rotated less than 180° in a clockwise direction.
(9) V. Cody in "Thyroid Research", E. Braverman and J. Robbins, Ed., Excerpta Medica, Amsterdam, 1975, p 290.
(10) Note: all crystallographic data for Figure 4 are accessible on the NIH PROPHET computer data files.
(11) V. Cody, W. L. Duax, and H. A. Hauptman, *Int. J. Pept. Protein Res.*, 5, 297 (1973).
(12) V. Cody and W. L. Duax, *Biophys. Biochem. Res. Commun.*, 52, 430 (1973).
(13) N. Camerman and A. Camerman, *Acta Crystallogr., Sect. B*, 30, 1832 (1974).
(14) V. Cody, unpublished results.
(15) In order to develop a consistent numbering scheme for all thyroactive compounds, a standard orientation was chosen such that the C^{α} atom was viewed above the inner phenyl ring plane. The inner phenyl ring was then numbered clockwise and the outer phenyl ring is numbered in such a manner that the C(2′) atom is always distal to the inner ring. In this manner, all the compounds are self-consistent and correlations between ϕ and ϕ' can be detected.

Structure–Activity Studies on Hallucinogenic Amphetamines Using Molecular Connectivity

Lemont B. Kier*

Department of Pharmaceutical Chemistry, Medical College of Virginia, Richmond, Virginia 23298

and Lowell H. Hall

Eastern Nazarene College, Quincy, Massachusetts 02170. *Received June 9, 1977*

A series of ring-substituted hallucinogenic amphetamines has been analyzed using molecular connectivity. A correlating equation has been found between potency and connectivity terms. The equation permits an interpretation of SAR. The equation is capable of predicting potency for amphetamines not in the list and mescalines and tryptamines.

The contributions of Shulgin and his colleagues have afforded an opportunity to study structure–activity relationships (SAR) among a fairly large list of hallucinogenic agents.[1] Specifically, the availability of over a score of ring-substituted amphetamines with comparative hallucinogenic potencies makes it possible to examine the structural features contributing to the activity.

Snyder and Merril were the first to use these data in a study employing semiempirical quantum mechanical calculations.[2] Their Hückel MO calculations showed some relationship between the energy of the highest occupied MO, E (HOMO), and the potency, established by Shulgin, among limited sets of amphetamine and tryptamine derivatives. Kang and Green calculated the same index using the INDO MO method for a set of 14 substituted amphetamines.[3] They found a modest correlation with potency.

These studies gave an indication that the substitution pattern on the amphetamine ring may influence the interaction with a receptor feature through forces of the van der Waals type. This has prompted us to analyze the interaction energies of 17 derivatives with model receptor molecules.[4] The correlation found, $r = 0.85$, was encouraging.

Barfknecht has pursued a different approach by attempting to relate partition coefficients of neutral amphetamines with potency.[5] The correlation with activity was modest. It is not certain how partition coefficients can be interpreted in terms of molecular structure of the amphetamine analogues.

In recent studies of biological SAR,[6-9] we have analyzed molecular structure in terms of the number and kinds of atoms, bonding type, and adjacency environment. This method, having its roots in topology, is called molecular connectivity.[9] The theory, formalism, and applications have been described in one source.[9] A condensed description of molecular connectivity calculations is given in Appendix I.

It is well known that measured property values vary with change in molecular structure. What is meant by change in structure? Structural change usually includes variation in the number and type of atoms, branching, cyclization, and change in bond types. Such properties as boiling point and molar refraction of normal alkanes and normal primary alcohols illustrate structural influence on properties. There is a linear relationship between these properties and the number of carbon atoms for these straight-chain molecules. The same trend is observed in homologous series of molecules which produce many interesting biological phenomena.

In these cases structural information necessary to establish structure–activity relationships is simply the number of carbon atoms. At the present time such relationships cannot be developed de novo from quantum mechanics and thermodynamics so that these properties could be predicted directly from the number of carbon atoms alone. However, a relating equation can be established by standard regression analysis.

Now consider branching in alkanes or alcohols. The information contained only in the number of carbon atoms is inadequate for establishing close relationships to properties. It is well known to organic chemists that chain branching within a set of isomers leads to lower boiling points. How may such a structural characterization, in this case branching, be represented in quantitative terms suitable for establishing useful relationships? The number of carbon atoms is easily quantitated, but how may branching be quantitated?

