The ^{13}C Nuclear Magnetic Resonance Substituent Chemical Shifts of Monosubstituted Cyclohexanes. Interpretation by Characteristic Vector Analysis

Romuald I. Zalewski
Department of General Chemistry, Academy of Economics, Poznań, Poland

The set of literature ^{13}C n.m.r. substituent chemical shifts (s.c.s.) of 43 monosubstituted cyclohexanes has been used in Characteristic Vector Analysis (CVA). All variability in the spectra generated by substituents and their configuration could be explained by three orthogonal characteristic vectors. Experimental s.c.s. are thus linear combinations of three independent phenomena acting at each carbon atom simultanously with varying degrees of intensity. The largest contribution, described by $V_{1,n}$ is associated with the polarizing power of the substituent through σ bonds. The contribution to s.c.s. by $V_{2,n}$ is due to a field-inductive effect of the substituent. The contribution associated with $V_{3,n}$ is significant only for C-γ and C-δ and results in an upfield or downfield shift due to the conformation of the substituent.

Correlations between reactivity parameters which apply to chemical equilibria and reaction rates and spectroscopic properties such as ^{13}C n.m.r. substituent chemical shifts (s.c.s.), are well known in the literature.[1] Equilibrium constants or rates of reactions can be satisfactorily predicted by measuring the chemical shift of the derivative in question, provided that the reactivity parameters σ, and the chemical shifts, reflect electronic ground-state perturbations at the nucleus.

In order to get a reasonable correlation, the Hammett equation has often been replaced by two-parameter equations.[1,2] The most frequent models used in n.m.r. spectroscopy are σ_I—σ_R by Taft[3] and Swain and Luptons $\mathscr{F}$—$\mathscr{R}$.[4] The latter was recently re-examined,[5] and improved $\mathscr{F}$ and $\mathscr{R}$ scales were quoted. The aim of these models was to achieve a separation of polar field and mesomeric contributions to the total substituent effects.

The significance of these treatments is in dispute because too much stress has been laid on the correlation coefficient as a criterion of goodness of fit, and not enough on the number of independent mechanisms of the transmission of substituent effects.

Recently a new interpretation of n.m.r. parameters in aromatic compounds was introduced, based on principal-component analysis.[6] It was applied to the analysis of 4-substituted styrenes,[6a,b] 2-substituted indanes,[6c] chalcones,[6d] phenylthienylacrylonitriles,[6e] monosubstituted benzenes,[6f] and triazenes,[6g] and showed that only one or two principal components are required to describe the s.c.s. variations. These principal components were shown to be related to various reactivity parameters of field-inductive and mesomeric character.

However, the analysis of a class of alicyclic compounds of various origins and structures, but with similar principal components, does not seem to have been attempted. Further development of ^{13}C n.m.r. requires a better theoretical understanding of the shielding mechanism[7] and a consistent picture of the effects of changes in substituents, and conformation, on chemical shift.[8] The latter approach has been begun by Dalling and Grant.[9] Interpretation of the significance of postulated relations and application of derived rules to unknown molecules of similar geometry should be based on rigorous standards of multiple regression analysis.[1a]

This paper is addressed to the problem of interpretation of substituent and conformation effects on the ^{13}C n.m.r. spectra of monosubstituted cyclohexanes using the concept of Characteristic Vector Analysis (CVA).

Mathematical Foundation of Characteristic Vector Analysis.[10]—A data set of measurements of n individuals with p variables is contained in the data matrix $\mathbf{D}$ ($n \times p$). It is useful to transform this data set into a new one in which the variables are mutually uncorrelated, and to know the number of linearly independent sources of variation in the system which represent the chosen variables. This method allows the orthogonalization of a set of measurements, and enables us to reach an 'economical' description.

