Substituent Effects on ^{13}C and ^{15}N Chemical Shifts in Triazenes Studied by Principal Components Multivariate Data Analysis

W. J. Dunn III,* C. Lins, G. Kumar, T. Manimaran and S. Grigoras

Department of Medicinal Chemistry, College of Pharmacy, University of Illinois at the Medical Center, 833 S. Wood, Chicago, Illinois 60680, USA

U. Edlund and S. Wold

Department of Organic Chemistry and Research Group for Chemometrics, Umeå University, S-901 87 Umeå, Sweden

Principal components analysis was applied to the ^{13}C and ^{15}N chemical shift data on a series of fifteen 1-(*para*-substituted-phenyl)-3-acetyl-3-methyltriazenes. It was found that the halogen-substituted triazenes formed a class, based on substituent effects, which was different from the remaining eleven triazenes. A one-component model described the halogen class, whereas a two-component model was necessary for a description of the second class. In the second class, substituents tended to cluster to form groups depending on their electronic character.

INTRODUCTION

We have recently shown 1-aryl-3,3-disubstituted triazenes (I) to have significant activity against African trypansomiasis.[1,2] In the design, synthesis and characterization of these compounds we have obtained ^{13}C and ^{15}N NMR spectra of a series of *para*-substituted 1-aryl-3-acetyl-3-methyltriazenes of structure I ($R_1 =$ CH_3, $R_2 = COCH_3$).

I
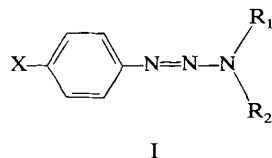
Since this represents a sufficient basis set[3] for studying the effect of substituents on the chemical shifts of the triazenes, we have submitted the data to statistical analysis. The method of analysis used was principal components analysis (PCA).[4]

The objective of the analysis was two-fold. It was of interest to determine if the method of pattern recognition, SIMCA, based on the strategy of deriving a principal components model for a class of compounds or objects[5] could be used to assist in making chemical shift assignments. It has been shown in previous studies[4,6] that this may be possible. In addition, it was of interest to explore the utility of principal components analysis in the modelling of ^{13}C and ^{15}N substituent chemical shifts (SCS) in perturbed and extended exocyclic aromatic systems. A variation of principal components analysis, factor analysis, has been used in recent reports[7] to interpret and model substituent chemical shifts in alkyl and aryl halides.

* Author to whom correspondence should be addressed.

TRADITIONAL APPROACH TO NMR DATA ANALYSIS

A solution to the problem of modelling SCS values is usually sought from the application of multivariable regression methods with chemical shift data as a dependent variable in the analysis. Independent variables are tentatively introduced into the analysis to test if relationships exist between them and the dependent variable.[8,9] Most such analyses result in at least two variables being required to satisfy predetermined statistical requirements for a level of fit of the data. Such an approach to the study of SCS must be made cautiously, as multivariable regression methods tend to overfit data[4,8,10] when applied this way.

PRINCIPAL COMPONENTS ANALYSIS

To avoid this difficulty, principal components analysis can be used to analyze the chemical shift data. Such an analysis begins with a data matrix such as that shown in Fig. 1. Here i is the variable index and k is the compound index.

The elements of the matrix are the measured chemical shift data y. In this study relative chemical shifts, $\Delta\delta_x = \delta_x - \delta_H$, are used.

If the compounds on which the data are measured satisfy the assumption that they are to some degree chemically similar, the matrix can be approximated by a principal components model [Eqn (1)].

$$y_{ik} = \alpha_i + \sum_{a=1}^{A} \beta_{ia}\theta_{ak} + \varepsilon_{ik} \qquad (1)$$

where A is the rank of the matrix or the number of product terms in the model and α_i is the mean value of variable i. Each product term in the model is

Figure 1. Data matrix for principal components analysis.

composed of a variable specific term, β, and a compound specific term, θ. The β terms are the singular vectors and the θ terms are the associated scores of the rows of the matrix. The difference in observed and predicted values of y are the residuals, ε_{ik}, which can be made arbitrarily small by including as many component terms into the model as required. Detailed discussions of the application of this method to chemical data have been published.[11,12] In the present analysis, however, we include only as many terms that have statistically significant predictive power (see below).

