

**Molecular Structures from
Diffraction and Spectroscopic Data**

Ewan Mitchell Brown

A thesis presented for the
Degree of Doctor of Philosophy
in the Faculty of Science
of the University of Edinburgh, 1994



DECLARATION

This thesis has not been submitted, in whole or in any part, for any degree at this or any other university. The work is original and my own, carried out under the direction of Prof.D.W.H.Rankin; where this is not so credit has been duly given.

To a famous ape

Acknowledgements

I would like to thank all the people at Edinburgh University who helped me throughout my time there. In particular, thanks go to David Reed for running the LCNMR spectra, Steve Henderson for assistance with practical work, Paul Brain for help with the electron diffraction analysis programs and Colin Pulham for sharing his knowledge of the theory and practice of vibrational analysis. Thanks also go to Naresh Mooljee for his invaluable advice and support in all aspects of computing and to Paul, Charles, Murray and Steve for moral support.

Special thanks go to my supervisor David Rankin for his assistance throughout and for his patience over the last few months.

Finally, I would like to express my gratitude to the Science and Engineering Research Council who funded this work.

Abstract

Modifications have been made to the programs used at Edinburgh for the analysis of the nuclear magnetic resonance spectra of molecules dissolved in liquid crystal solvents. In addition, an investigation has been made into the use of molecular mechanics force fields to calculate vibrational corrections to diffraction and spectroscopic data, necessary if a structural analysis is to be carried out using data from diverse sources. The technique of combined analysis has been applied to the structural determination of the three isomeric difluorobenzenes and of 2-chloropyrimidine, 3,6-dichloropyridazine and 2,6-dichloropyrazine. Data from electron diffraction, microwave spectroscopy and liquid crystal nuclear magnetic resonance spectroscopy have been used in these analyses, with varying degrees of success. The geometries obtained are largely consistent with results obtained for molecules of the same class.

Contents

Chapter 1 - Electron Diffraction

Introduction	2
Theory	6
Experimental	9
Computing.....	14
Limitations of Electron Diffraction	19
References.....	23

Chapter 2 - Liquid Crystal NMR Spectroscopy

Liquid Crystals	26
Theory.....	32
Experimental	36
Analysis of LCNMR Spectra.....	39
Limitations of LCNMR	43
References.....	47

Chapter 3 - Rotational Spectroscopy

Theory.....	50
Experimental.....	55
Limitations of Rotational Spectroscopy.....	58
References.....	59

Chapter 4 - Combined Analysis

Introduction.....	61
Electron Diffraction and the Shrinkage Effect.....	65
Vibrational Corrections to Direct Coupling Constants.....	68
Vibrational Corrections to Rotation Constants.....	69
Harmonic Force Field Analysis.....	70
Computing.....	76
Compatibility of Data.....	78
References.....	79

Chapter 5 - Modifications to the LCNMR Analysis Programs

Introduction	83
Input	84
Output	86
Calculation	88
Example - The LCNMR Analysis of 2 Chloropyridine	91
Conclusions	95
References	95

Chapter 6 - Using MM3 to Calculate Vibrational Corrections

Introduction	98
Calculation of Covariance Matrices	100
Examples.....	102
Conclusions	115
References.....	116

Chapter 7 - The Structures of the Difluorobenzenes

Introduction	119
Experimental	120
The Structure of o-difluorobenzene	122
The Structure of m-difluorobenzene.....	136
The Structure of p-difluorobenzene	145
Conclusions	153
References.....	159

Chapter 8 - Structures of Some Chlorine Substituted Heteroaromatic Compounds

Introduction	161
Experimental	163
2 Chloropyrimidine	165
3,6 Dichloropyridazine	175
2,6 Dichloropyrazine	184
Conclusions	191
References.....	193

Chapter 9 - Final Conclusions and Further Work..... 196

Appendix A - Fortran77 source code

A.I Example ED92 Model Subroutine

A.II Example ED92 Extra Data Subroutine

A.III Changes made to LCSIM

A.IV Modifications to MM3

Appendix B - Instructions for the LCNMR analysis programs

Appendix C - Notes on the use of MM3

Appendix D - Lecture courses and conferences attended

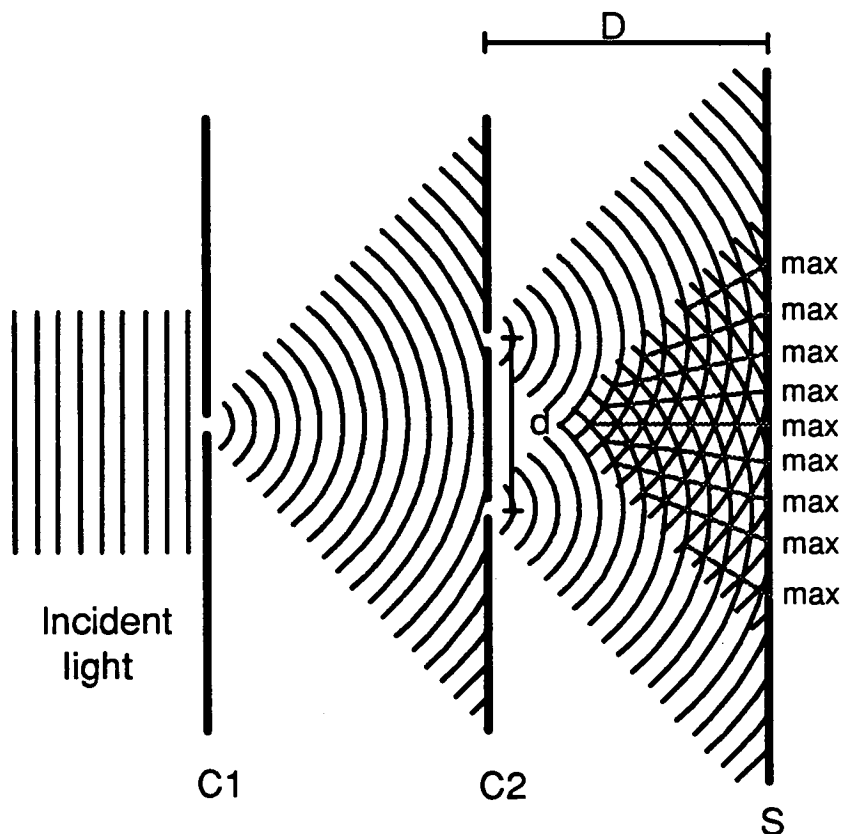
Chapter 1

Electron Diffraction

Introduction

In 1801 Thomas Young first demonstrated experimentally the diffraction and interference of light¹ thereby providing convincing evidence of its wave nature. A schematic representation of his experiment is shown in figure 1.1. The incident light passes through a pinhole made in a card (C1) and is diffracted, forming a spherical wavefront. The light then passes through a second card (C2) in which two pinholes are made. Spherical waves emerging from each of these holes interfere constructively and destructively leading to a pattern of alternating high and low intensity bands on the screen (S).

Figure 1.1 - Young's experiment



The distance between two adjacent maxima is given (for $D \gg d$) by

$$\Delta_{\max} = \frac{\lambda D}{d} \quad [1.1]$$

where λ is the wavelength of the light

D is the distance between card C2 and the screen (S)

and d is the distance between the two pinholes in card C2

Young's experiment can be used to calculate the wavelength of a light source (if d is known) or to determine the distance between two pinholes or slits (if λ is known). An important point here is that the interference pattern, which can be measured directly, can be used to determine quantities too small to measure directly with any degree of accuracy (d or λ).

Wave Particle Duality

In 1924 Louis de Broglie proposed his theory of wave-particle duality². The Compton and photoelectric effects both suggest that light behaves to some extent as if it consists of particles (photons). De Broglie took this idea one stage further, arguing that, if light can be seen to exhibit both wave-like and particle-like properties, it would seem logical that matter should also exhibit wave-like properties. He even went as far as to predict the properties of such waves, stating that for any particle travelling with a momentum, p , the corresponding wavelength (λ) is given by

$$\lambda = h/p \quad [1.2]$$

where h is Planck's constant.

It was not long before de Broglie's postulate was verified experimentally. Just as diffraction experiments had been used by Young to demonstrate the wave nature of light, the same principle was used to demonstrate the wave nature of a beam of electrons. In America, Davisson and Germer^{3,4} studied the scattering of electrons from the surface of nickel crystals and, in Scotland, Thomson and Reid⁵ passed high velocity electrons through various thin films and observed scattering patterns which could be accounted for by assuming the electron beam to be behaving as a wave. Furthermore, calculations based on the scattering patterns obtained and the known velocity of the electrons produced a wavelength consistent with that predicted in de Broglie's equation (1.2). The importance of this result is perhaps best underlined by the fact that Davisson and Thomson shared the Nobel Prize for Physics in 1937.

Young's experiment (described above) can in principle be used to determine either the wavelength of the light used or the separation of the slits; similarly, diffraction of electrons can be used to determine either the de Broglie wavelength or the separation of the diffraction centres (internuclear distances). Initially, such studies focused on the diffraction of electrons by the surfaces of crystals or by thin films. Such experiments had previously been carried out using X-rays. In 1930, Debye⁶ demonstrated that X-rays can also be diffracted by a gaseous sample. Because of the random orientation of molecules in the gas phase, the overall interference pattern appears as a series of concentric rings with the intensity distribution, from the centre to the edge of the pattern, a damped oscillating function. From the positions of the maxima and minima, structural parameters could be determined. The exposure times in such an experiment were, however, extremely long, due to the weak interaction

between X-rays and the electrons from which they are scattered. At the time, Herman F. Mark noted this result and immediately realised that a similar experiment using a beam of electrons rather than X-rays would considerably reduce the exposure times, because the interaction between the electron beam and the electric field of the atomic nuclei would be much stronger. With the help of Raimund Wierl, he set about designing such an experiment and soon started to produce diffraction patterns of some simple molecules^{7,8}. This can be considered as the "birth" of gas-phase electron diffraction (GED) and despite technical developments, the experiment remains largely unchanged to this day.

Theory

The total scattering intensity at any given point on the diffraction pattern can be written,

$$I_{\text{total}} = I_{\text{incoherent}} + I_{\text{inelastic}} + I_{\text{atomic}} + I_{\text{molecular}} \quad [1.3]$$

The term that is used in structural determination is $I_{\text{molecular}}$, the contribution from electrons scattered by pairs of nuclei. All other forms of scattering can be considered as background scattering. Incoherent scattering may arise from extraneous or double-collision scattering. Inelastic scattering corresponds to electrons which undergo a momentum change during collision. These terms can be eliminated by subtracting a polynomial function which passes through the mid-points of oscillations in the original radial scattering function. If the constituent atoms of the molecule under study are known, then the atomic scattering contribution can be calculated using tabulated scattering factors⁹. The data can therefore be reduced to a series of points representing only the molecular scattering. It is common practice to tabulate these as a function of the angular parameter

$$s = \frac{4\pi \sin(\theta/2)}{\lambda} \quad [1.4]$$

where θ is the scattering angle and is equal to two times the angle between the diffracted and undiffracted beams

and λ is the electron wavelength.

The advantage of using S , rather than θ , is that it is independent of the camera distance used and so data from different sources can be more easily compared and, if desired, combined.

The molecular scattering intensity can be derived by a quantum mechanical treatment¹⁰ and is given by

$$I_m(s) = A \sum_{\substack{i,j \\ i \neq j}} |f_i(s)| |f_j(s)| \cos[\eta_i(s) - \eta_j(s)] \exp(-\frac{1}{2} u_{ij}^2 s^2) \frac{\sin[s(r_{ij} - \kappa_{ij} s^2)]}{sr_{ij}} \quad [1.5]$$

where $f_i(s)$ is the atomic electron scattering amplitude for the i^{th} atom

$\eta_i(s)$ is the phase of the electron scattering for the i^{th} atom

u_{ij} is the mean amplitude of vibration between atoms i and j

r_{ij} is the distance between atoms i and j

κ_{ij} is an asymmetry constant (see below)

A can be considered to be a scaling factor and is given by

$$A = \frac{4\pi m^2 e^2}{h^4 \epsilon_0 s^4} \quad [1.6]$$

where m and e are the mass and charge of an electron

h is Planck's constant

and ϵ_0 is the permittivity of free space.

The asymmetry constant, κ_{ij} , mentioned above is related to the anharmonicity constant (a) of the Morse potential by the approximation

$$\kappa_{ij} = \frac{a_{ij}u_{ij}^4}{6} \quad [1.7]$$

For most bonded pairs of atoms a_{ij} can be assumed to be approximately 2.0 \AA^{-1} . For non-bonded pairs, however, a harmonic approximation is sufficient and a_{ij} can be assumed to be zero.

The phase term, $\cos[\eta_i(s) - \eta_j(s)]$, can be approximated using the equation

$$\eta_i(s) - \eta_j(s) = (a_i - a_j) + (b_i - b_j)s + (c_i - c_j)s^2 + (d_i - d_j)s^3 \quad [1.8]$$

The scattering factors, $f(s)$, and phase terms, $\eta(s)$, have been tabulated for all elements by Fink et al⁹. The constants a, b, c and d are determined by fitting a cubic function to the tabulated values of $\eta(s)$.

Experimental

By considering the schematic representation of Young's experiment, shown in figure 1.1, we can introduce the main components of the GED apparatus. This also serves to demonstrate the strong parallels between the two experiments.

To start with, in Young's experiment, we have the incident light. Clearly, in the GED experiment this is replaced with a beam of electrons. Electrons are emitted from a heated filament and accelerate, over a short distance, towards an anode held at a ground potential. The beam must clearly be as monochromatic as possible if the results are to be interpreted on the basis of a single wavelength and so the voltage must be as stable as possible. A practical requirement of the diffraction experiment is that the distance between maxima (Δ_{\max}) and the distance to the screen (D) are comparable (both must be accurately measurable on the laboratory scale). From equation 1.1 it can be seen that this is the case if the wavelength (λ) and the distance between diffraction sources (d) are comparable. Hence, if interatomic distances are to be measured, then a wavelength of the order of 10^{-10} m is required. From de Broglie's equation (1.2) a suitable velocity, and hence accelerating potential, can be calculated; electrons accelerated through a potential of 50 kV have an associated wavelength of approximately 0.0548 \AA (neglecting relativistic correction).

Once the beam passes the anode it is collimated and focused using a series of apertures and magnetic lenses, corresponding to the first card in Young's experiment. It is important that the beam is as narrow as possible as the

intersection with the sample must be occur in as small a volume as possible (approximating to a single point). This is difficult to achieve as the electrons are mutually repelled by electrostatic forces.

The equivalent of card C2 in figure 1.1 is the gaseous sample of the GED experiment. Instead of pinholes as diffraction sources we have the atomic nuclei of the molecules and the internuclear distances within the molecules determine the positions of the maxima of the diffraction pattern. The sample itself is introduced into the evacuated diffraction chamber through a fine nozzle; once more the sample beam must be as narrow as possible to reduce the intersection volume. A cold trap prevents the sample from filling up the chamber after it has passed through the electron beam. Various designs of nozzle may be used depending on the reactivity and volatility of the compound under study.

After passing through the sample, the diffracted beam continues through the diffraction chamber towards the screen. In the GED experiment, this usually consists of a photographic plate which records the intensity distribution of the scattered beam. The scattering intensity decreases sharply as the scattering angle increases ($I \propto s^{-4}$) but there is a limited range over which the photographic medium is sensitive; it is therefore necessary to attenuate the beam according to some known radial function. This is achieved by the addition of a rapidly rotating sector, placed immediately in front of the plate - a technique developed by Debye¹¹. The shape of the sector is designed to decrease the effective exposure time towards the centre of the plate, usually as a cubic function of s . To trap the portion of the beam which remains

undiffracted, a beam stop, consisting of an aluminium cylinder, is positioned at the centre of the sector. This, of course, prevents data being recorded for very small s but is necessary to prevent back reflection of the beam.

The Edinburgh GED Apparatus

The electron diffraction data used in this work were recorded using the Edinburgh University ED apparatus¹² which was originally constructed by Robert Jenkins of Cornell University, based on a design of Bauer and Kimura¹³. The set-up is largely as described above with a few additions.

To increase the range of scattering angles that can be measured, it is possible to position the sample inlet at various different distances from the camera (260 mm, 200 mm and 95 mm for high temperature data; 285 mm and 128 mm for room temperature data). The longer camera distances allow narrow angle scattering to be recorded with a good dispersion whereas the shorter camera distances record data at wider scattering angles, with the plate still a manageable size.

Behind the rotating sector there is a fluorescent screen which can be used for the visual inspection of the diffraction pattern before recording it photographically. Furthermore, by defocusing the electron beam to a wider diameter, an image of the sample nozzle is projected onto this screen thus allowing the alignment of the electron beam and the incoming sample.

The electrons used are accelerated by a potential of 44.5 kV and the beam is typically focused to a diameter of about 0.3 mm. The internal diameter of the

aluminium sample inlet nozzle is also 0.3 mm so the intersection volume is as small as possible. The sample itself is allowed to evaporate from an ampoule and effuse through the nozzle. Depending on the volatility of the compound under investigation the ampoule and inlet system may be kept cold with an ice or slush bath or warmed with a heating jacket. In many cases, however, the vapour pressure at room temperature is suitable.

Calibration of the apparatus is achieved by recording the diffraction pattern of a sample of benzene, prior to recording the pattern of the main sample. Because the structure of benzene has been very well determined^{14,15} the pattern it produces can be used to obtain accurate values for the electron wavelength and the camera distance. Exposure times of both the benzene and the main sample are bracketed in order to ensure that the intensities fall within the range of the photographic emulsion.

Once the plates have been recorded and developed it is necessary to convert the images to numerical data for use in the structural analysis. Historically, this was done by visual inspection of the plates - measuring the radii of the maxima in the pattern. There is clearly an inherent inaccuracy in this technique and the modern experiment generally uses a digital microphotometer to scan the plates, recording the optical density as a function of the radius. Plates recorded in Edinburgh are sent to the EPSRC laboratory at Daresbury where they are digitised using a Joyce-Loebl microdensitometer¹⁶. A computer is used to control the scanning process. First, the exact centre of the diffraction pattern is determined, after which, five optical density measurements are taken at each of one thousand evenly

spaced points, at a fixed radius. The scanning radius is then increased, by an amount corresponding to an integer value of s , and further measurements are taken. This process is repeated until the whole plate has been scanned. The large number of measurements taken for each value of s greatly increases the accuracy of the final optical density data.

Computing

The raw digitised data are received from Daresbury by direct transfer to the Edinburgh University mainframe computer, *festival*. Further analysis of the data is carried out using the program *ed92*^{17,18} which is used for both data reduction and structural analysis.

Data Reduction

The data reduction routines of *ed92* are used to convert raw optical density data (D) to molecular intensity data ($I_{\text{molecular}}$). Firstly, the total scattering intensities are obtained by correcting for the flatness of the plate and for the non-linearity of the photographic emulsion saturation (the blackness correction). At this stage, compensation must also be made for the presence of the rotating sector. This results in a total scattering intensity as defined in equation 1.3. The atomic scattering contribution can be calculated from theory and subtracted to leave just the molecular, inelastic and incoherent scattering terms. The inelastic and incoherent scattering intensity is removed by the subtraction of a smooth background function as described above. All that remains is the molecular scattering intensity data which are used in the structure refinement stage of the program.

Although it is the scattering intensity curve that is used in the ED analysis it is useful, at this stage, to introduce the radial distribution curve. This is obtained by carrying out a sine fourier transformation of the intensity curve. The advantage of this transformation is that the curve is more easily interpreted

visually. It is usually plotted as $P(r)/r$ against r , where $P(r)$ is the probability of finding a pair of nuclei of separation r . The curve consists of a number of Gaussian shaped peaks, each centred on an internuclear distance within the molecule; the width of the peak is related to the mean squared vibrational amplitude of the pair of atoms. The relative area under each peak is given by the expression

$$\text{Area} \propto \frac{n_{ij} Z_i Z_j}{r_{ij}} \quad [1.9]$$

where n_{ij} is the multiplicity of the internuclear distance

and Z_i, Z_j are the atomic numbers of atoms i and j .

Structure Analysis

The primary purpose of **ed92** is to determine the structure of the molecule under investigation. Parameters are refined, by a least-squares iterative algorithm¹⁹, to achieve the best fit of theoretical scattering intensities (calculated using equation 1.5) to the experimental intensities, derived in the manner described above. The parameters can be divided into two main groups, structural parameters and vibrational amplitudes, as well as overall scale factors.

Generally, the number of structural parameters is the minimum required to describe the positions of all the atoms in the molecule. These parameters may consist of bond lengths, angles, torsion angles etc. In cases where two or more bonds within a molecule are of similar length, it is often convenient to

define one parameter as the average bond length and other parameters as bond length differences. This means that the difference parameter(s) may be fixed, if correlation between the parameters is too high to allow simultaneous refinement. Calculation of the theoretical scattering intensities requires a knowledge of all the internuclear distances within the molecule. These are most easily obtained from a set of Cartesian co-ordinates and, to this end, a subroutine must be written which calculates co-ordinates from the chosen structural parameters. This routine may also include, implicitly, any assumptions made about the symmetry of the molecule. An example of such a subroutine (**COORD**) can be found in appendix A.I. The parameters are passed to the routine in the array **PAR** and Cartesian co-ordinates are returned to the main program in the arrays **X**, **Y** and **Z**.

