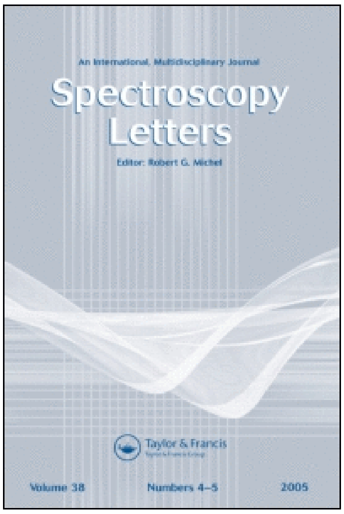

**Spectroscopy Letters**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597299

# The Prediction of the $^{13}$C NMR Signal Positions in Substituted Naphthalenes, Part 2: The Use of Statistical Substituent Chemical Shift (SSCS) Values

J. R. Nanney[a]; C. A. L. Mahaffy[b]
[a] Department of Mathematics, Auburn University at Montgomery, Montgomery, AL, USA [b] Department of Chemistry, Auburn University at Montgomery, Montgomery, AL, USA

**To cite this Article** Nanney, J. R. and Mahaffy, C. A. L.(2000) 'The Prediction of the $^{13}$C NMR Signal Positions in Substituted Naphthalenes, Part 2: The Use of Statistical Substituent Chemical Shift (SSCS) Values', Spectroscopy Letters, 33: 2, 255 — 267

**To link to this Article: DOI:** 10.1080/00387010009350075

**URL:** http://dx.doi.org/10.1080/00387010009350075

SPECTROSCOPY LETTERS, 33(2), 255–267 (2000)

# The Prediction of the $^{13}$C NMR Signal Positions in Substituted Naphthalenes, Part 2: The Use of Statistical Substituent Chemical Shift (SSCS) Values

**Key Words:** $^{13}$C NMR of naphthalenes, statstistical substituent chemical shift (SSCS) values, mathematical modeling

**J. R. Nanney**
**Department of Mathematics**
**Auburn University at Montgomery**
**Montgomery, AL 36117, USA**

**C.A.L. Mahaffy**
**Department of Chemistry**
**Auburn University at Montgomery**
**Montgomery, AL 36117, USA**

## Abstract

The prediction of the $^{13}$C NMR signals for derivatives of naphthalene has been investigated using statistical Substituent Chemical Shift (SSCS) values. For α-derivatives the model had a correlation coefficient of observed versus predicted line positions of $r=.98$ with an standard deviation of 2.1ppm while in the β case $r=.98$ with the standard deviation being 2.0ppm. Prediction of the 9 and 10 positions had an $r=.93$ with the standard deviation being 1.5ppm. The data base consisted of 5250 signals from 525 naphthalene derivatives.

www.dekker.com

## INTRODUCTION

In Part 1 [1] we reported the two mathematical models which utilized the Field (*F*), Resonance (*R*) and Charton's Steric Parameters (v) together with Molar Refractivity (*MR*)[2-10] to predict the $^{13}$C NMR line positions for substituted naphthalene derivatives. The first was for α and the second for β substituted compounds. In the α case the model had a correlatation coefficient of observed versus predicted line positions of $r=.973$ with an standard deviation of 2.2ppm while in the β case $r=.979$ with the standard deviation being 2.3ppm. The data base consisted of 3152 signals from 394 naphthalene derivatives.

We have previously reported studies using the same four parameters in the prediction of the $^{11}$B NMR of trigonal boranes[10], the $^{13}$C NMR of arenes[11-12], the nitrogen NMR spectra for derivatives of ammonia[13-14], and the $^{19}$F NMR signals for fluoroarenes[15-16] and fluoroarenetricarbonyl-chromium(0)complexes[17-18].

We have also investigated the use of the Statistical Substituent Chemical Shift (SSCS) values in predicting NMR line positions. An SSCS value is a measure of the shift in an NMR line position induced by substituting a particular group on a structure which acts as a standard. We have reported SSCS values for a number of systems: the $^{11}$B NMR spectra of trigonal boranes[19], the $^{13}$C NMR spectra of arenes[20-21], and the $^{19}$F NMR signals for fluoro-derivatives of arenes[22], six membered aromatic nitrogen heterocycles[23], arenetricarbonylchromium(0) complexes[24], and ethylenes[25].

In the present paper we report the application of the SSCS method to the $^{13}$C NMR spectra of substituted naphthalenes.

## RESULTS AND DISCUSSION

### The Database

We surveyed the literature from the period from 1970 until 1985. From this we extracted our data base in which:

1. The compound was a naphthalene derivative containing one or more groups. No heterocyclic compounds were considered. We did not permit compounds where an aromatic ring was directly attached to the naphthalene ring. The groups which were studied are listed in Table 2.

2. All signals were converted so as to use TMS as a reference.

The data base consists of 2100 α- and 2100 β-signals from 525 naphthalene derivatives.

### Statistical Analysis

SSCS methods work well only for compound types for which there is not much interaction between the substituted groups, since the basic idea of SSCS methods is to add the SSCS values for the substituents to the nmr value for the base compound. The models developed in Part 1 used interaction terms among the 4 parameters of the same substituent, but no interaction factors between different substituents. Because the predictions were quite accurate using such models, one can infer that SSCS methods should work well for substituted naphthalenes. Indeed they do as we show in the present paper.

Table 1 (In part). Compounds, observed and predicted 13C NMR from positions 1 - 10. The letters in the encoding for a compound represent one of the substituents in positions 1 - 8 listed in Tables 2, 3 and 4. The encoding F...E... for example represents 1-cyano-5-fluoronaphthalene. The letter coding is to make Tables 1 and 5 more compact. The group numbers in Tables 2, 3 and 4 are the numbers for those groups we have used in our other papers. The complete Table 1 has 525 compounds (5250 signals) and may be obtained from the authors upon request.