Beyond branching one encounters the occurrence of heteroatoms, multiple bonding, and cyclization, features which are not readily quantitated in a form suitable for SAR. Molecular connectivity attempts to describe quantitatively these kinds of structural features at the same level of information as an atom count, i.e., a numerical value which can be determined unambiguously for a given molecule and which is also transferrable from study to study.

Among isomers the number of atoms and bonds remains constant; hence, these numbers are inadequate for structural description. Molecular connectivity begins with this branching pattern by depicting the molecular structure as the familiar skeleton formula. Based on this molecular

graph the χ terms provide the information necessary for adequate structural description. The first-order index $^1\chi$ is a summation of terms; each term is a numerical value for each bond in the skeleton. The numerical value for each bond term is computed from the degree of branching at the two atoms which form the bond. Thus, the $^1\chi$ index is a count of the bonds weighted by the degree of branching.

It is now apparent that $^1\chi$ carries structural information: the number of atoms and the complexity of skeletal branching. For example, in chain lengthening $^1\chi$ increases with a constant increment, 0.5, with the addition of each methylene group. On the other hand, $^1\chi$ decreases with increasing complexity of branching. In the isomeric hexanes, for example, $^1\chi$ values are as follows: *n*-hexane, 2.914; 3-methylpentane, 2.808; 2-methylpentane, 2.770; 2,3-dimethylbutane, 2.643; 2,2-dimethylbutane, 2.561. It is significant that this ranking of isomers by $^1\chi$ corresponds exactly to the order of boiling points and heat of formation for the alicyclic hexanes and branched primary pentanols. Therefore, it appears that $^1\chi$ encodes, at the same level as an atom count, an amount of information significantly greater than a simple atom count. The connectivity index $^1\chi$ is, then, an index of molecular structure bearing sufficient information to provide adequate correlation for some properties of a wide variety of molecules.[9]

In the discussion up to this point we have shown a connectivity index which is based upon a numerical value for each bond. This decomposition of the molecular graph into fragments, called subgraphs, need not be limited to individual bonds. This process can be extended to include pairs of adjacent bonds, called paths of length two. For that matter, we can consider subgraphs of three, four, or more skeletal bonds describing paths, clusters, and chains in the molecular skeleton. Each of these represents structural features in the molecule, features which may relate to the various properties of the molecule. The details of each of these χ terms, $^m\chi_t$, are described in Appendix I, along with the method of calculation.

The structural meaning of each χ index may be interpreted in terms of the details of the structure of a given set of molecules. For the $^3\chi_p$ index, for example, the magnitude is a weighted count of all fragments or subgraphs consisting of three bonds joined as a path. The weighting of each contributor to $^3\chi_p$ reflects the degree of branching at each of the four atoms which determine the three bonds. Both the number of subgraphs and the weighting depends heavily on the structure. The number of $^3\chi_p$ subgraphs in *o*-xylene is 11 whereas it is 10 for the meta and para isomers. The weighting is different, however, in each isomer leading to unique values of $^3\chi_p$: 2.540 for *o*-xylene, 2.199 for *m*-xylene, and 2.305 for *p*-xylene. Thus, $^3\chi_p$ reflects something of the difference in the pattern of branching in these three isomers.

There is contained in $^3\chi_p$ for the xylenes structural information which may play a significant role in a relationship to properties of xylenes. It is then possible by a regression analysis search to isolate a small number of χ terms which, in concert, are highly related to property values.[9] The molecular connectivity terms which appear in such significant correlations describe the dependence of the property on molecular structure. Hence, the $^m\chi_t$ terms are not the properties themselves but they describe numerically the significant structural characteristics to which the property is related.

χ terms are derived from adjacency relations in the molecular skeleton, relations reflecting the branching pattern, and long known to parallel variation of properties with molecular structure. Hence, it appears that χ terms are numerical values which are fundamental in definition, as fundamental as the structural characteristics commonly referred to by such qualitative terms as branching, chain lengthening, cyclization, and others.

The appearance of a constellation of χ terms in an equation which correlates significantly with a property presents an opportunity to assess the important structural features, in the form of subgraphs, which relate to the property. That more than one χ term may be necessary for a significant correlation is not unexpected. The definition of a single structural characteristic for a set of molecules may not generally contain sufficient information to provide adequate description for a meaningful relation to properties. An analogy from geometry illustrates this point. Two parameters are necessary to define the area of rectangles and three parameters for trapezoids whereas only one is required for squares.

