Molecular Connectivity. 4. Relationships to Biological Activities

Lemont B. Kier,* Wallace J. Murray,

Massachusetts College of Pharmacy, Boston, Massachusetts 02115

and Lowell H. Hall

Eastern Nazarene College, Quincy, Massachusetts 02169. *Received April 3, 1975*

A simply computed number, encoding the molecular connectivity of a molecule, has been found to yield a correlation with biological activities of assorted compounds in several studies.

Molecular connectivity, the manner in which atoms are connected or branched in a molecule, is a fundamental characteristic of the structure. It is well known that chain isomers of a molecule have varying values of their physical and chemical properties. The frequently observed direct relationship between the number of atoms in a homologous series and a physical property is interrupted when a structural change introducing branching is encountered. In recent studies[1-3] we have employed a simple index derived from the connectivity of hydrocarbons suggested by Randić,[4] encoding in a single number this structural characteristic. In this series we have demonstrated a close correlation between this index, which we call the connectivity index, and several physical properties.[2,3]

The connectivity index has been found to correlate with a computed cavity surface area,[1] solubility,[2] and partition coefficient.[3] These are physical properties which have been traditionally linked to the surface area of a molecule. In this same series of studies we have found a correlation between χ and polarizability,[1] a physical property traditionally linked to the volume of a molecule. Other examples to be published reveal further correlations with properties linked to both surface area and volume of a molecule.

In the initial study in this series, we found a close correlation between χ and the nonspecific local anesthetic activity of a widely diverse series of molecules.[1] This series of molecules had previously been found to have activities correlating with polarizability.[5] In a subsequent study we found a correlation between χ and a biological activity previously linked to solubility.[3]

The finding of a single number, derived from molecular connectivity, which is capable of correlating with both volume and surface area-related physical properties, leads us to believe that this characteristic is contributing in a significant way to both (geometric) aspects of a molecule. The finding of a relationship between χ and a biological property encourages us to consider the generality of this relationship in a wider range of biological molecules. This consideration forms the basis of this study.

Calculation of the Connectivity Index. To calculate the value of χ for a molecule, write down the molecular skeleton, ignoring the hydrogens. The first row atoms are treated identically. A number, δ_i, is assigned to each atom in the molecule which equals the number of bonds connected to that atom. Thus, δ_i = 1, 2, 3, or 4 for carbon atoms. A value for each bond in the molecule, C_{ij}, is computed from each pair of bonded atoms by $C_{ij} = (\delta_i\delta_j)_k^{-1/2}$. Bond k is formed between atoms i and j. Finally, χ is computed for the molecule as $\chi = \Sigma C_{ij}$. In the case of a cyclic compound, there is one more bond than in the corresponding straight chain isomer. Accordingly, the value of one ring C_{ij} must be subtracted to arrive at χ (1).

A. A Study of Tadpole Narcosis. A nonspecific narcotic effect of a wide diversity of compounds on tadpoles has been demonstrated[6] (Table I). We have computed χ for this series and have found a correlation with the activity. The results are shown in Table I. A statistical analysis yields the following results, in which n is the number of cases, r is the correlation coefficient, and s is the standard deviation.

$$\log 1/C = 0.922\ (\pm 0.049)\ \chi - 0.931\ (\pm 0.158)$$
$$n = 36,\ r = 0.956,\ s = 0.297$$

B. A Study of Fungus Toxicity. A nonspecific toxic effect on the Madison 517 fungus has been demonstrated by a wide variety of molecules.[7] We have computed the χ for this series, omitting seven molecules which are difunctional. The results are shown in Table II. A statistical analysis yields the following results.

$$pC = 0.775\ (\pm 0.032)\ \chi - 1.077\ (\pm 0.119)$$
$$n = 45,\ r = 0.965,\ s = 0.263$$

C. A Study of Succinate Oxidase Enzyme Inhibitors. A diverse group of molecules has been demonstrated to have a range of activity in inhibiting, at the 15–20% level, the enzyme succinate oxidase from bovine muscle.[8] We have found that χ correlates with the activity of these molecules. The results are shown in Table III. A statistical analysis yields the following results.

$$pC = 0.916\ (\pm 0.073)\ \chi - 1.582\ (\pm 0.174)$$
$$n = 13,\ r = 0.966,\ s = 0.169$$