Vibrational amplitudes (u_{ij}) are obtained from a normal co-ordinate analysis of the molecule (see below) or are simply estimated on the basis of previous results. In either case, they are entered for every unique interatomic distance in the molecule, along with the multiplicity (n_{ij}) and the anharmonicity constant (a_{ij}). Wherever possible, these amplitudes should also be refined to fit the ED data. The problem of introducing a large number of extra refining parameters can be reduced somewhat by constraining groups of similar amplitudes to fixed ratios while refining only one amplitude in each group.

The refinement routines attempt to minimise the **R** factor

$$R_G = (D'WD / I'WI)^{\frac{1}{2}} \quad [1.10]$$

where \mathbf{D} is the difference vector, the elements of which represent the difference between theoretical and experimental intensities for each value of s .

and \mathbf{I} is a vector containing the scattering intensities.

\mathbf{W} , is an off-diagonal square matrix which is used to weight the data. The elements are determined by the following equations.

$$\text{i) } w_{ii} = (s_i - s_{\min}) / (s_{w1} - s_{\min}) \quad \text{for } s_{\min} \leq s_i \leq s_{w1}$$

$$\text{ii) } w_{ii} = 1 \quad \text{for } s_{w1} \leq s_i \leq s_{w2}$$

$$\text{iii) } w_{ii} = (s_{\max} - s_i) / (s_{\max} - s_{w2}) \quad \text{for } s_{w2} \leq s_i \leq s_{\max}$$

$$\text{iv) } w_{ij} = 0 \quad \text{for } |i - j| \neq 1$$

$$\text{v) } w_{ij} = 0.5(w_{ii} + w_{jj})(p/h)_k \quad \text{for } |i - j| = 1$$

where s_{\min} , s_{\max} represent the minimum and maximum extents of the data

s_{w1} , s_{w2} are weighting points, chosen by visual inspection of the intensity data or by using established default values.

and (p/h) is a correlation parameter²⁰.

This weighting scheme allows for the fact that the data is truncated at either end, by gradually decreasing the weight given to data at values of s outwith the range s_{w1} to s_{w2} . Furthermore, if two or more data sets are to be used then this allows them to be spliced together cleanly. The inclusion of off-diagonal elements helps to account for correlation between adjacent data points.

When the ED data has been prepared and the parameters have been given suitable initial values, refinement can begin. This is largely an interactive process in which parameters may be introduced into the refinement a few at a time until R_G settles into a minimum. It is important to examine the parameters at the end of each refinement stage to ensure that they still have "sensible" values. Clearly, some parameters will be less well determined than others and, indeed, it is often the case that one or more parameters cannot be refined at all.

If all goes well with the refinement, a final structure is obtained with as many simultaneously refining parameters as possible. A "good" R-factor for an electron diffraction study is considered to be between 5 and 10 percent ($R_G=0.05$ to 0.1). A correlation matrix can be printed, showing the degree of correlation between the refining parameters. The program also calculates estimated standard deviations (e.s.d.) for the refining structural and vibrational parameters; these can be used to evaluate the significance of the results.

Limitations

Suitability of Sample

If a compound is to be studied by GED it must, of course, be stable in the gas phase. More specifically, it should be possible to produce a vapour pressure of at least 1Torr at a temperature at which decomposition of the sample is not significant. It is for this reason that early GED work centred on small, light molecules which could be run at relatively low temperatures. In addition, molecules containing no heavy atoms generally require higher vapour pressures to produce good quality data, due to the relatively low scattering power of lighter nuclei.

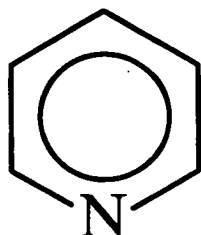
Light Atoms Poorly Determined

From equation 1.9 it is apparent that the contribution of light atoms to the scattering pattern is less than the contribution of heavier atoms in the same molecule. This is a particular problem in the case of hydrogen; consequently it is extremely difficult to determine the positions of hydrogen atoms in a molecule where heavier elements are present. The situation is improved somewhat if a large number of light atoms are related by symmetry such that the multiplicity factor becomes more significant. Consider for example the molecule $\text{Si}(\text{Si}(\text{CH}_3)_3)_4$; the peak in the radial distribution curve corresponding to the two-bond H...H distance would have an insignificant area were it not for the fact that the distance is repeated 36 times within the molecule.

Similar Distances Difficult to Resolve

In a molecule where two or more interatomic distances are similar, it can be difficult to determine these distances with any degree of certainty. This corresponds to a situation where two or more peaks in the radial distribution curve overlap. In the worst cases it may be impossible to solve the structure, uniquely, by ED alone. In the molecule pyridine, for example, there are three distinct bonded distances within the ring (2 C-C and 1 C-N), all of which lie under one peak in the radial distribution curve. Similarly, there are four distinct two-bond distances in the ring and three distinct three-bond distances, but only two further corresponding peaks in the radial distribution curve. In the case of interatomic distances involving hydrogen, this problem is compounded by the relatively low scattering power of the hydrogen nuclei (see above).

Figure 1.2 - Pyridine



This remains one of the most severe restrictions in the applicability of ED to structural analysis of larger molecules. Often it is impossible to obtain a structure without making a number of assumptions based on the structures of related molecules or chemical intuition.

Limitations of Data Quality

In spite of the many precautions taken in the ED experiment to obtain the best possible quality of data, no data set will ever be perfect. There are many sources of error including

- a) Errors in calibration of electron wavelength and apparatus dimensions
- b) The finite volume of the intersection of the electron and molecular beams (assumed in the theory to be a single point)²¹.
- c) Error in the "blackness correction".
- d) Uncertainties of the sector function and background subtraction.

A more exhaustive list of the errors arising in the ED experiment can be found in the investigation by Kuchitsu²². Modern techniques can go some of the way towards improving the quality of the data recorded. For example, using an electronic electron counting detector as the primary data collection method, rather than the traditional photographic plate, eliminates the errors associated with the "blackness correction" and the microphotometry stage of the experiment. Furthermore, by removing the limitation of the dynamic range of photographic media, there is no longer any need for a rotating sector; yet more errors are excluded. Undoubtedly, however, such technology introduces a whole new set of error sources which must also be considered. In any case, many of the errors associated with the photographic method can be eliminated by the use of benzene calibration, as described above.

Limitations in the Theory

If the aim of the ED refinement is to match a calculated scattering curve to the experimental curve, it is clearly of great importance that the theory, by which the former curve is calculated, is satisfactory. The theory expressed in equations 1.5 to 1.8 is the result of a number of approximations, for example

- a) Atoms are assumed to be spherical
- b) Scattering by any more than two nuclei is ignored.
- c) Multiple-electron scattering is ignored
- d) Electron exchange is ignored
- e) Relativistic effects are not taken into account
- f) Approximations are made in the treatment of anharmonicity (eq. 1.7)

These, and other theoretical limitations, are examined more fully in the review by Bartell²³.

It is probable that the current level of theory is sufficient for the analysis of data recorded by today's experimental methods but this may not always be the case. It is therefore important that approximations in the theory are not forgotten; as the quality of data increases, there may be a need to reassess the validity of such assumptions.

- ¹D.Halliday, R.Resnick, *Fundamentals of Physics*, 2nd Edition, Wiley & Sons, (1986)
- ²L. de Broglie, *Phil.Mag.*, 47, (1924), 446
- ³C.Davisson & L.H.Germer, *Nature*, 119, (1927), 558
- ⁴C.Davisson & L.H.Germer, *Phys.Rev.*, 30, (1927), 705
- ⁵G.P.Thomson & A.Reid, *Nature*, 119, (1927), 890
- ⁶P.Debye, *Lecture to "The Bunsen Society"*, (1930)
- ⁷H.Mark & R.Wierl, *Naturwiss.*, 18, (1930), 205
- ⁸R.Wierl, *Ann.Physik.*, 8, (1931), 521
- ⁹A.W.Ross, M.Fink & R.Hilderbrandt, "*International Tables for Crystallography*",
ed. A.J.C.Wilson, Kluwer Academic Publishers, Dordrecht, Boston and London,
Vol.C, (1992), 245
- ¹⁰I.Hargittai, *Stereochemical Applications of Gas-phase Electron Diffraction*, Part A, Ch.1,
eds. I.Hargittai & M.Hargittai, VCH, New York, (1988), 1
- ¹¹P.P.Debye, *Phys.Z.*, 40, (1939), 66
- ¹²C.M.Huntley, G.S.Laurenson & D.W.H.Rankin, *J.Chem.Soc., Dalton Trans.*, (1980), 954
- ¹³S.H.Bauer & K.Kimura, *J.Phys.Soc.Japan*, 17, (1962), 300
- ¹⁴W.V.F.Brooks, S.J.Cyvin & P.C.Kvande, *J.Phys.Chem.*, 69, (1965), 1489

- ¹⁵T.Iijima, M.Kimura & K.Tamagawa, *J.Mol.Struct.*, 36, (1976), 243
- ¹⁶S.Cradock, J.Koprowski & D.W.H.Rankin, *J.Mol.Struct.*, 71, (1981), 217
- ¹⁷A.S.F.Boyd, G.S.Laurenson & D.W.H.Rankin, *J.Mol.Struct.*, 71, (1981), 217
- ¹⁸D.M.Bridges, J.M.Freeman, G.C.Holywell & D.W.H.Rankin,
J.Organomet.Chem., 32, (1971), 87
- ¹⁹B.Beagley, J.M.Freeman, G.C.Holywell & D.W.H.Rankin, *J.Chem.Soc. (A)*, (1971), 785
- ²⁰Y.Murata & Y.Morino, *Acta.Cryst.*, 20, (1966), 605
- ²¹B.Miller & M.Fink, *J.Mol.Struct.*, 48, (1978), 363 ; 373
- ²²K.Kuchitsu, "*Molecular Structures and Vibrations*", ed. S.J.Cyvin,
Elsevier, Amsterdam, (1972), Ch. 10
- ²³L.S.Bartell, *Stereochemical Applications of Gas-phase Electron Diffraction*, Part A, Ch.2,
eds. I.Hargittai & M.Hargittai, VCH, New York, (1988), 55

Chapter 2

Liquid Crystal Nuclear Magnetic Resonance Spectroscopy (LCNMR)

Liquid Crystals

It is estimated that around 0.5% of all pure organic compounds exhibit a liquid crystalline phase¹. The discovery of such a mesophase is attributed to the botanist Friedrich Reinitzer² who described the substance cholesteryl benzoate as having 'two melting points'. He sent samples to the physicist Otto Lehmann who had previously reported³ that he had observed some compounds which seemed to melt over a wide range. Lehmann had not considered this to be due to the existence of a separate phase; rather he had suggested that the transition from solid to liquid was occurring over an extended temperature range. However, after many subsequent experiments, studying the melting of various compounds using a heating stage microscope, Lehmann concluded that he was indeed observing a 'new' phase of matter (the liquid crystalline phase). Substances exhibiting this phase are known as 'liquid crystals' or mesogens.

An important discovery in determining the nature of such mesophases was that they are birefringent⁴ (the refractive index measured using horizontally polarised light is different to that measured using vertically polarised light). This demonstrated the anisotropy of the liquid crystalline phase.

Since that time, our understanding of solid-liquid mesophases has increased greatly. Although they are fluid phases, there is a degree of local ordering of the molecules which is not present in a truly liquid phase. This local ordering may occur to varying degrees in different liquid crystals. The various phases have been classified into several distinct types.

The Nematic Mesophase

This phase, in which the ordering is purely orientational (i.e. there is no positional ordering), usually occurs in compounds consisting of rod-like molecules, which tend to align with their long axes parallel to one another. The time-averaged orientation is commonly described by a vector called the director. An idealised representation of such a phase can be seen in figure 2.1. It is important to remember that, at any one time, molecules may deviate from alignment with the director and that it is the time-averaged picture we are seeing. It should also be noted that, under normal conditions, there may exist any number of different directors representing regions of local anisotropy within the bulk sample.

The Chiral Nematic Mesophase

In the nematic phase, described above, the molecules tend to align parallel to one another. If, however, the molecules tend to align at a slight angle to one another then the director follows a helical pattern as we move through the sample. This is the chiral nematic, or cholesteric, mesophase and is usually exhibited by chiral mesogens. It is possible, however, to induce a chiral nematic phase, in a sample which is ordinarily nematic, by adding a small amount of a chiral impurity.

Smectic Mesophases

These differ from the nematic mesophase in that they possess a degree of positional ordering as well as orientational ordering. In general the molecules arrange themselves in layers (if a time-averaged view is taken).

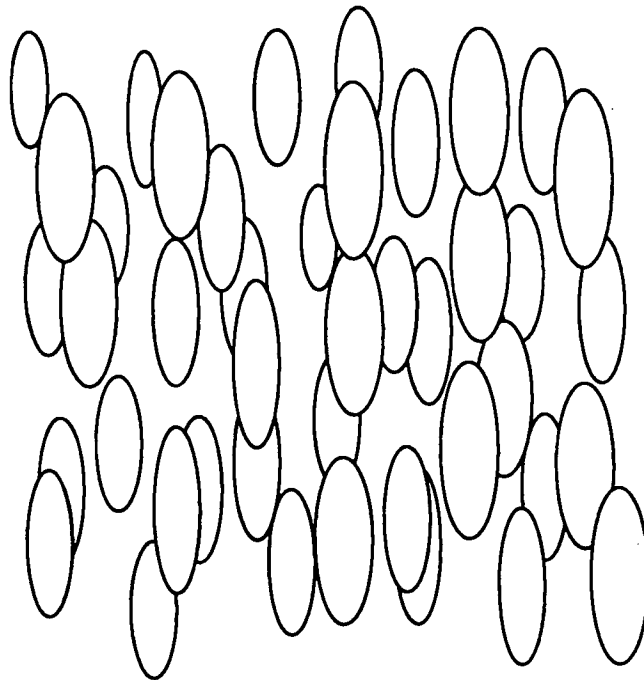


Figure 2.1

A time averaged representation of the nematic mesophase

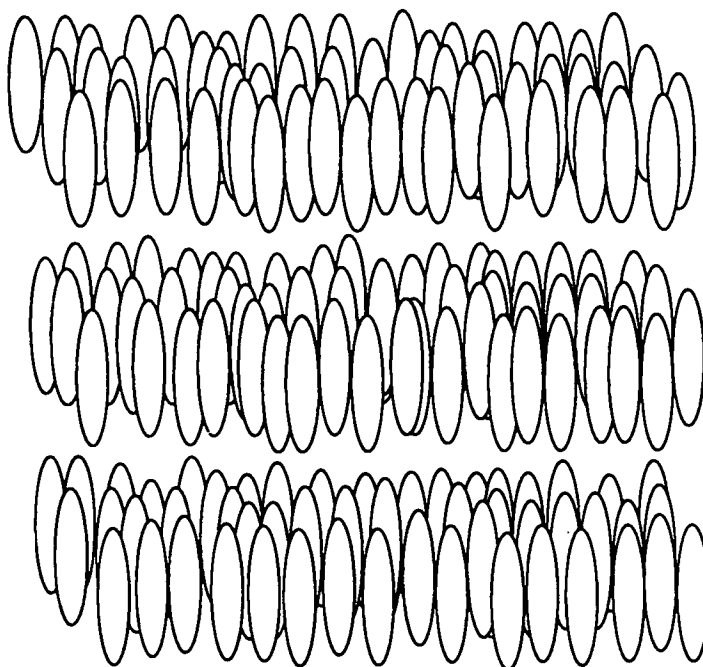


Figure 2.2

A time-averaged representation of the smectic A mesophase

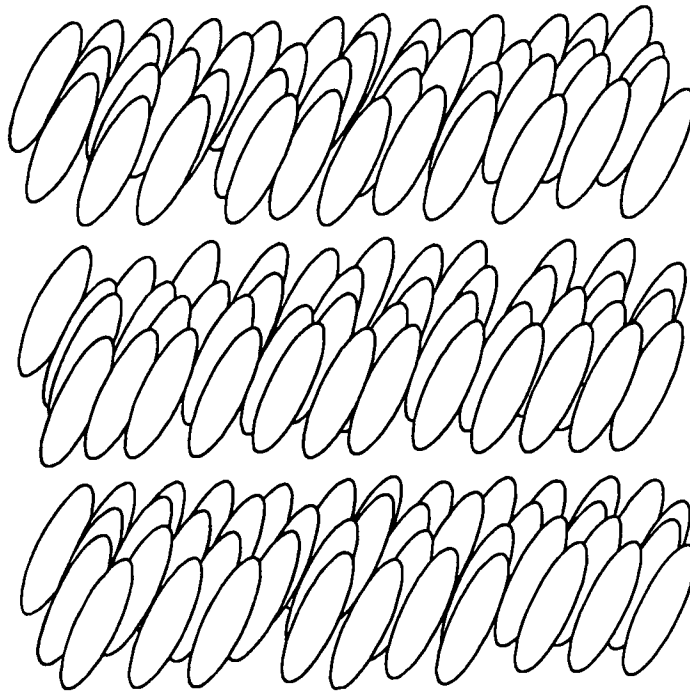


Figure 2.3

A time-averaged representation of the smectic C mesophase

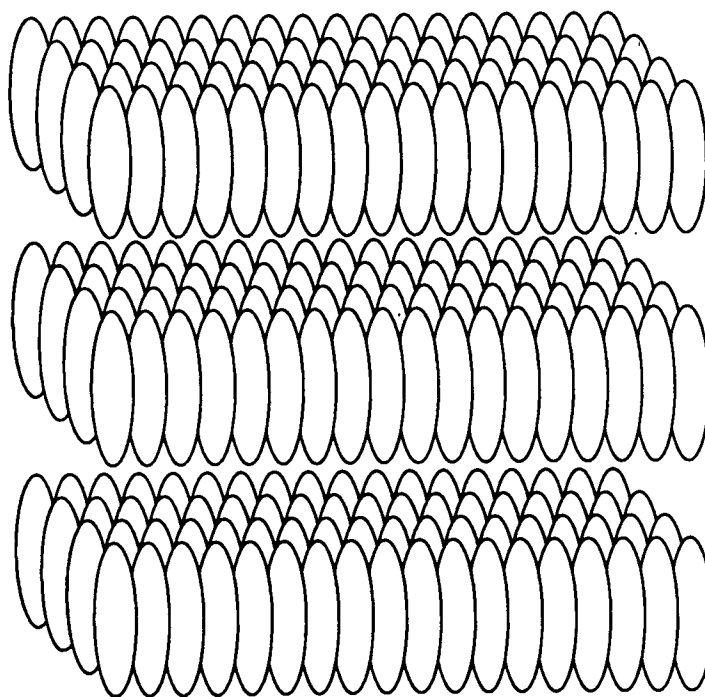


Figure 2.4

A time-averaged representation of the smectic B mesophase

Smectic phases have been subdivided into numerous classes according to the differences in the positional and orientational order. The smectic A mesophase shows no positional ordering within the layers and the director is perpendicular to them (figure 2.2). The smectic C mesophase is similar but the director is at an angle other than ninety degrees to the layers (figure 2.3). The smectic B mesophase differs from smectic A in that there is a degree of positional order within each layer (figure 2.4).

Compounds most likely to exhibit mesogenic properties are those with one molecular axis of very different length to the other two. This implies either a rod-like or disc-like shape. The phases of discotic liquid crystals are different to those described above and will not be discussed further here. Molecules with a generally rod-like shape include substituted biphenyls, bicyclohexyls and terphenyls. The molecules often have a rigid end, such as a biphenyl group, and a flexible end, usually a straight-chain alkyl or alkoxy group.

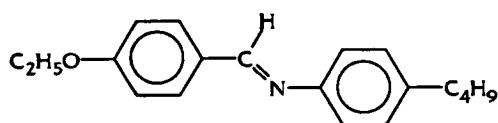
Much research has been carried out into what causes a compound to have a liquid crystalline mesophase and how to predict the existence or otherwise of such a phase. As yet, however, the only way of being certain is to synthesise the molecule and study its thermal properties. Some typical liquid crystals can be seen in figure 2.5.

It is mentioned above that, under normal conditions, the ordering within a liquid crystal is only short range. However, if the sample is placed in a modest electric or magnetic field the local directors align with each other, thus making the bulk sample anisotropic with one overall director. It is this

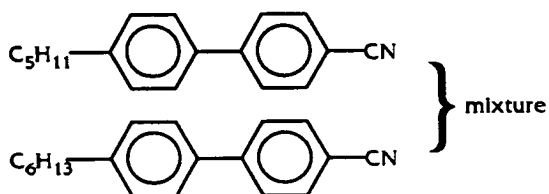
effect that is exploited in many of the applications of liquid crystals including liquid crystal displays (LCD) and liquid crystal nuclear magnetic resonance spectroscopy (LCNMR).

Figure 2.5 - Some typical liquid crystals

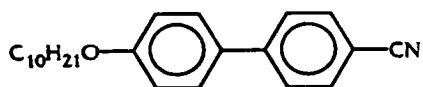
EBBA (nematic)



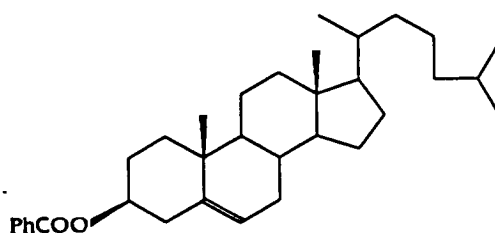
E5 (nematic)



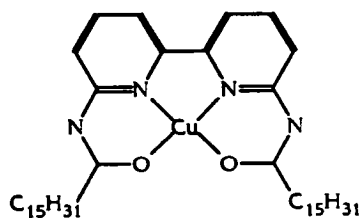
M30 (smectic A)



cholesteryl benzoate
(chiral nematic)



$Cu(LC_{16})$ (discotic⁵)



Theory

The NMR spectra of liquid crystals and of molecules dissolved in liquid crystals were first recorded by Saupe and Englert^{6,7}. Their most important conclusion, so far as the structural chemist is concerned, is that information about the structure of molecules dissolved in a liquid crystal can be obtained by an analysis of the NMR spectra of such solutions.