METHODS

### Synthesis and NMR measurements

The triazenes in this study were prepared using standard methods.[1,13] This involved preparing the diazo cation from the appropriate aniline and coupling this with methylamine. The resulting monomethyltriazene was treated with acetyl chloride in pyridine to yield the N-acetyltriazene. The triazenes were purified by crystallization or by preparative high-performance liquid chromatography using silica gel with hexane–diethyl ether as the mobile phase. Structures were verified using standard spectroscopic and analytical methods.

The ^{13}C NMR spectra were obtained as approximately 2 M solutions in chloroform-d. Spectra were proton decoupled and measured at 1.4 T on a Varian T 60 A spectrometer equipped with a Nicolet Instruments TT-7 FT attachment. Carbon-13 spin–lattice relaxation measurements (T_1) and proton-decoupled ^{15}N NMR spectra were obtained at 5.9 T on a Bruker WM-250 multinuclei instrument, using tunable 10 and 15 mm probe heads, respectively. In the ^{13}C shift studies, 4K data points were collected over a spectral width of 3 kHz. In the ^{15}N shift measurements a spectral window of 18.5 kHz was digitized by 16K

data points. A $\pi/4$ pulse width, using a pulse repetition delay of 30 s, was used in the ^{15}N experiments. Doping the samples with an inert, shiftless paramagnetic relaxation reagent,[14] $Cr(dpm)_3$, substantially improved the sensitivity. No attempts were made to determine the operating relaxation mechanisms, dipole–dipole or chemical shift anisotropy. Carbon-13 chemical shifts were reported relative to TMS. Nitrogen-15 chemical shifts were determined relative to $^{15}NH_4NO_3$ contained in an external capillary tube, and converted to a scale giving $\delta(^{15}NH_4NO_3) = 20.68$.[15]

Assignments

Carbon-13 shift assignment was straightforward, except for the OPh (**1**) and Ph (**2**) derivatives. However, a spin–lattice relaxation study could easily differentiate between the C-2, C-3 and the C-2', C-3' carbons. A larger motional freedom, vibrational and rotational, of the phenyl moiety resulted in significantly longer T_1 values for the C-2' and C-3' carbons. Considering the ^{15}N chemical shifts, the N-10 signal could be easily identified from its unique high-field position (see below). As a second-row element, nitrogen is subject to similar structural and electronic effects as carbon. Thus, a variable-by-variable plot against C-β in styrenes[16] easily distinguished the N-8 and N-9 resonances.
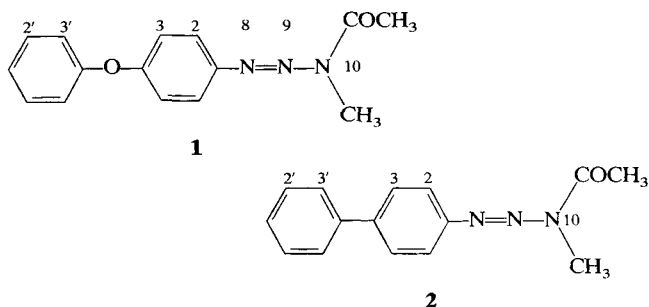
Data Analysis

The starting point in the principal components analysis is a data matrix (Fig. 1). It is often necessary to scale the data initially. This is done in the present case by subtracting from each element in the matrix the mean of its column and then dividing by the standard deviation. This gives each column a mean of zero and standard deviation of one, and has the effect of preventing a position(s) with a large variation in chemical shift from heavily weighting the analysis.

The next phase of the analysis involves approximation of the standardized matrix by a principal components model. This requires the determination of the effective rank of the matrix or determining the number of significant product terms, $A = 1, 2, \ldots$, in Eqn (1). This is done using a cross validation technique.[17] In this step elements of the matrix are systematically deleted and a model derived from the remaining matrix. The deleted elements are then predicted from the model and their residuals calculated. This is done until all elements have been deleted once, and only once.

The variance is then calculated for the predictions and compared with that of the data. If the variance is reduced after correction for differences in degrees of freedom, then the inclusion of that term is justified. This method ensures model stability and gives it optimum predictive capability.

