

APPLICATION OF FACTOR ANALYSIS TO THE
STUDY OF ^{13}C NMR CHEMICAL SHIFTS.¹

Kenneth B. Wiberg and William E. Pratt
Department of Chemistry, Yale University, New Haven, Conn. 06520

William F. Bailey
Department of Chemistry, University of Connecticut, Storrs, Conn. 06268

Although the subject of numerous investigations, the origins of ^{13}C nmr chemical shifts are still not well understood.² In order to simplify the problem, we have examined a wide variety of halogen substituted (F,Cl,Br,I) compounds. The similarity of the C-X bond dipoles³ should minimize contributions from field effect at reasonable distance from the site of substitution. The data were obtained in 2M chloroform-d solutions at 67.88 MHz.

In analyzing the data, we wished to use an unbiased method of determining the number of independent factors involved in producing the observed chemical shifts. For this reason, we have employed factor analysis⁴ using the difference in chemical shift between the substituted compound and the unsubstituted parent hydrocarbon (differential chemical shifts) as the input data. In factor analysis, the $m \times n$ data matrix \underline{D} (where m is the number of positions and $n = 4$, the number of substituents used) is multiplied by its transpose, \underline{D}^\dagger , to give a square correlation matrix, \underline{C} . Diagonalization of \underline{C} gives an eigenvector matrix \underline{B} and the eigenvector array λ :

$$\begin{aligned}\underline{D}^\dagger \underline{D} &= \underline{C} \\ \underline{B}^{-1} \underline{C} \underline{B} &= \lambda\end{aligned}$$

Since the eigenvectors are orthogonal:

$$\underline{B}^{-1} \underline{C} \underline{B} = \underline{B}^{-1} \underline{D}^\dagger \underline{D} \underline{B} = \underline{B}^\dagger \underline{D} \underline{D} \underline{B} = \underline{H}^\dagger \underline{H} = \lambda$$

where $\underline{D} \underline{B} = \underline{H}$. Now, the data matrix may be written as $\underline{D} = \underline{B}^{-1} \underline{H}$ or $\underline{D} = \underline{M} \underline{H}$ where $\underline{M} = \underline{B}^{-1}$. In this way, the data matrix has been factored into two new matrices \underline{M} and \underline{H} . The matrix \underline{M} may be associated with the molecular positions, and the matrix \underline{H} , may be associated with the substituents.

The eigenvalues, λ , indicate the relative importance of the columns in \underline{M} . An analysis of the data for 62 sets of halogen induced chemical shifts indicated that there were two major factors having $\lambda = 34$ and 7, as well as a smaller third factor, $\lambda = 0.3$. The other λ 's were vanishingly small. Using these three factors, the differential chemical shifts may be written as

$$\Delta\delta_x = h_{1x}m_1 + h_{2x}m_2 + h_{3x}m_3$$

where x indicates the halogen used, h_{kx} are the elements of the \underline{H} matrix and m_k are the elements of the \underline{M} matrix. With these three terms, the experimental data are reproduced with an rms error of 0.42 ppm. The largest deviations were observed with the α -positions. When the α -positions were treated separately, they gave slightly different \underline{M} and \underline{H} matrices, and the

use of 3 factors gave an rms error of 0.13 ppm. The β, γ, δ and ϵ positions when treated as a group gave an rms error of 0.10 ppm.

An examination of the halogen property matrices suggested that the two principal halogen property vectors were approximately a constant term and an arithmetic progression (1,2,3,4). The test vector, 1,0,0,1, was found to be satisfactory as the third term. The original \underline{H} matrix was rotated to give the "best" fit to the test vectors. The results are summarized in Table I. It can be seen that the vectors for the α -positions are quite similar to those for the other positions.

Table I

| Halogen Parameters from Factor Analysis | | | | | | | | | |
|---|--------------|-----|-----|--------------------|-------|--------|-------------------------|-------|--------|
| Halogen X = | Test Vectors | | | Calculated Vectors | | | | | |
| | | | | α | | | β, γ, δ | | |
| | | | | a_1 | a_2 | a_3 | a_1 | a_2 | a_3 |
| F | 1.0 | 1.0 | 1.0 | 1.000 | 1.000 | 0.999 | 1.000 | 0.989 | 1.008 |
| Cl | 1.0 | 2.0 | 0.0 | 0.951 | 2.001 | -0.169 | 0.997 | 2.160 | -0.110 |
| Br | 1.0 | 3.0 | 0.0 | 1.068 | 2.998 | 0.235 | 1.004 | 2.791 | 0.143 |
| I | 1.0 | 4.0 | 1.0 | 0.974 | 4.001 | 0.909 | 0.999 | 4.061 | 0.958 |