In an NMR experiment where an isotropic solvent is used, the appearance of the spectrum is determined by the chemical shifts (ω) of the nuclei observed and by the indirect coupling constants (J) between each observed nucleus and the other spinning nuclei in the molecule. There exists, however, a third type of parameter - direct coupling constants (D) - which become important in the LCNMR experiment. Whereas the indirect couplings are transmitted through the bonds within the molecule, the direct couplings are transmitted through space. They depend solely upon the distance between the nuclei in question and the average orientation of the vector joining these nuclei with respect to the magnetic field of the spectrometer.

In the conventional NMR experiment, direct couplings average to zero because of the relatively rapid rotation of the molecules on the NMR timescale. For this reason, only the indirect component of each coupling is observed. When the solvent is a liquid crystal, however, the environment of the solute molecules is anisotropic due to the alignment of the solvent molecules in the magnetic field of the spectrometer. Although the solute molecules still rotate rapidly, there is no longer an equal probability for all orientations. The molecule is said to be partially oriented. It follows from

this that the direct couplings no longer average to zero, as the rotation of the molecule is non-random.

The magnitude of the direct coupling is determined by the equation,

$$D_{ij} = -\frac{\mu_0 \hbar \gamma_i \gamma_j}{8\pi^2} \cdot \frac{S_{ij}}{r_{ij}^3} \quad [2.1]$$

where D_{ij} is the direct coupling between nuclei i and j

μ_0 is the permittivity of a vacuum

\hbar is Dirac's constant

γ_i & γ_j are the magnetogyric ratios of nuclei i and j respectively

r_{ij} is the internuclear distance

and

$$S_{ij} = \frac{1}{2} \langle 3 \cos^2 \vartheta_{ij} - 1 \rangle \quad [2.2]$$

where ϑ_{ij} is the angle between the internuclear vector and the direction of the spectrometer's magnetic field.

It can be seen from equation 2.1 that for each direct coupling constant measured, there are two unknown variables (r_{ij} and S_{ij}) which leaves us unable to determine either structural or orientational information. It is possible however to work around this problem if we choose to describe the average orientation of the whole molecule by way of an orientation tensor rather than explicitly describing the average orientation of each nuclear pair.

This leads to the equation,

$$D_{ij} = -\frac{\mu_0 \hbar \gamma_i \gamma_j}{8\pi^2 r_{ij}^3} (S_{zz} \cos^2 \vartheta_{ijz} + S_{yy} \cos^2 \vartheta_{ijy} + S_{xx} \cos^2 \vartheta_{ijx} + 2S_{xz} \cos \vartheta_{ijx} \cos \vartheta_{ijz} + 2S_{yz} \cos \vartheta_{ijy} \cos \vartheta_{ijz} + 2S_{xy} \cos \vartheta_{ijx} \cos \vartheta_{ijy}) \quad [2.3]$$

where $\vartheta_{ij\alpha}$ are the angles between the internuclear vector and the axes of the molecule

and $S_{\alpha\beta}$ are the elements of the orientation tensor,

$$S = \begin{pmatrix} S_{xx} & S_{yx} & S_{zx} \\ S_{xy} & S_{yy} & S_{zy} \\ S_{xz} & S_{yz} & S_{zz} \end{pmatrix} \quad [2.4]$$

which are determined by,

$$S_{\alpha\beta} = \frac{1}{2} \langle 3 \cos \vartheta_\alpha \cos \vartheta_\beta - \delta_{\alpha\beta} \rangle \quad [2.5]$$

where ϑ_α is the angle between the molecular α -axis and the magnetic field

and $\delta_{\alpha\beta}$ is the Kronecker delta (unity if $\alpha = \beta$, otherwise zero)

We have now improved upon equation 2.1 in that once the elements of the orientation tensor (orientation parameters) have been determined, any further couplings observed will yield structural information. In fact there are effectively only five independent orientation parameters because the orientation tensor is both symmetric and traceless. The picture improves further if the symmetry of the spin system under study is taken into account.

As the symmetry increases, certain orientation parameters become zero. Table 2.1 shows the independent orientation parameters that must be determined for any given symmetry.

Table 2.1

**Non-zero independent orientation parameters
for spin systems of various symmetries.**

Point Group	Orientation Parameters
C_1, C_i	$S_{zz}, (S_{xx}-S_{yy}), S_{xy}, S_{xz}, S_{yz}$
C_2, C_{2h}, C_s	$S_{zz}, (S_{xx}-S_{yy}), S_{xy}$
C_{2v}, D_2, D_{2h}	$S_{zz}, (S_{xx}-S_{yy})$
symmetric tops	S_{zz}
spherical tops	all parameters zero

Experimental

The LCNMR experiment is largely the same as any other NMR experiment. A standard 5 ml NMR tube may be used with the liquid crystal in place of the usual NMR solvent. Non-volatile samples may be weighed into the tube first and the tube then degassed on a vacuum line. Volatile samples may be measured out after the solvent is degassed using standard vacuum line techniques. In either case a typical sample is ~0.2 mmol. The tube may be sealed under vacuum then heated and shaken until the sample is dissolved. If necessary, the tube may contain a sealed capillary holding a suitable locking solvent such as acetone-d₆.

In choosing a suitable liquid crystal solvent, several factors must be considered. First, it is important that the particular phase of the liquid crystal is suitable. Chiral nematic mesophases are inappropriate due to the helical variation of the director throughout the sample although, if the external magnetic field is large enough, the helical arrangement may be broken, leading to a normal nematic phase⁸. In principle, smectic phases may be used for LCNMR. However, complications may arise if the sample is to be rotated due to the high viscosity of the solvent. On the whole the nematic phase is considered the most useful for LCNMR and indeed, in all the experiments discussed below, nematic liquid crystals were used.

The director in a nematic phase may align either parallel or perpendicular to the external field. One or other of these types of nematogen may be used for LCNMR depending on the particular spectrometer used - specifically the direction of the spinning axis with respect to the field. Spectrometers based on an electromagnet, generate a field perpendicular

to the tube axis, whereas, spectrometers in which a superconducting magnet is used generate a field parallel to the tube axis. As it is desirable to spin the sample (to minimise the effects of field inhomogeneity) it is necessary to use a liquid crystal which aligns parallel to the spinning axis. The work described here was, therefore, restricted to the use of liquid crystals align parallel to the magnetic field (the NMR spectrometers at Edinburgh use superconducting magnets).

Another factor in choosing a suitable solvent is the temperature range of the mesophase. A wide range is clearly more convenient as it is easier to ensure that the spectrum is run in the appropriate phase. In particular, it is often convenient to record spectra at room temperature. The actual temperature of the phase must also be suitable for the sample in question; the sample must clearly be soluble at the temperature of the experiment. Often, mixtures of different compounds are used to give solvents which best fit these criteria.

Finally, it is most important that the liquid crystal does not react with the solute used. In most cases this is not a problem but care should, nonetheless, be taken.

Because direct coupling constants can be very large (thousands of Hertz), LCNMR spectra are very often second order - arising when the coupling between nuclei is comparable to or greater than the difference between their chemical shifts. Furthermore, unlike in conventional NMR experiments, coupling may occur between magnetically equivalent nuclei. These facts together mean that the spectra obtained often have a very

large number of lines. To achieve the best resolution of such spectra (particularly if satellite peaks are to be observed), it is often best to carry out the experiment on a high field spectrometer. At Edinburgh the Bruker WH360 MHz and Varian VXR600 MHz are ideal for this purpose.

Analysis of LCNMR spectra

Once an LCNMR spectrum has been obtained, it is necessary to analyse it to determine the values of the direct couplings. The spectrum of the liquid crystal solvent itself is generally broad and featureless due to the large number of couplings involved. The solute, however, gives rise to relatively sharp peaks which can be measured accurately in frequency and intensity.

First order spectra

In cases where the spectrum is first order, the calculation of direct couplings is trivial. The spectrum can be analysed in largely the same way as is applied to first order spectra in isotropic solvents. Coupling constants, T_{ij} , are obtained which represent the total coupling, both direct and indirect. Where nuclei i and j are equivalent, such that in normal circumstances no indirect coupling is observed between them,

$$T_{ij} = 3D_{ij} \quad [2.6]$$

Where indirect couplings are observable however,

$$T_{ij} = 2D_{ij} + J_{ij} \quad [2.7]$$

Where necessary, J_{ij} can be measured in an isotropic solvent; hence D_{ij} can be evaluated in either case. This approach is complicated by the fact that both direct and indirect coupling constants can be positive or negative. Although the sign of J_{ij} is often known there is no simple way of determining the sign of D_{ij} initially. The problem can be simplified if some assumptions are made about the structure of the molecule which can indicate approximate values and relative signs of the direct couplings. In most

cases, however, there are still two equally likely solutions - differing in the signs of all direct couplings. It is possible to determine the relative signs of the total couplings, T_{ij} , by carrying out double resonance experiments^{9,10}. An alternative approach is to measure T_{ij} at two temperatures, thereby obtaining values for ΔD_{ij} (assuming J_{ij} to be temperature independent, to a reasonable approximation). In practice, however, it is often easier to simply refine structures to fit each set of data and see which gives the most realistic results.

Second order spectra

In all but the simplest of spin systems, LCNMR spectra tend to be second order. This is due to the large size of T_{ij} relative to the difference in chemical shift between nuclei. There are two main methods of obtaining coupling constants from such spectra - analytical and numerical.

Analytical solutions exist for many spin systems (AB_2 , ABC , AB_2C , A_2B_2X , $AA'A''A'''$ etc.) and can be found in the excellent book by Emsley and Lindon¹. These consist of tables of equations relating spectral parameters to line frequencies and intensities. Once lines have been identified, it is often possible to solve these equations simultaneously to obtain a unique set of parameters.

Where an analytical solution is not possible, iterative numerical methods can be used to refine spectral parameters to fit the observed transition frequencies. Many computer programs exist for this purpose and, although the description below refers specifically to the programs used at Edinburgh, the principals behind all such programs are the same. Many are derived from standard NMR analysis programs (e.g. **LAOCOON II** and **LAOCN 3**)¹¹

modified to incorporate direct coupling constants (e.g. **LEQUOR**¹², on which the Edinburgh programs are based).

To obtain reasonable starting values for the parameters to be refined, the interactive program, **lcsim** (see appendix B), can be used. This program uses approximate co-ordinates for the spinning nuclei to calculate approximate direct couplings. Orientation parameters are estimated on a trial and error basis. This largely relies upon visual comparison of simulated and experimental spectra. The chemical shifts and indirect couplings used are generally assumed to be transferable from NMR experiments using isotropic solvents. The validity of this assumption is discussed below. As well as approximate values for the direct couplings, the program also gives each peak an identification number which the refinement program, **sliquor**, can recognise and use to determine the best fit (see appendix B).

Sliquor refines spectral parameters (D's, J's and ω 's) to fit the observed spectrum and as the fit improves, more lines can often be identified and included in the refinement. Because no assumptions are made about the geometry of the molecule under study or the values of the orientation parameters, it should in principle be possible to fit the calculated spectrum to the experimental spectrum very accurately. In practice however, uncertainties in peak positions and in the values of assumed parameters can lead to a poorer fit. It may not be possible to refine all parameters simultaneously (particularly the indirect couplings). This is especially true in the analysis of satellite sub-spectra, where many of the peaks may be obscured by those of the parent spectrum or are unobservable due to their

low intensity. Despite these problems, however, it is often possible to obtain direct couplings with uncertainties of less than 0.1%. Details of modifications made to **lcsim** and **sliquor** can be found in subsequent chapters and a full description of their use in the analysis of LCNMR spectra can be found in Appendix B.

Limitations of LCNMR

As a technique for molecular structure determination, LCNMR has a number of limitations, some of which have been touched on above. These will now be discussed in more detail.

Restricted to spin- $\frac{1}{2}$ nuclei

Strictly speaking, LCNMR does not give information about molecular structure but instead information about spinning nuclei. Normally, direct couplings can only be obtained for nuclei with spin- $\frac{1}{2}$, so only the positions of spin- $\frac{1}{2}$ nuclei, or nuclei with a reasonably abundant spin- $\frac{1}{2}$ isotope, can be determined. In most of the structural studies to date ^1H or ^{19}F spectra have been analysed, often with further information obtained from ^{13}C and ^{15}N satellites.

Number of determinable parameters

Because direct couplings are required to evaluate the orientation parameters as well as structural parameters, there is a restriction on the number of structural data that can be obtained. For an experiment to yield any structural information, it must hold that

$$\text{NDC} > \text{NOP} \quad [2.8]$$

where **NDC** is the number of measurable direct couplings

and **NOP** is the number of independent orientation parameters required (see Table 2.1).

For example, no structural information can be obtained from the LCNMR spectrum of dichloromethane. Two direct couplings can be measured (D_{HH} and D_{CH}) but there are two independent orientation parameters for a C_{2v} spin system.

A complete structural determination of a given molecule can be made only if

$$\text{NDC} - \text{NOP} \geq \text{NSP} \quad [2.9]$$

where NSP is the number of independent structural parameters.

Furthermore, it follows from equation 2.1 that absolute internuclear distances can never be obtained from direct couplings. The overall scale of a molecule can never be separated from the magnitude of the orientation parameters. Ratios of distances and hence angles can be determined but if a complete, scaled structure is required then a scaling factor must be obtained by some other technique.

Solvent dependency of J and ω

It is common practice to use chemical shifts and indirect coupling constants measured in isotropic solvents in the analysis of LCNMR spectra. The possibility that such parameters may vary between solvents cannot be ignored. It is usually necessary to refine chemical shifts to fit the LCNMR spectrum (due to the anisotropy of ω^{13}) but this may not be possible for indirect couplings which are highly correlated with their associated direct couplings. To obtain values closest to those in the LCNMR experiment, spectra can be recorded using the liquid crystal solvent at a temperature at which it is in an isotropic phase. Ideally this should be done at several

temperatures and extrapolated back to the temperature of the LCNMR experiment.

Indirect coupling anisotropy

It is stated above that indirect coupling constants are independent of orientation. This, however, is an oversimplification. Indirect couplings do in fact have an anisotropic component (J^{aniso}) which upon partial orientation may become observable. Unfortunately it is not usually possible to separate J^{aniso} from the direct dipolar coupling and so neither can be evaluated independently. By making some structural assumptions it is sometimes possible to estimate the magnitude of J^{aniso} and this has been done for several systems^{14,15}. The encouraging conclusion from this work is that J^{aniso} for light nuclei is usually small (often negligible) relative to the direct dipolar coupling and so can be ignored¹³. This is particularly true if at least one of the nuclei is ^1H . These conclusions are supported by theoretical calculations^{16,17} of the size of J^{aniso} . The problem can therefore be worked around by excluding couplings between heavy nuclei from the structural refinement.

Correlation between vibration and orientation

Molecular vibration is fast on the NMR timescale and so the observed direct couplings are in fact time-averaged over all vibrations according to the equation

$$\langle D_{ij} \rangle_{\text{vib}} = -\frac{\mu_0 \hbar \gamma_i \gamma_j}{8\pi^2} \langle S_{ij} \cdot \bar{r}_{ij}^{-3} \rangle_{\text{vib}} \quad [2.10]$$

If it is assumed that molecular reorientation is slow relative to the period of vibration then the two can be separated to give

$$\langle D_{ij} \rangle_{\text{vib}} = -\frac{\mu_0 \hbar \gamma_i \gamma_j}{8\pi^2} \langle S_{ij} \rangle_{\text{vib}} \langle r_{ij}^{-3} \rangle_{\text{vib}} \quad [2.11]$$

This assumption is implicit in the theory described above but, more recently, its validity has been questioned. If there is some correlation between vibration and molecular orientation then equation 2.11 no longer holds and it becomes impossible to separate the averaging of the structural term from the averaging of the orientational term. As the molecule vibrates, its preferred orientation changes and in some cases reorientation may be fast enough to vary over the course of a vibration.

The most compelling evidence for the existence of this effect is the non-zero direct couplings of highly symmetric molecules (T_d , O_h) in certain solvents. Snyder and Meiboom¹⁸ noted that tetramethylsilane dissolved in *p,p'*-di-*n*-hexyloxyazoxybenzene exhibited appreciable direct couplings. In a truly tetrahedral molecule these couplings should average to zero.

A theory to describe and correct for such correlation has been developed by Lounila et al¹⁹ and has been applied to the structures of benzene²⁰ and the methyl halides²¹. Experimentally this involves recording LCNMR spectra in a number of solvents or mixtures of solvents, each of which must be analysed to obtain a set of direct couplings. These data are then analysed together to obtain a correction for the correlation. However, the need to record and analyse so many spectra for each structural analysis makes such work impractical for all but the simplest systems. A better approach is to use solvents in which the correlation between vibration and orientation is minimal. To determine which solvents are best in this respect,

their effect on the spectrum of a tetrahedral molecule such as $^{13}\text{CH}_4$ can be investigated - the most suitable solvents being those in which D_{CH} is closest to zero. Although the approach of Lounila et al¹⁹ is, in principle, more satisfactory, the vast majority of structures determined to date have involved no such correction and in most cases the results have been in good agreement with structures determined by other techniques.

References

- ¹J.W.Emsley & J.C.Lindon, "NMR Spectroscopy using Liquid Crystal Solvents",
Permagon Press, Oxford, (1975), 1
- ²F.Reinitzer, *Monatsch.*, 9, (1888), 421
- ³O.Lehmann, *Ann.Physik.*, 25, (1908), 852
- ⁴P.J.Collings, "*Liquid Crystals: Nature's Delicate Phase of Matter*",
Adam Hilger, Bristol, (1990), 66
- ⁵T.Kuboki, K.Araki, M.Yamada & S.Shiraishi, *Bull.Chem.Soc.Japan*, 67(4), (1994), 948
- ⁶A.Saupe & G.Englert, *Phys.Rev.Letters*, 11, (1963), 462
- ⁷G.Englert & A.Saupe, *Z.Naturforsch.*, 19a, (1964), 172
- ⁸M.Panar & W.D.Phillips, *J.Amer.Chem.Soc.*, 90, (1968), 3880
- ⁹P.D.Blair, *Ph.D. Thesis*, University of Edinburgh, (1984)
- ¹⁰I.W.Anderson, J.E.Bentham & D.W.H.Rankin, *J.Chem.Soc., Dalton Trans.*, (1973), 1215

- ¹¹S.Castellano & A.A.Bothner-By, *J.Chem.Phys.*, 41, (1964), 3863
- ¹²P.Diehl, H.P.Kellerhals & W.Niederberger, *J.Magn.Reson.*, 4, (1971), 352
- ¹³J.Lounila & J.Jokisaari, *Prog.NMR.Spectrosc.*, 15(3), (1982), 249
- ¹⁴J.Gerritsen & C.MacLean, *J.Magn.Reson.*, 5, (1971), 44
- ¹⁵G.J.den Otter & C.MacLean, *Chem.Phys.*, 3, (1974), 119
- ¹⁶M.Barfield, *Chem.Phys.Letters*, 4, (1970), 518
- ¹⁷R.Ader & A.Loewenstein, *J.Amer.Chem.Soc.*, 96, (1974), 5336
- ¹⁸L.C.Snyder & S.Meiboom, *J.Chem.Phys.*, 44, (1966), 4057
- ¹⁹J.Lounila & P.Diehl, *J.Magn.Reson.*, 56, (1984), 254
- ²⁰J.Lounila & P.Diehl, *Mol.Phys.*, 52(4), (1984), 827
- ²¹J.Lounila, P.Diehl, Y.Hiltunen & J.Jokisaari, *J.Magn.Reson.*, 61, (1985), 272

Chapter 3

Rotational Spectroscopy

Theory

The moment of inertia of a molecule about a particular axis is given by

$$I = \sum_i m_i r_i^2 \quad [3.1]$$

where m_i is the mass of atom i

and r_i is the perpendicular distance between atom i and the axis.

It is clear that, by determining a moment of inertia, information about the molecular structure can be deduced (the masses being known). A molecule has an infinite number of moments of inertia corresponding to the infinite number of possible axes through its centre of gravity. Experimentally, however, it is only possible to determine a maximum of three moments of inertia - the principal moments of inertia, I_A , I_B and I_C . By convention I_A is the smallest of the three and I_C the largest. It can be shown that the A and C axes must be mutually perpendicular and the third principal axis, B, is perpendicular to both A and C. Of course, two or more of these moments of inertia may be equal, whether by chance or due to molecular symmetry. The subsequent classification of molecules is summarised in table 3.1. In addition, the moments of inertia of a planar molecule are related by equation 3.2, and so only two independent observations can be measured.

$$I_C = I_A + I_B \quad [3.2]$$

In most cases, rotational spectroscopy can be used to determine moments of inertia. This involves measuring the differences between rotational energy levels for a which, as we shall see, depend upon the values of the principal moments of inertia.