In an analysis such as this it cannot be assumed that the chemical shift data at each position in the compounds will vary in a systematic way. There must be some criterion established for detecting variables which contain mostly noise or random error. This is done by defining modelling power, $MPOW = 1 - s_i/s_{iy}$, where s_i is the standard deviation of variable i after being fitted to the PC model and s_{iy} is the standard deviation of variable i before being fitted to the PC model. As MPOW approaches 1 the variable is well predicted, while $MPOW = 0$ indicates that a variable contains little or no systematic variation. The variation is then mostly noise and can be deleted from the analysis.

The reduced matrix is then approximated by a new PC model by repeating the above steps. Caution should be used in determining the number of components in the model, as the ratio of the number of variables (M) to component terms (A) must not approach 1.

From the analysis, the residuals, ε_{ik}, can be evaluated to determine if chemical shift assignments are internally consistent. High residuals for the chemical shift assignments at a particular position could indicate that incorrect assignments have been made.

From the ε_{ik} values for each compound a residual standard deviation, RSD, can be calculated for the fit of that compound to the model. The ε_{ik} values can also be used to calculate a standard deviation for the data set being analysed. The RSD of each compound, when compared with that of the data set, is a measure of the distance of each object from the class. Those compounds with RSD larger than twice that of the class can be treated in two ways: (1) they can be treated as outliers to the data set and removed from the analysis or (2) they can be placed in a new class and analysed separately. The latter alternative requires at least four compounds in order to constitute a statistically significant class.
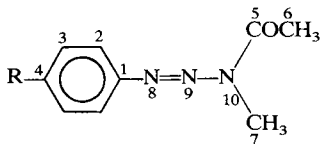
RESULTS

An examination of the data in Table 1 shows that the chemical shifts at positions C-5, C-6 and C-7 are essentially constant. These were not included in the analysis.

The data were treated initially as one class and scaled. The cross validation procedure indicated that two components were predictively significant. The tentative assignments given for the ipso and para position for the $N(CH_3)_2$-, tert-Bu- and OCH_3-triazenes were confirmed by comparing the residuals after reversing the signal assignments.

Based on MPOW, the chemical shifts of the ipso carbon and the carbon meta to it were low. These were deleted and the analysis redone using the C-2, C-4, N-8, N-9 and N-10 shielding differentials. The result was a two-component model explaining a total of 65% of the standard deviation. The first component explained 49% while the second explained 16%.

The θ_{ak} values and the residual standard deviation (RSD) resulting from this analysis are given in Table 2. The RSD of the class was found to be 0.34 and by comparing it with that of the Cl, Br and I analogs it was seen that these halogens are poorly approximated

Table 1. Differential chemical shifts[a] for the triazenes I

I

	C-1	C-2	C-3	C-4	C-5	C-6	C-7	N-8	N-9	N-10
H	148.7	121.7	128.8	128.8	173.4	21.6	27.0	408.2	446.1	208.9
$N(CH_3)_2$	−10.4	1.1	−17.3	21.7	−1.3	−0.3	−0.7	1.5	−14.6	−3.5
OCH_3	−6.5	1.2	−14.9	31.2	1.1	−0.2	−0.3	−0.7	−7.0	−1.9
OC_6H_5	−4.7	1.7	−10.2	29.2	−1.0	−0.1	−0.1	−1.4	−3.6	−0.7
CH_3	−2.2	−0.1	0.6	9.9	−0.8	0.0	−0.2	1.0	−2.8	−1.0
C_2H_5	−2.0	−0.1	−0.6	16.3	−0.7	−0.1	−0.1	0.5	−2.8	−0.9
F	−3.7	1.6	−13.2	33.8	−0.8	−0.1	−0.2	−3.2	−1.5	−0.4
C_6H_5	−1.3	0.4	−1.4	12.5	−0.9	−0.1	−0.1	−1.4	−0.9	0.4
Br	−1.2	1.5	3.1	−6.4	−0.9	0.0	0.2	−4.3	0.7	0.8
I	−0.6	1.7	9.2	−34.8	−0.9	0.0	0.2	−3.8	0.9	1.0
$COCH_3$	2.9	−0.1	0.2	7.9	−1.0	−0.2	−1.0	−4.4	5.4	2.6
t-C_4H_9	−2.4	−0.4	−3.2	23.1	−1.0	−0.1	−0.2	1.4	−2.2	−0.5
$COOCH_3$	2.7	−0.4	1.3	0.9	−1.1	−0.3	0.0	−4.3	5.0	2.3
Cl	−1.6	1.2	0.1	5.4	−1.0	−0.2	0.0	−4.8	0.3	0.4
COC_6H_5	2.5	−0.3	−0.6	8.3	−0.9	−0.1	0.2	−3.9	5.1	2.6

[a] The estimated accuracy is +0.1 ppm. All chemical shifts are relative, except for the H compound.

