

What follows are scanned copies of the article

**Automatic Analysis of n.m.r. Spectra: A Practicle Application
to the Spectra of Oriented 2,4-Dichlorobenzaldehyde and 2-Chlorobenzaldehyde**

by Peter Diehl and Juerg Vogt

which appeared in

Organic Magnetic Resonance, Vol.8, pp.638-642, 1976.

Better quality copies might be still available from the Journal.

The paper is based on a method for automatic analysis of NMR spectra developed by the above Authors and Stan Sykora published in 1975 (reference 1) and available at www.ebyte.it/stan/SS_Papers.html - 20.

Automatic Analysis of n.m.r. Spectra; A Practical Application to the Spectra of Oriented 2,4-Dichlorobenzaldehyde and 2-Chlorobenzaldehyde

P. Diehl* and J. Vogt

Department of Physics, University of Basel, Klingelbergstr. 82, 4056 Basel, Switzerland.

(Received 11 June, 1976; accepted 29 June, 1976)

Abstract—The capabilities of a novel computer program for the 'automatic analysis of n.m.r. spectra' [*J. Magn. Reson.* **19**, 67 (1975)] are demonstrated in the analysis of the experimental spectra of oriented 2,4-dichlorobenzaldehyde and 2-chlorobenzaldehyde. In both cases solutions are found by starting from parameters which do not allow a line assignment. The fitting procedure is described in detail and the conclusion is reached that the program is a typical search program which is able to find, fully automatically, parameters which are unknown in principle. However, for reasons of efficiency, a certain amount of interaction between the computer and the spectroscopist is desirable. The structural results indicate that in both molecules, which are assumed to be planar, the *O-trans*-configuration vastly predominates and that the benzene rings are considerably distorted.

INTRODUCTION

AS POINTED out in the first paper on automatic analysis,¹ the computer analysis of high resolution n.m.r. spectra often represents a tedious, time-consuming and expensive problem and always requires a great deal of prior knowledge of parameters. This preknowledge is only partially available for spectra of oriented molecules. Their degrees of order are usually unknown and, due to intramolecular motion, some geometrical details may also be unpredictable. For such spectra the existing methods of computer analysis with line assignment must usually fail. We therefore devised a completely different approach to the problem by using integral transforms of the spectrum and fitting these iteratively in a least squares fit.¹ In this process it is important to start with a limited number of transforms in order, initially, to fit only rough features of the spectrum and then to step up this number progressively for a fit of the finer details. Preliminary tests with theoretical spectra proved that the method worked quite well.¹ In the present study we demonstrate its abilities in the analysis of two experimental spectra of oriented molecules.

The molecules 2,4-dichlorobenzaldehyde and 2-chlorobenzaldehyde, which were chosen for the analysis, have one plane of symmetry so that their order is characterized

by three independent and unknown *S*-values. Although several studies⁴⁻¹¹ have indicated that the aldehyde group is predominantly in the *O-trans*-position (see Fig. 1), no information on the aldehyde proton coordinates could be located in the literature. The preknowledge on the 6 and 10 direct couplings, respectively and the 3 and 4 chemical shifts, respectively, is therefore limited.

EXPERIMENTAL

The compounds used were commercially available. The spectra of a 15 mol% solution of 2,4-dichlorobenzaldehyde in the nematic phase of EBBA [*N*-(*p*'-ethoxybenzylidene)-*p*-*n*-butylaniline] and of an 18 mol% solution of 2-chlorobenzaldehyde in the nematic phase of Merck Phase IV were recorded at 25 °C on a Bruker HX-90 Fourier-transform spectrometer with external D₂O-lock. The pulse width was 2.5 μs. 256 and 100 scans, respectively, were accumulated into a 16 k-computer. The spectral widths were 6250 Hz and 4000 Hz, with an average linewidth of 6 Hz.

SPECTRAL ANALYSIS

The automatic program

The version of the automatic program¹ which was used for the present study (copies are available to anyone interested) consists of a two-step procedure, starting with a set of weighted Fourier transforms (sine and cosine basis) together with the first 6 spectral moments. This combination, called basis 1, is followed by a set of Fourier transforms exclusively (basis 2). The second step succeeds the first automatically as soon as the variation of the parameters is less than a prefixed limit. The number of transforms used initially is roughly 50 and 100 for bases 1 and 2, respectively.