| Compound | Position | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | | 7 | | 8 | | 9 | | 10 | |
| | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED |
| F...E... | 110.1 | 109.4 | 133.4 | 132.8 | 125.3 | 124.9 | 125.7 | 125.6 | 158.7 | 159.3 | 110.3 | 109.9 | 128.6 | 127.7 | 120.9 | 121.0 | 133.4 | 133 .6 | 123.3 | 122.2 |
| L....... | 147.5 | 150.5 | 124.5 | 122.6 | 125.2 | 125.2 | 135.3 | 132.5 | 129.5 | 128.8 | 128.1 | 128.3 | 130.1 | 129.9 | 123.3 | 123.7 | 125.5 | 124.9 | 135.2 | 134.1 |
| Z....... | 128.0 | 127.7 | 131.2 | 129.7 | 125.5 | 124.3 | 134.1 | 132.4 | 129.4 | 129.5 | 126.9 | 125.7 | 128.3 | 127.1 | 126.7 | 125.8 | 132.3 | 131.4 | 134.9 | 133.5 |
| F....... | 110.6 | 110.0 | 133.5 | 132.3 | 126.0 | 124.8 | 134.2 | 132.8 | 129.7 | 128.2 | 128.3 | 127.0 | 129.5 | 128.4 | 125.3 | 125.2 | 133.8 | 132.4 | 132.8 | 132.2 |
| .L...... | 125.1 | 124.8 | 146.3 | 146.2 | 119.7 | 119.3 | 130.4 | 130.7 | 128.7 | 128.8 | 130.5 | 130.2 | 128.7 | 128.4 | 130.8 | 130.3 | 132.7 | 131.9 | 136.6 | 136.2 |
| .Z...... | 131.8 | 130.8 | 128.8 | 127.9 | 126.2 | 125.4 | 129.0 | 127.8 | 128.6 | 127.6 | 129.1 | 128.0 | 127.6 | 126.7 | 130.1 | 128.8 | 133.5 | 132.4 | 136.4 | 135.2 |
| .F...... | 134.8 | 133.7 | 110.0 | 109.3 | 127.0 | 126.3 | 130.1 | 128.9 | 128.8 | 127.7 | 129.8 | 128.9 | 128.4 | 127.4 | 129.2 | 128.4 | 133.1 | 132.2 | 135.4 | 134.4 |
| E....... | 159.5 | 159.0 | 109.8 | 109.4 | 126.0 | 125.9 | 124.2 | 123.7 | 128.1 | 127.3 | 127.3 | 127.0 | 126.6 | 126.7 | 120.9 | 120.8 | 124.3 | 123.6 | 135.7 | 134.8 |
| A....... | 122.7 | 123.1 | 129.7 | 128.4 | 125.7 | 126.2 | 127.7 | 127.3 | 128.1 | 128.2 | 126.4 | 127.0 | 126.9 | 128.1 | 127.0 | 127.0 | 132.0 | 131.8 | 134.6 | 134.5 |
| F....... | 108.8 | 110.0 | 131.0 | 132.3 | 123.5 | 124.8 | 131.8 | 132.8 | 127.3 | 128.2 | 126.1 | 127.0 | 127.1 | 128.4 | 123.5 | 125.2 | 131.4 | 132.4 | 130.7 | 132.2 |
| .E...... | 110.9 | 110.7 | 161.0 | 160.5 | 116.2 | 116.5 | 130.5 | 129.8 | 128.1 | 127.8 | 125.2 | 125.0 | 127.0 | 127.0 | 127.5 | 127.1 | 134.7 | 133.8 | 130.8 | 130.5 |
| .A...... | 130.0 | 129.6 | 119.9 | 121.0 | 129.4 | 127.5 | 129.6 | 129.7 | 127.9 | 128.1 | 126.3 | 126.6 | 126.9 | 127.1 | 127.1 | 127.5 | 134.3 | 134.0 | 131.6 | 132.7 |
| .O...... | 122.9 | 123.7 | 148.4 | 148.5 | 124.7 | 125.3 | 127.6 | 127.2 | 128.0 | 127.3 | 125.2 | 126.4 | 125.7 | 125.6 | 127.4 | 127.9 | 134.0 | 134.1 | 132.3 | 132.2 |
| .T...... | 133.8 | 133.6 | 137.8 | 137.8 | 129.8 | 129.8 | 127.0 | 127.0 | 128.1 | 128.1 | 126.2 | 126.2 | 125.7 | 125.7 | 128.1 | 128.1 | 133.1 | 133.6 | 133.8 | 133.6 |
| F...E... | 110.1 | 109.4 | 133.4 | 132.8 | 125.3 | 124.9 | 125.7 | 125.6 | 158.7 | 159.3 | 111.3 | 109.9 | 128.6 | 127.7 | 120.9 | 121.0 | 133.4 | 133.6 | 123.3 | 122.2 |
| .C...E.. | 107.6 | 110.0 | 142.1 | 138.8 | 118.1 | 118.6 | 127.1 | 128.3 | 109.6 | 110.3 | 157.1 | 157.2 | 115.3 | 117.4 | 126.5 | 127.5 | 130.5 | 131.7 | 126.8 | 128.2 |
| .J...... | 106.2 | 107.9 | 158.1 | 156.6 | 119.1 | 118.3 | 129.9 | 129.0 | 128.1 | 127.8 | 124.0 | 123.9 | 126.8 | 126.6 | 127.2 | 127.1 | 135.3 | 134.2 | 129.6 | 129.3 |
| .ME..... | 124.5 | 126.5 | 131.1 | 130.4 | 157.0 | 157.0 | 111.2 | 111.0 | 125.5 | 126.3 | 127.7 | 128.0 | 124.7 | 125.7 | 128.2 | 128.3 | 128.2 | 128.6 | 134.7 | 134.8 |
| .Z....E. | 129.4 | 130.0 | 127.9 | 128.4 | 123.9 | 123.9 | 127.1 | 127.7 | 129.2 | 129.5 | 117.6 | 116.0 | 159.7 | 160.7 | 111.3 | 111.6 | 132.4 | 132.6 | 131.6 | 132.0 |
| .Z.....E | 132.9 | 123.6 | 127.3 | 128.1 | 125.5 | 126.0 | 126.8 | 127.3 | 122.6 | 123.3 | 127.2 | 127.4 | 109.3 | 109.7 | 158.2 | 159.9 | 122.0 | 122.4 | 135.2 | 136.3 |
| .Z...... | 131.7 | 130.8 | 128.6 | 127.9 | 125.8 | 125.4 | 128.4 | 127.8 | 128.1 | 127.6 | 128.6 | 128.0 | 126.9 | 126.7 | 129.6 | 128.8 | 133.0 | 132.4 | 136.1 | 135.2 |
| Z....E.. | 127.6 | 127.5 | 128.7 | 128.2 | 124.8 | 124.8 | 133.5 | 131.6 | 110.6 | 112.3 | 159.3 | 159.8 | 116.7 | 117.1 | 128.0 | 127.7 | 130.7 | 128.3 | 134.2 | 133.7 |
| Z.....E. | 125.8 | 126.9 | 130.8 | 130.2 | 123.0 | 122.8 | 132.4 | 132.3 | 130.0 | 131.4 | 115.6 | 115.7 | 160.7 | 161.2 | 109.3 | 108.6 | 131.7 | 131.6 | 130.2 | 130.4 |
| Z....... | 126.5 | 127.7 | 129.5 | 129.7 | 123.4 | 124.3 | 132.3 | 132.4 | 127.4 | 129.5 | 125.0 | 125.7 | 126.5 | 127.1 | 124.8 | 125.8 | 130.3 | 131.4 | 132.8 | 133.5 |
| .J....E. | 104.2 | 107.1 | 156.9 | 157.1 | 116.8 | 116.7 | 128.0 | 128.8 | 128.6 | 129.7 | 112.4 | 113.9 | 159.7 | 160.6 | 108.9 | 109.9 | 134.3 | 134.3 | 124.6 | 126.2 |
| ........ | 128.0 | 127.9 | 125.9 | 126.5 | 125.9 | 126.5 | 128.0 | 127.9 | 128.0 | 127.9 | 125.9 | 126.5 | 125.9 | 126.5 | 128.0 | 127.9 | 133.5 | 133.6 | 133.5 | 133.6 |
| BB...... | 129.4 | 130.0 | 131.7 | 131.7 | 126.6 | 125.9 | 128.1 | 128.4 | 127.9 | 128.8 | 127.2 | 127.2 | 127.7 | 129.2 | 124.6 | 124.8 | 130.3 | 130.2 | 132.6 | 134.4 |
| MK...... | 140.6 | 140.6 | 132.2 | 132.2 | 129.9 | 129.9 | 126.5 | 126.5 | 129.1 | 129.1 | 124.1 | 124.1 | 125.0 | 125.0 | 125.0 | 125.0 | 133.7 | 133.7 | 132.0 | 132.0 |
| M.M..... | 147.1 | 141.2 | 113.9 | 120.0 | 147.5 | 139.0 | 116.5 | 125.8 | 127.8 | 128.5 | 127.1 | 127.0 | 126.0 | 128.4 | 121.2 | 123.1 | 134.2 | 133.8 | 124.9 | 127.7 |
| K..K.... | 132.1 | 133.4 | 126.0 | 126.7 | 126.0 | 126.7 | 132.1 | 133.4 | 124.4 | 125.8 | 125.1 | 125.6 | 125.1 | 125.6 | 124.4 | 125.8 | 132.5 | 133.5 | 132.5 | 133.5 |
| B..B.... | 130.9 | 131.1 | 125.9 | 126.3 | 125.9 | 126.3 | 130.9 | 131.1 | 125.0 | 126.5 | 127.7 | 128.8 | 127.7 | 128.8 | 125.0 | 126.5 | 131.6 | 132.1 | 131.6 | 132.1 |
| A..A.... | 122.5 | 122.5 | 130.0 | 128.1 | 130.0 | 128.1 | 122.5 | 122.5 | 127.7 | 127.3 | 128.0 | 128.6 | 128.0 | 128.6 | 127.7 | 127.3 | 132.8 | 132.6 | 132.8 | 132.6 |
| E..X.... | 156.0 | 157.4 | 109.1 | 111.6 | 122.7 | 128.3 | 135.4 | 137.0 | 121.2 | 126.6 | 125.9 | 125.9 | 125.6 | 125.9 | 124.2 | 122.6 | 126.4 | 124.7 | 131.8 | 130.1 |
| B......B | 130.4 | 129.7 | 131.8 | 128.6 | 127.2 | 126.5 | 129.7 | 127.9 | 129.7 | 127.9 | 127.2 | 126.5 | 131.8 | 128.6 | 130.4 | 129.7 | 127.8 | 129.6 | 138.2 | 134.7 |
| H......H | 96.0 | 98.7 | 145.1 | 143.5 | 128.1 | 127.6 | 132.1 | 133.0 | 132.1 | 133.0 | 128.1 | 127.6 | 145.1 | 143.5 | 96.0 | 98.7 | 136.9 | 135.3 | 133.0 | 135.3 |
| .BB..... | 127.0 | 127.4 | 130.1 | 131.3 | 130.1 | 131.3 | 127.0 | 127.4 | 126.8 | 127.1 | 128.8 | 127.6 | 128.8 | 127.6 | 126.8 | 127.1 | 132.3 | 132.4 | 132.3 | 132.4 |
| .JJ..... | 106.2 | 109.0 | 149.4 | 148.3 | 149.4 | 148.3 | 106.2 | 109.0 | 126.2 | 127.0 | 124.0 | 124.0 | 124.0 | 124.0 | 126.2 | 127.0 | 129.1 | 129.8 | 129.1 | 129.8 |
| .J...A.. | 105.7 | 108.1 | 157.8 | 156.7 | 119.6 | 118.8 | 128.3 | 128.5 | 129.5 | 128.5 | 117.0 | 118.4 | 129.5 | 127.6 | 128.3 | 128.9 | 132.9 | 133.2 | 129.9 | 129.7 |
| .E....E. | 110.3 | 109.9 | 162.0 | 161.0 | 115.3 | 115.0 | 130.4 | 129.7 | 130.4 | 129.7 | 115.3 | 115.0 | 162.0 | 161.0 | 110.3 | 109.9 | 135.5 | 136.0 | 127.6 | 127.4 |
| BBB..... | 129.3 | 131.4 | 130.4 | 131.7 | 131.4 | 130.8 | 127.3 | 126.8 | 127.3 | 128.0 | 127.9 | 128.2 | 127.7 | 129.4 | 124.8 | 124.8 | 130.0 | 130.4 | 132.0 | 132.9 |
| LK...K.. | 148.2 | 149.2 | 130.0 | 130.1 | 128.2 | 126.6 | 127.1 | 131.1 | 126.5 | 127.5 | 136.8 | 135.9 | 130.8 | 131.2 | 121.1 | 122.3 | 123.1 | 123.8 | 132.8 | 133.0 |
| AGA..... | 106.2 | 106.6 | 146.9 | 155.2 | 110.5 | 112.0 | 131.3 | 130.5 | 127.1 | 127.7 | 124.9 | 125.3 | 127.9 | 128.9 | 125.6 | 125.8 | 131.6 | 131.6 | 129.5 | 130.4 |
| AC...A.. | 103.6 | 105.6 | 142.3 | 142.4 | 118.5 | 118.3 | 127.5 | 128.1 | 129.8 | 129.5 | 116.4 | 118.1 | 129.8 | 130.1 | 130.7 | 126.5 | 129.4 | 132.1 | 126.7 | 129.2 |
| BBBB.... | 130.1 | 130.5 | 130.3 | 131.2 | 130.3 | 131.2 | 130.1 | 130.5 | 125.4 | 125.7 | 128.7 | 129.9 | 128.7 | 129.9 | 125.4 | 125.7 | 130.0 | 130.9 | 130.0 | 130.9 |
| BBBBBBBB | 128.7 | 128.3 | 135.0 | 134.6 | 135.0 | 134.6 | 128.7 | 128.3 | 128.7 | 128.3 | 135.0 | 134.6 | 135.0 | 134.6 | 128.7 | 128.3 | 129.4 | 128.1 | 129.4 | 128.1 |
| P......P | 146.3 | 149.1 | 111.5 | 114.5 | 126.7 | 126.5 | 119.3 | 123.1 | 119.3 | 123.1 | 126.7 | 126.5 | 111.5 | 114.5 | 146.3 | 149.1 | 117.7 | 122.4 | 138.0 | 138.4 |
| L..LL... | 150.6 | 155.9 | 124.8 | 123.2 | 127.6 | 124.7 | 147.8 | 150.8 | 145.9 | 147.1 | 129.3 | 127.8 | 130.8 | 130.4 | 129.1 | 129.1 | 127.0 | 125.8 | 117.4 | 116.6 |
| J..L..J. | 158.6 | 159.7 | 104.0 | 104.5 | 125.3 | 121.2 | 139.7 | 141.9 | 125.3 | 124.5 | 122.4 | 121.2 | 159.8 | 157.6 | 102.6 | 103.7 | 127.4 | 130.4 | 122.0 | 118.9 |