D. A Study of Thymidine Phosphorylase Inhibitors. A number of 1-substituted thymidine derivatives have been reported to inhibit the enzyme thymidine phosphorylase.[9] The values of χ for the series was found to correlate with the 50% inhibition concentration. The results are shown in Table IV. A statistical analysis yields the following results.

$$\log 1/C = 0.373\ (\pm 0.051)\ \chi - 3.415\ (\pm 0.325)$$
$$n = 11,\ r = 0.924,\ s = 0.207$$

E. Study of Adenosine Deaminase Inhibitors. A series of adenosine derivatives has been reported to be inhibitors of the enzyme adenosine deaminase.[10] The value of χ relates to the inhibitory activity of these molecules. The results are shown in Table V. A statistical analysis yields the following results.

$$\log I/S_{0.5} = 0.449\ (\pm 0.025)\ \chi - 2.801\ (\pm 0.151)$$
$$n = 8,\ r = 0.991,\ s = 0.082$$

F. Study of Butyrylcholinesterase Inhibitors. A series of piperidinecarboxylic acid amide derivatives has been found to inhibit butyrylcholinesterase.[11] The χ value of each member of the series was calculated and found to correlate with the inhibitory potency. The results are shown in Table VI. A statistical analysis yields the following results.

$$pI_{50} = 0.585\ (\pm 0.025)\ \chi + 0.617\ (\pm 0.241)$$
$$n = 7,\ r = 0.995,\ s = 0.062$$

Table I. Relationship of χ to Tadpole Narcosis

Compd	χ	Log 1/C obsd[a]	Log 1/C calcd
Methanol	1.000	0.24	0.001
Ethanol	1.414	0.54	0.38
Propanol	1.914	0.96	0.84
Butanol	2.414	1.42	1.29
Octanol	4.414	3.40	3.12
Isopropyl alcohol	1.732	0.89	0.70
Isobutyl alcohol	2.270	1.35	1.16
tert-Butyl alcohol	2.000	0.89	0.91
Isoamyl alcohol	2.770	1.64	1.62
tert-Amyl alcohol	2.561	1.24	1.43
Thymol	4.608	4.26	3.29
1,3-Dimethoxybenzene	4.363	3.35	3.07
1,4-Dimethoxybenzene	4.363	3.05	3.07
Acetone	1.732	0.54	0.67
2-Butanone	2.270	1.04	1.16
3-Pentanone	2.807	1.54	1.65
2-Pentanone	2.770	1.72	1.62
Acetophenone	3.804	3.03	2.56
Acetal	2.414	1.98	2.56
Ethyl ether	2.414	1.57	1.29
Anisole	3.432	2.82	2.22
Methyl acetate	2.270	1.10	1.16
Ethyl formate	2.414	1.15	1.29
Ethyl acetate	2.770	1.52	1.62
Ethyl propionate	3.307	1.96	2.11
Propyl acetate	3.270	1.96	2.07
Ethyl butyrate	3.807	2.37	2.56
Ethyl isobutyrate	3.680	2.24	2.45
Butyl acetate	3.770	2.30	2.53
Isobutyl acetate	3.626	2.24	2.40
Ethyl valerate	4.307	2.72	3.02
Amyl acetate	4.770	2.72	2.99
Butyl valerate	5.307	3.60	3.93
Methyl carbamate	2.270	0.57	1.16
Ethyl carbamate	2.770	1.39	1.62
Phenyl carbamate	4.318	3.19	3.03

[a] Reference 6.

Discussion

The six studies described here show a correlation between χ and the biological activity. The molecules and molecular fragments examined in each study have in common a probable minimal electrostatic role in their interaction with a receptor or enzyme active site. Rather a comparison of these molecules and fragments indicates a role of nonspecific interaction of the van der Waals type with complimentary features on their receptors. These studies and previous studies in this series lead us to believe that χ describes a structural characteristic which governs geometric features basic to the physical properties that we have encountered. The simplicity of the calculation of χ makes it possible to use this index in many circumstances.

It has been suggested that the success of the connectivity index χ, in relating to the interactions in this study, is that the connectivity matrix, which is related to the value of χ, is related to simple Huckel-type secular equations. From this, χ may be viewed as a summation of bond orders of the molecules studied. This in turn implies a relationship to atom–bond (polarization) and bond–bond (dispersion) interaction characteristics. As a consequence χ would be expected to relate to biological activities dependent primarily on these forces.