CVA thus provides a method of constructing new variables which are pairwise uncorrelated. Each characteristic vector is a linear combination of the observed variables, and these linear functions are chosen to be orthogonal. Thus, the characteristic vectors are composites of the original variables. The number of characteristic vectors existing in characteristic vector space is equal to, or less than, the number of variables in the data set, p. The first characteristic vector is defined as that linear combination of variables which has the maximum variance of all linear functions derivable from the given variables. The second characteristic vector is the linear combination of variables having the maximum variance of all linear functions of the given variables that are orthogonal to the first characteristic vector, and so on. The coefficients of the characteristic vectors are called characteristic vector loadings, contained in matrix $\mathbf{L}$ ($k \times p$), while the magnitudes of characteristic vectors upon each of the individuals are called CV scores, contained in matrix $\mathbf{V}$ ($n \times k$). k Denotes the number of orthogonal characteristic vectors, and usually $k \ll p$.

The transformation of experimental data d_i into V_i, which have the properties of mutual orthogonality and maximum variance, is achieved by defining V_i as the linear combination of d_i [equation (1)] or in matrix notation [equation (2)].

$$V_{1,i} = a_{1,1}d_{1,i} + a_{1,2}d_{2,i} + \cdots\cdots + a_{p,1}d_{p,i} \quad (i = 1, 2, \cdots\cdot n) \qquad (1)$$

$$V_{1,i} = \mathbf{D}a_1 \qquad (2)$$

V_1 is a vector of n elements (first row in $\mathbf{V}$) and a_1 is a vector of p elements (first column of eigenvector matrix $\mathbf{A}$ calculated from variance–covariance matrix).

The sum of squares of the elements of V_1 is given by equation (3) with a_1 chosen so as to maximize (3), under the constraint (4).

$$\sum_{i=1}^{n} v_i^2 = v_i v'_i = a_1\mathbf{D}'\mathbf{D}a_1 \qquad (3)$$

$$a_1 a'_1 = 1 \qquad (4)$$

It is true under these conditions that a_1 is a unit-length eigenvector corresponding to the largest eigenvalue in $\mathbf{D'D}$.[11] Vectors $a_2 \cdots a_p$ are defined in a similar way.

The variances of the V_i can be found from (5). After

$$V'V = \mathbf{A'D'DA} \qquad (5)$$

mathematical transformations, and using diagonal matrix $\mathbf{Q}$ of eigenvalues arranged in descending order of magnitude, we finally obtained equation (6).

$$V'V = \mathbf{IDI} = \mathbf{Q} \qquad (6)$$

Equation (6) shows that the variances of v_i are numerically equal to the elements q_i of diagonal matrix $\mathbf{Q}$ of eigenvalues in $\mathbf{D'D}$. The ratio of q_i elements to the trace of matrix $\mathbf{Q}$, *i.e.* equation (7), describes the percent of total data variability explained by the *i*th characteristic vector.

$$\frac{q_i}{Tr\mathbf{Q}} \times 100\% \qquad (7)$$

The elements of the variance–covariance matrix $\mathbf{A}$ can be converted into elements of characteristic vector loading matrix $\mathbf{L}$.

The transformation of experimental data d_i into v_i can be explained in geometrical terms, provided that only three properties are considered. If each d_i is treated as a geometrical vector, then the rows in data matrix $\mathbf{D}$ are the co-ordinates of n individuals, which can be visualized in Figure 1. It can be seen that the points tend to fall in a single plane. The position of an individual n needs only two co-ordinates rather than the usual three for its description. It is usual to define the new axes such that the first axis is aligned with the longest data dimension, and the second at right angles to it. Thus the projections of the points onto axes f_1 and f_2 are the elements of characteristic vectors V_1 and V_2 (CV scores) for these points. In Figure 1, cos α is the loading of the variable represented by the b_1 axis onto the characteristic vector represented by orthogonal axis f_1.