Because of the geometry of the naphthalene molecule, there are three types of signals which we call α signals (positions 1, 4, 5 and 8), β signals (positions 2, 3, 6 and 7) and γ signals (positions 9 and 10). Because of this the data base of signals must be divided into 3 databases resulting in three sets of SSCS values. We denote by $SA_{ij}$ the α SSCS value of group j = 1, 39 if it is in position i = 1, 8. These SSCS values are used for predicting α signals. With respect to the α-SSCS values the molecule is oriented so that the carbon yielding the nmr signal is in position 1. For computation purposes the groups attached to the naphthalene are encoded by number. The group numbers appear in Table 2. Since each compound has four *α signals, it is encoded four times in our α database each encoding putting one of the four α* positions in position 1. Part 1 [1] shows an example of the encoding of one compound in the database. Similarly the β SSCS values are denoted by $SB_{ij}$ , i = 1,8 j = 1,39 and the γ-SSCS values are denoted $SG_{ij}$ , i = 1,8 j = 1,39. With respect to the β-SSCS values the molecule is oriented so that the carbon yielding the signal is in position 2 and for the γ-SSCS values, position 9. There were 525 compounds in the study which resulted in three databases, 2100 α signals in the α database, 2100 β signals in the β database and 1050 γ signals in the γ database.

We will describe how the α- SSCS values were determined. The β- and γ-SSCS values were determined in much the same way. As we have explained in previous SSCS work, it is necessary to introduce the "existence" functions, X defined as follows:

$X_{ijk}$ = **1 if group j is in position i relative to the carbon yielding signal k; 0 otherwise**

Table 2. Group numbers, group names, and alpha SSCS values for groups in positions 1 - 8. These SSCS values are to be used for alpha signals with the carbon associated with the signal assumed to be in position 1. The numbered positions for SSCS values are numbered with respect to that assumption. The predicted signal from position 1 for naphthalene is 127.92 ppm.

| No. | Code | Name | Position | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | . | H | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 |
| 2 | A | Br | -4.79 | 1.72 | 1.83 | -.60 | .31 | .19 | -.43 | -.96 |
| 3 | B | Cl | 4.06 | -1.96 | 1.40 | -.90 | .86 | -.02 | -.78 | -2.31 |
| 4 | C | NH2 | 14.82 | -17.74 | 1.18 | -9.77 | 1.25 | -.45 | -2.32 | -2.97 |
| 5 | D | CH2CH3 | | -2.40 | -.35 | | | -1.31 | -.30 | |
| 6 | E | F | 31.10 | -17.20 | 1.90 | -4.21 | -.58 | -.16 | -.80 | -7.16 |
| 7 | F | CN | -17.94 | 5.79 | .95 | 4.87 | .25 | -.24 | .44 | -2.74 |
| 8 | G | OH | 23.72 | -18.38 | 1.45 | -8.66 | .19 | -.08 | -1.34 | -5.73 |
| 9 | H | I | -31.24 | 8.89 | 1.81 | 2.96 | 2.12 | -.32 | -.78 | 2.00 |
| 10 | I | CF3 | | -1.76 | 1.56 | | | .19 | .38 | |
| 11 | J | OCH3 | 28.08 | -20.01 | 1.06 | -8.48 | -.27 | -.13 | -.83 | -5.02 |
| 12 | K | CH3 | 7.17 | -1.05 | -.28 | -1.73 | .51 | -.26 | -1.12 | -2.60 |
| 13 | L | NO2 | 22.59 | -3.09 | 2.80 | 4.56 | .84 | .88 | 2.38 | -4.23 |
| 15 | M | COCH3 | 13.03 | -3.34 | .27 | 1.23 | .00 | -.82 | .56 | -3.96 |
| 16 | N | CH(CH3)2 | 13.73 | | | -1.14 | 1.44 | | | -1.80 |
| 17 | O | C(CH3)3 | 18.11 | -4.22 | -.75 | -.64 | 1.42 | -.61 | .03 | -.81 |
| 18 | P | N(CH3)2 | 22.81 | -21.15 | 1.08 | -4.49 | -.30 | -.15 | -1.32 | -1.66 |
| 20 | Q | CH2Br | 4.08 | -1.40 | .06 | 1.38 | .58 | -.34 | .57 | -4.52 |
| 21 | R | CH2OH | 8.28 | | | .18 | .58 | | | -4.42 |
| 40 | S | NHCOCH3 | | -10.75 | .50 | | | -1.26 | -.20 | |
| 41 | T | SiMe3 | 9.88 | 5.65 | -.92 | 1.78 | 1.28 | .18 | .18 | .18 |
| 65 | U | SO3H | 11.34 | -2.21 | 1.32 | 5.52 | 2.52 | 1.62 | 1.76 | -1.48 |
| 72 | V | GeMe3 | 12.48 | 5.08 | -.72 | 1.18 | 1.18 | -.22 | -.12 | .38 |
| 73 | W | PbMe3 | 22.18 | 8.28 | -.52 | .18 | 1.08 | -.12 | -.42 | 2.68 |
| 74 | X | SnMe3 | 13.29 | -.50 | .23 | -1.58 | 1.87 | -.71 | .55 | -.72 |
| 81 | Y | OCOMe | 18.68 | -9.41 | 1.23 | -2.02 | .08 | -.85 | -.11 | -6.82 |
| 117 | Z | COOH | -.22 | 2.83 | -.09 | 4.51 | 1.63 | -.36 | .85 | -2.13 |
| 125 | a | CH2COOH | 1.48 | | | -.02 | .48 | | | -4.62 |
| 161 | b | CHO | 8.33 | 5.64 | .52 | 2.01 | 1.99 | -.34 | 2.06 | -2.60 |
| 163 | c | COOMe | -.92 | 2.99 | -.01 | 5.48 | .78 | -.85 | 1.49 | -1.72 |
| 263 | d | CONMe2 | 6.88 | | | .98 | .38 | | | .18 |
| 271 | e | CH2CMe3 | 7.98 | .70 | -.88 | -1.32 | .58 | -.40 | -.36 | -2.72 |
| 276 | f | CH2PbMe3 | 12.78 | -4.82 | -.32 | -3.92 | .88 | -.52 | -1.12 | -4.12 |
| 292 | g | CH2SnMe3 | 11.63 | -2.85 | -.04 | -3.72 | .65 | -.33 | -1.07 | -3.62 |
| 293 | h | CMe2OH | 15.58 | | | .58 | .98 | | | -.62 |
| 306 | i | CMe=CH2 | 16.78 | | | -.72 | .38 | | | -2.12 |
| 308 | j | COtBu | 1.18 | | | 4.78 | .68 | | | -1.72 |
| 326 | k | FORMYL | 5.48 | 5.52 | .35 | 6.57 | .00 | -.49 | 2.10 | -2.92 |
| 331 | l | CH2P(O)(OEt)2 | .51 | .98 | .28 | -.13 | .96 | .08 | -.02 | -2.11 |