Table II. Relationship between Connectivity Index and Log Fungus Minimum Toxic Dose (pC)

Compd	χ	pC obsd[a]	pC calcd
Methyl alcohol	1.000	−0.24	−0.30
Ethyl alcohol	1.414	−0.04	0.01
Propyl alcohol	1.914	0.44	0.40
Butyl alcohol	2.414	0.87	0.79
Pentyl alcohol	2.914	1.38	1.18
Hexyl alcohol	3.414	1.83	1.56
Heptyl alcohol	3.914	2.32	1.95
Octyl alcohol	4.414	2.86	2.34
Nonyl alcohol	4.914	3.18	2.73
Decyl alcohol	5.414	3.57	3.11
Isopropyl alcohol	1.732	0.24	0.26
sec-Butyl alcohol	2.270	0.60	0.68
tert-Butyl alcohol	2.000	0.46	0.47
sec-Amyl alcohol	2.770	1.08	1.07
2-Methylbutyl alcohol	2.807	1.19	1.09
3-Methylbutyl alcohol	2.770	1.25	1.07
3-Pentyl alcohol	2.807	1.01	1.09
tert-Pentyl alcohol	3.125	1.44	1.34
2-Ethylbutyl alcohol	3.346	1.73	1.51
1-Methylheptyl alcohol	4.270	2.49	2.23
2-Ethylhexyl alcohol	4.345	2.55	2.29
Diphenylmethyl alcohol	5.877	2.57	3.47
Phenylethyl alcohol	3.432	1.57	1.58
3-Phenylpropyl alcohol	3.932	2.00	1.97
Ethyl ether	2.414	0.55	0.79
Propyl ether	3.414	1.55	1.56
Isopropyl ether	3.125	1.13	1.34
Butyl ether	4.414	2.54	2.34
Acetone	1.732	0.15	0.26
Methyl acetate	2.270	0.59	0.68
Ethyl acetate	2.770	0.80	1.07
Propyl acetate	3.270	1.23	1.45
Butyl acetate	3.770	1.69	1.84
Pentyl acetate	4.270	2.15	2.23
Heptyl acetate	5.270	2.60	3.00
Ethyl propionate	3.308	1.20	1.48
Ethyl butyrate	3.808	1.63	1.87
Ethyl caproate	4.808	2.59	2.64
Ethyl caprylate	5.808	3.39	3.42
Pentyl butyrate	5.308	2.85	3.03
2-Ethylbutyl acetate	4.702	2.36	2.56
1-Methylisopentyl acetate	4.520	2.14	2.42
Pentyl-*tert*-pentyl acetate	5.520	3.60	3.20
Isobutyl alcohol	2.270	0.77	0.68
2-Heptanone	3.770	1.94	1.84

[a] Reference 7.

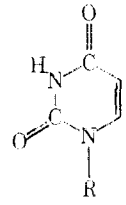
This explanation is not completely satisfying to us. Bond order or bond charge is an index of molecular orbital theory derived from the eigenfunctions of a Hückel-type matrix. The connectivity matrix which we can write for each molecule is not formally solved for the eigenvalues or eigenfunctions in computing χ. On the other hand, we have considered that each C_{ij} term could be looked upon as a form of a solution of a two-center (bond) matrix formed from δ_i and δ_j. The localized nature of each $(\delta_i\delta_j)^{-1/2}$ term provides a minimal influence on other $(\delta_i\delta_j)^{-1/2}$ terms in computing χ. The bond orders computed in molecular orbital theory from the Hückel matrix compute the eigenfunctions from all terms simultaneously in the matrix. Moreover this ex-

Table III. Relationship between Connectivity Index and Log C for I_{15-20} Succinate Oxidase

Compd	χ	pC obsd[a]	pC calcd
Methanol	1.000	0.57	0.67
Ethanol	1.414	0.09	0.29
Propanol	1.914	0.12	0.17
Isobutyl alcohol	2.270	0.60	0.50
Isopentyl alcohol	2.770	0.85	0.96
Allyl alcohol	1.914	0.05	0.17
Acetone	1.732	0.10	0.01
2-Butanone	2.270	0.33	0.50
2-Pentanone	2.770	0.74	0.96
Aniline	2.894	1.05	1.07
Pyridine	2.500	0.62	0.71
Phenol	2.894	1.40	1.07
2-Cresol	3.304	1.60	1.45

[a] Reference 8.