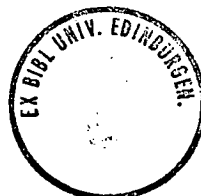
Table 3.1

Classification of molecules by their principal moments of inertia

Type	Definition	Examples
linear	$I_B = I_C, I_A = 0$	CO_2, HCN
spherical top	$I_A = I_B = I_C$	CH_4, SF_6
prolate symmetric top	$I_C = I_B > I_A$	$\text{CH}_3\text{I}, \text{CH}_3\text{SiH}_3$
oblate symmetric top	$I_C > I_B = I_A$	$\text{BF}_3, \text{benzene}$
asymmetric top	$I_C > I_B > I_A$	$\text{C}_2\text{H}_4, \text{CH}_2\text{Cl}_2$

Linear Molecules

A diatomic or linear polyatomic molecule has two equal principal moments of inertia, perpendicular to the molecular axis. The moment of inertia about the molecular axis is zero. The angular momentum, P , of such a molecule cannot be adequately expressed using classical mechanics and is in fact



quantised according to the equation¹

$$P = \hbar\sqrt{J(J+1)} \quad [3.3]$$

where $J=0,1,2,3\dots$ and is the rotational quantum number

and \hbar is Dirac's constant.

The kinetic energy of rotation is given by

$$E_r = \frac{1}{2}I\omega^2 \quad [3.4]$$

where I is the moment of inertia

and ω is the angular velocity.

As the angular momentum is related to the angular velocity by

$$P = I\omega \quad [3.5]$$

equations 3.3 to 3.5 can be combined to give

$$E_r = \frac{\hbar^2}{2I}J(J+1) \quad [3.6]$$

It is often convenient to express the energy equation in units of frequency giving the form

$$E_r = BJ(J+1) \quad [3.7]$$

B is known as the rotation constant (also in Hz) and is given by

$$B = \frac{\hbar}{4\pi I} \quad [3.8]$$

By determining the difference in rotational energies between levels with known values of J , the rotation constant and hence the moment of inertia can be evaluated.

An assumption has been made in deriving equation 3.7 which has not yet been discussed - namely that the internuclear distances within the molecule do not change as the rotational energy increases. This is known as the rigid rotor approximation. In fact it is not difficult to imagine that, as a molecule rotates more rapidly, centrifugal forces lead to a lengthening of the bonds. This centrifugal distortion can be taken into account by introducing a second term to the energy equation, giving

$$E_r = BJ(J+1) - DJ^2(J+1)^2 \quad [3.9]$$

where D is the centrifugal distortion constant.

More energy level differences must therefore be measured in order to determine the rotation constant, as D must also be evaluated.

Symmetric Tops

A molecule of which two of the principal moments of inertia are equal and the third is non-zero is described as a symmetric top. This occurs in molecules with a symmetry axis (rotation or rotation-reflection) of order greater than two. The total angular momentum is quantised as for a linear molecule but, in addition, a second quantum number, K , is required to describe the component of angular momentum about the symmetry axis. K may be any whole number less than or equal to J .

For a prolate symmetric top, in which I_A is the unique moment of inertia, the equation for rotational energy becomes

$$E_f = BJ(J+1) + (A - B)K^2 \quad [3.10]$$

Where A and B are the rotation constants

$$A = \frac{\hbar^2}{4\pi I_A} \quad [3.11]$$

$$B = \frac{\hbar^2}{4\pi I_A} \quad [3.12]$$

Taking into account centrifugal distortion, the equation becomes

$$E_f = BJ(J+1) + (A - B)K^2 - D_J J^2 (J+1)^2 - D_{JK} J(J+1)K^2 - D_K K^4 \quad [3.13]$$

It can be seen that three centrifugal distortion constants D_J , D_K and D_{JK} have been introduced.

The expressions for the rotational energy of an oblate symmetric top can be obtained by substituting C in place of A in equations 3.10 to 3.13.

Asymmetric Tops

A molecule in which all three principal moments of inertia are different is known as an asymmetric top. Unlike linear molecules or symmetric tops, the rotational energies of an asymmetric top cannot be written in closed form. Instead they are determined by solution of the appropriate wave equations. Computer programs are available for this purpose and yield the three rotation constants (A , B and C) which best fit the observed energies.

Experimental

Rotational spectroscopy measures the energy changes involved in transitions between two different rotational states. Two of the main techniques used are absorption spectroscopy and rotational Raman spectroscopy.

Absorption Spectroscopy

Rotational transition energies usually correspond to energies in the 5-6000 GHz region of the electromagnetic spectrum. This includes the far infra-red and microwave regions. By measuring the absorption of radiation over a suitable frequency range, such transitions can be observed.

In practice only broad-band radiation sources are available in the far infra-red region, so interferometers or grating spectrometers must be used. A consequence of this is that the resolution is not particularly high (~0.3 GHz at best). In the microwave region, tunable monochromatic sources, such as klystrons or backward-wave oscillators, are available². These allow rotational spectra of much higher resolution to be recorded and so rotation constants can be measured with extreme precision.

Rotational Raman Spectroscopy

In Raman spectroscopy, a molecule is excited to a virtual state by energy from an intense monochromatic source (usually a laser). This energy is soon released as the molecule returns to its original state. In some cases, however, there may be an increase or decrease in the frequency of the scattered radiation corresponding to a change in the rotational

(and/or vibrational) energy of the molecule. The difference in energies of the incident and scattered radiation is therefore equal to the difference in energy of the initial and final states of the molecule.

Other Techniques

In addition to absorption spectroscopy and rotational Raman spectroscopy, several other techniques exist by which rotation constants can be measured. These include vibration-rotation spectroscopy, which involves the analysis of rotational fine structure in the vibrational spectrum of a molecule, and Fourier-transform microwave spectroscopy. The latter is a relatively recent development, which relies upon the ability of modern computers to quickly perform the necessary data transformation³.

Selection Rules

The selection rules determining which transitions have non-zero intensity vary depending on the nature of the experiment. Where direct absorption or emission is being observed the selection rules are

- 1) $\Delta J = \pm 1$ ($\Delta K = 0$ for symmetric tops)
- 2) The molecule must possess a permanent dipole ($\mu \neq 0$)

The first rule leads to much simpler spectra than might be imagined from equations 3.9 and 3.13 as it greatly restricts the number of observable transitions. Rule 2 is in fact an approximation. The actual condition is that rotation of the molecule should result in an oscillating dipole. Tetrahedral molecules, such as CH_4 , do show very weak rotational spectra which arise

from the fact that centrifugal distortion leads to a effective reduction in the symmetry of the molecule.

In the case of rotational Raman spectroscopy the selection rules are more complex. For a linear molecule the selection rule is

$$\Delta J = 0, \pm 2$$

In the case of a symmetric top the rules are

- 1) $\Delta J = 0, \pm 1, \pm 2$

- 2) $\Delta K = 0$

Rule 1 also applies to an asymmetric top. However, as K is no longer a genuine quantum number, the remaining selection rules are extremely complicated. Once more, computer analysis can be used to yield the appropriate rotation constants.

Limitations of Rotational Spectroscopy

Number of Data

As mentioned above, a maximum of three rotation constants can be measured for a given molecule. In all but the simplest of cases, this is insufficient to carry out a complete structural determination. For example, the single rotation constant of HCN is not enough to determine both bond lengths. The situation is improved if data from molecules containing nuclei of different isotopes are also measured. Thus, by measuring the rotation constants of $\text{H}^{13}\text{C}^{14}\text{N}$ and $\text{H}^{12}\text{C}^{15}\text{N}$ as well as $\text{H}^{12}\text{C}^{14}\text{N}$, a structure can be found which fits all three data simultaneously. This approach is based on the assumption that the effect of isotopic substitution on the structure of the molecule is negligible. In most cases the assumption is valid, to a good approximation, but for light nuclei such as hydrogen, bond lengths may change appreciably. For accurate structural work, corrections for substitution effects on bond lengths and angles must be applied. A structure determined in this way (by measuring rotation constants of as many isotopically substituted species as possible) is often described as an r_s structure⁴.

Require Permanent Dipole

Determination of rotation constants by microwave spectroscopy, for which the results are the most accurate, usually requires that the molecule under study has a permanent dipole. Rotation constants can be obtained for molecules with no permanent dipole by the use of other techniques such as rotational Raman spectroscopy or by analysing the rotational fine structure

of their vibrational spectra. However, the experimental difficulties encountered and lower resolution of the spectra obtained make it less practical to carry out a full, accurate structural determination of such molecules by the use of rotation constants alone.

Some Atomic Positions Poorly Determined

The positions of atoms which have least effect on the moments of inertia of a molecule will be less well determined than those to which the moments of inertia are sensitive. From equation 3.1 it is apparent that where m_i or r_i is small, the contribution of atom i to the overall moment of inertia is small. Consequently, the positions of relatively light atoms (small m_i) or atoms close to one or more of the principal axes (small r_i) may be poorly determined⁵. This is a particular problem for atoms which lie close to the centre of gravity of the molecule.

References

¹J.M.Hollas, "*Modern Spectroscopy*", Wiley & Sons, Chichester, (1987)

²D.J.E.Ingram, "*Spectroscopy at Radio and Microwave Frequencies*", 2nd edition, Butterworths, London, (1967), 21

³C.N.Banwell, "*Fundamentals of Molecular Spectroscopy*", 3rd edition, M^cGraw-Hill, London, (1983), 26

⁴C.C.Costain, *J.Chem.Phys.*, **29**, (1958), 864

⁵C.C.Costain, *Trans.Am.Crystallogr.Assoc.*, **2**, (1966), 157

Chapter 4

Combined Analysis

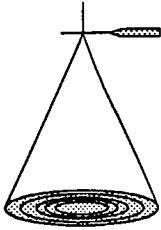
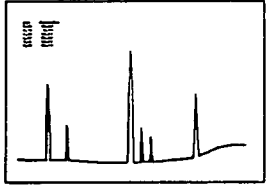
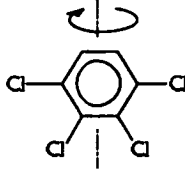
Introduction

Table 4.1 lists some of the advantages and disadvantages of the three structural techniques described above. It can be seen that the information obtainable from any one of these techniques, taken in isolation, is often insufficient to allow a complete structural determination of the molecule under investigation. However, the data from the different experiments are, in many ways, complementary. For example, electron diffraction is not good at determining the positions of light atoms in the presence of heavier atoms; LCNMR is particularly suited to locating hydrogen atoms. Microwave spectroscopy provides extremely accurate data, but often insufficient to determine the complete structure; LCNMR can supply more data but, even in the best cases, cannot provide absolute internuclear distances. It seems a logical approach to combine data from the different techniques, in order to obtain a complete and well determined structure. Before this can be done, however, it is important to ensure that the different data sets are compatible.

Definitions of structure

When we measure the geometry of a molecule, we are not determining the dimensions of a static structure. Instead, we are obtaining values which are averaged over all intramolecular motions. Even in its ground state, a molecule has a finite vibrational energy (the zero-point energy). It is not valid to assume that the average structure measured by one experimental technique represents the same average as that obtained by another.

Table 4.1 - Comparison of fluid phase structural techniques

Technique	Advantages	Disadvantages
<p style="text-align: center;">ED</p> 	<ul style="list-style-type: none"> • direct information on interatomic distances • gives very accurate structures for simple molecules 	<ul style="list-style-type: none"> • similar distances difficult to resolve • poor for large, non-symmetric molecules • light atoms in presence of heavier atoms poorly determined (especially hydrogen)
<p style="text-align: center;">LCNMR</p> 	<ul style="list-style-type: none"> • proton positions well determined • many observations uncorrelated • good for light atoms • no specialised apparatus required 	<ul style="list-style-type: none"> • information only for atoms with nuclear spin $\frac{1}{2}$ • sample must be soluble in a liquid crystal in its nematic temperature range • cannot determine absolute internuclear distances • complex second-order spectra may be difficult to interpret • anisotropy of indirect coupling, J, is a problem for couplings involving heavy atoms
<p style="text-align: center;">MW</p> 	<ul style="list-style-type: none"> • very accurate • very high resolution • can study individual vibrational states 	<ul style="list-style-type: none"> • molecule must have a permanent dipole moment • limited information without isotopic substitution • some atomic positions may be poorly determined

From the fundamental equations associated with each of the techniques discussed above, it can be seen that, to a first approximation

$$r_a \propto \langle 1/r \rangle^{-1} \quad [4.1]$$

$$r_d \propto \langle 1/r^3 \rangle^{-3} \quad [4.2]$$

$$r_o \propto \langle 1/r^2 \rangle^{-2} \quad [4.3]$$

where r_a is the apparent distance obtained from the ED experiment

r_d is the distance calculated from direct dipolar couplings

r_o is the distance derived from rotation constants

and r is the instantaneous internuclear distance.

In each case, the angle brackets represent time or ensemble averaging, over all intramolecular motions. It is apparent, from the above expressions, that distances measured by ED, LCNMR and rotational spectroscopy do not correspond to the same average geometry. If we are to combine data from these different sources, it becomes necessary to correct for the effects of vibrational motion, in order to reduce the results to a common basis.

The most fundamental definition of an internuclear distance is often considered to be that of the equilibrium distance, r_e . This corresponds to the separation at which the potential energy function (approximated for a diatomic molecule by the Morse potential¹) is at a minimum. Similarly, for a polyatomic molecule, the geometry corresponding to the minimum point on the potential energy surface, is described as the equilibrium (r_e) structure. Although this structure has a well defined physical meaning, it is, strictly speaking, a hypothetical concept; it describes a situation where the molecule is free from vibrational motion. In most cases, it is not possible to obtain an r_e structure experimentally and so a different definition of structure is required.

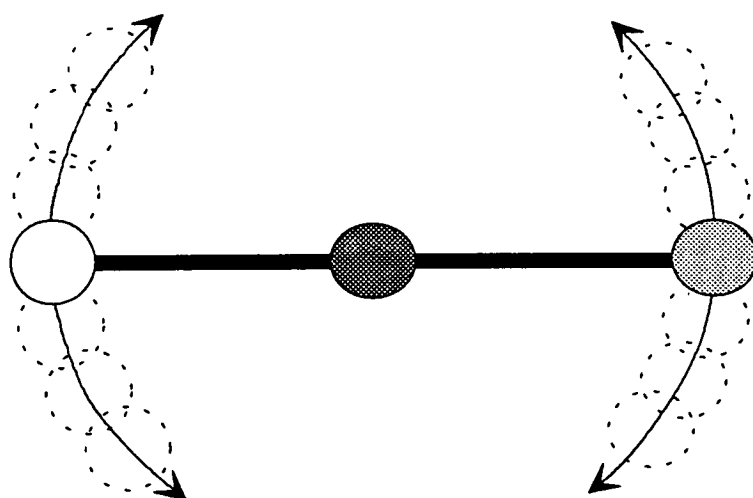
A more accessible alternative to the equilibrium structure is the r_α structure: the average nuclear positions over a Boltzmann distribution of vibrational states, at a given temperature. Closely related, is the r_α° (or r_z) structure, which can be defined as the average nuclear positions in the ground vibrational state. The different names merely reflect a difference in origin: the r_z structure is derived from ground state rotation constants. It is this type of structure that is commonly used in the combined analysis of data from different structural techniques or in the comparison of structures obtained by different methods.

Electron Diffraction and the Shrinkage Effect

If we consider the bending mode of a linear triatomic molecule (figure 4.1), it is apparent that, over the course of one vibration, the end atoms spend most of the time at a separation less than the equilibrium separation. Measuring such a structure by electron diffraction yields a non-bonded distance which is less than the sum of the two bonded distances^{2,3,4}. This observation is described by the theory of Hirota & Morino⁵ and is commonly known as the Bastiansen-Morino shrinkage effect. Similar effects can occur in any polyatomic molecule but are particularly significant for those with large amplitude vibrational modes. Unless corrections are made for such vibrations, the structures obtained in the ED determination, r_a or r_g , will be geometrically inconsistent (in the case of the linear triatomic, the apparent structure will be bent).

Figure 4.1

Origin of the shrinkage effect for a linear triatomic molecule



Fortunately, consideration of both the parallel and perpendicular amplitudes of vibration, for each pair of atoms, can yield the geometrically consistent r_α structure. The expressions relating the different types of structure are as follows^{6,7}:

$$r_g = r_a + \frac{u^2}{r_e} \approx r_a + \frac{u^2}{r_a} \quad [4.4]$$

where r_g is the average internuclear distance

and u^2 is the mean square parallel amplitude of vibration.

The r_g structure can then be converted to the r_α structure by taking the perpendicular amplitudes into account

$$r_\alpha = r_g - K_T - \delta r \quad [4.5]$$

where δr is a term accounting for centrifugal distortion.

K_T is the mean square perpendicular amplitude of vibration, at the temperature of the experiment, and is given by

$$K_T = \frac{\langle (\Delta x)^2 + (\Delta y)^2 \rangle}{2r_e} \quad [4.6]$$

Δx & Δy are the components of the instantaneous displacement of one atom from its equilibrium position (where the z-axis lies along the internuclear vector). Because the correction term, K_T , is small compared to r , it is reasonable to use r_a in place of r_e in the denominator of the above expression.

r_{α}° distances can be calculated from r_g distances using the equation⁸

$$r_{\alpha}^{\circ} = r_g - \frac{3a}{2}(u_T^2 - u_0^2) - K_T \quad [4.7]$$

where a is an anharmonicity constant

u_T & u_0 are the root mean square amplitude of vibration at the experimental temperature (T) and at absolute zero.

This equation differs from previously published versions^{6,7} which have been shown to be incorrect.

To a good approximation, the value of the anharmonicity constant, a , can be taken to be 2.0 Å for bonded distances and zero for non-bonded distances. The precise value is not important, as the correction term in which it is used is always very small. Parallel and perpendicular vibrational amplitudes (u_T , u_0 and K_T) and the centrifugal distortion correction (δr) can be obtained from a harmonic force field analysis (see below).

Vibrational Corrections to Direct Coupling Constants

The approach used in calculating the vibrational corrections to LCNMR data is that of Sýkora et al⁹. Observed direct coupling constants, D^o , are related to the corrected values, D^α , by the equation

$$D^\alpha = D^o - d^h \quad [4.8]$$

From these D^α values, r_α structural parameters can be determined. The harmonic correction term, d^h , is given by

$$d^h = -\frac{\mu_o \hbar \gamma_i \gamma_j}{8\pi^2} \text{Tr}(S\Phi^h) \quad [4.9]$$

where $\text{Tr}()$ denotes the trace of the enclosed tensor product

and S is the orientation tensor.

The elements of the tensor Φ^h are given by

$$\Phi_{\alpha\beta}^h = \left[C_{\alpha\beta} - 5 \sum_{\gamma} \zeta_{\gamma} (C_{\alpha\gamma} \zeta_{\beta} + C_{\beta\gamma} \zeta_{\alpha}) + \frac{5}{2} \zeta_{\alpha} \zeta_{\beta} \sum_{\gamma\delta} C_{\gamma\delta} (7\zeta_{\gamma} \zeta_{\delta} - \delta_{\gamma\delta}) \right] / r^5 \quad [4.10]$$

where ζ_{α} is the cosine of the angle between the internuclear vector and the α -axis of the molecular co-ordinate system

$\delta_{\gamma\delta}$ is the Kronecker delta

$C_{\alpha\beta}$ are the elements of the covariance matrix.

The summations are carried out for all $\gamma, \delta = x, y, z$. The covariance matrix can be obtained from a harmonic force field analysis of the molecule, as described below.

Vibrational Corrections to Rotation Constants

Observed ground state rotation constants, B_o , can be used to calculate a structure, r_o . The precise physical meaning of this structure is unclear; furthermore, the structure obtained varies depending on which isotopomer is used. For these reasons, it is difficult to compare the r_o structure to those obtained by other techniques, or to use B_o rotation constants as extra data in a combined analysis. However, B_o can be related to B_e (rotation constants which lead to an r_e structure) by the equation^{7,10,11}

$$B_e = B_o + \sum_i \frac{1}{2} d_i \alpha_i \quad [4.11]$$

where d_i is the degeneracy of the i^{th} vibrational mode

and α_i is the vibrational contribution to the rotation constant of the i^{th} mode.

α_i can be separated into a harmonic component (α_i^h) and an anharmonic component (α_i^a). The harmonic contribution can be obtained from a harmonic force field analysis but, for all but the simplest of molecules, the anharmonic contribution cannot be determined and the r_e structure is not attainable. For this reason it is common to calculate the B_z rotation constants (corresponding to the r_z structure) using the equation

$$B_z = B_o + \sum_i \frac{1}{2} d_i \alpha_i^h \quad [4.12]$$

As mentioned above, an r_z structure is fully equivalent to an r_α^o structure. B_z rotation constants can, therefore, be used to supplement ED data, if the structural refinement is carried out on an r_α^o basis (using equation 4.7).

Harmonic Force Field Analysis

Perhaps surprisingly, it can be shown that molecular vibrations may be described, to a close approximation, using classical mechanics¹². Many of the features of such an analysis are best introduced by looking at the simplest case: that of a diatomic molecule. It is convenient to define an axis system with the origin at the molecular centre of mass and both atoms lying on the z-axis. The vibration can then be described in terms of a displacement, q , which is the deviation of the bond length from its equilibrium value at any one time. If it is assumed that the system obeys Hooke's law (i.e. that q is proportional to the force, F , required to cause the displacement) then we can write

$$F = -fq \quad [4.13]$$

where f is known as the force constant of the bond.