A more detailed account of the geometrical model is given by Morrison[11b] and Davies.[12]

Choice of Data (*Chemical Shifts Data Matrix*).—The ^{13}C n.m.r. spectral data of 43 monosubstituted cyclohexanes were selected from the literature.[13–17] The spectra were recorded in almost uniform conditions, $CFCl_3$ being the solvent. The temperature was kept low (180 K) to provide chemical shifts for axial and equatorial conformers. Only 10 of the compounds reported in ref. 17 were measured at room temperature, and the conformation ratio was not reported. Among monosubstituted cyclohexanes were 17 pairs of axial–equatorial conformers. Substituents available cover the wide range of chemical properties, as shown in Table 1.

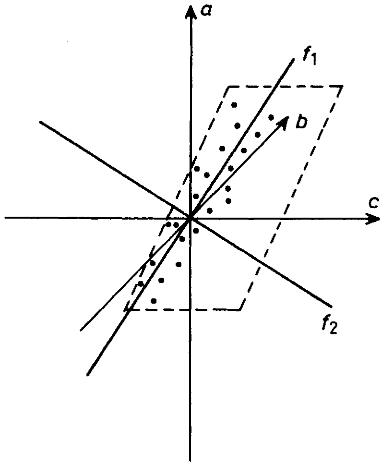

Figure 1. Geometrical explanation of CVA original axes *a*, *b*, *c* could be replaced by characteristic vector axes f_1 and f_2

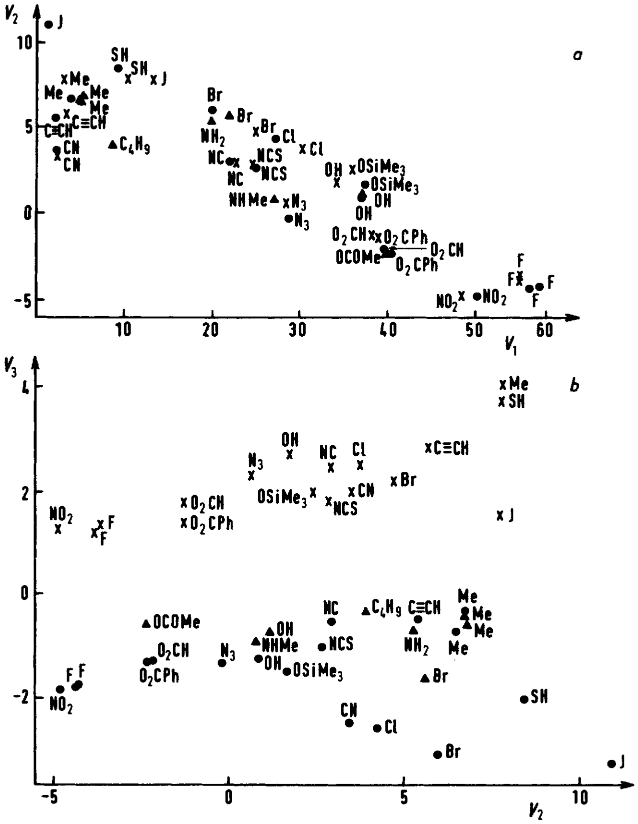

Figure 2. *a*, plot of V_2 *verus* V_1 and *b*, plot of V_3 *versus* V_2: ×, axial conformers; ●, equatorial conformers; ▲, undeclared conformation

CVA Processing of Chemical Shift Data Matrix.—The initial data matrix of the chemical shifts of 43 monosubstituted cyclohexanes for six carbon atoms was constructed (43 × 6). Since the α-C(2) and C(6) and β-C(3) and C(5) carbon atoms were equivalent, the matrix format was then reduced to four chemical shifts per one compound, *i.e.* C(1), C(2), C(3), and C(4). CVA calculations were performed, and each element of the data matrix could be represented by equation (8) in which $\varepsilon_{n,p}$

$$d_{n,p} = \bar{d}_n + V_{1,n}l_{1,p} + V_{2,n}l_{2,p} + V_{3,n}l_{3,p} + \varepsilon_{n,p} \qquad (8)$$