Since χ terms are structure descriptors, a given χ term may appear in the equations relating to several physical properties and biological activities. A particular χ term, therefore, is not exclusively related to any particular property. As an example, let us refer to the appearance of $^3\chi_p$ in the correlation equation for density.[10] In this present study, as shown later, the $^3\chi_p$ term is also found to be an important descriptor. It does not follow that the property density bears any necessary relationship to hallucinogenic potency. The significance of the appearance of $^3\chi_p$ in the equation is its structural interpretation, as discussed in connection with each property.

Hallucinogenic Amphetamines. In the present study, we have considered 23 substituted amphetamines whose activity has been expressed as concentration relative to mescaline to produce an hallucinogenic effect.[5] We have converted the data in Table I to a molar basis and have used the log of this molar value. Excluded from this list are a few compounds whose activity has been expressed by an indefinite "less than" value. These excluded compounds are utilized in Table II for predictions of potency.

All $^m\chi_t$ terms have been computed for each molecule using a high speed computer program which is part of a computing system recently prepared in our laboratory. A total of 18 χ terms was examined. A multivariate search of connectivity terms was conducted in the regression analysis using a program system which considers all possibilities of two, three, or more variables. The best three-variable relationship is

$$\log \mu = \frac{45.16\ (\pm 7.30)}{^3\chi_p} + 1.288\ (\pm 0.20)\ ^6\chi_p - \frac{4.298\ (\pm 0.19)}{^4\chi_{pc}{}^v} - 5.592\ (\pm 2.32) \quad (1)$$

$$r = 0.920,\ S = 0.251,\ n = 23,\ F = 35.0\ (p < 0.0001)$$

The observed and predicted values are shown in Table I. Also in this table are listed the numerical values of the connectivity indices. Based on the appropriate *F* test, the addition of each of the variables in the regression analysis is statistically significant. The *F* for the addition of the term $1/^3\chi_p$ to the one-variable equation based on $1/^4\chi_{pc}{}^v$ is 33.6 as compared to the *F* value tabulated for $p < 0.0001$, 14.8. For the addition of the third variable, $^6\chi_p$, $F = 7.6$ as compared to the tabulated *F* value for $p < 0.015$, 7.6. Thus, the addition of the second variable is significant above the 99.9% level and the addition of the third above the 98.5% level. Values for *E* (HOMO) and for the Hammett σ were also included in the regression search but

Table I. Substituted Amphetamines and Predicted Hallucinogenic Activity

No.	Ring position and group					$^3\chi_p$	$^6\chi_p$	$^4\chi_{pc}{}^v$	Exptl[a] log μ	Calcd[a] log μ
	2	3	4	5	6					
1			OCH_3			3.348	1.034	0.469	0.59	0.55
2	OCH_3		OCH_3			4.124	1.508	0.642	0.67	0.87
3	OCH_3			OCH_3		4.124	1.683	0.638	0.87	1.06
4		OCH_3	OCH_3	OCH_3		4.808	1.830	0.739	0.37	0.55
5	OCH_3	OCH_3		OCH_3		4.830	2.083	0.765	0.63	1.01
6	OCH_3		OCH_3		OCH_3	4.853	2.031	0.798	1.03	1.12
7	OCH_3	OCH_3			OCH_3	4.933	2.045	0.810	1.14	1.06
8	OCH_3		OCH_3	OCH_3		4.892	2.058	0.785	1.26	1.00
9	OCH_3	OCH_3	OCH_3	OCH_3		5.629	2.363	0.917	0.86	0.92
10		$-OCH_2O-$				4.203	1.705	0.576	0.41	0.21
11		OCH_3	$-OCH_2O-$			4.925	2.252	0.707	0.43	0.62
12	$-OCH_2O-$		OCH_3			5.043	2.272	0.756	0.48	0.80
13	OCH_3	$-OCH_2O-$				5.027	2.197	0.753	1.00	0.71
14	OCH_3		$-OCH_2O-$			4.993	2.317	0.751	1.08	0.71
15	OCH_3	OCH_3	$-OCH_2O-$			5.746	2.749	0.885	0.75	1.09
16	OCH_3	$-OCH_2O-$		OCH_3		5.761	2.906	0.887	1.13	1.29
17	OCH_3		OC_2H_5	OCH_3		5.027	2.285	0.768	1.22	0.93
18	OCH_3		Br	OCH_3		4.574	1.762	1.157	2.71	2.92
19	OCH_3		CH_3	OCH_3		4.574	1.762	0.888	1.89	1.85
20	OCH_3		C_2H_5	OCH_3		4.892	2.058	0.910	2.01	1.70
21	OCH_3		n-C_3H_7	OCH_3		5.027	2.285	0.880	1.94	1.60
22	OCH_3		n-C_4H_9	OCH_3		5.296	2.436	0.880	1.63	1.34
23	OCH_3		n-C_5H_{11}	OCH_3		5.546	2.548	0.880	1.09	1.09

[a] Molar basis, ref 1 and 5, converted by multiplying by the ratio of molecular weights of amphetamine to mescaline.