Newton's second law of motion (force=mass x acceleration) gives

$$F = \mu \frac{d^2q}{dt^2} \quad [4.14]$$

where μ is the reduced mass of the molecule

and $\frac{d^2q}{dt^2}$ is the second derivative of the displacement, q , with respect to time, t (i.e. the acceleration).

Combining equations 4.13 and 4.14 gives the second order differential equation

$$\mu \frac{d^2q}{dt^2} + fq = 0 \quad [4.15]$$

This can be solved by the substitution of

$$q = A \cos(2\pi\nu t + \rho) \quad [4.16]$$

to give the familiar equation of simple harmonic motion

$$\nu = \frac{1}{2\pi} \sqrt{\frac{f}{\mu}} \quad [4.17]$$

From equation 4.16 it can be seen that q varies periodically with a frequency of ν , corresponding to the vibrational frequency of the molecule. ρ is simply a phase factor and A is the maximum amplitude of vibration.

It is convenient to define the force constant, f , in terms of the potential energy, V . The potential energy of such a system can be found by the integration of the exerted force with respect to the displacement, hence

$$V = -\int_0^q (fq) dq \quad [4.18]$$

This leads to the expression

$$V = \frac{fq^2}{2} \quad [4.19]$$

which can be twice differentiated, with respect to q , to give the definition of the force constant

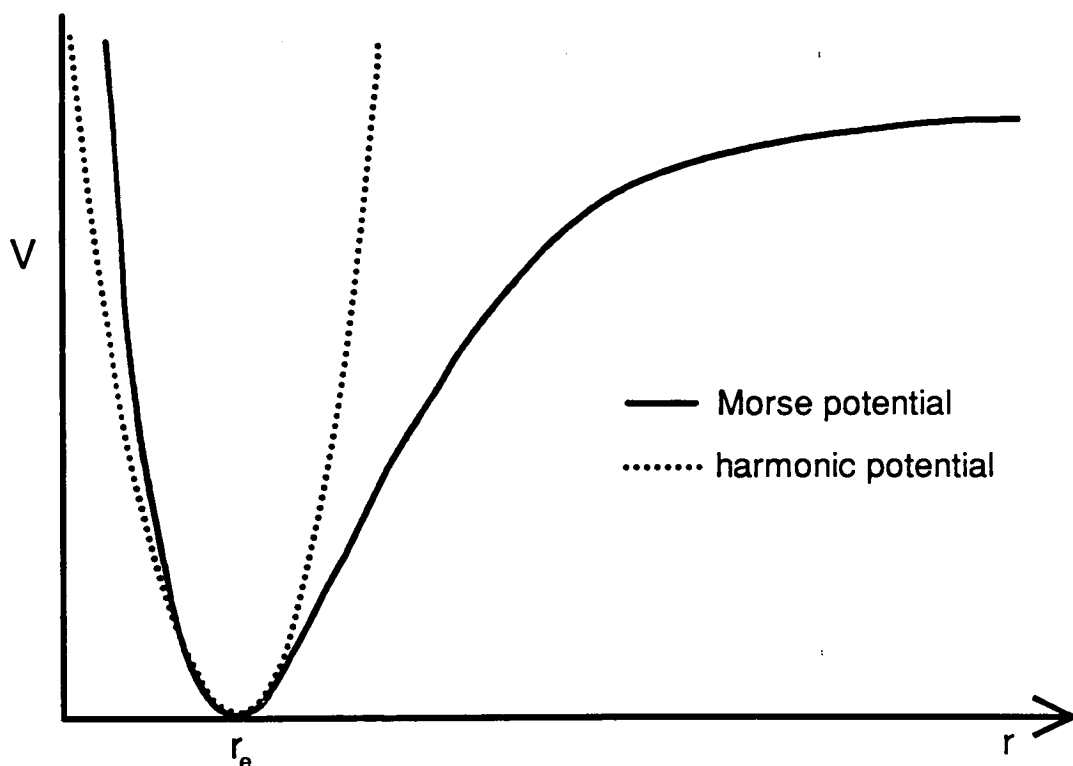
$$f = \frac{d^2V}{dq^2} \quad [4.20]$$

Hence, a knowledge, in terms of force constants, of the potential function of a molecule can be used to determine information about the vibrational motion of the molecule.

The above equations describe a diatomic molecule. Furthermore, it has been assumed that the atom pair behaves as a simple harmonic oscillator (this is implicit in equation 4.13). It transpires that this approximation is good for many modes of vibration in a wide range of molecules. Figure 4.2 shows the potential energy of a diatomic molecule, as a function of internuclear separation (r). The dotted line represents the potential function obtained by the harmonic approximation (see equation 4.19). The two curves deviate for small r , where the true curve is steeper due to interatomic repulsion, and for large r , where the true curve tends asymptotically to the bond dissociation energy. However, close to the equilibrium separation (r_e) the two curves are very similar; the harmonic approximation is good.

Figure 4.2

The potential function of a diatomic molecule



The extension of the above theory to the vibrational analysis of polyatomic molecules increases the mathematical complexity of the problem but the principles are largely the same. It becomes convenient to formulate the expressions using matrix notation and, in most cases, it is necessary to carry out the calculations using a computer. The most commonly adopted system is that of Wilson et al.^{13,14} and many computer programs have been developed for the routine vibrational analysis of polyatomic molecules^{11,15}.

A molecule with N atoms has $3N-6$ vibrational degrees of freedom ($3N-5$ in the case of a linear molecule). The vibrational motion of the molecule can, therefore, be resolved into $3N-6$ vibrational modes. These modes can be described in terms of internal co-ordinates which are divided into five basic types:

- bond stretches
- angle bends
- linear angle bends
- torsions
- trigonal out-of-plane bends

The set of internal co-ordinates, \mathbf{q} , is related to the Cartesian displacement co-ordinates, \mathbf{x} , by the expression

$$\mathbf{q} = \mathbf{B}\mathbf{x} \quad [4.21]$$

The elements of the matrix, \mathbf{B} , can be calculated using equations for each of the five types of internal co-ordinate, listed by Wilson et al.¹³ It is often convenient to define the vibrational motion in terms of symmetry

co-ordinates, \mathbf{Q} , which are linear combinations of the internal co-ordinates and are usually grouped according to symmetry. The coefficients for the linear combination are contained in a matrix, \mathbf{U} , such that

$$\mathbf{Q} = \mathbf{U}\mathbf{q} \quad [4.22]$$

The \mathbf{B} -matrix is used to calculate a matrix, \mathbf{G} , with the equation

$$\mathbf{G} = \mathbf{B}\mathbf{M}^{-1}\mathbf{B}' \quad [4.23]$$

where \mathbf{B}' denotes the *transpose* of \mathbf{B}

and \mathbf{M}^{-1} is the inverse of a diagonal matrix, \mathbf{M} , whose elements are the atomic masses (occurring three times for each atom).

Each internal co-ordinate (or symmetry co-ordinate) has an associated force constant. In addition, it is necessary to define a set of off-diagonal force constants, or interaction constants. This gives a more general definition of the force constants

$$f_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j} \quad [4.24]$$

where $i=j$ for the diagonal force constants (cf. equation 4.20)

and $i \neq j$ for the interaction constants.

These force constants become the elements of a symmetric matrix, \mathbf{F} .

If the \mathbf{G} and \mathbf{F} matrices are known, the vibrational frequencies can be calculated by solving the equation

$$\mathbf{G}\mathbf{F}\mathbf{L} = \mathbf{L}\mathbf{\Lambda} \quad [4.25]$$

where \mathbf{L} contains the eigenvectors of the product $\mathbf{G}\mathbf{F}$

and $\mathbf{\Lambda}$ is a diagonal matrix whose elements are the corresponding eigenvalues, λ_i .

The vibrational wavenumbers (in cm^{-1}) are related to the eigenvalues by

$$\lambda_i = 4\pi^2 c^2 \omega^2 \quad [4.26]$$

When carrying out a vibrational analysis, the vibrational frequencies are usually known and it is the force constants which must be determined. Unfortunately, there may be no unique solution to equation 4.24 if the \mathbf{F} -matrix is not known. For a molecule with n vibrational modes there are $n(n-1)/2$ unknown elements of the \mathbf{F} -matrix and so the problem is seriously underdetermined.

The use of symmetry co-ordinates, rather than internal co-ordinates, improves the situation as the \mathbf{F} -matrix can then be "symmetry blocked"; interaction constants between modes of different symmetry can be equated to zero. Additional observations can be used to constrain the force constants further. These may include the vibrational frequencies of isotopically substituted species, Coriolis coupling constants (arising from the interaction of rotation and vibration) or experimentally determined vibrational amplitudes. However, in most cases, the \mathbf{F} -matrix is arrived at by a combination of limited refinement and trial-and-error. For this reason, the results of the force-field analysis are often unreliable.

Computing

Parallel and perpendicular amplitudes of vibration, covariance matrices and vibrational corrections to rotation constants can be obtained from vibrational analysis programs such as **GAMP**¹⁶ and **asym20**¹¹. Direct coupling constants are corrected from D^o to D^α using the program **bmgy**¹⁷ which is based on equations 4.8 to 4.10 above. B_o rotation constants are converted to B_z by including the correction term as described in equation 4.12.

The combined analysis of data from LCNMR, rotational spectroscopy and electron diffraction can be carried out using the program, **ed92**, described in chapter one. The basic requirement is that the program can use the model molecular geometry to calculate theoretical values for the extra data. The "intensity" vector, **I** is extended to include the extra observations and the vector, **D**, can now also include differences between theoretical and experimental values of the extra data (see equation 1.10). The weighting matrix, **W**, is extended with diagonal terms only, each of which is proportional to the reciprocal square of the estimated uncertainty of the datum in question. During the least squares refinement, the estimated standard deviation of each ED data set is calculated and can be used as an uncertainty for the ED data. The weights of the extra data are then scaled so that the weights of both ED and non-ED data are proportional to the reciprocal squares of their uncertainties.

The calculation of theoretical values for the non-ED data is carried out in a subroutine called **EXTRA**, which must be written for the molecule under study. An example of such a subroutine can be found in appendix A.II.

The molecular geometry is passed into EXTRA as Cartesian co-ordinates in the arrays X(),Y() and Z(); the theoretical values for the extra data are returned in the array, E(). Library functions exist for the calculation of direct dipolar coupling constants (ED92XN) and rotation constants (ED92XM). The parameters for these functions are explained by comments in the source code.

The corrections to the ED distances, from r_a to r_α , are carried out by **ed92**. Parallel and perpendicular mean amplitudes of vibration, obtained from the harmonic force field analysis, are entered into the program and used as described in equations 4.4 to 4.6. To refine an r_α^o structure it is necessary to include the anharmonic term, $\frac{3a}{2}(u_T^2 - u_o^2)$, as defined in equation 4.7. This is done by simply adding this term to K_T before the values are entered. The program is then instructed to refine the structure as r_α and the analysis is continued.

Compatibility of Data

It has been shown above that it is possible to make the necessary vibrational corrections to combine data from ED, LCNMR and rotational spectroscopy to obtain, in principle, a self-consistent structure. However, it is important to consider the compatibility of such data further.

The structures derived from ED and rotational spectroscopy clearly correspond to the same physical geometry. In both cases the sample is in the gas phase and so the individual molecules can, to a very good approximation, be considered to be free from any external forces. This compatibility has been shown by the many successful combined analyses of this type^{18,19,20}.

LCNMR data, however, are recorded with samples in solution and so the question arises as to whether the structure is the same as that in the gas phase. There is no simple answer to this question, other than to say that it depends upon the particular molecule under study and the solvent used. Diehl and Niederberger²¹ showed that the structure of benzene, obtained by an LCNMR analysis, agreed, to within experimental error, with the r_g structure determined by ED. Similar results have been obtained for a number of other molecules, including pyridine²² and cyclopropane²³. On the other hand, an LCNMR analysis of norboradiene using the liquid crystal EBBA²⁴ produced a structure significantly different from that determined using Merck Phase IV as a solvent²⁵. Clearly, at least one of these structures must differ from the gas phase structure and so combined analysis, in this case, would be unwise.

Fortunately, it seems that the LCNMR structures of fairly rigid molecules tend to show no significant distortion in changing from the gas phase to solution. This is particularly true if care is taken to choose a liquid crystal solvent which shows a low degree of interaction with the solute. These factors were involved in the decision to study a range of substituted aromatic and heteroaromatic compounds at Edinburgh, by the combined analysis of ED, LCNMR and rotational data^{26,27,28}. The success of these studies is perhaps the best indication of the compatibility of data from the different techniques.

References

¹P.M.Morse, *Phys.Rev.*, **34**, (1929), 57

²A.Almenningen, O.Bastiansen & T.Munthe-Kaas, *Acta Chem.Scand.*, **10**, (1956), 261

³A.Almenningen, O.Bastiansen & M.Traetteberg, *Acta Chem.Scand.*, **13**, (1959), 1699

⁴A.Almenningen, O.Bastiansen & M.Traetteberg, *Acta Chem.Scand.*, **15**, (1961), 1557

⁵E.Hirota & Y.Morino, *J.Chem.Phys.*, **23**, (1955), 737

⁶I.Hargittai, "*Stereochemical Applications of Gas-phase Electron Diffraction*", Part A, Ch.1, eds. I.Hargittai & M.Hargittai, VCH, New York, (1988), 35

⁷A.G.Robiette, "*Molecular Structure by Diffraction Methods*", Volume 1, Chapter 4, eds. G.A.Sim & L.E.Sutton, Chem.Soc.SPR, London, (1973)

⁸S.Cradock, P.B.Liescheski, D.W.H.Rankin & H.E.Robertson, *J.Am.Chem.Soc.*, **110**(9), (1988), 2759

- ⁹S.Sykora, J.Vogt, H.Bösiger & P.Diehl, *J.Magn.Reson.*, 36, (1979), 53
- ¹⁰G.Graner, "*Accurate Molecular Structures: Their Determination and Importance*", Chapter 4, eds. A.Domenicano & I.Hargittai, Oxford University Press, (1992), 65
- ¹¹L.Hedberg & I.M.Mills, *J.Mol.Spectr.*, 160, (1993), 117
- ¹²P.Gans, "*Vibrating Molecules*", Chapman & Hall, London, (1971)
- ¹³E.B.Wilson, J.C.Decius & P.C.Cross, "*Molecular Vibrations*", McGraw-Hill, New York, (1955)
- ¹⁴S.Califano, "*Vibrational States*", Wiley & Sons, London, (1976)
- ¹⁵W.F.Murphy, *Can.J.Chem.*, 69, (1991), 1672
- ¹⁶S.Cradock, G.S.Laurenson & D.W.H.Rankin, *J.Chem.Soc., Dalton Trans.*, (1981), 187
- ¹⁷S.Cradock, private communication
- ¹⁸K.Kuchitsu & M.Nakata, "*Stereochemical Applications of Gas-phase Electron Diffraction*", Part A, Ch.7, eds. I.Hargittai & M.Hargittai, VCH, New York, (1988), 227
- ¹⁹R.K.Heenan, A.G.Robiette, *J.Mol.Struct.*, 54, (1979), 135
- ²⁰K.Tamagawa, T.Iijima, M.Kimura, *J.Mol.Struct.*, 30, (1976), 243
- ²¹P.Diehl & W.Niederberger, *J.Magn.Reson.*, 9, (1973), 495
- ²²J.W.Emsley, J.C.Lindon & J.Tabony, *J.Chem.Soc., Faraday Trans.*, 71, (1975), 579
- ²³N.J.D.Lucas, *Mol.Phys.*, 22, (1971), 233

²⁴E.E.Burnell & P.Diehl, *Canad.J.Chem.*, 50, (1972), 3566

²⁵J.W.Emsley & J.C.Lindon, *Mol.Phys.*, 29, (1975), 531

²⁶S.Cradock, P.B.Liescheski, D.W.H.Rankin & H.E.Robertson, *J.Amer.Chem.Soc.*,
110(9), (1988), 2759

²⁷P.B.Liescheski & D.W.H.Rankin, *J.Mol.Struct.*, 196, (1989), 1

²⁸S.Cradock, P.B.Liescheski & D.W.H.Rankin, *J.Magn.Reson.*, 91, (1991), 316

Chapter 5

Modifications to the LCNMR Analysis Programs

Introduction

As mentioned in chapter two, it is often necessary to use computer programs in the analysis of LCNMR spectra. However, the programs used in Edinburgh (**lequor** and **sliquor**, described above) had become somewhat outdated and, indeed, proved inadequate in the analysis of some of the more complex spectra recorded. The spectra of molecules with low symmetry, such as 2-chloropyridine¹, could not be assigned successfully. Similar problems had been encountered in the LCNMR analysis of molecules containing more than three or four spin- $\frac{1}{2}$ nuclei, for example methyl silane² and o-difluorobenzene³. It was decided, therefore, to update the programs, making changes in the following three areas:

Input: to make general improvements to the input procedures.

Output: to take advantage of newly available computing equipment in producing a high resolution graphical output, with hardcopy options.

Calculation: to account for vibrational effects when calculating the observed direct coupling constants (D^o).

It was hoped that such changes would allow a reanalysis of the unsolved spectra listed above and also pave the way for future work on more complex systems.

Because of the extent of the modifications made to **lequor**, and to avoid confusion with older versions of the program reported in the literature^{4,5}, it was decided to rename it **lcsim** (from **LCNMR simulation**).

Input

A considerable source of frustration when using the program **lequor** was the inflexibility and inconsistency of the input procedure. The program would crash at the slightest input error and the whole input would have to be retyped. Furthermore, the user had to input what seemed to be unnecessary information. For example, after entering the nuclear types (^1H , ^{13}C etc.) it was still necessary to input the spin multiplicity of each nucleus.

To overcome these problems a stand-alone program, **makeicsim**, was written to produce input files for **icsim**. Wherever possible the input to **makeicsim** is in free format and, therefore, less prone to crashing. There are also a number of error checks included to reduce the chances of mistakes being made. The program uses a lookup table, **mult()**, to obtain the multiplicity of each nucleus. The source code for **makeicsim** can be found in the INPUT section of appendix A.III.

Two further changes were made to the **icsim** input procedure. A function was added to convert any characters to upper case so that the interactive section of the program became case-insensitive. A more important addition was the ability to use a numbering scheme of the user's choice (in previous versions of the program the nuclei had to be numbered consecutively, starting at one). A consistent numbering scheme greatly reduces the chance of human error when entering the various spectral parameters by removing the need to convert between the **icsim** numbering and the user's numbering.

Two arrays, **uti** and **ltu**, were introduced to allow the program itself to perform the translation. These are automatically set up by the program, such that

uti(n) gives the equivalent internal index number of the nucleus with user index number **n**.

and **ltu(n)** gives the equivalent user index number of the nucleus with internal index number **n**

These two arrays act like dictionaries translating input from and output to the user, respectively. Similar changes were made to the refinement program **sliquor** and, wherever possible, consistency between the two programs was maintained.

Output

Previous modifications to **lequor** and **sliquor** made it possible to produce a graphical output of the simulated spectrum⁵. However, this output was extremely crude, consisting of little more than columns of asterisks representing the peaks in the spectrum (see figure 5.1). Using an eighty column terminal the resolution of such a plot is inadequate for simulation of some of the complex second-order spectra produced in the LCNMR experiment. Modern terminals allow a much more sophisticated graphical output; in the case of the SUN terminals available in the Chemistry Department, a horizontal resolution of more than 1000 pixels is possible. Both programs were modified to take advantage of this technology. At the same time, an additional option was added, which allows a representation of the experimental spectrum to be displayed on the same plot as the calculated spectrum. This makes possible the direct visual comparison of observed and theoretical line positions.

The source code for the graphical output routines⁶ can be found in the OUTPUT section of Appendix A.III. The routines make extensive use of the UNIRAS graphics libraries⁷. These allow lines and text to be plotted to an X-window on the screen or to any other output device for which a suitable driver is available, including the departmental graph plotter and laser printer. This has the advantage that it is unnecessary to write separate routines for screen output and hardcopy output (the subroutine **xplot** performs both functions).

Input of the experimental spectrum is via a text file called **lcsim.exp** (or **sliq.exp** in the case of **sliquor**), each line giving the frequency and

intensity of one peak. A similar file is created by the program (called **lcsim.calc** or **sliq.calc**) containing the calculated frequencies and intensities. If required, the experimental intensities can be scaled to match the calculated intensities using the program **scale**; this program also allows a minimum intensity threshold to be applied to the experimental lines. The calculated lines are plotted above the frequency axis whereas the experimental lines are plotted below the axis. In the case of **lcsim**, the current values of the orientation parameters are listed beside the plot. Examples of the plots possible with **lcsim** and **sliquor** are shown in figures 5.2 and 5.3 respectively.

The routines **xopen** and **xclose** open and close the X-window on the screen, in which the plot will appear. The size and position of this window is determined by a command line option when running the program. For the sake of convenience, the program is run from a shell script which contains default values for the window geometry. The shell script also determines the type of terminal being used and passes this information to the program. In this way, output to an X-window is not attempted unless supported by the terminal.

The routine **xdraw** allows the user to select a hardcopy device and then calls the appropriate subroutines to create a plot file which can later be sent to that device. This feature is available irrespective of the type of terminal being used.

As a result of these modifications it becomes much easier and quicker to compare the calculated and experimental spectra, visually. This is vital when attempting to determine approximate values for the orientation parameters by trial and error.