Table II

| Molecular Parameters from Factor Analysis | | | | | |
|---|----------|-------|--------|-------|----------------|
| Molecule | Position | b_1 | b_2 | b_3 | Rms Error, ppm |
| Ethyl | α | 91.51 | -26.19 | 8.21 | 0.10 |
| 1-Propyl | α | 84.70 | -24.43 | 8.63 | 0.07 |
| 1-Butyl | α | 86.08 | -24.77 | 8.80 | 0.08 |
| 2-Propyl | α | 85.18 | -21.25 | 7.26 | 0.20 |
| 3-Pentyl | α | 71.10 | -16.68 | 7.99 | 0.06 |
| t-Butyl | α | 80.30 | -15.99 | 5.91 | 0.19 |
| 3-Ethyl-3-Pentyl | α | 58.26 | - 8.49 | 7.18 | 0.02 |
| Cyclohexyl | α | 75.38 | -18.73 | 7.41 | 0.00 |
| 1-Bicyclo[2.2.2]octyl | α | 77.92 | -14.64 | 6.65 | 0.01 |
| Ethyl | β | 8.79 | 1.46 | -0.91 | 0.11 |
| 1-Propyl | β | 7.44 | 1.01 | -0.93 | 0.11 |
| 2-Propyl | β | 4.68 | 2.72 | -1.06 | 0.08 |
| t-Butyl | β | 1.09 | 3.90 | -1.20 | 0.06 |
| 1-Butyl | β | 7.51 | 0.94 | -0.86 | 0.12 |
| 3-Pentyl | β | 4.52 | 1.82 | -1.09 | 0.09 |
| 3-Ethyl-3-Pentyl | β | 1.74 | 2.64 | -1.02 | 0.06 |
| Cyclohexyl | β | 4.36 | 2.29 | -1.36 | 0.11 |
| 1-Bicyclo[2.2.2]octyl | β | 3.44 | 2.97 | -1.26 | 0.09 |
| 1-Propyl | γ | -9.12 | 1.95 | -0.08 | 0.07 |
| 1-Butyl | γ | -8.51 | 1.68 | 0.32 | 0.08 |
| 3-Pentyl | γ | -6.53 | 1.59 | 0.14 | 0.06 |
| 3-Ethyl-3-Pentyl | γ | -5.10 | 1.30 | 0.29 | 0.07 |
| Cyclohexyl | γ | -5.33 | 1.42 | -0.34 | 0.06 |
| 1-Bicyclo[2.2.2]octyl | γ | 0.47 | 0.81 | -0.06 | 0.14 |

After rotation of the halogen property matrix to give the "best" fit to the test vectors, a corresponding rotation of the molecular parameters gives a set of vectors which may be used to study the structural dependence of the chemical shifts. The vectors for a group of

related compounds are given in Table II. The calculated differential chemical shifts are given by

$$\Delta\delta_x = a_{1x}b_1 + a_{2x}b_2 + a_{3x}b_3$$

where the a_{kx} are the elements of the rotated \underline{H} matrix and the b_k are the elements of the rotated \underline{M} matrix.

Since the a_{1x} terms are essentially unity, the b_1 's correspond to intercepts and the b_2 's correspond to slopes of plots of $\Delta\delta_x$ against the a_{2x} terms. The b_3 terms represent curvatures in these plots. The first molecular factor (b_1) may be considered as the chemical shift difference between a hydrogen as a substituent and a hypothetical halogen having the same principal quantum number. The a_{2x} terms are linearly related to a variety of experimental quantities (C-X bond polarizabilities, C-X bond ionization potentials, first electronic transition of the methyl halides) which are related to the "freeness" of the valence electrons. The second molecular parameter, b_2 , gives the sensitivity of the given position in the molecule to this halogen parameter. The third parameter is more difficult to interpret. At the α -carbon, the b_3 values are positive and are close to 7. With the β -carbons, the b_3 values have the opposite sign, and again have similar magnitudes (-1.08 ± 0.13). The γ -position leads to a range of b_3 values, both positive and negative. The values appear to be related to conformation.