For the analysis, the automatic program needs the frequencies as well as the intensities of the observed transitions. To a first approximation the measured signal height may be used. It is, however, more favourable to determine the peak area which may be obtained

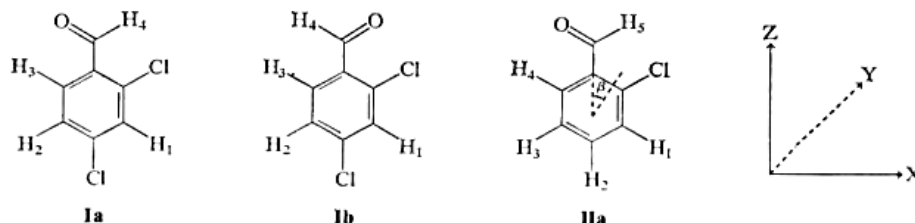


FIG. 1. Ia: *O-trans*-2,4-dichlorobenzaldehyde; Ib: *O-cis*-2,4-dichlorobenzaldehyde; IIa: *O-trans*-2-chlorobenzaldehyde.

*Author to whom correspondence should be addressed.

from the spectrum integral or, if there are broad background signals (liquid crystal solvents), by approximating the signal area as a triangle. Even for overlapping transitions this procedure gives reasonable results if the scale is sufficiently expanded (1.5 Hz per mm or less).

2,4-Dichlorobenzaldehyde

The experimental spectrum of the 4-spin system 2,4-dichlorobenzaldehyde consists of 24 lines. As a start for the automatic program, approximate values of the parameters were used from the benzaldehyde molecule² (column A of Table 1). A test showed that a line assignment with these starting parameters was impossible.

TABLE 1. 90MHz N.M.R. PARAMETERS OF ORIENTED 2,4-DICHLOROBENZALDEHYDE (FOR NUMBERING SEE FIG. 1). THE INDIRECT COUPLINGS ARE TAKEN FROM THE LITERATURE:¹³ $J(12) = 1.78$ Hz, $J(13) = 0.24$ Hz, $J(23) = 7.54$ Hz, $J(24) = 0.80$ Hz. ALL VALUES ARE IN HZ

Parameter	A	B	C	D
ν_1	200.0	-64.7	-70.9	-70.7 ± 0.4
ν_2	200.0	-80.9	-100.8	-96.5 ± 2.4
ν_3	200.0	-153.9	-140.4	-145.2 ± 2.4
ν_4	-300.0	-386.9	-372.4	-372.4 ± 0.5
D_{12}	0.0	57.4	58.2	51.3 ± 2.5
D_{13}	0.0	-38.3	-32.6	-26.0 ± 2.6
D_{14}	-100.0	-214.1	-209.8	-209.6 ± 0.4
D_{23}	-1000.0	-1707.0	-1721.5	-1721.9 ± 0.5
D_{24}	-100.0	-40.8	-112.1	-113.5 ± 4.9
D_{34}	-500.0	-335.4	-246.8	-244.8 ± 5.6

A: Automatic analysis program: parameters of the start-spectrum.

B: Automatic analysis program: parameters after one run. Number of transforms of basis 1:40. Number of transforms of basis 2:100.

C: Automatic analysis program: parameters after one more run. All lines with equal intensity ($\alpha = 0$ in Eqn 13 of Ref. 1). 200 transforms of basis 1, 400 transforms of basis 2.

D: UEA program: final parameters with 24 lines assigned.

A single run with 40 transforms of basis 1 and 100 of basis 2 led to the parameters given in column B. This solution turned out to be close enough to the experimental spectrum to allow line assignment. We tried, however, to proceed further with a run in which the imprecise intensity information was dropped ($\alpha = 0$ in Eqn 13 of Ref. 1). This is equivalent to fitting the transition frequencies exclusively. For this run the number of transforms was increased to 200 and 400 in bases 1 and 2, respectively, in order to create a more pronounced minimum in the r.m.s. error of the least squares fit. The resulting parameters are shown in Table 1, column C.

Finally the lines were assigned and a conventional program (UEA)³ was used to determine the parameters and their errors as shown in column D (r.m.s. of the fit: 0.8 Hz). The relatively large errors in D_{24} and D_{34} indicate that the 24 observed lines only poorly define the 10 parameters. This must also be the reason for the disagreement between the two solutions of columns C and D for all parameters which have large errors.