Table 2. Statistical data from the principal components analyses of the chemical shift data

		θ_1			θ_2			RSD				
	X	One-class analysis	Two-class analysis	Classes separate	One-class analysis	Two-class analysis	Classes separate	One-class analysis[a]	Two-class analysis[b]	Class 1[c]	Class 2[d]	σ
1	N(CH₃)₂	4.77	4.61	4.26	0.37	0.05	0.07	0.27	0.42	0.38	—	−0.83
2	OCH₃	2.57	2.67	2.48	0.75	1.04	1.00	0.12	0.19	0.17	—	−0.27
3	O-ϕ	1.28	1.60	1.60	0.55	1.46	1.49	0.16	0.39	0.45	—	−0.03
4	CH₃	0.75	0.26	0.15	−1.27	−1.29	−1.33	0.01	0.14	0.18	—	−0.03
5	C₂H₅	0.70	0.31	0.15	−1.03	−0.97	−1.06	0.01	0.17	0.14	—	−0.15
6	F	0.70	−2.76	+3.04	1.36	—	—	0.27	0.19(0.28)[e]	—	0.13	0.06
7	C₆H₅	−0.21	−0.44	−0.39	−0.31	0.03	0.06	0.01	0.18	0.22	—	−0.01
8	Br	−1.22	0.61	−0.97	0.25	—	—	0.51	0.13	—	0.22	0.23
9	I	−1.60	2.38	−2.10	−0.44	—	—	0.73	0.22	—	0.26	0.18
10	COCH₃	−2.61	−2.93	−2.65	0.46	0.77	0.84	0.30	0.01	0.01	—	0.50
11	tert-Bu	0.88	0.42	0.21	−1.04	−1.12	−1.36	0.33	0.43	0.27	—	−0.15
12	COOCH₃	−2.52	−2.99	−2.71	0.30	0.36	0.48	0.21	0.25	0.22	—	0.45
13	Cl	−0.90	−0.22	+0.03	0.65	—	—	0.50	0.30	—	0.24	0.23
14	CO-ϕ	−2.39	−2.78	−2.52	0.36	0.55	0.57	0.39	0.18	0.21	—	0.43
15	H	−0.21	−0.74	−0.57	−0.97	−0.88	−0.76	0.26	0.35	0.23	—	0.00

[a] Standard deviation of class = 0.34.
[b] Standard deviation of class 1 = 0.27; and of class 2 = 0.22.
[c] Compare with class standard deviation of 0.25.
[d] Compare with class standard deviation of 0.22.
[e] Standard deviation for fit to class 1.

by the one class model. The F analog (RSD = 0.27) is well within that of the set as a whole.

Hence the halogens, including F, were placed in a separate class. This was done to compare how the chemical shifts of these analogs were modelled relative to the other eleven compounds. The non-halogen class is now referred to as class 1 and the halogens as class 2. In this analysis class 1 was again approximated by a two-component model, while class 2 was modelled by a one-component model. In addition, all chemical shifts were now significant based on MPOW. The RSD values are given for the fit of the classes to their respective models in Table 2.

If the variables of the members of one class are fitted to the model of the other class and, also, in a converse manner, the distance between classes can be obtained. The result is the distance matrix shown in Fig. 2, and it is seen that the classes are well separated in the data space. This strongly supports the separation of the two classes.

An examination of the RSD values for fit of the F analog to the two class models reveals an interesting result. This compound is well fitted to both models, which suggests that the two classes have this compound in common.