Calculation

Theoretical direct coupling constants are calculated by **lcsim** using internuclear distances from an approximate structure. Clearly, the closer this structure is to the "true" structure, the better the predicted couplings will be. However, the original version of the program failed to take molecular vibration into account and so, even using the best available structure, the direct coupling constants would not correspond to the experimental values. In other words, the spectrum was simulated using D^α constants, calculated from an r_α structure, whereas observed spectra derive from D^0 coupling constants, related to D^α according to equations 4.8 to 4.10. In many cases, the effect of vibrational averaging is small but in the analysis of complex spectra, in particular satellite spectra, it should not be ignored. For this reason, it was decided to incorporate vibrational corrections into **lcsim**.

The correction from D^α to D^0 is essentially the reverse of that carried out by the program **bmgv**, described above. The covariance matrix is obtained from a harmonic force field analysis of the molecule and is contained in a file along with the components of the internuclear vectors. This information is all that is required to evaluate the elements of the tensor Φ^h , as defined in equation 4.10. The remainder of the calculation is carried out using equations equivalent to 4.8 and 4.9.

The subroutine **DDCNMR**, which calculates the direct dipolar coupling constants, was modified to carry out the necessary vibrational corrections. The relevant source code can be found in Appendix A.III. Much of the calculation section was copied directly from **bmgv** and it was at this time

Figure 5.2

lcsim simulation of the ^1H LCNMR spectrum of 2 chloropyridine

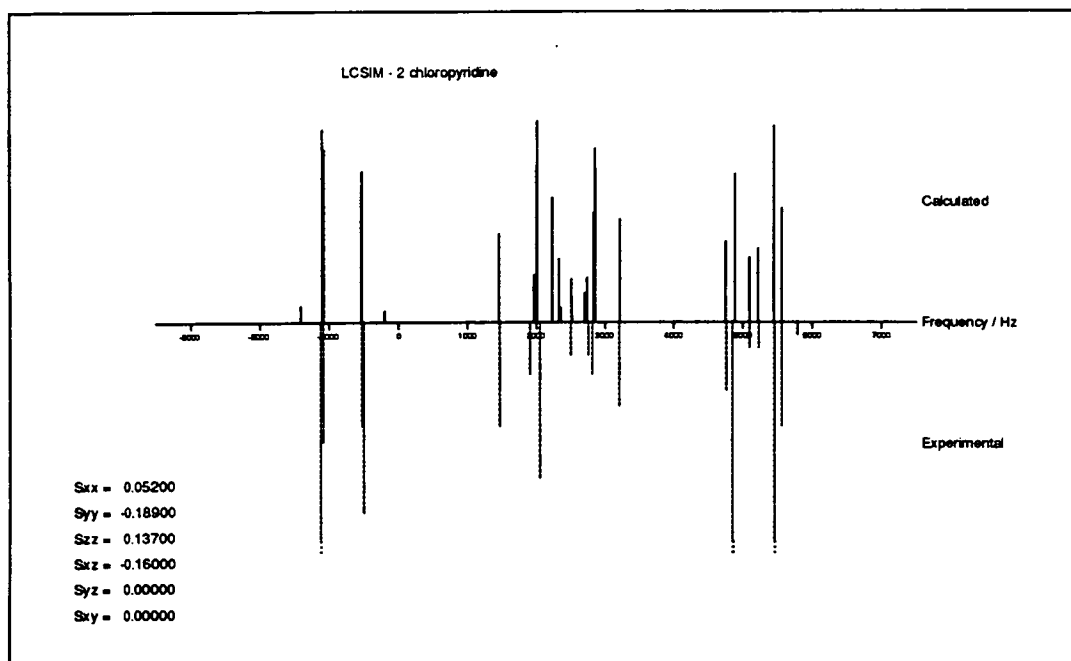
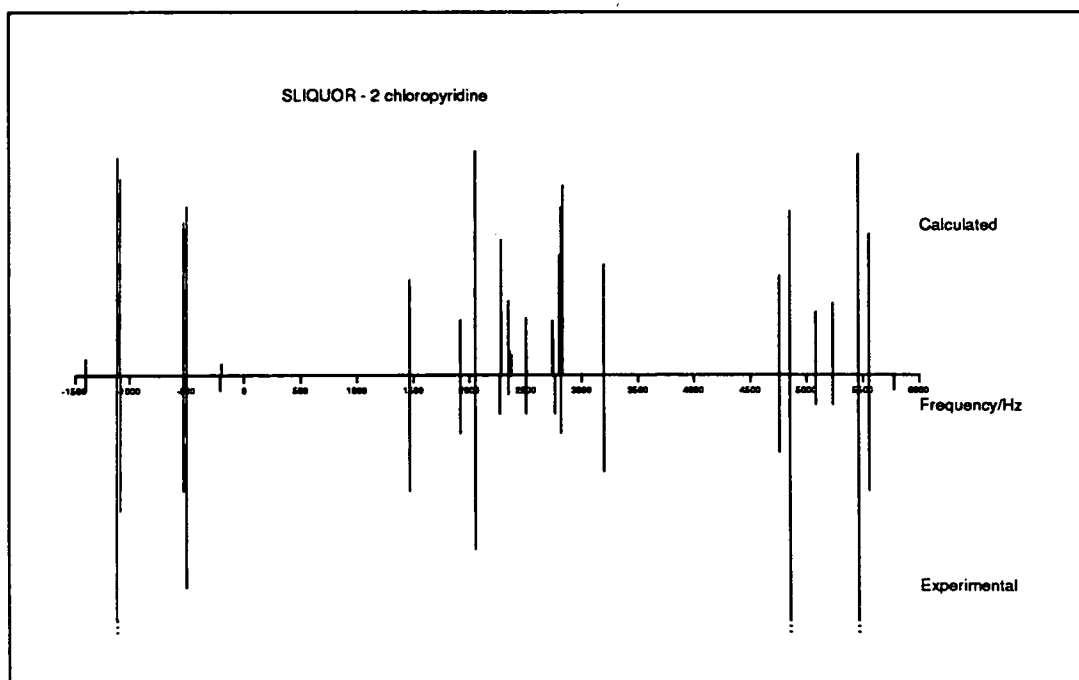


Figure 5.3

sliquor simulation of the ^1H LCNMR spectrum of 2 chloropyridine



Example - The LCNMR Analysis of 2 Chloropyridine

The LCNMR spectrum of 2 chloropyridine was recorded on the Edinburgh University Bruker WH360 MHz spectrometer using the liquid crystal solvent E7, which is nematic at room temperature. Previous attempts to analyse this spectrum failed¹ due largely to the fact that there are three unknown orientation parameters required to calculate the direct dipolar coupling constants. Even a limited search looking at all possible combinations of orientation parameters in the range of +0.2 to -0.2 with a step size of 0.05 would produce 729 (i.e. 9^3) simulated spectra. When it is considered that, in principle, orientation parameters can take any value between -0.5 and +1.0, it is clear that the problem is by no means trivial. Using the output of **lequor** to identify those simulations which resembled the experimental spectrum proved impossible, due mainly to the poor resolution of the simulated plot. It is quite conceivable that a close fitting simulation may have been overlooked (compare the theoretical spectrum in figure 5.1 with the experimental spectrum in figure 5.4).

A reanalysis of the same spectrum using **icsim**, incorporating the improvements described above, was carried out. The ring structure was estimated from a comparison of similar molecules and by use of the superposition method, described by Brookman¹, in which angles are calculated by averaging the equivalent angles in simpler molecules (in this case, chlorobenzene and pyridine). The C-H bonds were assumed to bisect the external ring angles and to be 1.085 Å in length. Initial values for the proton chemical shifts (ω) and the indirect coupling constants (J) were taken from the analysis by Cox and Bothner-By⁹. A vibrational analysis was

carried out using a modified version of Allinger's molecular mechanics program MM3¹⁰ (the modifications are described in a subsequent chapter).

Figure 5.4 - the ¹H LCNMR spectrum of 2 chloropyridine in E7

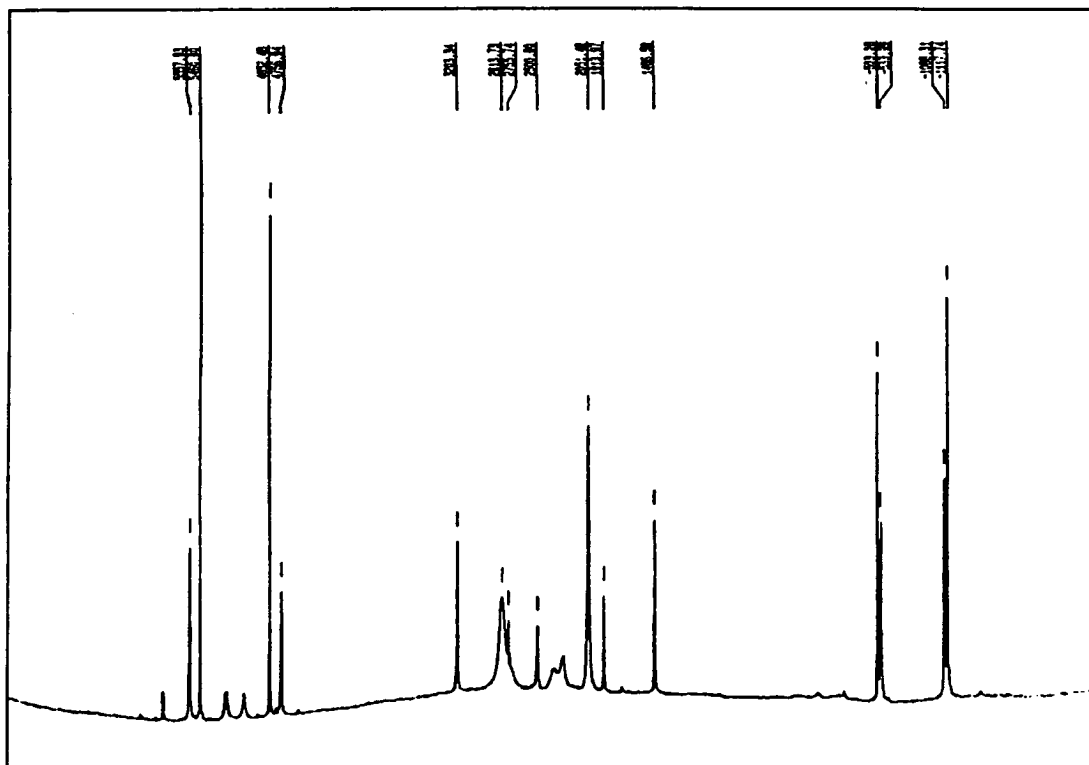
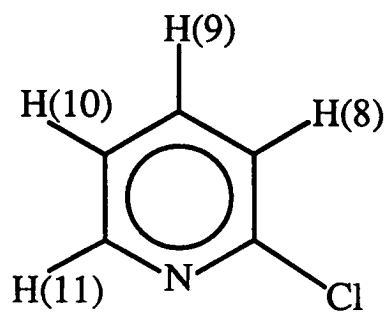


Figure 5.5 - the numbering of the hydrogen atoms in 2 chloropyridine



Once the data had been entered, **lcsim** was used to determine approximate values for the orientation parameters, and hence the direct coupling constants. The improved quality of the graphical output, in particular the inclusion of the experimental lines on the same plot, allowed a much quicker evaluation of the simulated spectra. Furthermore, it was possible to develop a "feel" for the effect of each of the orientation parameters and to predict what changes were required to alter certain features of the simulated spectrum (e.g. the overall width or the spacing of a particular pair of lines). By a combination of systematic search and a reasoned approach, orientation parameters were derived which gave a close enough simulation of the spectrum to allow the assignment of some of the individual lines (see figure 5.2).

The improvements to the spectrum simulation program had already proved effective but the analysis was continued to determine accurate values for the direct coupling constants. This involved the use of the program, **sliquor**, which refines approximate spectral parameters, derived from **lcsim**, to improve the fit of theoretical to experimental line positions. As the refinement progresses, more lines can be assigned until, all being well, the whole spectrum is assigned. The refined spectral parameters can then be considered as experimentally determined and should closely reproduce the experimental spectrum.

In the case of 2 chloropyridine, this process was not entirely successful. A total of 17 lines were unequivocally assigned and could be reproduced theoretically, with an rms deviation of just 2.6 Hz (figure 5.3). However, problems arose due to the presence of three very broad lines in the central region of the spectrum, at around 2500, 2348 and 2269 Hz. Line

broadening often occurs in such systems, due to the presence of the quadrupolar ^{14}N nuclei, but in this case it was exacerbated by the fact that there are many overlapping peaks in this region. Unfortunately, this meant that accurate experimental frequencies could not be determined for a number of the theoretical lines in the simulated spectrum. Consequently, it proved impossible to refine the direct coupling constants $D_{9,11}$ and $D_{10,11}$ which depend strongly on the positions of these lines (the numbering of the hydrogen atoms is shown in figure 5.5). Nonetheless, the remaining four direct coupling constants and the four chemical shifts were all successfully refined giving the values shown in table 5.1. The indirect coupling constants could not be refined and were assumed to be unchanged from isotropic values.

Table 5.1 - results of the LCNMR analysis of 2 chloropyridine

Parameter	Initial value (from lcsim)	Refined value (from sliquor)
ω_8	2324.9	2327.4(17)
ω_9	2166.8	2165.3(83)
ω_{10}	2167.4	2171.3(83)
ω_{11}	2546.1	2551.0(29)
$D_{8,9}$	-322.4	-315.0(47)
$D_{8,10}$	-294.8	-333.7(48)
$D_{8,11}$	-257.3	-226.5(17)
$D_{9,10}$	-1945.4	-1950.4(8)
$D_{9,11}$	-221.392	-
$D_{10,11}$	188.314	-

Conclusions

The successful analysis of the spectrum of 2-chloropyridine demonstrates that the changes made to the LCNMR analysis programs are genuinely useful ones. In particular the improved graphical output greatly reduces the difficulty in determining approximate values for the orientation parameters. The fact that the assignment is incomplete is due to the limitations of the spectrum rather than limitations in the analysis.

The inclusion of vibrational corrections in **lcsim** makes it worthwhile to obtain as good an initial structure as is available to increase the chances of producing a good simulation. This is of particular importance where the couplings are large in which case the corrections can amount to several hundred Hertz.

The modified programs have also been successful in the reanalysis of the spectra of *o*-difluorobenzene (see below) and methyl silane², both of which were previously unsolved, and in the LCNMR analysis of thiazole¹¹.

References

¹C.A.Brookman, Ph.D.Thesis, University of Edinburgh, (1993), 174

²P.J.Mathieson, unpublished results

³A.Reive, Honours project, University of Edinburgh, 1992

⁴P.Diehl, H.P.Kellerhals & W.Niederberger, *J.Magn.Reson.*, **4**, (1971), 352

⁵S.Cradock, P.B.Liescheski & D.W.H.Rankin, *J.Magn.Reson.*, **91**, (1991), 316

⁶these routines were written jointly with P.J.Mathieson

⁷UNIRAS Ltd., "*FGL/GRAPHICS User Guide - Version 6*", Manchester Computing Centre,
1st edition IUCR reprint, (1989)

⁸R.Wasser, M.Kellerhals & P.Diehl, *Magn.Reson.Chem.*, 27, (1989), 335

⁹R.H.Cox & A.A.Bothner-By, *J.Phys.Chem.*, 73(8), (1969), 2465

¹⁰P.Aped & N.L.Allinger, *J.Am.Chem.Soc.*, 114, (1992), 1

¹¹H.Gierens, Undergraduate Project Report, University of Edinburgh, 1994

Chapter 6

Using MM3 to Calculate Vibrational Corrections

Introduction

The importance of vibrational analysis in the combined analysis of data from different structural techniques has been emphasised above. Often, this can be one of the most severe limitations to the accuracy of the structure obtained. Determination of a unique harmonic force field is not usually possible due to the relatively small number of observable data, compared to the number of unknown force constants. For this reason, there are usually large uncertainties associated with any results calculated using the such a force field.

Naively, one might assume the force constants associated with similar internal co-ordinates or normal modes to be transferable between different molecules. This would allow these constants to be fixed in the refinement, thereby reducing the problem of underdetermination of the force field. In practice, however, this is not the case. The force constant for a particular bend in one molecule can be significantly different to that for a similar bend in another molecule. At best, the force constants for previously studied molecules can be used to obtain initial values for the force field refinement of a new molecule. This variation is largely due to the lack of explicit consideration of van der Waals interactions between atoms. In order to fit the observed vibrational frequencies, force constants are adjusted until they compensate for the effects of such interactions. Because van der Waals interactions can vary considerably between molecules, the final values of the force constants may also vary and consequently are not completely transferable.

An alternative approach would be to calculate the van der Waals interactions separately and then use these to calculate the effective force constants from a set of "pure" force constants (which should now be transferable between molecules). This method is employed in molecular mechanics calculations¹ such as those performed by the program **MM3**^{2,3}. In addition to a set of transferable force constants, the program uses a set of structural parameters, defining the "natural" values of the various bond lengths and angles. The current **MM3** force field includes sufficient data to perform calculations over a wide range of molecules, although it is best suited to studies of organic molecules. If necessary, extra parameters can be added to the existing set; values can be estimated by comparison with those of similar internal co-ordinates. However, these parameters will be less reliable, as the "official" parameters have usually been tested for a number of molecules.

The use of the **MM3** force field to calculate vibrational corrections to ED, rotational spectroscopy and LCNMR data would greatly simplify the process of combined analysis. In its original form, the program calculates the parallel and perpendicular amplitudes of vibration necessary for the refinement of an r_{α}° structure by electron diffraction (equations 4.4 to 4.6). The harmonic contributions to the rotation constants, required to convert B_0 rotation constants to B_z (equation 4.11), are also calculated. However, the covariance matrices, used to correct LCNMR couplings from D° to D^{α} (equations 4.7 to 4.9), are not calculated by the standard program. It was therefore necessary to modify the program to include such calculations.

Calculation of Covariance Matrices

Sýkora et al⁴ have shown that the elements of the covariance matrix for a pair of atoms, *i* and *j*, are given by

$$C_{\alpha\beta}^{ij} = \langle \Delta_{\alpha} \Delta_{\beta} \rangle_{\text{vib}} \quad [6.1]$$

where Δ_{α} and Δ_{β} are components of the instantaneous excursion of the internuclear vector from the equilibrium position ($\alpha, \beta = x, y, z$). These components are related to $u_{i\alpha}^{(v)}$, the mass-weighted Cartesian components of the normal-coordinate vector of the v^{th} vibrational mode, by the equation

$$\Delta_{\alpha}^{ij} = \sum_{v=1}^{3N} (u_{i\alpha}^{(v)} - u_{j\alpha}^{(v)}) Z_v \quad [6.2]$$

where Z_v is the amplitude of the normal co-ordinate of the v^{th} vibrational mode. The terms $u_{i\alpha}^{(v)}$, closely related to the elements of the eigenvector, **L**, of equation 4.24, are calculated by **MM3** and stored in the array **eigvec()**. The normal co-ordinate amplitude, Z_v , depends only upon the frequency of the v^{th} vibrational mode and the temperature, *T*. This relationship is conveniently expressed in the form

$$\langle Z_v^2 \rangle_{\text{vib}} = \frac{A}{\omega_v} \coth(B\omega_v T^{-1}) \quad [6.3]$$

where ω_v is the frequency expressed in wavenumbers (cm^{-1})

and *A* & *B* are constants given by

$$A = \frac{h}{8\pi^2 c} \quad [6.4] \quad \text{and} \quad B = \frac{\bar{c} h}{2k} \quad [6.5]$$

where \bar{c} is the velocity of light (in cm s^{-1}) and *k* is the Boltzmann constant.

The above expressions can be combined to give the working equation

$$C_{\alpha\beta}^{ij} = \sum_{v=1}^{3N} (u_{i\alpha}^{(v)} - u_{j\alpha}^{(v)})(u_{i\beta}^{(v)} - u_{j\beta}^{(v)}) \frac{A}{\omega_v} \coth(B\omega_v T^{-1}) \quad [6.6]$$

which is incorporated into the **MM3** subroutine **cyvin()**. Extracts from the source code for this routine can be found in Appendix A.IV. The lines which were added to calculate the covariance matrices are in upper case. The routine already included nested loops, over all normal modes (variable **L**) and all atom pairs (variables **i** and **j**), and so the calculation of the covariance matrices could conveniently be included at this point.

The amplitude term, $\langle Z_v^2 \rangle_{\text{vib}}$, is calculated and stored in the variable **FWT**. Inside a further pair of nested loops (effectively $\alpha, \beta=x,y,z$) the eigenvector terms are evaluated and the resulting contributions to the covariance matrix elements are summed with each iteration of the normal mode loop. Finally, the covariance matrix for atom pair **i,j** and the components of the internuclear vector are written to the file **dcmm3**, in a format which can be read directly by the program **bmgv** (described above). A second file, **covar.mm3**, is created and contains only the covariance matrices. This was used in the development and testing of the modifications and these lines could be deleted if necessary.

Examples

The changes made to **MM3** were tested by comparing the resulting vibrational corrections with those obtained by other methods. This also helped to give an indication of the suitability of the **MM3** force field for such calculations.

Vibrational Amplitudes

To obtain an r_{α} molecular geometry it is necessary to carry out corrections involving mean square parallel and perpendicular amplitudes of vibration (see equations 4.4 to 4.6). In general, the parallel amplitudes can be refined to fit the ED data and so the absolute values obtained from the vibrational analysis are not the final values used in the structural analysis. More important are the relative values of the amplitudes associated with similar interatomic distances. These can not usually be refined separately; instead they are tied together with fixed ratios and refined as a group. Amplitudes obtained using the **MM3** force field have been used in the structural analysis of various aromatic compounds⁵ and the results are consistent with similar studies using force fields obtained from a normal co-ordinate analysis. A more direct comparison can be made by comparing the amplitudes calculated by the two methods for molecules with well determined force fields. Table 6.1 shows the results of such calculations for the molecules formaldehyde, methyl fluoride and dichloroethene.