2-Chlorobenzaldehyde

The experimental spectrum of the 5-spin system 2-chlorobenzaldehyde contains roughly 90 distinct transitions which define the 15 parameters much more accurately than in the case of 2,4-dichlorobenzaldehyde. The spectrum is shown in Fig. 2.

As starting parameters for the automatic program,

values of benzaldehyde² were again used in combination with results of the preceding analysis of 2,4-dichlorobenzaldehyde (see column A of Table 2 and Fig. 2).

As Fig. 2 demonstrates, a line assignment is impossible at this stage. In the first computer run transition amplitudes were used as intensity information, whereas in the subsequent runs the transition area was approximated by triangles. The results reported in column B were obtained with 50 transforms of basis 1 and 100 of basis 2. (Chemical shifts were kept identical for all nuclei). In three successive runs from an improved start (column C) with 60 and 100 transforms, respectively, the data of column D were obtained. Already at this stage 34 lines could in principle be assigned. The automatic program was again applied in two steps with 90 and 150 transforms, respectively, for basis 1 and 100 and 200 transforms, respectively, for basis 2 and led to the parameters of column E (see also Fig. 2). Line assignment was now possible for 86 lines and with a conventional program the parameters of column F were derived (r.m.s. error: 0.24 Hz).

For the solution of the present problem six successive runs of the automatic program would have been sufficient. However, in principle, it may be necessary to start at the beginning with a variety of runs with different parameters in order to see whether an improvement of the fit has been obtained, i.e. whether a case which is suitable for further investigation is detected. An example of an unsuccessful fit is shown in Fig. 2 and in column G of Table 2. Although its intensity distribution is similar to that of the experimental spectrum, the parameters obtained are quite unreasonable.

It may also happen that an impurity signal blocks the progress of the automatic fit which then proceeds very slowly and demonstrates obvious difficulties to reproduce this line. In such a case the interaction with the program, that is inspection of intermediate data, helps considerably to avoid trouble and to speed up the process.

MOLECULAR GEOMETRY AND ORIENTATION

Rotational isomerism in 2-chlorobenzaldehyde has been studied by dipole measurements,⁴⁻⁷ infrared,⁸ ¹H n.m.r.^{9,10} and ¹³C n.m.r.¹¹ spectroscopies. These investigations show that the molecule exists predominantly in the planar *O-trans*-configuration (Fig. 1). This is believed to be due to repulsive dipole-dipole interactions between the halogen and the oxygen in the *O-cis*-rotamer and also to the possibility of intramolecular hydrogen bonding in the *O-trans*-form. Moreover, the planar *O-cis*-form is sterically hindered and an aldehyde group rotation of about 30° out of plane would be required in order to relieve the strain.⁴ An infrared⁸ study of the vapour state proved the planar *O-trans*-form to predominate whereas the structure of the second form could not be specified.

On the assumption of a planar form, the present n.m.r. study of the oriented species shows a clear preference for the *O-trans*-rotamer in both molecules. Unfortunately the possibility of non-planar forms cannot be investigated. Of the direct couplings between the benzene ring and the aldehyde proton, only the one in the *ortho* position is sensitive to the rotation. The problem with one coupling for three coordinates is consequently underdetermined.