Table II. Predictions of Hallucinogenic Potency[a]

No.	2	3	4	5	6	$^3\chi_p$	$^6\chi_p$	$^4\chi_{pc}{}^v$	Exptl μ	Calcd[b] μ
				Amphetamines						
24	OCH_3	OCH_3	OCH_3			4.91	1.78	0.79	2	2-7
25		OCH_3	OCH_3			4.10	1.41	0.61	2	1-4
26	OCH_3	$-OCH_2O-$				4.92	2.25	0.71	1	2-7
27	OC_2H_5		$-OCH_2O-$			5.13	2.45	0.73	7	2.5-8
28	OCH_3		OCH_3	OC_2H_5		5.03	2.30	0.77	7	4.8-15.1
				Mescalines						
29			OCH_3			3.24	0.97	0.37	1	0.01-0.02
30		OCH_3	OCH_3			3.99	1.31	0.57	0.2	0.6-2.0
31		OCH_3	OCH_3	OCH_3		4.70	1.70	0.64	1	0.3-0.8
32	OCH_3		OCH_3	OCH_3		4.78	1.89	0.68	1	0.2-0.7
33	OCH_3	$-OCH_2O-$				4.91	2.03	0.65	5	0.3-1.1
34		OCH_3		$-OCH_2O-$		4.82	2.12	0.60	1	0.2-0.6
				Tryptamines						
35	4-Hydroxy-*N*,*N*-dimethyltryptamine					5.00	2.09	0.92	32	23-75
36	6-Hydroxy-*N*,*N*-dimethyltryptamine					4.93	2.22	0.90	25	26-51

[a] Calculated from the equation. Activity from ref 2. [b] Range of potency computed using the equation standard deviation.

neither showed up in a significant correlation individually or in combination with $^m\chi_t$ terms.

The equation was derived from 18 χ terms; $^0\chi$, $^1\chi$, $^3\chi_p$, $^4\chi_p$, $^6\chi_p$, $^4\chi_{pc}{}^v$, and their squares and reciprocals. The set of variables in the equation is the result of a search through all possible sets of three χ terms, a total of 816 sets. A criticism could be raised at this time that some set of three χ terms from this list may give significant correlation with random numbers. We have carefully analyzed this question. We have run our search program against 200 sets of random numbers, printing out the best correlation found for sets of three χ terms. By this technique, 163 200 sets of three χ terms were correlated against random numbers. Only 4% of the r values for the best three-variable correlations lay in the range of 0.6–0.77 and 0.77 was the highest value found. More detail on this procedure is given in Appendix II. This study clearly establishes that these χ terms are not producing random correlations at a significant level, especially compared to the correlation of the χ terms against the biological activity.

In an additional analysis, the three χ variables were replaced by random numbers and the regression against the activity was carried out. This procedure was repeated 100 times. The average r value obtained is 0.353. No r value exceeded 0.75.

Topliss advanced a caution about producing chance correlations from regression searches, based on correlation of sets of random numbers correlated against random numbers.[11] This present study indicates that χ terms, defined in a specific manner to depend upon molecular structure and not in a random fashion, do not significantly correlate with random numbers. Hence, the potential problem described by Topliss is avoided in this study. The set of χ terms does not correlate significantly with sets of random numbers nor do sets of random numbers, in place of the χ terms, correlate well with the biological activity.

Discussion

The statistical analysis of the relationship reveals that there is unexplained by the equation 20% of the activity (in log μ). This compares well with the 25% variation in the activity from this testing technique, as estimated by Shulgin.[1] Thus a correlation much above 0.9 may have a diminished significance in view of this uncertainty. Even

so, the attainment of a correlation coefficient of 0.92 in this study marks a new high in relating structure with activity among such a large list of hallucinogenic amphetamines. The correlation is high enough and the list sufficiently broad that we have an opportunity to draw some conclusions about the salient structural features contributing to the activity.

We can derive a number of generalizations about the impact of structural variation on activity from the equation. For this discussion, we consider each of the connectivity terms separately and describe their predicted effect on activity due to change in structure. Structural change will usually influence the numerical value of each term. However, we will focus our attention on the most influential terms on activity responding to a structural change. The χ terms are given in Table I.