Table 6.1

Comparison of vibrational amplitudes calculated from experimentally determined force fields (using **asym20**) and from semi-empirical molecular mechanics force fields (**MM3**)

Molecule	Nuclei	U asym20 Å	U mm3 Å	K asym20 Å	K mm3 Å
formaldehyde 	1,2	0.0374	0.0392	0.00043	0.00045
	1,3	0.0780	0.0800	0.00884	0.00880
	2,3	0.0912	0.0940	0.00297	0.00305
	3,4	0.1188	0.1216	0.00514	0.00522
methyl fluoride 	1,2	0.0464	0.0462	0.00037	0.00045
	1,3	0.0766	0.0779	0.01086	0.01149
	2,3	0.1014	0.1078	0.00387	0.00400
	3,4	0.1272	0.1290	0.00978	0.01061
1,1 dichloroethene 	1,2	0.0423	0.0433	0.00365	0.00393
	1,3	0.0477	0.0514	0.00194	0.00216
	2,5	0.0755	0.0769	0.01970	0.01661
	5,6	0.1180	0.1223	0.02410	0.01948
	3,4	0.0654	0.0724	0.00018	0.00023
	4,5	0.0947	0.0991	0.00564	0.00524
	4,6	0.1451	0.1429	0.00580	0.00502
	1,5	0.0985	0.0979	0.01211	0.01103
2,4	0.0607	0.0676	0.00077	0.00089	

For the **asym20** calculations, the force fields for formaldehyde, methyl fluoride and dichloroethene are taken from references 6, 7 and 8 respectively.

These results show a good agreement between the amplitudes calculated by the two methods. The magnitude of the vibrational correction is usually very small compared to the interatomic distance and so slight errors in the calculated amplitudes are insignificant. This assumption could break down in the case of molecules with large amplitude motions. However, for the small, fairly rigid molecules described in the remainder of this thesis, amplitudes calculated using the **MM3** force field are accurate enough.

Corrections to LCNMR Data

The suitability of **MM3** for the calculation^d of vibrational corrections to direct dipolar coupling constants was investigated by calculating corrections for a number of molecules for which D^α constants were available in the literature. In general, the agreement was remarkably good, particularly for the class of compounds currently under study at Edinburgh (i.e. small-ringed aromatic molecules).

Table 6.2 shows a comparison of results for chlorobenzene, as calculated by Diehl⁹, Craddock et al¹⁰ and by **MM3**; all calculations relate to the original NMR data of Diehl⁹. Two main conclusions can be drawn from these collected results. Firstly, the corrections in all three cases are similar, for most of the couplings. Secondly, there is as much disagreement between the two sets of published data as there is between either of these sets and the results determined using **MM3**. If anything, the results from **MM3** are closer to those of Diehl. This is not entirely unexpected because the force constants used by Diehl are generalised ones which, like those of the **MM3** force field, can be applied to a range of similar molecules. It is not

possible to draw any firm conclusions as to which corrections are "best" but it can be said that the use of the **MM3** force field produces corrections which are as good as those obtained by other methods.

These results are possibly clearer when presented in graphical form. Figure 6.1 shows the corrections made in each of the three studies as a percentage of the relevant coupling constant. It is immediately apparent that the **MM3** corrections are very similar to those of the published data. In figure 6.2, a similar graph is presented showing the vibrational corrections to the couplings of pyrimidine. This is based on the original LCNMR data of Diehl et al.¹¹, a combined analysis by Rankin et al.¹² and new corrections calculated using the **MM3** force field. Once more the corrections calculated using **MM3** are very similar to those of Diehl et al. However, in this case, the corrections from reference 12 are consistently larger than those from the other two studies.

Similar comparisons have been carried out for a number of different molecules and the general conclusion is that the **MM3** force field is suitable for vibrational corrections of this kind. Once more, it is likely that problems would be encountered if molecules which undergo large amplitude vibrations were considered. However, this limitation also applies to vibrational analysis by any other method.

Table 6.2 - Vibrational corrections (d^h) to LCNMR data for chlorobenzene

Coupling	D° (ref.9)	d^h (ref.9)	d^h (ref.10)	d^h (MM3)
$D_{1,8}$	41.67	4.32	4.39	4.80
$D_{2,8}$	879.88	9.53	11.61	8.74
$D_{3,8}$	276.32	2.43	3.28	2.40
$D_{4,8}$	55.47	1.41	1.60	1.23
$D_{1,9}$	56.59	1.11	1.31	1.17
$D_{2,9}$	278.13	2.56	2.36	2.78
$D_{3,9}$	846.16	9.43	10.10	9.83
$D_{4,9}$	42.35	4.25	3.66	4.96
$D_{8,9}$	873.79	1.60	2.81	1.65
$D_{1,10}$	59.13	1.13	0.86	1.10
$D_{2,10}$	76.57	1.46	0.95	1.46
$D_{3,10}$	244.10	2.32	1.38	2.59
$D_{4,10}$	2509.79	8.69	8.91	8.64
$D_{8,10}$	132.29	1.38	1.34	1.29
$D_{9,10}$	295.87	2.62	1.74	2.77
$D_{2,11}$	20.07	0.70	0.80	0.75
$D_{3,11}$	12.19	0.98	0.90	1.15
$D_{8,11}$	38.12	0.71	1.23	0.71
$D_{9,11}$	20.77	1.40	1.11	1.49
$D_{2,12}$	12.53	1.04	1.44	1.04
$D_{3,12}$	20.61	0.63	1.07	0.58
$D_{8,12}$	21.03	1.43	1.71	1.43

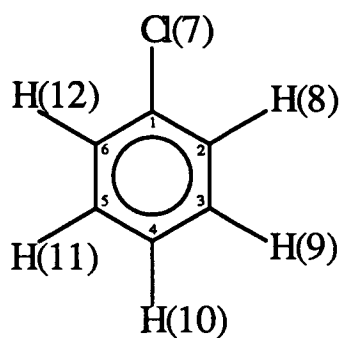


Figure 6.1 - Relative corrections to chlorobenzene coupling constants

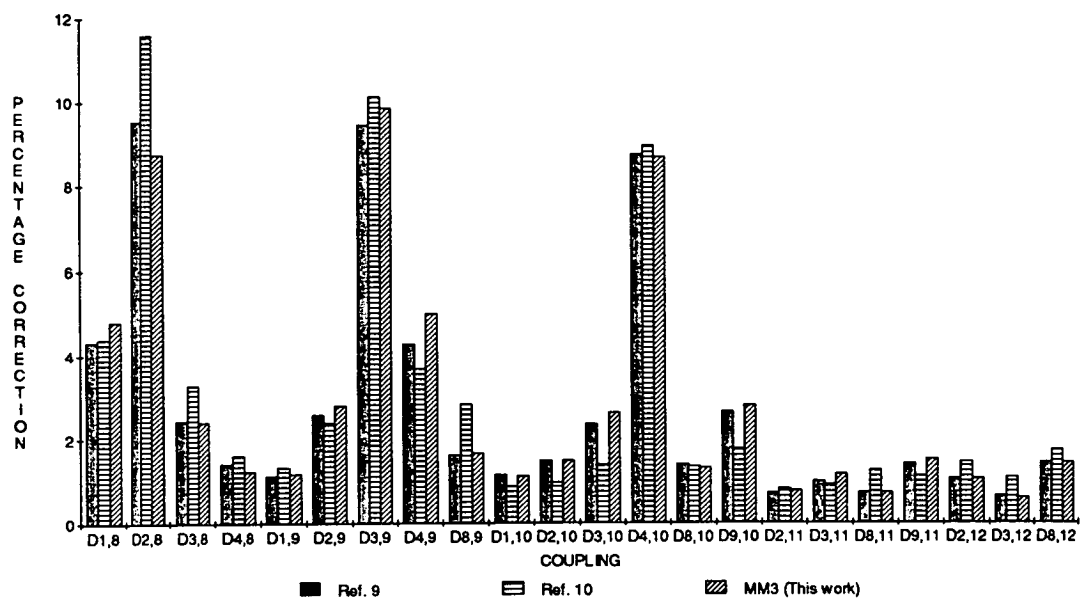
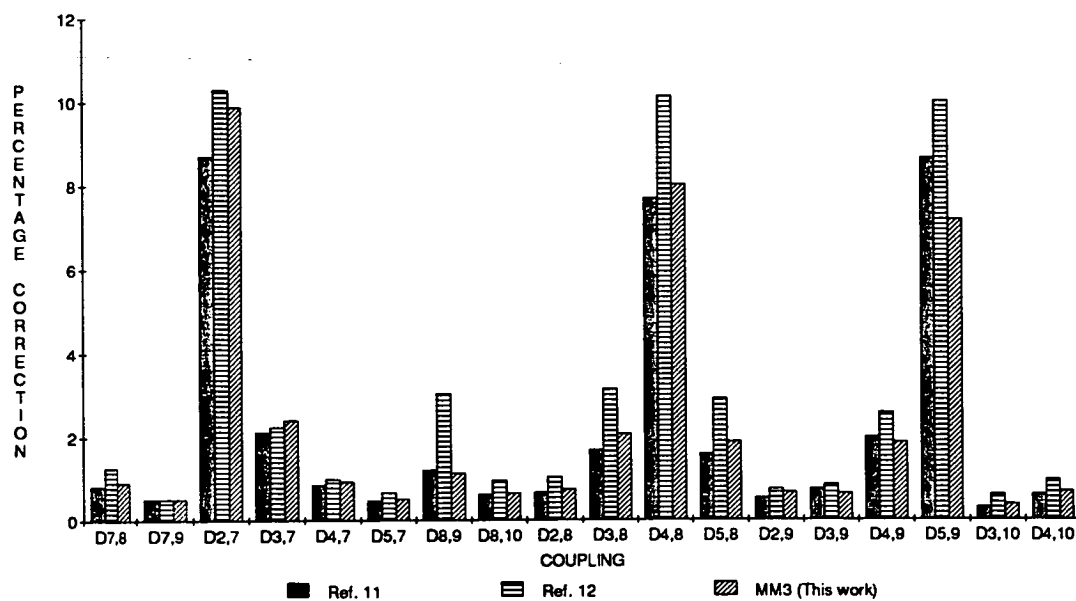


Figure 6.2 - Relative corrections to pyrimidine coupling constants



It is important to include estimated uncertainties for any data used in a combined analysis, in order to give the appropriate relative weights to each of the data during the least squares refinement (see Chapter 4). The uncertainties of LCNMR data are a combination of the uncertainties of the experimentally determined couplings (as calculated by the program **sliquor**) and the uncertainties due to the vibrational corrections. In most cases it is the latter which is dominant so it is important to make an estimate of these uncertainties. In previous studies^{10,12,13} this was done by making variations to the force field, while retaining a reasonable fit to the observed vibrational frequencies, and using the dispersion of the results obtained as an estimate of the uncertainty due to the vibrational correction. With the **MM3** force field this is not a practicable solution as most of the force parameters are an integral part of the program and cannot easily be varied. However, a survey of the results from a number of previous studies^{10,12,13} indicates that, to a first approximation, the uncertainty is proportional to the size of the correction term. The uncertainty in the corrected coupling constants (D^α) can initially be estimated to be approximately ten percent of the correction (d^h). This produces uncertainties which are reasonably consistent with those used in previous combined analyses. However, this estimation procedure may be refined by carrying out combined structural analyses on a number of molecules, using LCNMR data obtained using a variety of solvents (as well as ED and rotational data). Inconsistencies between experimentally determined coupling constants and values calculated in the structural analysis should give some indication as to which coupling constants should be given larger uncertainties in future analyses (see chapter 7).

In cases where the vibrational correction is very small the error in the observed coupling constant may become significant and should also be taken into consideration.

Corrections to Rotation Constants

The unmodified version of **MM3** calculates the corrections necessary to convert B_0 rotation constants to B_z (as required for a combined analysis). Nonetheless, it was decided to check the accuracy of such corrections by comparison with existing data. Initial results in this area were not at all encouraging, with some of the corrections made by **MM3** being very different to those made by conventional normal co-ordinate analysis programs. The precise reason for this discrepancy is unclear but it seems reasonable to conclude that corrections to rotation constants are much more sensitive to the quality of the force field used. However, if the **MM3** force field is to be used in the combined analysis of ED, LCNMR and rotational data, it is important that this problem is overcome or, at least, understood to the extent that the uncertainties in the corrections can be estimated.

Equation 4.11 shows how the correction to a rotation constant is related to the harmonic contribution terms, α_i^h . A knowledge of how these terms are calculated, in the harmonic force field analysis, should give an insight into why corrections calculated using the **MM3** force field differ from those calculated using experimentally determined force fields. The general expression describing the calculation of these terms, for a particular inertial axis, can be written¹⁴

$$\alpha_i^h = -\frac{2B^2}{\omega_i} \left\{ 3A_i + \sum_j \frac{1}{2} \zeta_{ij}^2 \left(\frac{(\omega_i + \omega_j)^2}{\omega_j(\omega_i - \omega_j)} - \frac{(\omega_i - \omega_j)^2}{\omega_j(\omega_i + \omega_j)} \right) \right\} \quad [6.7]$$

where A_i is the second derivative of the moment of inertia with respect to the normal co-ordinate, Q_i

and ζ_{ij} is the Coriolis coupling constant between modes i and j .

Clearly, the correction is strongly dependent on the vibrational frequencies (ω) of the molecule. One of the problems with **MM3** is that its principal function is the calculation of molecular structures. Consequently, the parameters of the force field are chosen to best reproduce experimental geometries rather than vibrational frequencies. Unfortunately, the net result is a program which does neither job particularly well (although, considering the simplicity of the principles involved, it is remarkable that it performs as well as it does). In an attempt to overcome this problem, the program was modified slightly to allow experimental frequencies to be read in from a file and used in the subsequent calculations. In terms of equation 6.7, this means that A_i and ζ_{ij} are unchanged but that experimentally determined values are used for the frequency terms, ω_i and ω_j . It should be remembered that there is little theoretical justification for using this inconsistent mixture of experimental frequencies and a force field which corresponds to different theoretical frequencies. This can only be justified if it can be shown that the corrections calculated in this way are closer to those calculated by other methods. A similar change was made which allows the program to read an experimentally determined structure from a file. This means that the vibrational analysis can be carried out on the basis of an experimental rather than a theoretical geometry.

The first molecule to be used to test these modifications was methyl fluoride, mainly because its harmonic force field has been well determined⁷. Corrections to the two rotation constants were calculated using **asym20** and compared to those calculated by **MM3**, with and without experimental frequencies and structure being used. The results are summarised in table 6.3 and are denoted respectively **MM3^F**, **MM3^S**, **MM3** or **MM3^{F,S}** depending on whether experimental frequencies, structure, neither or both are used (all values are in MHz).

Figure 6.3

Vibrational corrections to the rotation constants of methyl fluoride (MHz)

Axis	Constant	asym20 ⁷	MM3	MM3 ^F	MM3 ^S	MM3 ^{F,S}
A	155959	-541	-1084	-1020	-579	-546
B = C	25847	-77	-81	-73	-82	-75

These results seem to justify the slightly questionable methods used to obtain them. By using both experimental frequencies and an experimental geometry, the corrections calculated by **MM3** are in very good agreement with those obtained from the normal co-ordinate analysis. It would also appear that the improvement is largely due to the use of an experimental structure. However, this is just one example and should not be taken in isolation. Similar calculations were carried out for a number of other molecules to see if this method presents a genuine solution to the problem. The results are presented in table 6.4. In the case of formaldehyde and dichloroethene, the corrections calculated by **MM3** were compared with

values obtained using experimentally determined force fields (using the program **asym20**). The values of the rotation constants given for these molecules are approximate and serve only to give an indication of the relative size of the correction. For the remaining examples published rotation constants are used and the corrections are compared to published values, calculated by the program **gamp**.

Table 6.4 - Vibrational corrections to rotation constants (MHz)

Molecule	Refs.	Axis	R.Const.	asym20 or gamp	MM3	MM3 ^F	MM3 ^S	MM3 ^{F,S}
formaldehyde	6	A	288787	1431	3545	3602	3720	3787
		B	39151	140	126	122	125	120
		C	34515	35	65	58	66	59
dichloroethene	8	A	7536	12.4	14.2	11.6	15.0	12.4
		B	3431	4.72	3.25	2.81	3.20	2.77
		C	2357	0.81	1.03	0.90	1.01	0.88
pyridazine	15,20	A	6242.95	3.45	4.80	2.96	4.84	2.97
		B	5961.09	3.14	2.80	1.99	2.78	1.99
		C	3048.70	0.59	0.48	0.60	0.48	0.60
pyrimidine	12,15	A	6276.86	1.26	4.17	2.90	4.37	2.93
		B	6067.18	1.92	2.46	1.85	2.36	1.76
		C	3084.49	0.66	0.56	0.53	0.57	0.53
furan	16,17	A	9446.96	6.87	8.36	7.53	8.34	7.51
		B	9246.61	7.44	5.13	4.67	5.16	4.70
		C	4670.88	2.14	1.31	1.27	1.31	1.27
thiophene	18,19	A	8041.77	5.89	4.63	5.84	4.83	6.11
		B	5418.12	4.24	2.90	2.68	2.78	2.58
		C	3235.77	1.24	0.90	0.81	0.90	0.81

The results of these calculations are by no means as clear cut as those obtained for methyl fluoride. In some cases, using experimental frequencies and geometry improves the calculated corrections but, in just as many cases, the corrections show little change or even get worse. It should be remembered that there is often a large uncertainty associated with the corrections calculated by the normal co-ordinate analysis method and, for this reason, it is impossible to say for certain what represents a "good" result. Nonetheless, it seems that the uncertainties in the corrections calculated by **MM3** can be very large indeed. On average, however, there seems to be a general improvement when experimental frequencies are used in the calculation. Using an experimental geometry seems to make little difference for these molecules, unlike the results obtained for methyl fluoride. For this reason, it seems more practical, and to some degree more satisfactory, to use only experimental frequencies in the calculation of the corrections, particularly in the case of the aromatic molecules described in this subsequent chapters.

Once more, an appreciation of the uncertainties in the vibrational corrections is important if the results are to be used in a structural analysis. The deviation of the results obtained by the different methods of calculation is an indication of the uncertainty in each of the correction terms. Using the values presented in table 6.4, it would seem reasonable to assume the uncertainty in the vibrational correction to be approximately fifty percent of the correction itself. This is considerably larger than uncertainties estimated using normal co-ordinate analysis programs^{10,12,20} and may, at first, seem unacceptable. However, when the size of the correction terms relative to the rotation constants is considered (typically ~0.05%), it is clear

that corrected rotational data can still contribute useful structural information to combined analyses.

Summary

It has been shown that the **MM3** force field can be used to make the vibrational corrections necessary to reduce data from diverse structural techniques to a common basis. This is particularly useful for molecules of low symmetry which can prove extremely difficult to analyse using conventional normal co-ordinate analysis programs. Notes on the use of **MM3** for such calculations can be found in Appendix C.

Parallel and perpendicular amplitudes are reasonably insensitive to the quality of the force field used. In any case, parallel amplitudes can often be refined during the combined analysis. Amplitudes which can not be refined independently can be refined in groups, as described above, in which case it is the ratios of these amplitudes that are of greater importance than the absolute values. Slight errors in the calculated perpendicular amplitudes are of little significance, as the values involved are usually extremely small.

Corrections to LCNMR data can now be calculated using a combination of the programs **MM3** and **bmgv**. An estimation of the uncertainties of the direct coupling constants can be made by combining the uncertainties from the spectral analysis with those associated with the vibrational correction calculation. The latter can be assumed to be 10% of the value of the correction term, d^h , down to a minimum threshold of about 0.1 Hz.

It is undeniable that **MM3** is less suitable for the calculation of vibrational corrections to rotation constants than for the calculation of either amplitudes of vibration or corrections to LCNMR data. However, in the absence of a reliable force field for normal co-ordinate analysis, **MM3** can be used to make such corrections, as long as the increased uncertainty is

taken into account in the subsequent structural analysis. This uncertainty is estimated to be about 50% of the correction term (which typically corresponds to a 0.025% uncertainty in the associated rotation constant).