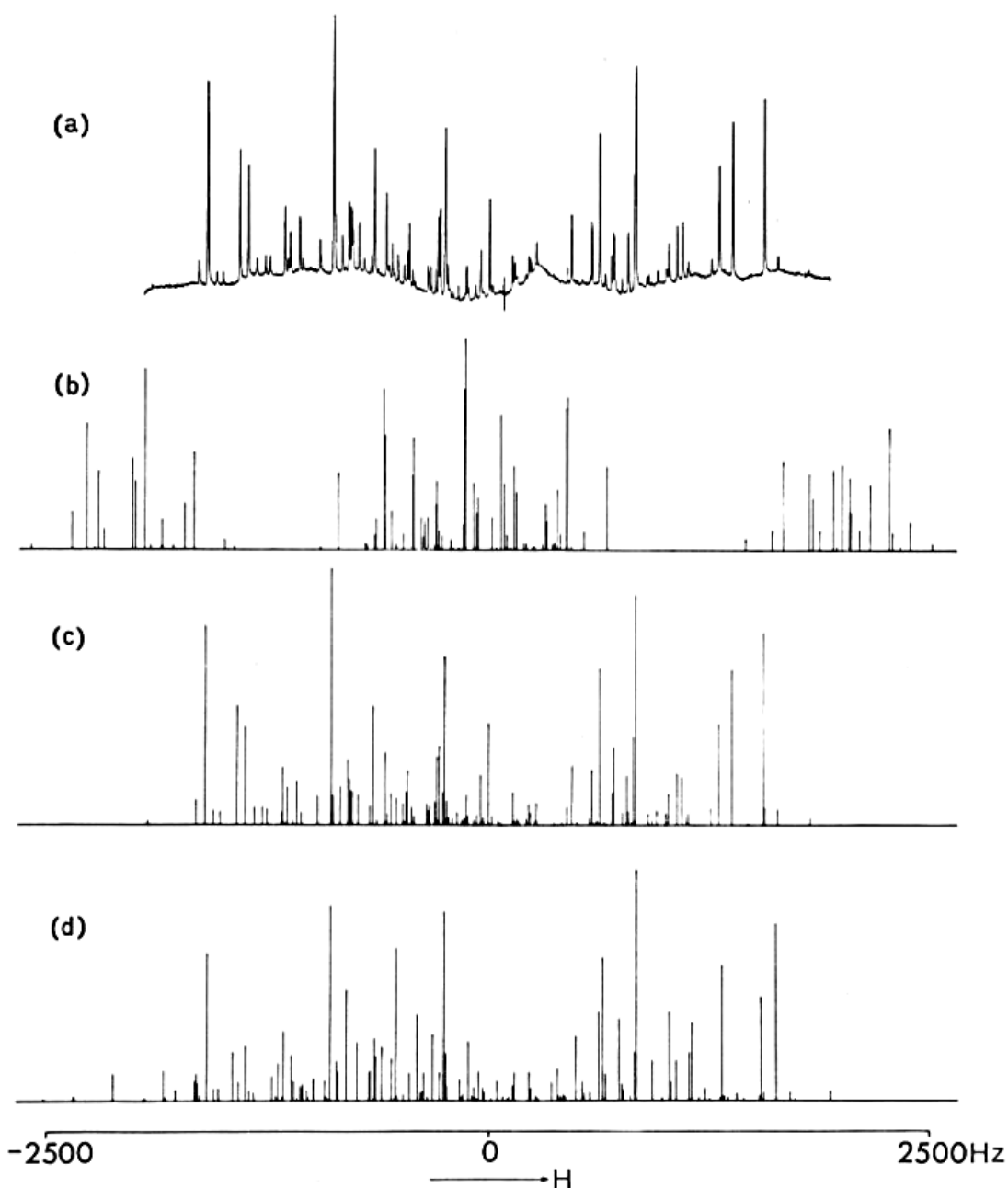


FIG. 2. 90 MHz n.m.r. spectrum of oriented 2-chlorobenzaldehyde. (a) Experimental spectrum: concentration, 18 mol% in Merck Phase IV; temperature 25 °C; linewidth 6 Hz. (b) Start-spectrum for the automatic analysis program (Column A in Table 2). (c) Final spectrum of the automatic analysis program (Column E in Table 2). (d) Unsuccessful fit with the automatic analysis program (Column G in Table 2).

The 6 direct couplings of 2,4-dichlorobenzaldehyde define the 3 orientation parameters and the position of the aldehyde proton in the molecular plane. The 10 direct couplings of 2-chlorobenzaldehyde, furthermore, provide detailed information on the benzene ring distortion. The results for 2-chlorobenzaldehyde obtained by the computer program SHAPE¹² are summarized in Table 3.

The distance ratios r_{12}/r_{23} and r_{31}/r_{23} are significantly

larger than those for the undistorted benzene ring. This demonstrates the strong interaction of both, proton 4 with the oxygen atom and proton 1 with its chlorine neighbour.

The positive off-diagonal order element S_{xz} shows the contribution of the chlorine substituent to the orienting forces. Diagonalization of the tensor leads to an angle β of 11.4° between the molecule fixed z-axis and the principal axis of the order tensor (Fig. 1).