Here i = 1,8 varies over the positions, j = 1,39 varies over 39 groups used in the study and k = 1, 2100 varies over the 2100 α signals in the database. Note: Because we have tried to be consistent with the way we have numbered groups in our papers, the actual groups are not numbered 1 through 39 in Table 2, but rather use the group numbers we have used in previous papers. If all the $SA_{ij}$ were known and also c, the nmr signal for naphthalene, then because of the assumed additivity (not quite realized) of SSCS values the k'th signal would be predicted to be $YPRED_k$ where $YPRED_k$ is given by the formula:

$$YPRED_k = c + \sum_{i=1}^{8}\sum_{j=1}^{39} X_{ijk} SA_{ij} \qquad (k=1,2100)$$

It is the $SA_{ij}$, of course, which are not known and are to be determined. Substituting the observed k'th signal, into the formula for $YPRED_k$ yields a system of 2100 equations in 312 unknowns, the unknowns being the $SA_{ij}$. A few groups did not appear in certain positions in the database, and so the computer

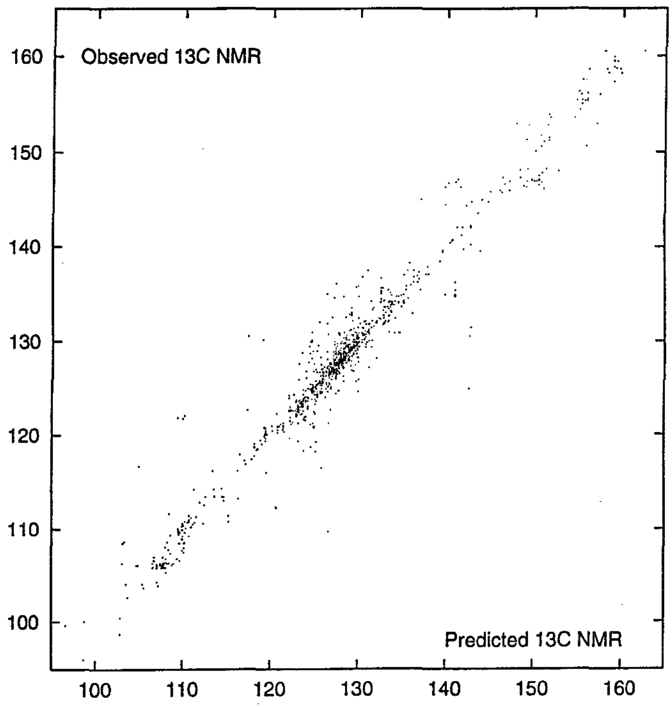

Figure 1. Correlation between 2100 pairs of observed and predicted alpha signals. r = .98

program reduced the number of unknowns to 272. This accounts for the 40 blank entries in Table 2. Such a system of equations is virtually certain to be inconsistent and cannot therefore be solved by usual methods. We computed the best solution in the least squares sense by computing the values of the 272 unknowns which would minimize SSE, the sum of the squares of the errors of prediction.

$$SSE = \sum_{k=1}^{2100} (YOBS_k - YPRED_k)^2$$

This results in 272 α SSCS values. Using these α SSCS values to predict the 2100 signals results in a correlation between observed and predicted nmr signals of r = .98. The standard deviation is s = 2.1 ppm which gives 1.4 ppm as the average error of prediction. This correlation is shown in Figure 1.

Applying the same procedure for the 2100 β signals resulted in 272 β SSCS values in Table 3. The correlation between observed and predicted nmr signals for β signals is r = .98 with a standard deviation of s = 2.0 ppm. The average error of prediction for β signals is 1.3 ppm. This correlation is shown in Figure 2.

Table 3. Group numbers, group names and beta SSCS values for groups in positions 1- 8. These SSCS values are to be used in the prediction of beta signals with the carbon associated with the signal assumed to be in position 2. The numbered positions for SSCS values in the table are numbered with respect to that assumption. The predicted signal from position 2 in naphthalene is 126.49 ppm.

| No. | Code | Name | Position | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | . | H | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 |
| 2 | A | Br | 1.95 | -5.49 | 1.05 | -.34 | .46 | .08 | .57 | 1.63 |
| 3 | B | Cl | .34 | 4.87 | -.04 | -.52 | .54 | .14 | 1.01 | 1.74 |
| 4 | C | NH2 | -14.46 | 13.83 | -8.38 | 1.21 | -1.00 | -3.34 | .92 | -.76 |
| 5 | D | CH2CH3 | | 15.39 | .71 | | | -.54 | -.28 | |
| 6 | E | F | -17.05 | 34.03 | -10.00 | -.64 | .53 | -1.52 | .49 | .19 |
| 7 | F | CN | 5.83 | -17.18 | -.23 | -1.73 | .49 | 2.42 | .86 | 1.86 |
| 8 | G | OH | -17.33 | 25.65 | -8.66 | .01 | -.10 | -2.19 | .68 | -.90 |
| 9 | H | I | 13.50 | -35.13 | 8.16 | .52 | .60 | .66 | .74 | 3.52 |
| 10 | I | CF3 | | 1.72 | -4.52 | | | 2.55 | 1.65 | |
| 11 | J | OCH3 | -20.77 | 30.09 | -8.23 | 1.13 | -.48 | -2.56 | .10 | -.77 |
| 12 | K | CH3 | .71 | 8.42 | 1.64 | -.51 | -.83 | -.90 | -.27 | -.05 |
| 13 | L | NO2 | -3.89 | 19.73 | -7.19 | -1.28 | 1.84 | 3.69 | 1.93 | 3.38 |
| 15 | M | COCH3 | -3.03 | 13.95 | -3.48 | -1.42 | -.21 | 1.05 | .70 | .87 |
| 16 | N | CH(CH3)2 | -2.71 | | | 1.77 | -1.50 | | | -1.22 |
| 17 | O | C(CH3)3 | -3.22 | 22.05 | -1.15 | -1.47 | -1.87 | -.10 | -.86 | -1.76 |
| 18 | P | N(CH3)2 | -9.01 | 22.67 | -9.66 | .69 | -.65 | -4.19 | -.03 | -2.97 |
| 20 | Q | CH2Br | .91 | 8.47 | 1.36 | -1.49 | -.69 | .50 | -.04 | -.29 |
| 21 | R | CH2OH | -1.49 | | | -1.19 | -.89 | | | -.49 |
| 40 | S | NHCOCH3 | | 9.24 | -6.19 | | | -.54 | .32 | |
| 41 | T | SiMe3 | 6.61 | 11.31 | 3.31 | -.99 | -1.39 | -.29 | -.79 | -1.29 |
| 65 | U | SO3H | -.89 | 12.89 | -3.09 | .84 | 2.12 | 2.51 | .59 | 4.14 |
| 72 | V | GeMe3 | 5.51 | 13.11 | 3.21 | -.89 | -1.19 | -.79 | -.79 | -1.39 |
| 73 | W | PbMe3 | 8.11 | 19.81 | 6.91 | -.69 | -.69 | -.89 | -.99 | -1.19 |
| 74 | X | SnMe3 | 2.42 | 13.81 | 1.58 | 2.19 | -.82 | 1.65 | -.34 | -1.10 |
| 81 | Y | OCOMe | -8.49 | 22.23 | -5.33 | -1.19 | -.19 | -.26 | .38 | -.19 |
| 117 | Z | COOH | 3.23 | 1.43 | -1.06 | -2.18 | -.75 | 1.53 | .22 | .63 |
| 125 | a | CH2COOH | 1.31 | | | -1.39 | -.99 | | | -.29 |
| 161 | b | CHO | 6.55 | 7.90 | -2.42 | 2.92 | 1.00 | 2.57 | .69 | 1.89 |
| 163 | c | COOMe | 3.91 | 1.06 | -1.06 | -1.79 | -.09 | 2.08 | .48 | 1.31 |
| 263 | d | CONMe2 | -2.69 | | | -1.39 | -.19 | | | .41 |
| 271 | e | CH2CMe3 | 2.41 | 10.97 | 3.10 | -1.29 | -1.29 | -1.03 | -.57 | -1.29 |
| 276 | f | CH2PbMe3 | -2.89 | 15.61 | .31 | -.99 | -1.19 | -2.59 | -.89 | -1.49 |
| 292 | g | CH2SnMe3 | -3.03 | 14.21 | -1.35 | -.76 | -1.14 | -2.35 | -.63 | -1.80 |
| 293 | h | CMe2OH | -3.89 | | | -1.89 | -1.29 | | | -1.49 |
| 306 | i | CMe=CH2 | -1.99 | | | -1.29 | -.89 | | | -.69 |
| 308 | j | COtBu | 3.31 | | | -1.89 | -.39 | | | .91 |
| 326 | k | FORMYL | 9.77 | 7.72 | -2.16 | -3.62 | .28 | 2.57 | .52 | 2.77 |
| 331 | l | CH2P(O)(OEt)2 | 1.60 | 4.01 | 2.11 | .29 | -.50 | -.49 | .01 | -.16 |

Predicting the γ signals was somewhat problematic. In Part 1 we gave no prediction equation. It appears to us that many of the γ signals are misinterpreted. Because they are close together (about 130 ppm), it is difficult to tell which is 9 and which is 10. Errors of measurement are probably a significant percentage of the actual signal variation, and so we did not expect to get a high correlation, but we expected better than what we got. There is almost no correlation between γ signals and the substituted groups, which seems to say that the signals are affected by something, but it is not the substituted groups. This clearly cannot be true.