denotes the residue unexplained by the CVA model at a particular number of the characteristic vector; other symbols were explained previously. The variances of $\varepsilon_{n,p}$ were calculated for each column and each row of the reconstituted data matrix with k = 1, 2, and 3 characteristic vectors. The magnitude of $w_{\varepsilon,k} \leqslant (\text{S.D.})^2$ indicates a satisfactory fit of the experimental data to the CVA model. The plots of $V_{2,n}$ *versus* $V_{1,n}$ and $V_{3,n}$ *versus* $V_{2,n}$ were printed in addition to matrices of characteristic

Table 1. V_1, V_2, and V_3 characteristic vector scores of equatorial (1—26) and axial (27—43) monosubstituted cyclohexanes

Subst.	No.	$\bar{d}_n$	$V_{1,n}$	$V_{2,n}$	$V^x_{3,n}$	Ref.	No.	$\bar{d}_n$	$V_{1,n}$	$V_{2,n}$	$V^x_{3,n}$	Ref.
CN	1	26.8225	2.2397	3.4763	−0.2453	13	27	25.5525	2.1931	3.4345	1.9910	13
C≡CH	2	28.1075	2.1852	5.3955	−0.5043	13	28	26.5975	3.2592	5.6741	2.8209	13
Me	3	30.7650	4.4712	6.8119	−0.5906	16*a*						
	4	30.4500	4.3504	6.5808	−0.4495	16*b*						
	5	31.3000	3.8251	6.5872	−0.3261	15	29	27.5950	3.1149	7.7740	4.0520	15
	6	30.9675	4.3783	6.4553	−0.7468	14						
C_4H_9	7	31.2500	8.6411	3.9126	−0.3463	17						
NC	8	34.1475	21.9247	2.8961	−0.5491	13	30	31.9100	22.6847	2.9198	2.4494	13
NCS	9	35.0200	24.8912	2.6049	−1.0319	13	31	32.7575	24.6522	2.8616	1.7975	13
N_3	10	35.3525	28.8402	−0.0451	−1.3226	13	32	33.1750	28.4221	0.6345	2.2928	13
O_2CH	11	38.3600	39.5662	−2.2243	−1.3181	13	33	36.0675	38.3230	−1.2295	1.7774	13
O_2CPh	12	38.6350	40.2161	−2.3022	−1.3377	13	34	36.1975	38.7582	−1.3132	1.3930	13
NO_2	13	41.5700	50.2982	−4.7986	−1.8531	14	35	39.5025	48.5306	−4.8337	1.2623	14
SH	14	32.0575	9.2236	8.400	−2.0473	13	36	28.9200	10.2679	7.7723	3.7250	13
I	15	31.9125	1.2359	10.8907	−3.2794	15	37	31.3150	10.3286	7.7305	1.5458	15
Br	16	35.3750	21.8926	5.5752	−1.6557	17						
	17	36.7050	19.9213	5.9482	−3.0989	15	38	35.0550	25.4115	4.6720	2.1957	15
NH_2	18	34.4750	19.8617	5.3167	−0.7082	17						
NHMe	19	35.8250	27.0133	0.7656	−0.9283	17						
Cl	20	37.8675	27.2603	4.2361	−2.5995	15	39	35.6075	30.2001	3.7377	2.4864	15
OH	21	38.8750	36.9344	1.1545	−0.7593	17						
	22	39.6050	36.1001	0.8411	−1.2565	14	40	36.3900	34.1961	1.7532	2.7362	14
OCOMe	23	38.4000	39.7319	−2.3241	−0.6046	17						
$OSiMe_3$	24	39.4300	37.3512	1.6569	−1.5372	13	41	36.3925	35.9431	2.4225	1.9821	13
F	25	43.4350	56.2483	−4.3556	−1.8000	13	42	41.1350	55.0299	−3.6416	1.3082	13
	26	43.4700	56.5710	−4.3050	−1.7775	15	43	41.2075	55.0731	−3.8192	1.2061	15