One of the principal conclusions derived from this study may be seen by examining the results for the series:

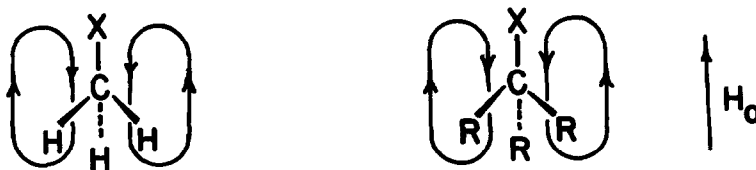
ethyl, 2-propyl, t-butyl

1-propyl, 3-pentyl, 3-ethyl-3-pentyl

1-propyl, cyclohexyl, 1-bicyclo[2.2.2]octyl

The first group represents increasing methyl substitution, and the second is formed by increasing ethyl substitution. The third group is similar to the second, except that the alkyl groups are "tied back" by ring formation. In each case, increasing alkyl substitution leads to a decrease in intercept for the α -position as might be expected for the inductive effect of the alkyl groups.

The slopes (b_2) at the α -position are of particular interest. In each series, increasing alkyl substitution leads to a decrease in slope. The effect is clearly related to the bulk of the alkyl group since three ethyl groups lead to a much larger effect than three methyl groups. This observation may be explained as follows. Consider the systems:



If X is an iodine, the applied magnetic field, H_0 , leads to electron circulation which produces an induced field in the indicated direction. This will lead to a large upfield shift for methyl iodide as compared to methyl fluoride. This effect will, of course, be averaged over all orientations. However, electron circulation will be minimal when the C-X bond is perpendicular to the magnetic field.⁵ When alkyl groups are introduced at the α -position, their greater

bulk will result in an attenuation of the induced magnetic field, and the difference between the fluoride and iodide will decrease.

Another observation deals with the relationship between the b_1 and b_2 parameters for the α , β and γ positions (Fig. 1 and 2). Both the series of increasing methyl substitution and increasing ethyl substitution give linear relationships among the parameters, suggesting a relationship between the mechanisms producing the differential chemical shifts at the three positions. The full data along with a more detailed interpretation of the results will be presented elsewhere.

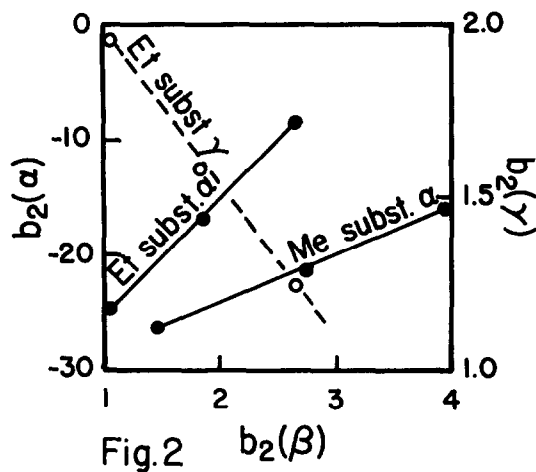
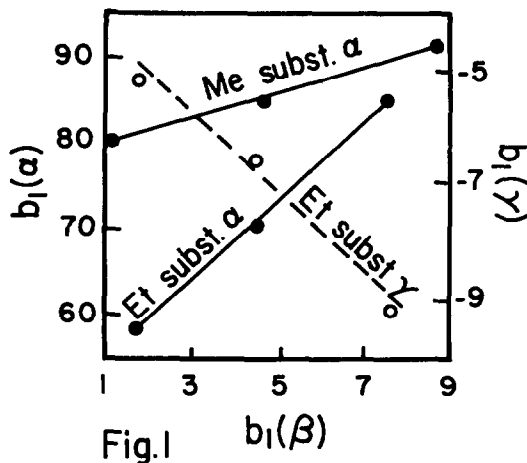


Figure 1. Relationship between b_1 (intercept) terms for α , β and γ positions.

Figure 2. Relationship between b_2 (slope) terms for α , β and γ positions.

References

1. This investigation was supported by Grant CHE-7419380 from the National Science Foundation. We wish to acknowledge the support of the Southern New England High Field NMR Facility made possible by a grant from the Biotechnology Resources Program of the National Institutes of Health (RR-798).
2. Cf. J.B. Stothers, "Carbon-13 NMR Spectroscopy," Academic Press, New York, N.Y. 1972, Chapter 5.
3. C.P. Smyth, "Dielectric Behavior and Structure," McGraw-Hill, New York, N.Y. 1955, p. 269.
4. P.H. Weiner, E.R. Malinowski and A.R. Levinstone, *J. Phys. Chem.*, **74**, 4537 (1970); P.H. Weiner and D.G. Howery, *Anal. Chem.*, **44**, 1189 (1972).
5. H. Spiesecke and W.G. Schneider, *J. Chem. Phys.*, **35**, 722 (1961).

(Received in USA 29 September 1978)