Analysing the data as a two-class problem is useful if the objective of the analysis is to show that the data for the two classes behave differently. It was also of interest to determine how the SCS for each class behaved when analysed independently of the other. In the two-class analysis the analysis is anchored to the mean values of the variables determined over both
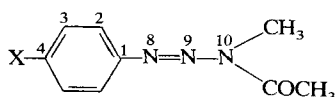
	Class	
Class	1	2
1	1.00	4.30
2	4.30	1.00

Figure 2. Distance matrix for the two-class analysis.

classes. This has the effect of obscuring the effect of substituents on the chemical shifts within each class.

The classes were then analysed separately with separate scaling with the following results. (1) Class 1 was fitted to a two-component model with the components accounting for 47% and 21% of the standard deviation, respectively. All variables were significant except the *ipso* shift. (2) Class 2 was fitted to a one-component model which accounted for 75% of the standard deviation. On the basis of a low MPOW, the shifts for the *meta* carbon and the first exocyclic nitrogen (N-8) were deleted from the analysis. The statistical data are given in Table 2 and Table 3.

Table 3. β_{ia} values

Class	β	C-1	C-2	C-3	C-4	N-8	N-9	N-10
1	β_{i1}	−0.45	0.35	−0.40	—	−0.34	−0.45	−0.44
	β_{i2}	−0.02	0.58	−0.41	—	−0.66	0.10	0.24
2	β_{i1}	−0.45	—	−0.45	0.44	—	−0.44	−0.45

DISCUSSION

This analysis was approached entirely from an empirical point of view. Initially the data were placed in one class and fitted to a single model to test for internal consistency of chemical shift assignments.

This treatment detected that the halogens Cl, Br and I were poorly fitted by the empirical PC model, suggesting that these compounds belonged to a distinctly different class. This was substantiated when the data were analysed as a two-class problem.
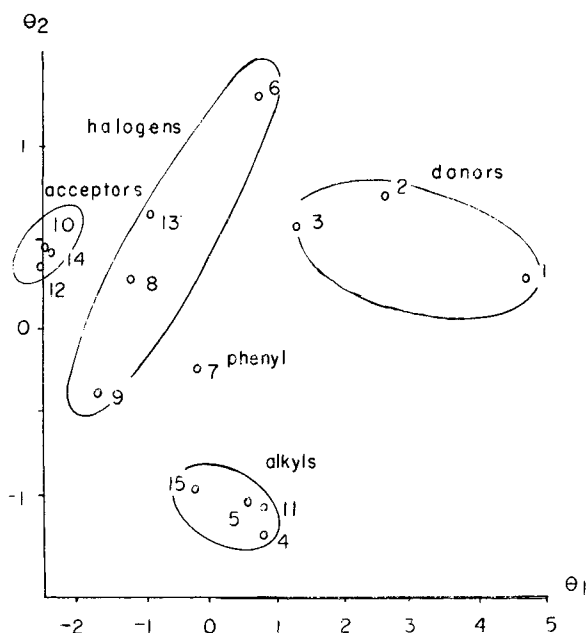
Figure 3. Plot of the two principal components, θ_1 and θ_2, for the one-class analysis.

The last result can be explained as follows. In Fig. 3 we have plotted the θ_1 and θ_2 components, resulting from the initial single model analysis. A grouping into four subclasses, donors, acceptors, alkyl and halogens, is suggested. However, the number of substituents should exceed 4–5 in each class in order to manifest statistically such clustering. Recently, a similar analysis of a data set formed by seven electronic and steric substituent constants for 28 of the most common substituents clearly showed such a grouping into four subclasses.[18] This clustering was confirmed by comparing the variance for the four classes with the total variance of the complete data set by an F-test procedure (variance ratio 6.25, $F_{crit} = 1.3$).

A similar grouping into four subclasses was observed in a multivariate data analysis of ^{13}C NMR shift data of more than seventy monosubstituted benzenes.[19] Even in this analysis, the halogens and donors had the fluorine compound in common. Hence, a reasonable explanation for the present data analysis result is that the halogen subclass has an extension (sign of β_{ia}) different from the two-dimensional plane connecting the remaining three subclasses.