Table 2. Main results of CVA calculation for classes A + E, A, and E

				Residual variance in one- and two-component model[a]			
Class	Compd.	TV_1 (%)	TV_2[a] (%)	C(1)	C(2)	C(3)	C(4)
A + E	1—43	97.41	99.64	0.582 165 0.003 538	15.710 210 0.003 002	4.593 660 1.578 710	4.028 676 1.878 945
E	1—26	97.24	99.95	0.708 833 0.000 278	17.622 338 0.003 762	1.968 183 0.240 155	4.379 429 0.203 821
A	27—43	98.05	99.95	0.578 151 0.000 126	12.658 196 0.052 697	6.959 858 0.120 950	0.377 059 0.346 429

[a] The inclusion of V_3 yields 100% TV and 0.000 000 residual variance.

vector scores V, characteristic vector loadings L, and reconstituted data matrices with V_1, V_2, and V_3.

Results and Discussion

The initial data matrix of the ^{13}C chemical shifts of 43 monosubstituted cyclohexanes (equatorial, axial, and mixed conformation) was used in CVA calculations. The results show that three characteristic vectors were significant, explaining 97.41, 2.23, and 0.36% of the total data variance, respectively. The unexplained residue variance, $w_{\varepsilon,2}$ = 0.003 538 and 0.003 002 for C(1) and C(2) respectively, was much lower than the squared standard deviation of chemical shift measurements (S.D.).[2] This indicates that only two characteristic vectors are required to describe the chemical shift variance of C(1) and C(2) in all 43 compounds. On the other hand, $w_{\varepsilon,2} > 1.55$ for C(3) and C(4) is not acceptable and calls for a third characteristic vector.

In Figures 2a and b we show two planar projections of the three-dimensional space. Visually, (Figure 2b) all compounds seem to fall into two clusters, A-axial with positive $V_{3,n}$ and E-equatorial, including 10 compounds of mixed conformation with negative $V_{3,n}$. In addition, Figure 2a shows the tendency to a linear pattern of negative slope between V_2 and V_1. The simple ordination procedures[18] applied to points of Figure 2a allow us to distinguish four classes of substituents. The division was done according to the first atom in the substituent of the cyclohexane ring, *i.e.* C, N, O, or halogen.

However the analysis of data for individual clusters does not provide a significantly better CVA model. The best results were achieved when all compounds were divided into axial cyclohexanes (class A) and equatorial cyclohexanes (class E). The main results of CVA calculations for all monosubstituted cyclohexanes (class A + E) and for classes A or E are reported in Table 2.

% TV is higher by 0.3 and residual variances are significantly lower, especially for carbon atoms C(3) and C(4) in classes A nd E as compared with one class, A + E. However, the third characteristic vector must be included into the model to lower the residual variance. CV scores for a particular substituent are practically independent of class for V_1; V_2 are more dependent on class and V_2 are very strongly dependent. In Table 1, the values of V_1, V_2, and V_3 are reported for class A and E compounds.

It can be seen from Figure 2 and the data in Table 1 that values of $V_{1,n}$ are distributed over a range of 60 units, and values of $V_{2,n}$ over a range of 17 units. In addition the tendency to linear dependence of negative slope can be observed. The

Table 3. Data of characteristic vector loadings for axial (A) and equatorial (E) cyclohexanes