An examination of the predicted 9 and 10 signals compared to the reported observed signals seemed to indicate that a large number of them had been interpreted in reverse order. We decided to develop an algorithm that would determine the values of $SG_{ij}$ which minimized the SSE, and at the same time switch signals 9 or 10 if that would lower SSE. Developing a working algorithm to meet these criteria was perhaps the most difficult achievement in this work. The algorithm we developed is of the iterative type and will produce an answer only in case it converges, which it does.

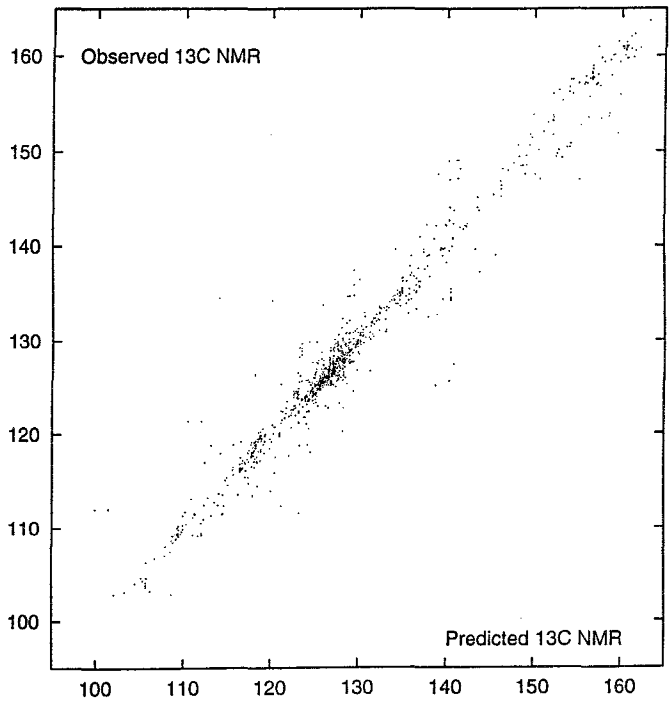

Figure 2. Correlation between 2100 pairs of observed and predicted beta signals. r = .98

This algorithm did reverse the order of many of the pairs of $\gamma$ signals. Ultimately we calculated a set of 272 $\gamma$ SSCS values for use in predicting $\gamma$ signals. It should be noted that in using these SSCS values for prediction the molecule should be oriented so that the carbon yielding the signal is in position 9. Because of symmetry, there are only two orientations and therefore two encodings for $\gamma$ signals. It should also be noted that symmetry dictates that $SG_{1j} = SG_{8j}$, $SG_{2j} = SG_{7j}$, etc., so that each group has only 4 different $\gamma$ SSCS values. In predicting $\gamma$ signals, one should also be aware that because of the reinterpretation algorithm used for calculating $\gamma$ SSCS values, we are not as sure of their reliability as we are for $\alpha$ and $\beta$ SSCS values. However comparison of similar groups show a reasonable pattern to the $\gamma$ SSCS values. The correlation between predicted and observed $\gamma$ signals is $r = .93$ with a standard deviation of $s = 1.5$. Thus the average error prediction for $\gamma$ signals is about 1.0 ppm, comparable to that for $\alpha$ and $\beta$ signals. This correlation is shown in Figure 3.

We tested the predictive value of this method on a group of uninterpreted naphthalenes from [56] One of these uninterpreted naphthalenes was 1-hydroxy-2,4-dichloronaphthalene. We will use it as an example to show how the SSCS calculations are to be done. First the $\alpha$ signals receive 4 encodings. From table 1, 2 or 3 we see that the OH group is group number 8, while Cl is number 3. Thus the four encodings for $\alpha$ signals are :

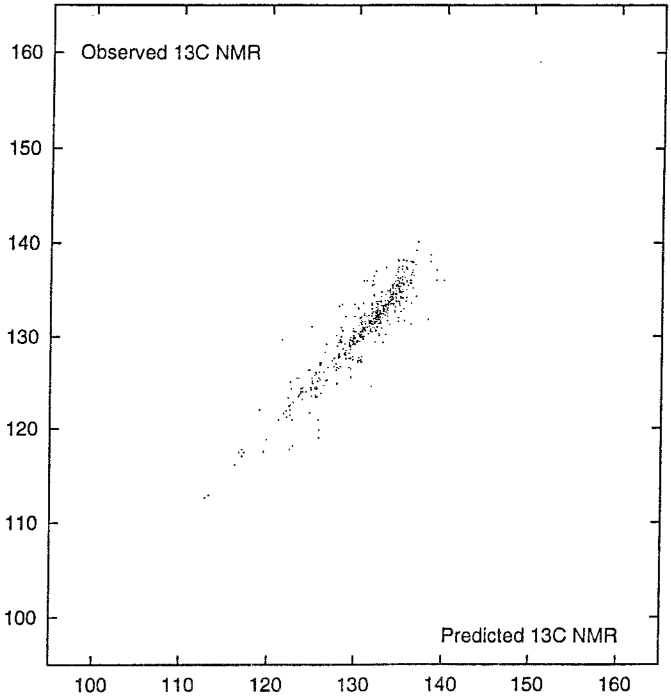

Figure 3. Correlation between 1050 pairs of observed and predicted gamma signals. r = .93

```
8 3 1 3 1 1 1 1 * *
3 1 3 8 1 1 1 1 * *
1 1 1 1 3 1 3 8 * *
1 1 1 1 8 3 1 3 * *
```

The computation for the first α signal is:

$$
\begin{aligned}
\mathrm{YPRED} &= SA_{1,8} + SA_{2,3} + SA_{3,1} + SA_{4,3} \\
&\quad + SA_{5,1} + SA_{6,1} + SA_{7,1} + SA_{8,1} + 127.92 \\
&= 23.72 - 1.96 + 0.00 - 0.90 \\
&\quad + 0.00 + 0.00 + 0.00 + 0.00 + 127.92 \\
&= 148.8 \text{ ppm}
\end{aligned}
$$

For the second α signal:

Table 4. Group numbers, group names and gamma SSCS values for groups in positions 1 - 8. These values are to be used for predicting gamma signals with the assumption that the carbon associated with the signal is in position 9. The numbered positions in the table are numbered with respect to that assumption. The predicted signal from position 9 in naphthalene is 133.64 ppm.

| No. | Code | Name | Position | | | |
|---|---|---|---|---|---|---|
| | | | 1,8 | 2,7 | 3,6 | 4,5 |
| 1 | . | H | .00 | .00 | .00 | .00 |
| 2 | A | Br | -1.82 | .39 | -.95 | .81 |
| 3 | B | Cl | -2.02 | -1.45 | .20 | .51 |
| 4 | C | NH2 | 1.32 | 1.22 | -5.68 | -8.33 |
| 5 | D | CH2CH3 | | -.09 | -1.50 | |
| 6 | E | F | -10.01 | .19 | -3.12 | 1.19 |
| 7 | F | CN | -1.22 | -1.43 | .81 | -1.44 |
| 8 | G | OH | .59 | .76 | -4.45 | -8.44 |
| 9 | H | I | .84 | 1.64 | -1.07 | .84 |
| 10 | I | CF3 | | -.83 | 1.17 | |
| 11 | J | OCH3 | -4.24 | .51 | -4.31 | -1.67 |
| 12 | K | CH3 | -.63 | .32 | -1.44 | .52 |
| 13 | L | NO2 | -8.73 | -1.71 | 2.51 | .46 |
| 15 | M | COCH3 | -.85 | -1.89 | .99 | -4.01 |
| 16 | N | CH(CH3)2 | -.26 | | | -.20 |
| 17 | O | C(CH3)3 | -1.60 | .45 | -1.47 | 2.04 |
| 18 | P | N(CH3)2 | -5.62 | 2.02 | -6.24 | 2.37 |
| 20 | Q | CH2Br | -2.84 | -.76 | -.61 | .06 |
| 21 | R | CH2OH | -1.34 | | | -1.34 |
| 40 | S | NHCOCH3 | | .21 | -2.70 | |
| 41 | T | SiMe3 | 1.96 | -.08 | -.08 | 1.96 |
| 65 | U | SO3H | -1.88 | 1.00 | 1.00 | -1.53 |
| 72 | V | GeMe3 | 2.06 | -.29 | -.29 | 2.06 |
| 73 | W | PbMe3 | 3.11 | -.14 | -.14 | 3.11 |
| 74 | X | SnMe3 | 3.30 | -.13 | -1.17 | 1.03 |
| 81 | Y | OCOMe | .86 | -1.57 | -.41 | -6.94 |
| 117 | Z | COOH | -2.24 | -1.22 | 1.52 | -.15 |
| 125 | a | CH2COOH | -.14 | | | -1.94 |
| 161 | b | CHO | -.37 | 5.42 | -3.87 | -2.97 |
| 163 | c | COOMe | .46 | -1.27 | 2.19 | -1.94 |
| 263 | d | CONMe2 | -.24 | | | -4.14 |
| 271 | e | CH2CMe3 | -.34 | -1.23 | -.71 | .26 |
| 276 | f | CH2PbMe3 | -2.94 | -1.44 | -1.44 | .06 |
| 292 | g | CH2SnMe3 | -2.89 | .32 | -3.10 | .83 |
| 293 | h | CMe2OH | -.79 | | | -.79 |
| 306 | i | CMe=CH2 | -2.64 | | | .26 |
| 308 | j | COtBu | .36 | | | -2.14 |
| 326 | k | FORMYL | -5.42 | 6.29 | -4.89 | 1.28 |
| 331 | l | CH2P(O)(OEt)2 | -1.00 | .26 | -.94 | .78 |

$$
\begin{aligned}
\text{YPRED} &= SA_{1,3} + SA_{2,1} + SA_{3,3} + SA_{4,8} \\
&\quad + SA_{5,1} + SA_{6,1} + SA_{7,1} + SA_{8,1} + 127.92 \\
&= 4.06 + 0.00 + 1.40 - 8.66 \\
&\quad + 0.00 + 0.00 + 0.00 + 0.00 + 127.92 \\
&= 124.7 \text{ ppm}
\end{aligned}
$$

This should be sufficient for the reader to see how the other 8 signals are predicted. In these two examples the SSCS values are from Table 2, since these are α signals. For β signals and γ signals use SSCS values from Tables 3 and 4 respectively.