The inclusion of halogens in the data sets of SCS often meets with failure, especially using two-parameter models. In fact, in most reports three of the four halogens are tested; iodine is usually not included, or is deleted from the analysis owing to poor fit.[20] In a PC analysis similar to the present study of the ^{13}C differentials of 15 2-substituted indenes,[4] the two included halogens, Br and Cl, were classified as outliers. In a ^{13}C NMR study of 4-substituted styrenes[16] it was noted that the halogens behaved anomalously by using the F and R model of Swain and Lupton.[30] In fact, it was recommended that 'these substituents should be avoided in investigations of SCS or at very least not included in limited data sets.' Similar limitations were proposed for the basis sub-

stituent set used for the analysis of electronic substituent effects.[21] Two halogens may be included, but not both Br and Cl.[22] The anomalous properties of halogens have also been noted, theoretically, in MO calculations. Thus, it was not possible to confirm the overlap and backdonation theory, i.e. no correlation was found between overlap and σ_R, independently of the choice of the basis MO set.[23] In another theoretical study to support the σ_R^0, σ_I model only one halogen was included, interestingly only the fluorine compound.[24]

The cluster pattern above can also be a plausible explanation for the increased use of three-parameter models, such as F, R and Q models.[25] It will always be possible to explain four clusters reasonably well with a three-dimensional hyperplane. The major reason that the different behavior of halogens has not been stressed more frequently is probably that multiple regression analysis tends to obscure outliers.

The chemical shifts of ^{13}C and ^{15}N are determined by a total screening constant $\sigma_{tot} = \sigma_{loc}^D + \sigma_{loc}^P + \sigma_{OTHER}$, where σ_{loc}^D derives from magnetically induced local electronic circulations, σ_{loc}^P is a measure of deviations from spherical symmetry of the electron distribution and σ_{OTHER} accounts for anisotropy, field effects, solvent effects, etc.[15,26,27] Although solvent effects are more important for ^{15}N than for carbon, it is generally assumed that σ_{loc}^P is the dominating term, and attempts to account for systematic variation in chemical shifts are based on variations in this term.[28]

Assuming A (a function of various bond orders) and ΔE (the average excitation energy) to be constant, the linearity between charge and shielding stems from the dependence of $\langle r^{-3} \rangle_{2p}$ on the charge.[28] By adding electrons an expansion of the 2p orbitals is expected.

$$\sigma_{loc}^P \propto \frac{A}{\Delta E} \langle r^{-3} \rangle_{2p} \qquad (2)$$

To exemplify Eqn 2 for ^{15}N, the nitrogen signal will appear at lower field if[15] (1) low-lying excited states exist involving electrons on the nitrogen atom (decreasing ΔE), (2) the electron density in the 2p orbitals decreases or electrons are in orbitals with larger s character (decreasing r), (3) there is multiple bonding to nitrogen (A increases).

For the chemical shifts to be modelled for the two classes suggests that it may be possible to correlate the PCA parameters with parameters not related to NMR. If such correlations could be found they could provide information about the mechanism by which substituents exert their effect(s). To test this, correlation between θ_1 from the PCA for the class 1 compounds and charge density was attempted. Using MNDO (QCPE 353) optimized geometries of a small set of the triazenes, point charges at the atoms in five of the class 1 triazenes were calculated using CNDO/2 methods (QCPE 141). These are given in Table 4. The five analogs span the range in θ_1 for this class, and a plot of the first eigenvector, θ_1', of the charge density matrix against θ_1 from the chemical shift matrix is given in Fig. 4. This shows that the first component to the chemical shifts at these positions in the triazenes are strongly related to the charge at these positions. This is to be expected if the first component is due to

Table 4. CNDO/2 net charge[a]

Position

X	1	2	3	4	5	6	8	9	10
H	-0.0289	0.1092	-0.0364	0.0138	-0.0114	0.0141	-0.1341	0.0474	-0.0759
COCH$_3$	-0.0359	0.1215	-0.0424	0.0204	-0.0384	0.0310	-0.1369	0.0533	-0.0743
CH$_3$	-0.0223	0.0994	-0.0297	-0.0047	0.0289	-0.0045	-0.1310	0.0427	-0.0767
OCH$_3$	0.0042	0.0800	-0.0110	-0.0582	0.1806	-0.0405	-0.1259	0.0354	-0.0778
N(CH$_3$)$_2$	0.0037	0.0655	-0.0032	-0.0655	0.1488	-0.0659	-0.1207	0.0262	-0.0796

