-1.00
Log (Partition Coefficient)
0.00
2.00
4.00
6.00
Connectivity Index, $^1\chi$

Figure 2

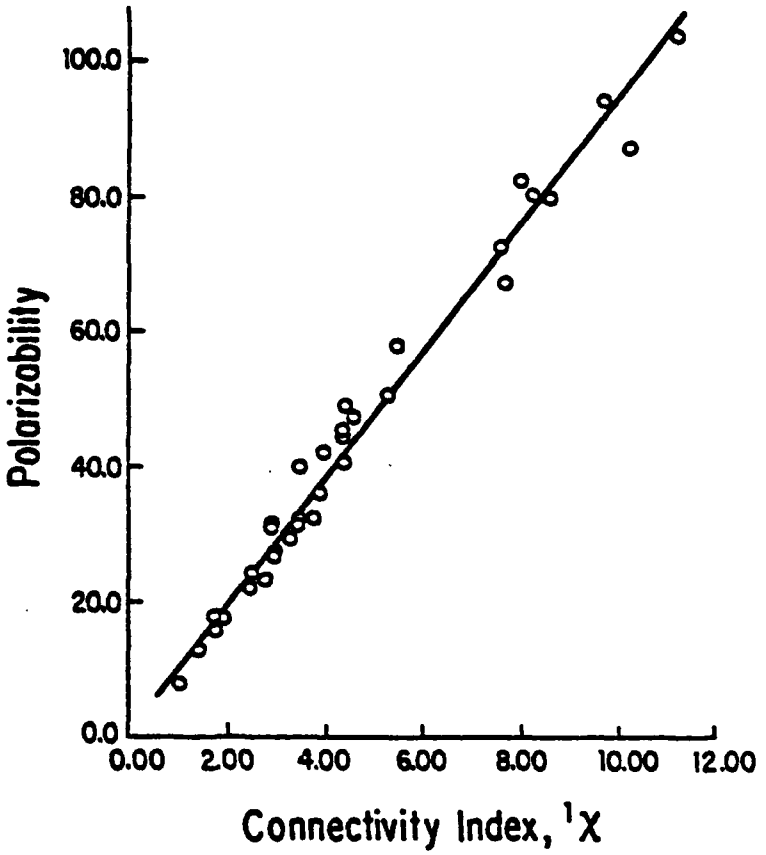

Polarizability
100.0
80.0
60.0
40.0
20.0
0.0
0.00
2.00
4.00
6.00
8.00
10.00
12.00
Connectivity Index, $^1\chi$

Figure 3

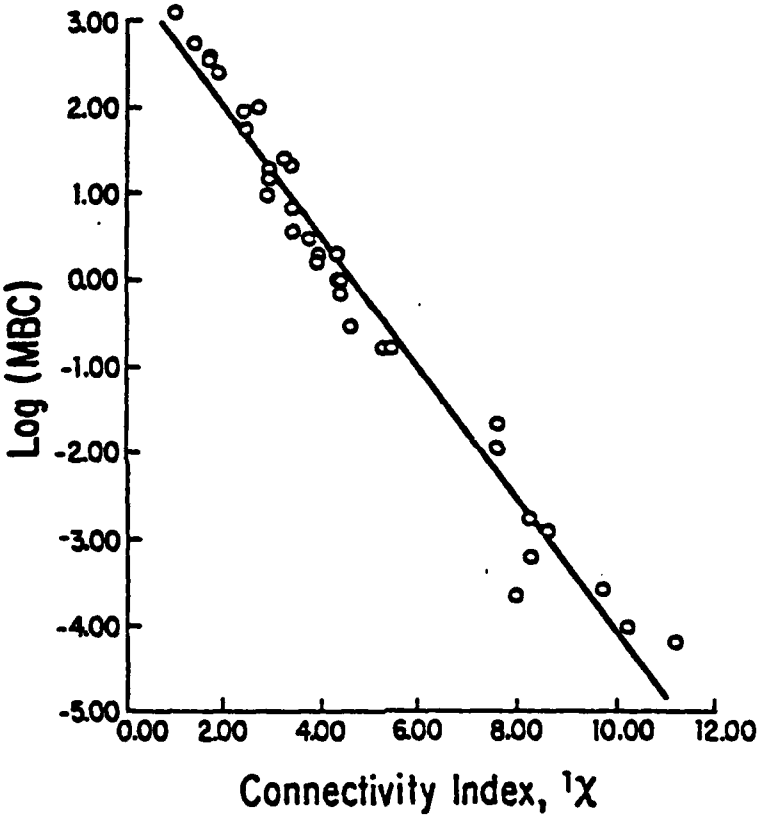

3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-5.00
Log (MBC)
0.00
2.00
4.00
6.00
8.00
10.00
12.00
Connectivity Index, $^1\chi$

Figure 4

END

FILMED

5-86

DTIC