After prediction of the ten signals, we "freeze" the encoding for the compound at the first α encoding for purposes of listing which predicted signal goes with which carbon. You will see the two above computed predicted signals in positions 1 and 4 in the frozen encoding.

Table 5. Interpretation of 53 uninterpreted substituted naphthalenes from [56]. The observed signals from [56] have been matched with signals predicted from the SSCS values in Tables 2, 3 and 4 by minimizing SSE. The SSE column shows the minimum SSE value over all 10! possible permutations of the observed signals. The letter encoding of the compounds is explained in Table 1.

| Compound | Position | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | | 7 | | 8 | | 9 | | 10 | | |
| | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | PRED | OBS | SSE |
| K....... | 135.1 | 134.2 | 127.2 | 126.5 | 126.0 | 125.5 | 126.2 | 125.7 | 128.4 | 128.5 | 125.7 | 125.5 | 126.4 | 126.3 | 125.3 | 124.1 | 133.0 | 132.6 | 134.2 | 133.6 | 3.8 |
| .K...... | 126.9 | 127.1 | 134.9 | 135.3 | 128.1 | 128.1 | 127.6 | 127.5 | 127.7 | 127.6 | 125.6 | 124.9 | 126.2 | 125.8 | 126.8 | 126.8 | 134.0 | 133.6 | 132.2 | 131.6 | 1.4 |
| .D...... | 125.5 | 124.9 | 141.9 | 141.7 | 127.2 | 127.3 | 127.6 | 127.5 | 126.6 | 127.0 | 125.9 | 125.5 | 126.2 | 125.7 | 127.6 | 127.7 | 133.6 | 133.7 | 132.1 | 131.9 | 1.1 |
| K.K..... | 134.8 | 134.9 | 128.8 | 128.9 | 134.4 | 133.8 | 125.1 | 124.7 | 127.3 | 127.8 | 125.4 | 125.2 | 125.5 | 125.5 | 125.1 | 123.8 | 131.6 | 130.8 | 134.5 | 133.9 | 3.4 |
| A....... | 123.1 | 122.7 | 128.4 | 129.8 | 126.2 | 126.0 | 127.3 | 127.2 | 128.2 | 128.2 | 126.9 | 126.6 | 128.1 | 127.8 | 127.0 | 127.0 | 131.8 | 131.9 | 134.4 | 134.5 | 2.3 |
| .B...... | 126.0 | 126.0 | 131.4 | 131.5 | 126.4 | 126.5 | 129.3 | 129.4 | 127.9 | 127.7 | 126.6 | 126.7 | 127.5 | 127.0 | 127.1 | 126.8 | 132.2 | 131.6 | 133.8 | 134.0 | .8 |
| .A...... | 129.6 | 129.5 | 121.0 | 119.7 | 127.5 | 127.7 | 129.8 | 129.8 | 128.1 | 129.1 | 126.6 | 126.1 | 127.1 | 126.8 | 127.5 | 126.9 | 134.0 | 134.4 | 132.7 | 131.7 | 4.5 |
| Q....... | 132.0 | 133.1 | 127.4 | 127.6 | 125.0 | 125.3 | 129.3 | 129.7 | 128.5 | 128.7 | 125.8 | 126.1 | 126.2 | 126.5 | 123.4 | 123.6 | 130.8 | 130.9 | 133.7 | 133.9 | 1.8 |
| A..K.... | 121.4 | 120.6 | 127.9 | 129.4 | 126.9 | 126.3 | 134.5 | 134.2 | 125.6 | 124.4 | 126.9 | 126.8 | 127.3 | 126.8 | 127.5 | 127.6 | 132.3 | 131.7 | 133.8 | 133.7 | 5.4 |
| J....... | 156.0 | 155.3 | 105.7 | 103.7 | 127.6 | 125.8 | 119.4 | 120.1 | 127.7 | 126.3 | 126.0 | 125.5 | 125.7 | 125.1 | 122.9 | 121.9 | 129.4 | 127.4 | 132.0 | 134.4 | 21.7 |
| .J...... | 107.9 | 105.6 | 156.6 | 157.5 | 118.3 | 118.6 | 129.0 | 128.9 | 127.8 | 127.6 | 123.9 | 123.5 | 126.6 | 126.3 | 127.1 | 126.6 | 134.1 | 134.5 | 129.3 | 129.3 | 7.0 |
| .A...J.. | 129.5 | 129.5 | 118.4 | 116.9 | 127.6 | 128.3 | 128.9 | 129.5 | 108.1 | 105.6 | 156.7 | 157.7 | 118.8 | 119.7 | 128.6 | 128.4 | 129.7 | 129.9 | 133.2 | 132.9 | 11.4 |
| G....... | 151.6 | 151.1 | 109.2 | 108.7 | 126.5 | 126.4 | 119.3 | 120.7 | 128.1 | 127.6 | 126.4 | 125.8 | 125.6 | 125.2 | 122.2 | 121.4 | 134.2 | 134.7 | 125.2 | 124.3 | 5.0 |
| GK...... | 150.6 | 148.4 | 117.6 | 116.3 | 128.1 | 128.9 | 119.0 | 120.1 | 127.8 | 127.5 | 125.5 | 125.3 | 125.3 | 125.2 | 121.1 | 120.8 | 134.6 | 133.3 | 123.8 | 124.1 | 10.2 |
| G..B.... | 150.7 | 150.4 | 108.6 | 108.6 | 126.8 | 126.0 | 123.3 | 124.4 | 125.8 | 125.4 | 128.1 | 127.5 | 126.1 | 125.7 | 123.0 | 122.0 | 134.7 | 131.5 | 123.2 | 123.5 | 14.4 |
| GB.B.... | 148.8 | 147.9 | 113.5 | 113.7 | 126.8 | 126.4 | 124.7 | 123.8 | 125.8 | 126.4 | 128.3 | 127.3 | 127.1 | 126.5 | 122.3 | 121.5 | 133.3 | 129.6 | 123.4 | 122.7 | 18.2 |
| G..J.... | 143.2 | 146.6 | 110.3 | 107.1 | 105.7 | 104.1 | 147.3 | 147.8 | 123.1 | 124.9 | 125.6 | 125.4 | 125.1 | 125.1 | 121.9 | 122.0 | 132.6 | 125.8 | 121.0 | 121.3 | 74.0 |
| AG...... | 104.8 | 106.1 | 154.1 | 150.5 | 117.5 | 117.1 | 128.8 | 128.1 | 128.1 | 127.7 | 124.8 | 124.0 | 128.8 | 129.2 | 125.6 | 125.2 | 132.6 | 132.2 | 130.0 | 129.6 | 16.7 |
| .G...A.. | 109.7 | 109.5 | 152.2 | 153.5 | 118.4 | 117.1 | 128.9 | 129.7 | 129.6 | 129.7 | 118.8 | 118.7 | 128.2 | 127.9 | 128.4 | 128.9 | 133.4 | 132.9 | 129.6 | 129.9 | 4.7 |
| AG...A.. | 104.9 | 106.0 | 154.2 | 150.8 | 118.1 | 117.9 | 128.3 | 128.3 | 129.9 | 130.5 | 119.3 | 118.2 | 129.9 | 130.0 | 127.5 | 127.1 | 131.6 | 130.9 | 130.4 | 130.8 | 14.9 |
| .G....J. | 108.7 | 108.8 | 152.2 | 153.9 | 115.3 | 115.2 | 129.2 | 129.5 | 128.9 | 129.2 | 116.1 | 116.2 | 157.3 | 158.1 | 106.6 | 104.7 | 134.9 | 135.9 | 124.9 | 124.3 | 6.5 |
| GG...... | 133.3 | 128.3 | 134.8 | 137.4 | 117.8 | 118.3 | 120.7 | 118.7 | 128.0 | 127.2 | 124.2 | 124.4 | 126.3 | 125.6 | 120.8 | 122.6 | 135.0 | 140.0 | 120.8 | 120.7 | 64.9 |
| G.G..... | 153.1 | 155.8 | 100.5 | 100.2 | 152.1 | 154.3 | 100.9 | 100.9 | 126.8 | 125.5 | 127.1 | 126.2 | 123.4 | 121.0 | 122.1 | 119.8 | 129.8 | 135.4 | 126.0 | 122.0 | 72.8 |
| R....... | 136.2 | 137.7 | 125.0 | 124.1 | 125.3 | 125.3 | 128.1 | 127.1 | 128.5 | 128.2 | 125.6 | 125.4 | 126.0 | 125.7 | 123.5 | 123.6 | 132.3 | 133.1 | 132.3 | 130.6 | 7.8 |
| P....... | 150.7 | 150.7 | 117.5 | 113.8 | 127.2 | 125.7 | 123.4 | 122.8 | 127.6 | 128.2 | 125.8 | 125.0 | 123.5 | 124.0 | 126.3 | 125.6 | 128.0 | 128.7 | 136.0 | 134.7 | 20.0 |
| G...C... | 152.9 | 153.0 | 108.2 | 107.7 | 125.7 | 123.7 | 116.3 | 112.8 | 142.9 | 144.1 | 111.9 | 107.7 | 126.8 | 125.5 | 112.4 | 109.6 | 125.9 | 124.0 | 126.5 | 125.0 | 51.4 |