The greater the number of adjacent ring substituents, the greater the value of $^3\chi_p$. Compare compound **4** with **7** in Table I. This connectivity term is also increased (due to two extra subgraph terms) when a methylenedioxy group replaces a pair of methoxyl groups in the same position. The contribution to the calculated activity is a net reduction. Compare **14** with **8**, **4** with **11**, and **9** with **15**.

Ring substituents longer than two major atoms (as in a methoxyl group) result in an increased value of $^3\chi_p$, hence a lower calculated activity. Compare **17** with **8** and **20** with **21–23**. Longer ring substituents would concurrently increase the $^6\chi_p$ term but sixth-order subgraph values are small relative to $^3\chi_p$ subgraph values; hence the effect on activity is governed primarily by the $^3\chi_p$ index.

Increasing the number of substituents tends to increase the $^3\chi_p$ term contributing to a reduction in the calculated log μ value. Compare **8** with **9** and **3** with **5**. However, this effect is concurrent with an increase in the other two χ terms which tend to increase the calculated activity. The net effect of increasing the number of substituents appears to be a balance resulting in a general prediction of the activity being less sensitive to the number of substituents than to their type and placement. Compare **3**, **5**, and **9** for roughly comparable activities in spite of different numbers of substituents.

The effect of the $^6\chi_p$ term is comparatively modest based on the small coefficient for the term and the low numerical values calculated for $^6\chi_p$ relative to the third- and fourth-order χ terms. Analogues with substituents in the 2 and 6 ring positions, such as methoxyl groups, have an extra $^6\chi_p$ subgraph; hence they contribute to an increased value of log μ. Compare **5** and **4** and also **14** and **11**. Pairs of ring substituents like methoxyl groups positioned para to each other contribute an extra subgraph term of this order; hence, the numerical value of $^6\chi_p$ is increased, elevating the calculated activity. Compare **3** with **2** and **5** with **4**.

More substituents and/or longer substituents have a modest effect of increasing $^6\chi_p$ although the same structural effects increase the $^3\chi_p$ term to a greater extent, offsetting the influence on the calculated activity.

The $^4\chi_{pc}{}^v$ term is increased by additional ring substituents, resulting in an increase in calculated log μ. As pointed out earlier, the number of substituents influences all χ terms leading to a complex effect based more on type and position than number. The $^4\chi_{pc}{}^v$ term is increased considerably when a ring substituent has a δ^v value less than one. Thus compound **18** is very active, reflecting $\delta^v = 0.250$ for the bromine. Substituents longer than two major atoms have no more subgraph terms for $^4\chi_{pc}{}^v$ but one subgraph has a $\delta^v = 2$ rather than $\delta^v = 1$. As a consequence the $^4\chi_{pc}{}^4$ is lower, decreasing the calculated log μ. Compare **21** with **20**, **23** with **22**, and **17** with **8**.

Equation 1 meets two tests at this point in the study. First, it affords a good correlation with the potency, accounting for over 80% of the variation in the data. Secondly, it allows for a considerable amount of SAR interpretation, compatible with the evidence and in agreement with previous empirical observations.[1]

A third test of the significance of the equation is the ability to use it to predict the activity of molecules not used in its derivation. To attempt this, we have considered three sets of molecules: the amphetamines with indefinite values of activity, a number of mescaline derivatives, and molecules outside of the phenylalkylamine class. These three sets are found in Table II along with the computed connectivity terms.

The amphetamines in Table II, entries **24–28**, are molecules with indefinite statements of potency. The computed values, from eq 1, are converted to mescaline units (μ) and expressed as a range, based on the standard deviation of the equation. The agreement must be considered to be reasonably good, especially if it is noted that three of these values (compounds **24–26**) are at the low end of the activity range used in computing eq 1. The predictions of the activity of compounds **27** and **28** are very good.

The second set of molecules predicted in Table II are six mescaline derivatives. A consideration of these compounds represents a departure from the basic structures used in developing the equation. The mescalines are phenethylamines, lacking the α-methyl group of the amphetamines. The predictions are very good. Evidently the absence of the α-methyl group is a deterent to activity as seen in the comparisons of these molecules with the correspondingly substituted amphetamines in Table I. The connectivity descriptions of the mescalines, relative to the amphetamines, reveal lowered values for each connectivity term. The reduction in the $^4\chi_{pc}{}^v$ term, however, is about 20%, thereby constituting the major influence in a negative sense on the calculated activity.