TABLE 2. 90 MHz N.M.R. PARAMETERS OF ORIENTED 2-CHLOROBENZALDEHYDE (FOR NUMBERING SEE FIG. 1). THE INDIRECT COUPLINGS ARE TAKEN FROM THE LITERATURE:¹⁴ $J(12) = 8.141$ Hz, $J(13) = 1.055$ Hz, $J(14) = 0.473$ Hz, $J(15) = 0.035$ Hz, $J(23) = 7.433$ Hz, $J(24) = 1.839$ Hz, $J(25) = 0.050$ Hz, $J(34) = 7.758$ Hz, $J(35) = 0.784$ Hz, $J(45) = 0.078$ Hz. ALL VALUES ARE IN HZ

Parameter	A	B	C	D	E	F	G
v_1	0.0	-3.1	0.0	-17.5	4.6	4.7 ± 0.3	180.6
v_2	0.0	-3.1	0.0	-6.8	-20.5	-19.3 ± 0.3	180.6
v_3	0.0	-3.1	0.0	-21.8	-6.7	-4.4 ± 0.3	180.6
v_4	0.0	-3.1	0.0	-21.8	-49.8	-52.7 ± 0.3	180.6
v_5	-300.0	-301.3	-300.0	-277.0	-286.6	-289.1 ± 0.1	-1004.1
D_{12}	-200.0	-495.0	-500.0	-540.0	-552.9	-552.4 ± 0.2	-597.6
D_{13}	50.0	-213.0	-50.0	-96.0	-69.3	-75.0 ± 0.3	-98.0
D_{14}	-30.0	51.6	0.0	-50.8	-49.5	-48.0 ± 0.2	-129.3
D_{15}	-150.0	-101.4	-100.0	-24.4	-80.4	-80.9 ± 0.2	114.8
D_{23}	-200.0	-166.4	-200.0	-378.8	-376.4	-375.7 ± 0.3	47.0
D_{24}	-150.0	-84.4	-100.0	-80.5	-106.9	-101.8 ± 0.3	-325.3
D_{25}	-100.0	-81.3	-100.0	-107.8	-50.5	-50.7 ± 0.3	-273.9
D_{34}	-1300.0	-739.6	-700.0	-676.4	-680.9	-681.8 ± 0.2	-387.3
D_{35}	-100.0	-81.6	-100.0	-50.9	-63.1	-65.8 ± 0.2	17.9
D_{45}	-200.0	-345.9	-300.0	-270.6	-247.0	-243.3 ± 0.3	-338.4

A: Automatic analysis program. Parameters of the start-spectrum (see Fig. 2).

B: Automatic analysis program. Parameters after one run. 50 transforms of basis 1, 100 transforms of basis 2. Measure for the intensity: transition amplitudes.

C: Automatic analysis program. New starting parameters on the basis of B.

D: Automatic analysis program. Parameters after three successive runs. 60 transforms of basis 1, 100 transforms of basis 2. Measure for the intensity: triangle-area.

E: Automatic analysis program. Parameters after two successive runs starting with parameters D. 90 and 150 transforms respectively, of basis 1, 100 and 200 transforms, respectively, of basis 2. Measure for the intensity: triangle-area (see Fig. 2).

F: UEA program. Final parameters with 86 lines assigned.

G: Automatic analysis program. Unsuccessful fit. 50 transforms of basis 1, 100 transforms of basis 2. Measure for the intensity: transition amplitudes. (see Fig. 2).

TABLE 3. DISTANCE RATIOS AND ORIENTATION PARAMETERS OF ORIENTED 2-CHLOROBENZALDEHYDE (FOR NUMBERING SEE FIG. 1)

Parameter	N.m.r. value	Undisturbed ring
r_{12}/r_{23}	1.053 ± 0.003	1.000
r_{13}/r_{23}	1.694 ± 0.005	1.732
r_{14}/r_{23}	1.957 ± 0.003	2.000
r_{15}/r_{23}	2.068 ± 0.005	—
r_{24}/r_{23}	1.815 ± 0.005	1.732
r_{25}/r_{23}	2.594 ± 0.011	—
r_{34}/r_{23}	1.069 ± 0.003	1.000
r_{35}/r_{23}	2.297 ± 0.008	—
r_{45}/r_{23}	1.429 ± 0.004	—
S_{zz}^a	0.0436 ± 0.0006	—
S_{xz}^a	0.1074 ± 0.0011	—
S_{zx}^a	0.0136 ± 0.0004	—

^a On the basis of $r_{23} = 2.481$ Å (value for benzene ¹⁵).