| .GC..... | 110.7 | 106.6 | 143.8 | 145.9 | 131.7 | 138.3 | 111.6 | 107.9 | 125.5 | 124.4 | 125.2 | 122.6 | 123.8 | 121.1 | 126.1 | 125.3 | 128.7 | 127.3 | 130.4 | 129.4 | 98.8 |
| L....... | 150.5 | 146.4 | 122.6 | 123.0 | 125.2 | 125.0 | 132.5 | 134.2 | 128.8 | 128.5 | 128.3 | 127.2 | 129.9 | 129.3 | 123.7 | 123.9 | 124.9 | 124.0 | 134.1 | 134.6 | 22.9 |
| LK...... | 149.5 | 147.7 | 131.0 | 128.4 | 126.9 | 126.5 | 132.2 | 130.3 | 128.5 | 127.9 | 127.4 | 127.4 | 129.6 | 128.0 | 122.6 | 121.1 | 125.2 | 124.5 | 132.7 | 132.1 | 19.6 |
| B......L | 127.7 | 127.4 | 130.2 | 130.8 | 127.8 | 127.8 | 127.9 | 128.0 | 133.3 | 132.2 | 125.8 | 125.1 | 124.3 | 123.0 | 148.2 | 147.2 | 122.9 | 121.3 | 134.6 | 135.8 | 8.9 |
| J..L.... | 160.6 | 160.5 | 104.4 | 101.8 | 123.7 | 125.4 | 142.0 | 139.0 | 123.4 | 123.3 | 129.4 | 127.1 | 127.6 | 126.7 | 123.7 | 126.4 | 129.9 | 129.9 | 123.2 | 122.7 | 32.3 |
| GL...... | 148.6 | 155.5 | 128.9 | 127.7 | 119.3 | 119.2 | 122.1 | 120.1 | 129.0 | 128.2 | 130.1 | 131.3 | 127.5 | 125.0 | 124.6 | 124.9 | 132.5 | 137.3 | 127.7 | 127.0 | 85.5 |
| C..L.... | 147.3 | 153.6 | 110.8 | 105.3 | 123.8 | 123.3 | 140.7 | 131.7 | 124.9 | 123.5 | 128.9 | 130.2 | 127.6 | 127.9 | 125.8 | 124.8 | 135.4 | 130.3 | 116.6 | 120.6 | 198.7 |
| L.L..... | 153.3 | 147.0 | 115.4 | 117.4 | 144.9 | 143.3 | 129.4 | 129.4 | 131.1 | 130.8 | 130.3 | 129.6 | 133.6 | 133.2 | 124.6 | 123.6 | 127.4 | 127.1 | 132.4 | 133.1 | 48.7 |
| GL.L...G | 147.4 | 161.4 | 126.7 | 127.2 | 115.3 | 117.2 | 144.8 | 133.3 | 116.1 | 126.3 | 133.5 | 130.0 | 112.0 | 113.0 | 149.1 | 169.1 | 133.6 | 130.6 | 110.5 | 111.7 | 859.5 |
| M....... | 140.9 | 135.2 | 123.5 | 124.2 | 125.1 | 126.3 | 129.1 | 130.0 | 127.9 | 128.6 | 126.3 | 127.9 | 127.4 | 128.3 | 124.0 | 125.9 | 132.8 | 133.8 | 129.6 | 132.9 | 55.3 |
| .M...... | 124.6 | 126.7 | 140.4 | 135.5 | 123.0 | 123.8 | 128.2 | 129.5 | 127.1 | 127.7 | 127.5 | 128.3 | 127.2 | 128.3 | 128.5 | 130.0 | 131.8 | 132.4 | 134.6 | 134.4 | 36.2 |
| .M...K.. | 124.3 | 126.7 | 139.5 | 138.6 | 122.7 | 123.9 | 127.1 | 129.0 | 126.0 | 127.6 | 136.0 | 135.8 | 128.8 | 129.9 | 128.2 | 129.3 | 130.3 | 130.7 | 135.0 | 133.8 | 17.9 |
| MG...... | 122.6 | 119.8 | 149.1 | 163.9 | 116.4 | 114.8 | 130.6 | 131.9 | 127.8 | 128.5 | 124.1 | 124.3 | 128.0 | 129.5 | 122.6 | 123.7 | 133.5 | 137.3 | 125.2 | 128.0 | 256.5 |
| GM...... | 148.3 | 162.3 | 123.1 | 124.8 | 123.0 | 124.3 | 119.5 | 113.1 | 127.3 | 127.3 | 127.4 | 129.9 | 126.3 | 125.8 | 122.7 | 118.2 | 132.3 | 137.2 | 126.2 | 125.1 | 293.6 |
| bJ...... | 116.2 | 112.4 | 163.1 | 163.8 | 121.2 | 116.3 | 131.0 | 131.4 | 129.8 | 129.7 | 124.9 | 124.8 | 128.5 | 128.3 | 124.5 | 124.6 | 133.8 | 137.5 | 126.4 | 128.2 | 56.5 |
| bG...... | 117.9 | 111.8 | 158.7 | 164.0 | 120.8 | 118.7 | 131.4 | 132.1 | 129.8 | 128.9 | 125.3 | 124.1 | 129.1 | 128.8 | 124.0 | 120.6 | 134.0 | 138.4 | 126.2 | 127.5 | 104.3 |
| a....... | 129.4 | 131.3 | 127.8 | 127.3 | 125.1 | 125.3 | 127.9 | 127.8 | 128.4 | 128.3 | 125.5 | 125.5 | 126.2 | 125.9 | 123.3 | 123.8 | 133.5 | 133.4 | 131.7 | 131.9 | 4.3 |
| .ZG..... | 132.2 | 132.4 | 119.3 | 115.2 | 151.1 | 156.0 | 109.5 | 110.7 | 126.2 | 125.9 | 128.7 | 129.2 | 124.5 | 123.7 | 128.7 | 128.9 | 128.0 | 126.6 | 135.9 | 137.2 | 46.9 |
| Z....... | 127.7 | 127.7 | 129.7 | 129.8 | 124.3 | 124.8 | 132.4 | 132.8 | 129.6 | 128.5 | 125.7 | 125.4 | 127.1 | 127.4 | 125.8 | 126.1 | 131.4 | 130.6 | 133.5 | 133.4 | 2.4 |
| .Z...... | 130.8 | 130.5 | 127.9 | 127.9 | 125.4 | 125.2 | 127.8 | 127.8 | 127.6 | 127.5 | 128.0 | 128.2 | 126.7 | 126.5 | 128.8 | 129.0 | 132.4 | 132.1 | 135.2 | 134.9 | .4 |
| GZ...... | 154.5 | 160.5 | 110.6 | 105.9 | 125.4 | 124.2 | 119.2 | 118.0 | 127.8 | 127.4 | 127.9 | 129.0 | 125.8 | 124.8 | 123.0 | 123.2 | 133.0 | 136.7 | 126.7 | 125.5 | 78.7 |
| ZG...... | 109.3 | 105.7 | 155.4 | 162.8 | 115.7 | 118.9 | 133.9 | 135.7 | 129.5 | 128.6 | 123.5 | 123.2 | 127.8 | 127.9 | 124.5 | 125.0 | 132.2 | 131.8 | 129.0 | 128.0 | 84.6 |
| .ZG..... | 132.2 | 132.4 | 119.3 | 115.2 | 151.1 | 156.0 | 109.5 | 110.7 | 126.2 | 125.9 | 128.7 | 129.2 | 124.5 | 123.7 | 128.7 | 128.9 | 128.0 | 126.6 | 135.9 | 137.2 | 46.9 |
| GZ.G.... | 145.8 | 153.2 | 110.6 | 105.2 | 108.1 | 104.5 | 142.9 | 128.9 | 122.0 | 114.7 | 127.0 | 126.0 | 125.7 | 124.8 | 123.2 | 122.0 | 124.6 | 123.0 | 127.3 | 128.3 | 352.9 |
| .ZG.G... | 132.4 | 128.9 | 119.2 | 115.3 | 150.2 | 155.2 | 103.7 | 105.6 | 149.9 | 151.4 | 111.4 | 110.2 | 124.5 | 127.9 | 120.0 | 124.1 | 119.5 | 119.7 | 136.5 | 132.0 | 107.6 |
| Z......Z | 125.6 | 126.6 | 130.4 | 128.0 | 123.6 | 125.1 | 134.1 | 133.6 | 134.1 | 131.4 | 123.6 | 119.8 | 130.4 | 128.7 | 125.6 | 126.3 | 129.2 | 126.7 | 133.3 | 129.9 | 51.5 |
| .Y...... | 118.5 | 118.4 | 148.7 | 148.3 | 121.2 | 121.1 | 129.1 | 129.3 | 127.1 | 127.6 | 126.2 | 125.6 | 126.9 | 126.5 | 127.8 | 127.7 | 132.1 | 131.4 | 133.2 | 133.7 | 1.7 |

| **Position** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Group** | **8** | **3** | **1** | **3** | **1** | **1** | **1** | **1** | **•** | **•** |
| **YPRED** | **148.8** | **113.5** | **126.8** | **124.7** | **125.8** | **128.3** | **127.1** | **122.3** | **133.3** | **123.4** |

The compounds from [56] give ten nmr signals, but no indication of which one goes with which position. The predicted signals allow one to match these ten observed signals up with the closest predicted signals. This is difficult to do correctly manually. We wrote an algorithm which is programmable to match the observed with predicted signals so as to give the closest fit possible. It is a brute force method in that it simply has the computer look at all the 10! possible interpretations and pick out the best match.