The third set of molecules in Table II represents an attempt to use the equation to predict activities of two molecules which are major departures in structure from the phenylalkylamines. Specifically, we have predicted the activities of two tryptamines from the equation. The predictions are remarkably close to the experimental value in mescaline units, in spite of the fact that no tryptamine derivatives were used in deriving the equation.

Our connectivity analysis of amphetamine SAR considered ring methoxyl, methylenedioxy, and alkyl groups, their number, and position. The absence of these functional groups in compounds **35** and **36** suggests the possibility that a broader interpretation of the connectivity indices is possible in terms of salient structural features contributing to potency. Unfortunately, we are limited in the possibilities of exploring this aspect by the relative paucity of comparable test data and the inherent uncertainties in available data. We can conclude, however, that the molecular connectivity method has transcended structural similarities found in narrow chemical classes.

It should be stated quite clearly that the connectivity indices are purely numerical descriptions of facets of molecular structure under the definition of molecular connectivity. The indices are not physicochemical properties. The interpretation of the equation is truly SAR since structural characteristics are revealed which relate to activity. There is no connotation of a mechanism inherent in the equation or its interpretation. It is infor-

mation which, when analyzed as we have shown, imparts structural information to the medicinal chemist, possibly useful in further synthetic modification.

In conclusion, we can state that a structural description, using three molecular connectivity indices, results in a good correlation with the hallucinogenic potency of a list of amphetamines. These connectivity indices lend themselves to translation into distinct structural characteristics understandable by the organic chemist. Further, these same structural descriptors contained in an equation permit the prediction of potency of molecules both closely related to the set used in deriving the equation and also several molecules which have significant structural differences. These results give some fresh insight into some of the structure–activity relationships of hallucinogenic molecules.

Acknowledgment. This research was supported by a grant from the National Institute on Drug Abuse (DA-01353).

Appendix I

Calculation of Molecular Connectivity Indices. A. Graph. The calculation of the molecular connectivity indices begins with a consideration of the molecule, written down as a skeleton formula without hydrogen atoms. We refer to these drawings as hydrogen-suppressed *graphs*. The atoms are called *vertices*, while the lines joining the vertices, representing formal chemical bonds, are called *edges*.

The graph may be decomposed into uniform parts called *subgraphs*. Subgraphs composed of adjacent two-edge parts are called *paths* of length two. Subgraphs made up of three adjacent and consecutive edges are called subgraphs of path length three. Subgraphs composed of three edges joined at a common vertex, forming a star, are called *cluster* subgraphs. A subgraph composed of four edges, three of which share a common vertex but containing no cycle, is called a path/cluster subgraph. Subgraphs formed from edges adjacent to each other so that a triangle, square, pentagon, or higher enclosed figure is obtained are called *chain* subgraphs.

The number of edges in a subgraph is the *order* of the subgraph. Each subgraph order and type present in a molecular graph contribute to a number called a *molecular connectivity index* of the specified order, m, and type, t: ${}^{m}\chi_{t}$.

A graph may be decomposed into the vertices alone, ignoring the edges. This consideration is of the zero order of the graph. The computed molecular connectivity index of the graph for this type and order is the zero order.

B. Numerical Values Associated with Vertices. Numerical values are associated with each vertex in a graph. There are two prescriptions for these numbers depending upon the level of consideration in the study. This leads to the possibility of two different numbers associated with a particular vertex, depending on the original molecular structure from which the graph is derived. The two levels of consideration are the *connectivity level* and the *valence level*.

1. **Vertex Connectivity Value.** An integer is associated with each vertex in a graph which reflects the number of edges converging at that vertex. This is the connectivity value. The nature of the atom in the molecule giving rise to the graph vertex is not considered. The small Greek letter delta, δ, symbolizes this number.

2. **Vertex Valence Value.** An integer is associated with each vertex in a graph which reflects the nonhydrogenic valence of the atom depicted by that vertex. This is the valence value. It is symbolized by the small Greek letter δ (superscript v), δ^{v}. The valence δ is thus the number of valence electrons in the original atom, Z^{v}, minus the number of suppressed hydrogens at the vertex, h. The expression $\delta^{v} = Z^{v} - h$ leads to the associated numbers in the valence level of molecular connectivity.