		C(1)	C(2)	C(3)	C(4)
L_1	E	0.848 610 3	−0.120 075 3	−0.369 786 1	−0.358 748 9
	A	0.849 077 2	−0.141 746 8	−0.420 063 7	−0.287 266 8
L_2	E	−0.171 024 1	0.852 816 5	−0.266 618 5	−0.415 173 5
	A	−0.169 778 3	0.792 847 6	−0.583 986 8	−0.039 082 3
L_3	E	−0.024 725 7	0.051 027 2	−0.736 323 8	0.670 022 3
	A	−0.015 591 7	−0.318 277 5	−0.482 188 6	0.816 037 8

magnitude of $V_{1,n}$ increases when the electronegativity of the substituents increases, as a general rule. For example I^- of small electronegativity has a very small V_1 value, and F^- of high electronegativity indicates the highest V_1 value. On the other hand, the magnitude of $V_{2,n}$ decreases as the electronegativity of the substituent increases. Closer examination of the relation between $V_{1,n}$ or $V_{2,n}$ and electronegativity[19] indicates the random scattering of points. Only a few substituents, including halogens, amino groups, and hydroxy groups, form good straight lines. Alkyls, C≡CH, NHMe, OAc, NO_2, and NCS produce very large positive or negative deviations and for the other 5 substituents no data on electronegativity were available. The other substituent parameters reflecting inductive properties (*e.g.* σ_I, $\mathcal{F}$) gave even worse results.

Hence characteristic vectors are orthogonal, they represent unrelated phenomena. The explanation of their chemical background must be based on observed relations to electronegativity and the magnitude of characteristic vector loading presented in Table 3, and could be explained by existing theory.

Characteristic vector loadings for V_1 are positive and very large for C(1) but small and negative for all other carbon atoms. We believe that $V_{1,n}$ describes the polarization power of substituents through the σ bond in axial and equatorial cyclohexanes (and perhaps in other alicyclic systems) as visualized by ^{13}C n.m.r. Electronegative substituent-attracting electrons from carbon induce a partial positive charge at C(1) and then induce partial negative charges at C(2), C(3), and C(4) by σ-inductive nonalternant effect.[20,21] Thus CV loadings represent the magnitude and sign of those partial charges.

CV loadings for V_2 are positive and very large for C(2) but negative for all other carbons, C(3) being more negative. These facts correlate with the relationship between the magnitude of V_2 and electronegativity, showing that the high sensitivity associated with the first phenomenon is accompanied by weak activity of the second phenomenon, and *vice versa*. In our opinion, $V_{2,n}$ describes the inductive power of the substituent by field effect[22] as visualized by ^{13}C n.m.r. Field effects induce a positive charge at C(2) through space, and perturb σ-orbitals by a σ-inductive nonalternatant effect, inducing negative charges at C(3) and C(4). Field effects depend on the dipole moment of the substituent[22] and could be approximated by σ_I,[3] $\mathcal{F}$,[4,5–23] or other similar scale in many applications. However the field effect upon nuclear magnetic shielding depends upon the dipole moment, but with different geometric terms,[24] and the relationship between the above scales and $V_{2,n}$ is random. We believe that nonproportionality between V_1 and V_2 is caused by the nonproportional nature of field and inductive effects,[24a,d] especially for polyatomic substituents.

It is difficult to explain the nature of the third mechanism which contributes to the overall shift effect by moderate or weak shifts of the γ and δ carbons, depending on substituent and conformation.

Schneider and Hoppen[13] reported that electronegative substituents shield C-δ axial groups in all cases to a significantly lesser degree (1—2 p.p.m.). CNDO- and INDO-calculated charge densities fail to reproduce this shift. By contrast, frontier orbital treatment[25] qualitatively predicts an accumulation of electron density at C-δ by electronegative substituent. Recent force-field calculations[26] indicate no significant geometry changes upon equatorial halogen, alkoxy, and alkyl substituents in cyclohexane, but axial substituents cause a change in torsional angles.[27]

A specific role for axial substituents has been proposed in the investigation of C-γ shifts in γ-gauche conformations of hydrocarbons by sterically induced charge polarization of the C–H bond.[28] The coupled bond angle–torsional angle gauche effect theory[29] seems to be applicable in the compounds studied. Assuming that CV loadings L_3 will characterize the sign and relative magnitude of charge in the molecule we come to agreement with the sterically induced charge polarization concept.