[a] −, Electrons; reference point is neutral atom.
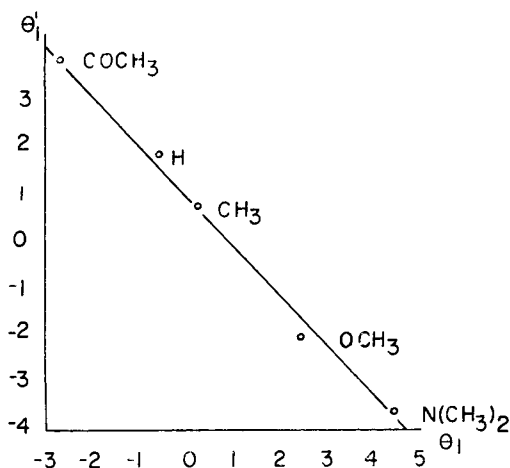

Figure 4. Plot of the first components, θ_1, of the analysis of the halogens against the first component, θ_1', of the charge density matrix.

Figure 5. Plot of the first components, θ_1, of the analysis of the halogens against the first major component, a_2, from the factor analysis of alkyl halide chemical shifts by Wiberg et al.[7]

the paramagnetic contribution. Similar correlations between Hammett sigma and other Hammett-like substituent constants were also successful for the class 1 compounds.

Attempts to relate θ_2, which contributes significantly (21%) to the SCS, were not successful and an explanation of this effect is difficult to give. For the 2-substituted indenes[4] a similar result was obtained. The first dominant component correlated well with σ_p or σ_R, but the second component showed no correlation with other σ parameters.

There is no relationship between θ_1 for the halogens and the Hammett σ constant. θ_1 increases regularly from −2.10 for I to 3.04 for F. Almost any non-NMR parameter which increases in this progression would give a good correlation. Deviations for heavy halogens have been attributed to spin–orbit coupling,[18] which can be mentioned as one explanation.

Wiberg et al.[7] have recently studied the SCS of the halogens in alkyl and aryl systems. In an attempt to interpret their factor analysis results they have come to this same conclusion. It was concluded that the major factor leading to variation in chemical shifts in
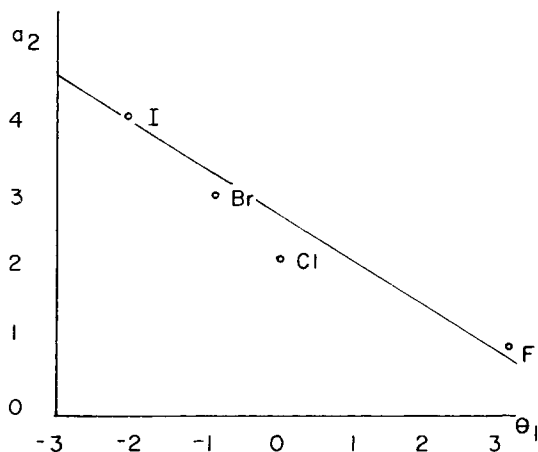
alkyl halides is related to 'freeness' of valence electrons about the halogen. A plot of their major factor against θ_1 for halogens is given in Fig. 5, and it can be seen that the two effects are very similar.

The fact that θ_1 for the halogens is not related to the Hammett σ constant suggests that it is a different effect than that in the non-halogen class. A periodic dependence of substituent effects on ring-carbon chemical shifts was proposed earlier.[29]

One interesting observation about the division of the classes as suggested by the SIMCA analysis is that all of the substituents bonded to the phenyl ring by a second row element fit in class 1. This includes the F analog. We are preparing the p-S-CH$_3$-triazene to determine where it will also fit into this classification scheme. These results will be reported later.