3. **Vertex Valence Values for Higher Row Atoms.** The derivation of the valence δ values for atoms beyond the second row in the periodic table leads to the same value for each family member, for example, seven for each halogen atom. Similarly, sulfur would be indistinguishable from oxygen. The relation which leads to appropriate δ^{v} values is $(Z - h)/(Z - Z^{v})$. The δ^{v} values are $\delta_{Cl}{}^{v} = 0.70$ and $\delta_{Br}{}^{v} = 0.25$.

C. Molecular Connectivity Indices. The molecular connectivity indices are computed for each type and order of subgraph into which the molecular graph can be decomposed. Each subgraph order and type retains the δ values for each vertex as found in the molecular graph. The level of molecular connectivity consideration dictates whether δ values or δ^{v} values are used in the computation.

1. **Zero-Order Molecular Connectivity Indices.** The indices of the zero order are computed from each graph by forming the reciprocal square root of each δ (or δ^{v}). These values are designated by the letter c and are called subgraph terms. Thus $c = (\delta)^{-1/2}$ or $c = (\delta^{v})^{-1/2}$. The molecular connectivity index for this order is obtained as a summation of all subgraph c terms from the graph. The index is symbolized by the large Greek letter χ. For the zero order, connectivity level index, the symbol becomes ${}^{0}\chi$. For the valence level, this index is ${}^{0}\chi^{v}$. The summary equation becomes ${}^{0}\chi = \Sigma c_i = \Sigma_{i=1}{}^{n}(\delta_i)^{-1/2}$ for order zero, connectivity level with n vertices.

2. **First-Order Molecular Connectivity Indices.** First-order molecular connectivity indices are calculated from each graph by considering all subgraphs of order one. These are, by definition, all of the edges in the graph. The value of c_{ij} for each edge is computed as the reciprocal square root of the product of the δ values associated with each edge. These subgraph values are then summed for each edge in the molecule. The summarizing expression becomes ${}^{1}\chi = \Sigma c_{ij} = \Sigma(\delta_i\delta_j)^{-1/2}$. For a valence level consideration, the δ^{v} is used.

3. **Second-Order Molecular Connectivity Indices.** Second-order molecular connectivity indices are calculated from each graph by considering all subgraphs of path length two, that is, all adjacent two-edge segments in the graph. The same procedure is used so that ${}^{2}\chi = \Sigma c_{ijk} = \Sigma(\delta_i\,\delta_j\,\delta_k)^{-1/2}$. The use of δ^{v} values leads to a valence level index symbolized as ${}^{2}\chi^{v}$.

4. **Third-Order Molecular Connectivity Indices.** Third-order molecular connectivity indices are computed for each of three different types of subgraphs of order three, when present. These types are paths, clusters, and chains, symbolized by ${}^{3}\chi_{p}$, ${}^{3}\chi_{c}$, and ${}^{3}\chi_{ch}$, respectively. The summary equations for the first two types are ${}^{3}\chi_{p}$ or ${}^{3}\chi_{c} = \Sigma c_{ijkl} = \Sigma(\delta_i\delta_j\delta_k\delta_l)^{-1/2}$. For the chain type index the equation is ${}^{3}\chi_{ch} = \Sigma c_{ijk} = \Sigma(\delta_i\delta_j\delta_k)^{-1/2}$, since only three edges formed by three vertices make up this subgraph. If the valence level of the index is being computed, the δ^{v} is used.

5. **Higher Order Molecular Connectivity Indices.** Fourth and higher order molecular connectivity indices may be calculated depending upon the size and complexity of the molecular graph. Using the symbol m for the order and t for the type, the general symbol for all possible molecular connectivity indices is ${}^{m}\chi_{t}$ or ${}^{m}\chi_{t}{}^{v}$. In fourth and higher order subgraph possibilities, the path/cluster type emerges as a source of another molecular connectivity index. The computation of this index follows the same

Table III. Correlations Involving Random Numbers. Number of R Values[a] in the Range

Range of R values	Best R value for replacement of act.[b]	R value for replacement of χ terms[c]
0.00–0.10	None	None
0.10–0.20	None	7
0.20–0.30	None	18
0.30–0.40	9	26
0.40–0.50	49	25
0.50–0.60	82	19
0.60–0.70	52	3
0.70–0.80	8	2
0.80–0.90	None	None
0.90–1.00	None	None

[a] Correlation coefficient r. [b] Out of 200 runs. [c] Out of 100 runs.

procedure as outlined before. The general equation for computation of a χ index of type t and order m is written as follows

$$^{m}\chi_t = \sum_{j=1}^{^{m}n_s} {}^{m}c_j = \sum_{j=1}^{^{m}n_s}\left[\prod_{i=1}^{m+1}(\delta_i)_j^{-1/2}\right]$$

where $^{m}c_j$ is the subgraph term for mth order subgraphs and $^{m}n_s$ is the number of mth order subgraphs. The symbol $\prod$ stands for "the product of". For chain terms, $^{m}\chi_{ch}$, only m terms are included in the square brackets, instead of $m + 1$, as explained in section 4.