CONCLUSIONS

Summarizing this practical experience with the program for an automatic analysis of n.m.r. spectra we may conclude that the fitting process, which in principle may be performed fully automatically in one run and without intermediate interaction with the person analysing the spectrum, should in practice, for reasons of efficiency, be divided into a few separate steps with operator interactions as follows.

Several first runs with various starting parameters and a small number of transforms (50 for basis 1, 100 for basis 2) are made. Of these the one which provides a solution reproducing the rough features of the experimental spectrum (similar grouping of transitions, similar spectral width and strong transitions) is selected. A new run is then performed which departs from the improved parameters just obtained. If necessary, this pro-

cess is repeated several times with a successively increasing number of transforms up to 200 and 400, respectively, for the bases 1 and 2.

If in a run the resultant spectrum shows poorer agreement than the preceding one, this usually indicates that the number of transforms used was too low.

As soon as a sufficient number of transitions can be identified and assigned, it is recommended that the imprecise intensity information is dropped and that the parameters are refined on the basis of line positions exclusively. This may be done either by assigning the lines with a conventional program or by using the automatic program with $\alpha = 0$ (Eqn 13 of Ref. 1). The latter method is similar in complexity to the line assignment, because avoiding the pitfalls of overlapping lines being treated as single transitions or of giving too much weight to very weak lines is equivalent to the identification of individual transitions.

In conclusion, one can say that the automatic analysis program is not a foolproof program, and it should not be used to analyse any spectrum without interaction with the person who performs the analysis. In its present version it is a typical search program which finds solutions in cases where there are several parameters, unknown in principle, where consequently trial and error starts with subsequent line assignment, i.e. conventional programs, must fail. The fact that the automatic analysis of the spectrum of oriented 2-chlorobenzaldehyde required only four days, instead of several weeks as one would have predicted from experience with similar spectra and conventional programs, allows us to be optimistic for the future of this novel approach.

Acknowledgement—We are grateful to Mr Bösigler for recording the spectra and to the Swiss National Foundation for the Advancement of Science for financial support.

REFERENCES

1. P. Diehl, S. Šykora and J. Vogt, *J. Magn. Reson.* **19**, 67 (1975).
2. P. Diehl, P. M. Henrichs and W. Niederberger, *Org. Magn. Reson.* **3**, 243 (1971).
3. J. A. Ferretti, R. K. Harris and R. B. Johannesen, *J. Magn. Reson.* **3**, 84 (1970).
4. C. L. Cheng, R. J. W. Le Fèvre, G. L. D. Ritchie, P. A. Goodman and P. H. Gore, *J. Chem. Soc. B* 1198 (1971).
5. H. Lumbroso and C. G. Andrieu, *Bull. Soc. Chim. Fr.* 1575 (1973).
6. C. T. Aw, H. H. Huang and E. L. K. Tan, *J. Chem. Soc. Perkin Trans. 2* 1638 (1972).
7. E. A. W. Bruce, G. L. D. Ritchie and A. J. Williams, *Aust. J. Chem.* **27**, 1809 (1974).
8. F. A. Miller, W. G. Fateley and R. E. Witkowski, *Spectrochim. Acta, Part A* **23**, 891 (1967).
9. G. J. Karabatsos and F. M. Vane, *J. Am. Chem. Soc.* **85**, 3886 (1963).
10. R. Wasylshen and T. Schaefer, *Can. J. Chem.* **49**, 3216 (1971).
11. T. Drakenberg, R. Jost and J. M. Sommer, *J. Chem. Soc. Perkin Trans. 2* 1682 (1975).
12. P. Diehl, P. M. Henrichs and W. Niederberger, *Mol. Phys.* **20**, 139 (1971).
13. D. G. De Kowalewski and V. J. Kowalewski, *J. Chem. Phys.* **37**, 1009 (1962).
14. D. G. De Kowalewski and S. Castellano, *Mol. Phys.* **16**, 567 (1969).
15. A. Langseth and B. P. Stoicheff, *Can. J. Phys.* **34**, 350 (1956).