We can see from the data in Table 3 the accumulation of negative CV loadings at C(3) and the accumulation of negative loadings weakened by positive CV loadings induced by the third mechanism at C(4).

Of key importance for a contribution by a third mechanism, however, is the sign and magnitude of the third characteristic vector. V_3 for axial conformers is always negative for equatorial conformers. V_3 determines the upfield or downfield shift at C(3) and C(4) associated with the third mechanism.

Since the sign of V_3 depends on the orientation of the substituent, we can imagine and examine in future a possible link between the third mechanism and the field gradient shift.[24b]

We have to keep in mind that the concept of substituent chemical shift presented now differs basically from the previous one.[13] Now we treat s.c.s. as a linear combination of three independent contributions at each carbon atom described by the general equation (8). The parameters of this equation for various substituents are presented in Table 1, and for carbon atoms in Table 3. The general applicability of these parameters to other alicyclic systems will be discussed in another paper.

References

1 (*a*) J. Shorter, 'Correlation Analysis of Organic Reactivity,' Research Studies Press, Wiley, New York, 1982; (*b*) D. F. Ewing, in 'Correlation Analysis in Chemistry; Recent Advances,' eds. J. Shorter and N. B. Chapman, Plenum Press, New York, 1978, Ch. 8; (*c*) S. Ehrenson, R. T. C. Brownlee, and R. W. Taft, *Prog. Phys. Org. Chem.*, 1973, **10**, 1; R. D. Topsom, *ibid.*, 1976, **12**, 1; (*d*) O. Exner, in 'Advances in Linear Free Energy Relationships,' ed. N. B. Chapman and J. Shorter, Plenum Press, London, 1972, ch. 1.

2 J. Shorter in 'Correlation Analysis in Chemistry; Recent Advances,' ed. J. Shorter and N. B. Chapman, Plenum Press, New York, 1978, ch. 4, p. 119.

3 R. W. Taft, *J. Am. Chem. Soc.*, 1953, **75**, 4231; R. W. Taft and I. C. Lewis, *ibid.*, 1958, **80**, 2436.

4 C. G. Swain and E. C. Lupton, *J. Am. Chem. Soc.*, 1968, **90**, 4328.

5 C. G. Swain, S. H. Unger, N. R. Rosenquist, and M. S. Swain, *J. Am. Chem. Soc.*, 1983, **105**, 492.

6 (*a*) U. Edlund and S. Wold, *J. Magn. Reson.*, 1980, **37**, 183; (*b*) M. Sjöström and U. Edlund, *ibid.*, 1977, **13**, 285; (*c*) B. Eliasson and U. Edlund, *J. Chem. Soc.*, *Perkin Trans. 2*, 1981, 403; (*d*) G. Mussumara, S. Wold, and S. Gronowitz, *Org. Magn. Reson.*, 1981, **17**, 118; (*e*) G. Mussumara, G. Scarlatta, and S. Wold, *Gazz. Chim. Ital.*, 1981, **111**, 499; (*f*) D. Johnels, U. Edlund, H. Grahm, S. Hellberg, M. Sjöström, S. Wold, S. Clementi, and W. J. Dunn III, *J. Chem. Soc.*, *Perkin Trans. 2*, 1983, 863; (*g*) W. J. Dunn III, C. Lins, G. Kumar, T. Monimaran, S. Grigoras, U. Edlund, and S. Wold, *Org. Magn. Reson.*, 1983, **21**, 450.

7 G. L. Nelson and E. A. Williams, *Prog. Org. Chem.*, 1976, **12**, 229; K. A. K. Ebraheem and G. A. Webb, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1977, **11**, 149.