Acknowledgements

REFERENCES

1. W. J. Dunn, J. Powers, J. B. Kaddu and A. R. Njogu, J. Pharm. Sci. **69**, 1465 (1981).

2. W. J. Dunn, M. Greenberg and J. Powers, J. Pharm. Sci. **71**, 126 (1982).

3. W. J. Hehre, R. W. Taft and R. D. Topsom, *Prog. Phys. Org. Chem.*, edited by R. W. Taft, Academic Press, New York, **12**, 159 (1976).
4. U. Edlund and S. Wold, *J. Magn. Reson.* **37**, 183 (1980); B. Eliasson and U. Edlund, *J. Chem. Soc. Perkin Trans. 2* 403 (1981).
5. S. Wold, *Pattern Recognition* **8**, 127 (1976).
6. M. Sjostrom and U. Edlund, *J. Magn. Reson.* **25**, 285 (1977).
7. K. B. Wiberg, W. E. Pratt and W. F. Bailey, *J. Org. Chem.* **45**, 4936 (1980); W. F. Bailey, E. A. Cioffi and K. B. Wiberg, *J. Org. Chem.* **46**, 4219 (1981).
8. J. Bromilow, R. T. C. Brownlee, V. O. Lopez and R. W. Taft, *J. Org. Chem.* **44**, 4766 (1979); J. Bromilow, R. T. C. Brownlee, D. J. Craik, M. Sadek and R. W. Taft, *J. Org. Chem.* **45**, 2429 (1980); R. T. C. Brownlee and D. J. Craik, *J. Chem. Soc. Perkin Trans. 2* 760 (1981); J. Bromilow, R. T. C. Brownlee, D. J. Craik, P. R. Fiske, J. E. Rowe and M. Sadek, *J. Chem. Soc. Perkin Trans. 2* 753 (1981).
9. G. A. Olah, D. W. Westerman and D. A. Forsyth, *J. Am. Chem. Soc.* **97**, 3419 (1975).
10. J. G. Topliss and R. P. Edwards, in *Computer Assisted Drug Design*, edited by E. C. Olsen and R. E. Christoferssen. Symposium Series No. 112, American Chemical Society, Washington, D.C. (1979).
11. S. Wold and M. Sjostrom, in *Chemometrics: Theory and Application*, edited by B. R. Kowalski, Symposium Series No. 52, American Chemical Society, Washington, D.C. (1977).
12. K. G. Joreskog, J. E. Klovar and R. A. Reyment, *Geological Factor Analysis*. Elsevier, Amsterdam (1976).
13. O. Dimroth, *Chem. Ber.* **38**, 670 (1905).
14. G. C. Levy, U. Edlund and J. G. Hexem, *J. Magn. Reson.* **19**, 259 (1975).
15. G. C. Levy and R. L. Lichter, *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy.* Wiley–Interscience, New York (1979).
16. G. K. Hamer, I. R. Peat and W. F. Reynolds, *Can. J. Chem.* **51**, 897 (1973).
17. S. Wold, *Technometrics* **20**, 397 (1978).
18. S. Clementi, U. Edlund, M. Sjostrom and S. Wold, *Acta Chem. Scand.* in press.
19. D. Johnels, S. Hellberg, S. Clementi, W. J. Dunn, U. Edlund, H. Grahn, M. Sjostrom and S. Wold, unpublished results.
20. J. R. Roarke and W. B. Smith, *J. Phys. Chem.* **73**, 1043 (1969).
21. S. Ehrenson, R. T. C. Brownlee and R. W. Taft, *Progr. Phys. Org. Chem.*, edited by R. W. Taft, Academic Press, New York, **10**, 1 (1973).
22. R. D. Topsom, *Progr. Phys. Org. Chem.*, edited by R. W. Taft, Academic Press, New York, **12**, 1 (1976).
23. P. Politzer and J. W. Timberlake, *J. Org. Chem.* **37**, 3557 (1972).
24. W. F. Reynolds, P. G. Nezey and G. K. Hamer, *Can. J. Chem.* **55**, 522 (1977).
25. W. B. Smith and T. W. Proulx, *Org. Magn. Reson.* **8**, 567 (1976).
26. R. Laatikaiinen, *Org. Magn. Reson.* **14**, 366 (1980).
27. D. F. Ewing, *Org. Magn. Reson.* **12**, 499 (1979).
28. M. Karplus and J. A. Pople, *J. Chem. Phys.* **38**, 2803 (1963).
29. I. Morishima, K. Endo and T. Yonezawa, *J. Chem. Phys.* **59**, 3356 (1973); N. Inamoto, S. Masuda, K. Tokumaru, K. Tori, M. Yoshida and Y. Yoshimura, *Tetrahedron Lett.* 3711 (1976).
30. C. G. Swain and E. L. Lupton, *J. Am. Chem. Soc.* **90**, 4328 (1968).