An explanation of the details of this method requires that we first define some notation. Let the 10 predicted signals be denoted by $YPRED_i$ $i=1,10$ and $Y_i$ $i = 1,10$ any arbitrary ordering of the 10 observed signals, which we will regard as the initial ordering. Then any reordering of the observed signals has the form $Y_{p(i)}$ $i=1,10$ where p is any one of the 10! permutations of the indices 1 through 10. For any such permutation p we can compute:

$$SSE(p) = \sum_{i=1}^{10} (\ YPRED_i - Y_{P(\,i\,)}\ )^2$$

The value of SSE(p) depends on the permutation p, and we need to find which permutation p minimizes SSE(p). This permutation gives the best fit and thus determines the order of observed signals which match the predicted signals best. The minimum SSE(p) can be found by computing SSE(p) for each of the 10! permutations and choosing the smallest one. This is a fair challenge to even a fast computer, but can be done.

By this method we were able to interpret the nmr signals of the 53 naphthalenes in [56]. These compounds together with our interpretation of the nmr signal are in Table 5. The number labeled SSE is the minimum of the SSE(p) described above. The smaller the value of SSE the better the fit of observed to predicted nmr. There is no standard for interpretation of SSE with respect to this method, because this is, as far as we know, a new algorithm. We believe that in this case since the average error of prediction is about 1.5 ppm per signal that an SSE under $10(1.5^2\ ) = 22.5$ represents a chemically useful interpretation. SSE under 100 *probably* indicates a correct interpretation.

## EXPERIMENTAL

Statistical computations were done on a Sun SPARC-10 using Statistical Analysis Software and on a Pentium Pro-200 computer running the Linux operating system.

## Acknowledgments

We thank the Auburn University at Montgomery Grant-in-Aid program, together with the Chemistry and Mathematics Departments for support.

## References

1. C.A.L. Mahaffy, J.R. Nanney and R.E. Jetton, *J. Molecular Structure (Theochem)*, submitted November 1997.
2. C. Hansch, A. Leo, S.H. Unger, K.H. Kim, D. Nikaitani,and E.J. Lien, *J. Med. Chem.*, *16* (1973) 1207.
3. C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979.

4. M. Charton, *J. Am. Chem. Soc., 97* (1975) 1552.
5. Idem., *ibid., 97* (1975) 3691 and *97* (1975) 3694.
6. Idem., *J. Org. Chem., 41* (1976) 2217.
7. Idem., in *Design of Biopharmaceutical Properties through Prodrugs and Analogues,* E. B. Roche (Ed.), Am. Pharm. Assoc., Washington, DC, 1977.
8. C. Hansch, A. Leo and R.W. Taft, *Chem. Rev. 91* (1991) 165.
9. C.G. Swain and E.C. Lupton, *J. Am. Chem. Soc., 90* (1968) 4328.
10. J.R. Nanney, R.E. Jetton and C.A.L. Mahaffy, *J. Molecular Structure (Theochem), 388* (1996) 19.
11. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, *Abstracts from the 4th Conference on Current Trends in Computational Chemistry,* Vicksburg, Mississippi, November, 1994, page 97.
12. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, *J. Molecular Structure (Theochem), 388* (1996) 139.
13. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, *Abstracts from the 6th Conference on Current Trends in Computational Chemistry,* Vicksburg, Mississippi, November, 1996.
14. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, *J. Molecular Structure (Theochem)*, 425 (1998) 29
15. J.R. Nanney, A.K. Traylor and C.A.L. Mahaffy, *J. Fluorine Chem., 64* (1993) 217.
16. J.R. Nanney, R.E. Jetton and C.A.L. Mahaffy, *J. Fluorine Chem., 80* (1996) 105.
17. B. Armstrong, A. Grier, J.B. Hamilton, H. Khuu, C.A.L. Mahaffy, J. Rawlings and J.R. Nanney, *Inorg. Chim. Acta, 191* (1992) 189.
18. J.R. Nanney and C.A.L. Mahaffy, *Inorg. Chim. Acta, 201* (1992) 55.
19. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, Electronic Journal of Theoretical Chemistry, 2 (1996) 24.
20. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, Abstracts from the 3rd Conference on Current Trends in Computational Chemistry, Vicksburg, Mississippi, November, 1994, page 97.
21. R.E. Jetton, J.R. Nanney, C.A.L. Mahaffy and C.C. Price, Structural Chem., 6 (1995) 301.
22. C.A.L. Mahaffy and J.R. Nanney, J. Fluorine Chem., 67 (1994) 67.
23. J.R. Nanney and C.A.L. Mahaffy, ibid., 68 (1994) 181.
24. C.A.L. Mahaffy and J.R. Nanney, ibid., 68 (1994) 165.
25. R.E. Jetton, J.R. Nanney and C.A.L. Mahaffy, ibid., 72 (1995) 121.
26. D.K. Dalling, K.H. Ladner, D.M. Grant and W.R. Woolfenden, *ibid, 99* (1977) 7142.
27. A. Bulbarela, H. Tlahuext, H.R. Morales, L. Cuéllar, G. Uribe and R. Contreras, *Magn. Reson. Chem., 24* (1986) 1093.
28. D. Doddrell, M. Barfield, W. Adcock, M. Aurangzeb and D. Jordan, *JCS Perkin II,* (1976) 402.
29. E. Breitmaier, G. Haas and W. Voelter, Atlas of Carbon-13 NMR Data, Heyden, London.
30. W. Adcock, W. Kitching, V. Alberts, G. Wickham, P. Barron and D. Doddrell, *Org. Magn. Reson. 10* (1977) 47.
31. L. Ernst, *Org. Magn. Reson., 9* (1977) 35.
32. M. Bullpitt, W. Kitching, W. Adcock and D. Doddrell, *J. Organometal. Chem., 116* (1976) 161.
33. M. Bullpitt, W. Kitching, W. Adcock and D. Doddrell, *ibid, 116* (1976) 187.
34. L. Ernst, *J. Mag. Reson., 22* (1976) *279 and 22* (1976) 402.
35. W. Kitching, M. Bullpitt, D. Gartshore, W. Adcock, T.C. Kohr, D. Doddrell and I.D. Rae, *J. Org. Chem., 42,* (1977) 2411.
36. M. Bullpitt, W. Kitching, D. Doddrell and W. Adcock, *ibid., 41* (1976) 760.
37. P.E. Hansen, *Org. Magn. Reson., 12* (1979) 109.
38. J. Seita, J. Sandstrom and T. Drakenberg, *Org. Magn. Reson., 11* (1978) 239.
39. L. Ernst, *J. Magnetic Reson., 20* (1975) 544.
40. L. Ernst, *Z. Naturforsch., 30b,* (1975) 788.
41. L. Ernst, *ibid., 30b,* (1975) 794.
42. L. Ernst, *Chem. Ber., 108* (1975) 2030.
43. N.K. Wilson and R.D. Zehr, *J. Org. Chem., 43* (1978) 1768.
44. H. Takai, A. Odani and Y. Sasaki, *Chem. Pharm. Bull., Japan, 26* (1978) 1966.
45. P. van de Weijer, C. Mohan and D.M.W. van den Ham, *Org. Magn. Reson., 10* (1977) 165.
46. U. Ewers, H. Günther and L. Jaenicke, *Angew. Chem., 87* (1975) 356.
47. W. Kitching, M. Bullpit, D. Doddrell and W. Adcock, *Org. Magn. Reson., 6* (1974) 289.
48. P.R. Wells, D.P. Arnold and D. Doddrell, *JCS Perkin 2,* (1974) 1745.
49. Wilson and J.B. Strothers, *J. Magnetic Reson., 15* (1974) 31.
50. D.H. Hunter and J.B. Strothers, *Can. J. Chem., 51* (1973) 2884.
51. B. Mechin, J.C. Richer and S. Odiot, *Org. Magn. Reson., 14* (1980) 79.
52. S.R. Johns and R.I. Willing, *Aust. J. Chem., 29* (1976) 1617.
53. W. Adcock, D.P. Cox and W. Kitching, *J. Organometal. Chem., 133* (1977) 393.

54. P.A. Claret and A.G. Osborne, *Org. Magn. Reson.*, *8* (1976) 147.
55. L. Ernst, *Org. Magn. Reson.*, *8* (1976) 161.
56. C.J. Pouchert and J. Behnke, The Aldrich Library of $^{13}$C and $^{1}$H FT NMR Spectra, The Aldrich Chemical Company, Milwaukee, WI., 1992.

Date Received: April 15, 1999
Date Accepted: November 1, 1999