Table IV. Inhibitors of Thymidine Phosphorylase



R	χ	Log 1/C obsd[a]	Log 1/C calcd
Methyl	3.698	2.30	2.04
Butyl	5.405	1.35	1.40
Isopentyl	5.592	1.30	1.33
Cyclopentyl	5.270	1.28	1.45
Isohexyl	6.092	1.17	1.14
Pentyl	5.904	1.15	1.21
$-(CH_2)_3Ph$	7.254	1.11	0.71
$-(CH_2)_2Ph$	6.754	0.80	0.90
$-CH_2Ph$	6.254	0.76	1.08
$-(CH_2)_4Ph$	7.754	0.60	0.53
$-(CH_2)_5Ph$	8.254	0.32	0.34

[a] Reference 9.

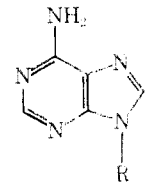
planation in no way accounts for the success of χ in correlating with a computed cavity surface area,[1] a property which is geometric in nature.

This is by no means a closed issue and it will be the subject of intensive study in our laboratory.

The generality of the use of χ to predict biological activity relating to nonspecific molecular features is supported by these findings.

Acknowledgment. We gratefully acknowledge the helpful discussions with M. Randić, Tufts University, Medford, Mass. We also acknowledge the use of the Computer Center facilities of Eastern Nazarene College, Quincy, Mass.

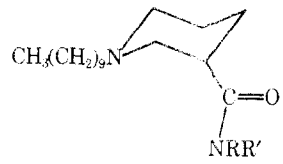
Table V. Inhibition of Adenosine Deaminase



R	χ	Log $I/S_{0.5}$ obsd[a]	Log $I/S_{0.5}$ calcd
Methyl	4.065	0.86	0.98
Ethyl	4.603	0.79	0.74
Propyl	5.103	0.52	0.51
Butyl	5.603	0.36	0.29
Pentyl	6.103	0.15	0.06
Hexyl	6.603	0.15	0.16
Heptyl	7.103	0.49	0.39
Octyl	7.603	0.62	0.61

[a] Reference 10.

Table VI. Butyrylcholinesterase Inhibitory Potencies



R	R′	χ	pI_{50} obsd[a]	pI_{50} calcd
H	H	8.166	4.21	4.16
H	CH_3	8.727	4.46	4.49
H	C_2H_5	9.227	4.86	4.78
CH_3	CH_3	9.109	4.66	4.72
CH_3	C_2H_5	9.647	5.01	5.03
C_2H_5	C_2H_5	10.184	5.28	5.34
C_3H_7	C_3H_7	11.184	5.98	5.93

[a] Reference 11.

References and Notes

(1) L. B. Kier, L. H. Hall, and W. J. Murray, *J. Pharm. Sci.*, in press.
(2) L. H. Hall, L. B. Kier, and W. J. Murray, *J. Pharm. Sci.*, in press.
(3) W. J. Murray, L. H. Hall, and L. B. Kier, *J. Pharm. Sci.*, in press.
(4) M. Randić, *J. Am. Chem. Soc.*, in press.
(5) D. Agin, L. Hersh, and D. Holtzman, *Proc. Natl. Acad. Sci. U.S.A.*, **53**, 952 (1965).
(6) E. Overton, "Studies on Narcosis", Fischer, Jena, Germany, 1901.
(7) R. H. Baechler, *Proc. Am. Wood-Preserv. Assoc.*, **43**, 94 (1947).
(8) F. Battelli and L. Stern, *Biochem. Z.*, **52**, 226 (1913).
(9) B. R. Baker and M. Kawaza, *J. Med. Chem.*, **10**, 302 (1967).
(10) H. J. Schaeffer, R. N. Johnson, E. Odin, and C. Hansch, *J. Med. Chem.*, **13**, 452 (1970).
(11) J. M. Clayton and W. P. Purcell, *J. Med. Chem.*, **12**, 1087 (1969).