8 G. E. Maciel, *Top. Carbon-13 NMR Spectrosc.*, 1974, **1**, 1.

9 D. K. Dalling and D. M. Grant, *J. Am. Chem. Soc.*, 1974, **96**, 1827.

10 T. W. Anderson, 'Introduction to Multivariate Analysis,' Wiley, New York, 1958; M. G. Kendall, 'A Course in Multivariate Analysis,' Griffin, London 1965; S. J. Press, 'Applied Multivariate Analysis,' Holt, Rinehard and Winston, New York, 1973; P. M. Mather, 'Computational Methods of Multivariate Analysis in Physical Geography,' Wiley, New York, 1976; J. L. Simonds, *J. Opt. Soc. Am.*, 1963, **53**, 968; H. Woud, Ph.D Thesis, McGill Univ., Montreal, 1974; E. R. Malinowski and D. G. Hovery, 'Factor Analysis in Chemistry,' Wiley–Interscience, New York, 1980.

11 (*a*) J. Johnston, 'Econometric Methods,' 2nd edn., McGraw-Hill, New York, 1972; (*b*) D. F. Morrison, 'Multivariate Statistical Methods,' McGraw-Hill, New York, 1967.

12 J. C. Davis, 'Statistics and Data Analysis in Geology with Fortran Programs,' Wiley, New York, 1973.

13 H. J. Schneider and V. Hoppen, *J. Org. Chem.*, 1978, **43**, 3866.

14 O. A. Subbotin and W. M. Sergeyev, *J. Chem. Soc.*, *Chem. Commun.*, 1976, 141.

15 O. A. Subbotin and W. M. Sergeyev, *J. Am. Chem. Soc.*, 1975, **97**, 1080.

16 (*a*) D. K. Dalling and D. M. Grant, *J. Am. Chem. Soc.*, 1967, **89**, 6612; (*b*) H. Jancke and H. Werner, *J. prakt. Chem.*, 1980, **322**, 247.

17 Le Roy, F. Johnson, and W. C. Jankowski, 'Carbon-13 NMR Spectra. A Collection of Assigned, Coded and Indexed Spectra,' Wiley, New York, 1972.

18 J. C. Gover, *The Statistician*, 1967, **17**, 13; P. M. Mather, 'Computional Methods of Multivariate Analysis in Physical Geography,' Wiley, New York, 1976, pp. 329—331.

19 J. E. Huheey, *J. Phys. Chem.*, 1965, **69**, 3284; 1966, **70**, 2086.

20 E. T. McBee, I. Serfaty, and T. Hidgins, *J. Am. Chem. Soc.*, 1971, **93**, 5711.

21 J. A. Pople and M. Gordon, *J. Am. Chem. Soc.*, 1967, **89**, 4253.

22 N. Bjerum, *Z. Phys. Chem.*, 1923, **106**, 219; F. H. Westheimer and J. G. Kirkwood, *J. Chem. Phys.*, 1938, **6**, 513.

23 P. G. Mezey and W. F. Reynolds, *Can. J. Chem.*, 1977, **55**, 1567.

24 (*a*) W. F. Reynolds, *J. Chem. Soc.*, *Perkin Trans. 2*, 1980, 985; (*b*) J. G. Batchelor, J. Feeney, and G. C. K. Roberts, *J. Magn. Reson.*, 1975, **20**, 19; (*c*) A. D. Buckingham, *Can. J. Chem.*, 1960, **38**, 300; (*d*) W. F. Reynolds, *Prog. Phys. Org. Chem.*, 1983, **14**, 165.

25 N. S. Zefirow, *Tetrahedron Lett.*, 1975, 1087.

26 N. D. Allinger, *Adv. Phys. Org. Chem.*, 1976, **13**, 1.

27 H. J. Schneider and W. Gschwendtner, *J. Org. Chem.*, 1982, **47**, 4216.

28 D. M. Grant and B. V. Chaney, *J. Am. Chem. Soc.*, 1967, **89**, 5315; H. J. Schneider and E. P. Weingand, *ibid.*, 1977, **99**, 8362.

29 D. G. Gorenstein, *J. Am. Chem. Soc.*, 1977, **99**, 2252.

Received 18*th January* 1985; *Paper* 5/100