Appendix II

The question of the quality of correlation against sets of random numbers was approached in two ways: (1) the biological data, halucinogenic activity, was replaced by sets of random numbers which were regressed against the χ terms of the regression equation; (2) the χ terms were replaced by sets of random numbers and regressed against the halucinogenic activity.

For the random number study the same computer program was used as in the search for the best set of χ terms. In this program a systematic search is made through all sets of n independent variables where n is selected by the program user. The program RFIND prints out correlation information for any set of n variables for which the correlation coefficient exceeds a predetermined value, RMIN.

For study I 200 sets of random numbers were used in place of the biological data. These numbers were generated by subroutine RANR as supplied with the software for the PRIME 300 computer. Tabulation was made of the highest value of r achieved for each set of random numbers used. For each set 816 triplets of χ terms were examined for correlation. These results are shown in Table III. No correlations approaching the significance of those with the biological data were found.

In study II the three χ terms were replaced by 100 sets of random numbers generated as above. The correlation coefficient r was tabulated for each and is given in Table III. Again no highly significant correlations were found.

References and Notes

(1) A. Shulgin, J. Sargent, and C. Narenjo, *Nature (London)*, 221, 537 (1969).
(2) S. H. Snyder and C. R. Merril, *Proc. Natl. Acad. Sci. U.S.A.*, 54, 259 (1965).
(3) S. Kang and J. P. Green, *Nature (London)*, 226, 645 (1970).
(4) T. DiPaolo, L. B. Kier, and L. H. Hall, *J. Theor. Biol.*, in press.
(5) C. F. Barfknecht, D. E. Nichols, and W. J. Dunn III, *J. Med. Chem.*, 18, 208 (1975).
(6) L. H. Hall, L. B. Kier, and W. J. Murray, *J. Pharm. Sci.*, 64, 1974 (1975).
(7) L. B. Kier, W. J. Murray, and L. H. Hall, *J. Med. Chem.*, 18, 1272 (1975).
(8) T. DiPaolo, L. B. Kier, and L. H. Hall, *Mol. Pharmacol.*, 13, 31 (1977).
(9) L. B. Kier and L. H. Hall, "Molecular Connectivity in Chemistry and Drug Research", Academic Press, New York, N.Y., 1976.
(10) L. B. Kier, L. H. Hall, M. Randič, and W. J. Murray, *J. Pharm. Sci.*, 65, 1226 (1976).
(11) J. G. Topliss and R. J. Costello, *J. Med. Chem.*, 15, 1066 (1972).

Quantitative Structure–Activity Relationships in Centrally Acting Imidazolidines Structurally Related to Clonidine

Pieter B. M. W. M. Timmermans* and Pieter A. van Zwieten

Department of Pharmacy, Division of Pharmacotherapy, University of Amsterdam, Plantage Muidergracht 24, Amsterdam, The Netherlands. *Received May 19, 1977*

The central hypotensive action of clonidine and 26 structurally related derivatives was quantified by means of an ED_{30} obtained from dose-response curves following intravenous administration to anesthetized, normotensive rats. Multiple regression analyses of the biological data yielded correlation equations comprising a relationship between hypotensive activity and molecular structure. In the equations the pharmacokinetics together with the actual engagement of the central α-adrenoceptor are accounted for. More detailed characteristics of this central α-adrenoceptor emerged from correlation studies in which new ED_{30} values, associated with brain concentrations, were employed. The use of this biological parameter at the α-adrenoceptor level allowed the presentation of a hypothetical working model for the mechanism of interaction between this receptive site and clonidine-like imidazolidines.

Clonidine [2-(2,6-dichlorophenylimino)imidazolidine hydrochloride, Figure 1] has been introduced into clinical medicine as an effective antihypertensive drug (Catapresan, Catapres). Its hypotensive action has been explained by a central mechanism. Clonidine is presumed to stimulate central α-adrenoceptors located at medullary sites. This brings about a decrease in peripheral sympathetic tone and an increase in vagal reflex activity (for reviews, see ref 1 and 2).

Since its discovery a number of authors have considered