References

- ¹N.L.Allinger, "Accurate Molecular Structures: Their Determination and Importance", Chapter 14, eds. A.Domenicano & I.Hargittai, Oxford University Press, (1992), 336
- ²available through the Quantum Chemistry Program Exchange
- ³P.Aped & N.L.Allinger, *J.Am.Chem.Soc.*, 114, (1992), 1
- ⁴S.Sýkora, J.Vogt, H.Bösiger & P.Diehl, *J.Magn.Reson.*, 36, (1979), 53
- ⁵C.A.Brookman, Ph.D. Thesis, The University of Edinburgh, 1993
- ⁶C.R.Pulham, private communication
- ⁷J.L.Duncan, D.C.M^cKean & G.K.Speirs, *Mol.Phys.*, 24(3), (1972), 553
- ⁸J.R.Durig, R.M.Jones, C.J.H.Schutte & G.Zerbi, *Pure & Appl.Chem.*, 57(1), (1985), 121
- ⁹P.Diehl & J.Jokisaari, *J.Mol.Struct.*, 53, (1979), 55
- ¹⁰S.Cradock, J.M.Muir & D.W.H.Rankin, *J.mol.Struct.*, 220, (1990), 205
- ¹¹P.Diehl, T.Bjorholm & H.Bösiger, *J.Magn.Reson.*, 42, (1981), 390
- ¹²S.Cradock, P.B.Liescheski, D.W.H.Rankin & H.E.Robertson, *J.Amer.Chem.Soc.*, 110(9), (1988), 2759

- ¹³S.Cradock, C.Purves & D.W.H.Rankin, *J.Mol.Struct.*, 220, (1990), 193
- ¹⁴M.Toyama, T.Oka & Y.J.Morino, *J.Mol.Spectrosc.*, 6, (1961), 472
- ¹⁵K.B.Wiberg, *J.Mol.Struct.*, 224, (1990), 61
- ¹⁶P.B.Liescheski & D.W.H.Rankin, *J.Mol.Struct.*, 196, (1989), 1
- ¹⁷D.W.Scott, *J.Mol.Spectrosc.*, 37, (1971), 77
- ¹⁸P.B.Liescheski & D.W.H.Rankin, *J.Mol.Struct.*, 178, (1988), 227
- ¹⁹D.W.Scott, *J.Mol.Spectrosc.*, 31, (1969), 451
- ²⁰S.Cradock, C.Purves & D.W.H.Rankin, *J.Mol.Struct.*, 220, (1990), 193

Chapter 7

The Structures of the Difluorobenzenes

Introduction

In many ways, the three isomeric difluorobenzenes are ideally suited for structural analysis by the combined analysis of ED, LCNMR and rotational data. The difficulty of determining their structures by ED alone is heightened, compared to other substituted aromatics, by the fact that the carbon-fluorine bond length is similar to the carbon-carbon bond length. This effectively reduces the number of distinct features in the radial distribution curve and leads to a high degree of correlation in the analysis. The molecules are ideally suited to LCNMR analysis, having six 100% abundant spin- $\frac{1}{2}$ nuclei. If ^{13}C satellites can be assigned then it is possible to determine direct coupling constants for all HH, FF, HF, CF and CH nuclear pairs. Indeed it is possible, in principle, to determine a complete, although unscaled, structure using LCNMR data alone. This has been done in all three cases using a variety of liquid crystal solvents^{1,2,3}. However, evidence of anisotropy of some of the indirect coupling constants between FF and CF nuclear pairs would suggest that it would be wise to exclude these values from the analysis. The use of ED data is therefore important both to make up for this loss of data and to provide an overall scale to the structure. Rotation constants of ortho and meta difluorobenzene, and of some of their ^{13}C isotopomers, have been measured by MW spectroscopy⁴. This provides a useful set of independent data which can also be used in the combined analysis. Unfortunately, the rotation constants of p-difluorobenzene cannot be measured by MW spectroscopy, as the molecule does not possess a permanent dipole. However, the higher symmetry of this isomer should mean that it presents less of a problem for a combined analysis of ED and LCNMR data alone.

Experimental

Samples of 99% pure *ortho*, *meta* and *para* difluorobenzene were obtained from the Aldrich Chemical Company and used without further purification. ED data were recorded using the Edinburgh apparatus using procedures described in chapter 1. The compounds were all sufficiently volatile to obtain a suitable vapour pressure at room temperature. In each case, data were obtained at long and short camera distances. Further experimental details can be found in table 7.1. The atom numbering used throughout this chapter is shown in figure 7.1.

In addition to the published LCNMR data, a complete set of HH, FF and HF direct coupling constants was obtained for *o*-difluorobenzene using the liquid crystal solvent E7.

Vibrational corrections

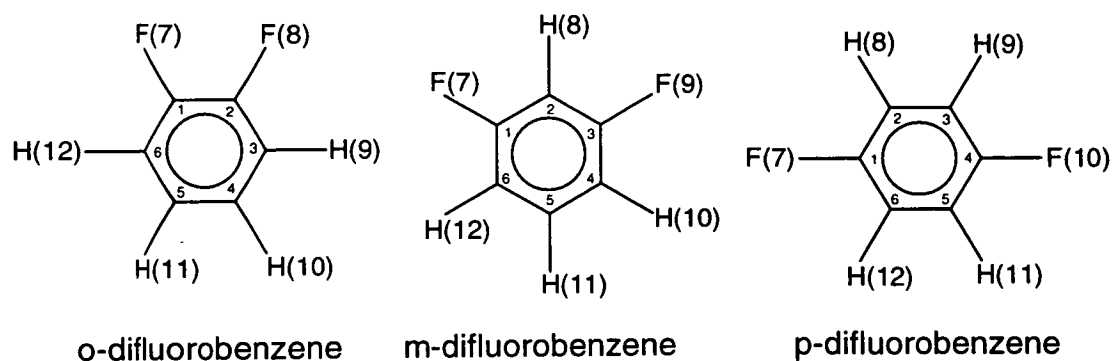
Data from the three techniques, ED, LCNMR and MW spectroscopy, were reduced to a common basis using the corrections described in chapter 4. The vibrational analysis was carried out using the modified version of the molecular mechanics program, **MM3**, described in chapter 6. Although vibrational corrections to the published direct dipolar couplings are presented in the original papers, new corrections were calculated, as it is a primary aim of this work to verify that **MM3** can be used for this purpose.

For the calculation of vibrational corrections to the rotation constants of *ortho* and *meta* difluorobenzene, experimentally determined vibrational frequencies⁵ were used in conjunction with the **MM3** force field, as described in chapter 6.

Table 7.1

Experimental details and weighting parameters
for the ED data obtained for the difluorobenzenes

	o-difluorobenzene		m-difluorobenzene		p-difluorobenzene	
Camera dist. / mm	285.78	128.24	285.96	128.22	285.96	128.22
Nozzle temp. / K	293	293	293	293	293	293
$\Delta s / \text{\AA}^{-1}$	0.2	0.4	0.2	0.4	0.2	0.4
$s_{\text{min}} / \text{\AA}^{-1}$	2.0	6.0	2.0	6.0	2.0	6.0
$s_{\text{w1}} / \text{\AA}^{-1}$	4.0	8.0	4.0	8.0	4.0	8.0
$s_{\text{w2}} / \text{\AA}^{-1}$	12.2	30.4	12.2	30.4	12.2	30.4
$s_{\text{max}} / \text{\AA}^{-1}$	14.4	35.6	14.4	35.6	14.4	35.6
Correl ^D parameter	0.4968	0.4155	0.4871	0.4036	0.4771	0.3475
Scale factor	0.819(6)	0.811(15)	0.712(5)	0.729(14)	0.728(3)	0.714(9)
Wavelength / \AA	0.05747	0.05742	0.05749	0.5749	0.5749	0.5749

Figure 7.1 - Atom numbering of the difluorobenzenes

The structure of *ortho*-difluorobenzene

Analysis of LCNMR spectra

The ^1H and ^{19}F NMR spectra of *o*-difluorobenzene, in the liquid crystal solvent E7, were recorded at room temperature and are shown in figures 7.2 and 7.3 (these plots were obtained using the modified version of **sliquor**, described above). The large number of lines, due to the presence of six spin- $\frac{1}{2}$ nuclei, made the analysis of these spectra far from trivial.

The analysis was carried out using the programs **lcsim** and **sliquor** as described in previous chapters, using chemical shifts and indirect coupling constants taken from the published data of Ernst et al.⁶ The ^{19}F spectrum was analysed first due to its relative simplicity. A systematic search for approximate values for the orientation parameters was carried out, with limited success; although many of the more intense lines in the spectrum could be identified, it was impossible to assign any of the smaller lines unambiguously. It transpired that the key to the assignment of this spectrum was the identification of a first order subspectrum involving the four most intense lines (marked 1,2,3 and 4 in figure 7.2). Experimenting with trial values for the orientation parameters showed that the separation of lines 1 and 2 (or lines 3 and 4) is exactly equal to $3D_{\text{FF}}$. Thus, D_{FF} can be measured directly from the spectrum and is equal to ± 404.89 Hz. Furthermore, because the vector between the two fluorine atoms lies parallel to the *y*-axis of the co-ordinate system used, the value of D_{FF} is dependent on only one orientation parameter (S_{yy}). By determining D_{FF} , the value of S_{yy} is also determined (for the assumed internuclear separation).

Figure 7.2 - Simulation of the ^{19}F spectrum of o-difluorobenzene in E7

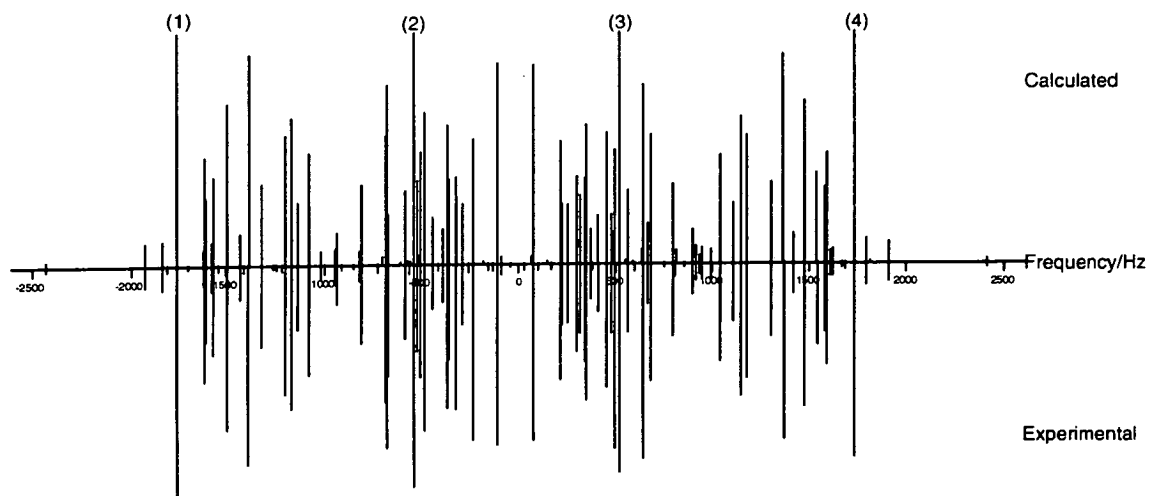
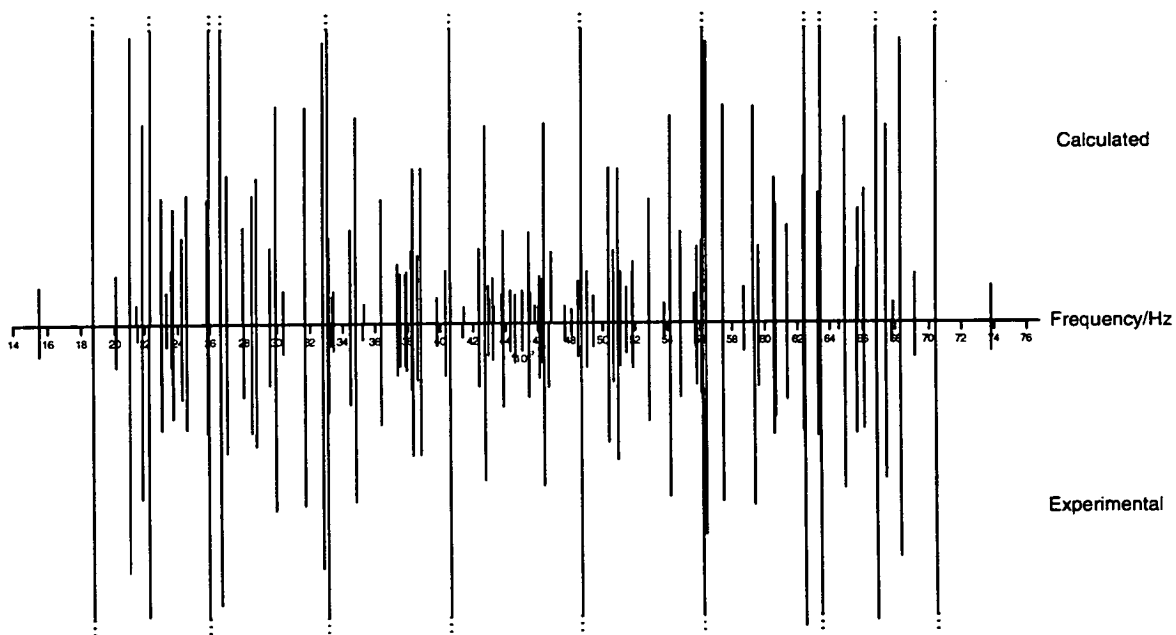


Figure 7.3 - Simulation of the ^1H spectrum of o-difluorobenzene in E7



Knowing the approximate value of S_{yy} it becomes a simple matter to vary S_{zz} until enough of the lines in the calculated and experimental spectra can be matched to allow refinement of the spectral parameters.

It must be remembered that there are still two possible solutions to the spectral assignment, depending on the sign of D_{FF} . Fortunately, during the refinement of the direct coupling constants it became apparent that D_{FF} is negative; the couplings obtained with D_{FF} negative reproduced the experimental spectrum with an r.m.s. deviation of just 0.31 Hz whereas with D_{FF} positive the r.m.s. deviation was 2.7 Hz. Figure 7.2 shows the final calculated spectrum, derived from refined direct coupling constants, and a representation of the experimental spectrum. In the final refinement a total of 76 lines were assigned; only lines arising from two or more overlapping peaks were left unassigned.

Once the ^{19}F spectrum had been analysed, assignment of the ^1H spectrum was relatively simple. Both spectra were recorded using the same sample at the same temperature and so, to a good approximation, the orientation parameters are the same for each spectrum. All that remains is to refine the calculated direct couplings to fit the experimental spectrum. In the final refinement a total of 87 lines were assigned and the r.m.s. deviation was 0.37 Hz. Once more, the only lines which were not assigned were those arising from overlapping peaks. The calculated and experimental spectra are shown in figure 7.3.

The final values obtained for the direct coupling constants are listed in table 7.2 with uncertainties in parenthesis (in units of the final digit). A vibrational analysis was carried out, using the modified version of **MM3** (described in chapter 6). Vibrationally corrected direct dipolar coupling

constants , D^α , can also be found in table 7.2, with uncertainties from both the spectral analysis and the vibrational correction terms taken into account.

Table 7.2

Direct coupling constants obtained from the LCNMR analysis of o-difluorobenzene in E7

Coupling	D°	d^h	D^α
D7,8	-404.89(5)	-0.144	-404.75(10)
D7,9	-112.57(5)	0.402	-112.97(10)
D7,10	-95.22(5)	0.267	-95.49(10)
D9,10	-894.62(10)	15.9	-910.5(16)
D7,11	-163.63(7)	0.860	-164.49(10)
D9,11	-132.38(13)	1.84	-134.22(24)
D10,11	-566.81(24)	18.5	-585.3(19)
D7,12	-775.58(6)	9.54	-785.1(10)
D9,12	-72.11(21)	0.681	-72.79(24)

Structure refinement of o-difluorobenzene

The molecule is assumed to be of C_{2v} symmetry and the geometry can be described by 11 independent structural parameters (see table 7.3). Initial values for these parameters were chosen such that all C-C bond lengths were 1.39 Å, both C-H bond lengths were 1.085 Å and the C-F bond length was 1.34 Å. All angles were initially assumed to be 120°. Parallel and perpendicular amplitudes of vibration were calculated for each of the 36 distinct internuclear distances within the molecule.

The refinement was started with only ED data being used. The key structural parameters were introduced one at a time until gradually the refinement converged, with R_G at 6.2% (see equation 1.10). At this stage, the rC-C difference parameters would not refine with reasonable standard deviations and so were fixed at zero. No attempt was made to refine the parameters relating to the C-H bonds. The ring showed little deviation from a regular hexagonal geometry although the angles at the fluorine substituted carbons opened out slightly to 120.6(2)°.

The refinement was then switched to an r_α^o basis which brought R_G down to 5.9% after a few refinements. At this point some vibrational amplitudes were allowed to refine, starting with the directly bonded C-C and C-F amplitudes, which were refined as a group with ratios fixed at the values calculated by **MM3**. This made little difference to the refinement and so further amplitudes were successively included until all those for heavy atom pairs were either refining or tied to a refining amplitude. An attempt was made to reintroduce the rC-C difference parameters to the refinement but they would still not refine to reasonable values. Similarly, the rC-H

difference parameter and the CCH angle deviations could not be refined at this stage. The final structure obtained using ED data alone is shown in table 7.3.

Rotation constants of the principal isotopic species were introduced as extra data, although initially these were given relatively low weight (i.e. high uncertainties). As the refinement progressed, this weight was gradually increased until the uncertainties matched those initially estimated from the vibrational corrections. It is interesting to note that even these two extra data were sufficient to reduce the estimated standard deviations of some of the parameters significantly. The rotation constants for the three ^{13}C isotopomers were then introduced one at a time, once more with low weighting initially. However, there was remarkably little deviation between the calculated and experimental values even before the refinement started and no problems were experienced in fitting the experimental rotation constants to within their respective uncertainties (see table 7.4). The introduction of the MW data also allowed the refinement of two of the rC-C difference parameters although the difference $r_{34} - r_{45}$ could still not be refined. As with the analysis using ED data alone, none of the parameters involving the hydrogen atoms could be successfully refined, other than the mean C-H bond length. The final structure obtained using ED and MW data is also shown in table 7.3.

As with the rotational data, LCNMR data were initially added with increased uncertainties, which were gradually reduced to their estimated values. Initially only the data recorded in the solvent ZLI1167 were used. The two independent orientation parameters, S_{yy} and S_{zz} , were allowed to refine and initially given the values obtained in the original LCNMR analysis¹.

Inclusion of the LCNMR data allowed the simultaneous refinement of all 13 structural and orientational parameters in addition to 7 amplitudes (or groups of tied amplitudes). Although the fit to the MW data remained good and R_G rose only slightly, to 6.5%, many of the calculated direct coupling constants differed from the observed values by as much as 3 or 4 standard deviations. This was particularly true of the directly bonded C-H couplings and some of the long range C...H couplings. These discrepancies cannot be attributed solely to underestimation of the errors due to the vibrational corrections as they also occur for a number of couplings for which the vibrational corrections are extremely small. Some possible explanations for such discrepancies are presented at the end of this chapter.

The fit to the observed coupling constants was improved slightly when $D_{7,8}$ was excluded from the refinement. This is the one of the worst fitting of the couplings and this can possibly be attributed to a significant anisotropy of the indirect coupling constant (see chapter 2).

A similar refinement was carried out using the LCNMR data obtained using the solvent ZLI1132 with similar results. However, significant structural differences were found in the difference parameter $r_{34} - r_{45}$ and the two C-H angle deviation parameters. With the ZLI1132 data in particular, the difference parameter was much larger than might be expected, refining to $-0.027(6)$ Å. It is not obvious why bonds C(3)-C(4) and C(4)-C(5) should be significantly different in length, neither being adjacent to the fluorine substituents. With the ZLI1167 data this value was considerably smaller, $-0.009(6)$ Å, which might seem more realistic. It is perhaps significant that most of the couplings involving C(4) do not fit well in either solvent, which suggests that the position of this atom is not at all well

determined. Unfortunately, this difference parameter is strongly correlated with the C(4)-H(10) angle deviation parameter which means that the position of H(10) is also poorly determined, and consequently many other calculated coupling constants are indirectly affected.

It was hoped that by refining the structure to fit both LCNMR data sets simultaneously, as well as the ED and MW data, some of the discrepancies between the two structures could be resolved. In addition, this should help to reduce the effect of any random errors in the LCNMR data. In fact it would be better still if LCNMR data recorded in other solvents could also be used but, with the current version of **ed92**, the number of non-ED data is limited to a maximum of 50. It is for this reason that the data obtained using the solvent E7 were not included in the structural analysis. It is anticipated that future versions of the program will allow a larger number of extra data to be used.

The final refinement is based on ED data at two camera distances, 8 rotation constants and 21 direct coupling constants from each of the two solvents (ZLI1167 and ZLI1132). In fact, the simultaneous use of both sets of LCNMR data did little to reduce the indeterminacy of the position of C(4). In table 7.2 two sets of results are presented, one with the parameter r34 - r45 refining and the second with it fixed at zero. On the whole, the latter structure seems more reasonable but it must be remembered that it is subject to an extra constraint. Table 7.5 shows the final values obtained for the amplitudes of vibration with estimated standard deviations for those which were refined during the analysis. Table 7.6 shows the least squares correlation matrix calculated during the final refinement. The final molecular geometry is shown in figure 7.4 and the corresponding molecular

scattering intensity and radial distribution curves are shown in figures 7.5 and 7.6 respectively.

Table 7.3 - Final parameters - *ortho*-difluorobenzene

Parameters [†]	ED	ED+MW	All data (1)	All data (2)
Structural				
(r _{1,2} +r _{2,3} +r _{3,4} +r _{4,5})/4	1.3895(10)	1.3894(10)	1.3931(10)	1.3922(10)
r _{1,2} - (r _{2,3} +r _{3,4} +r _{4,5})/3	-0.002(8)	0.013(9)	-0.002(4)	0.001(4)
r _{2,3} - (r _{3,4} +r _{4,5})/2	0.0 (fixed)	-0.11(9)	-0.018(4)	-0.017(4)
r _{3,4} - r _{4,5}	0.0 (fixed)	0.0 (fixed)	-0.019(5)	0.0 (fixed)
rC-F	1.338(4)	1.344(4)	1.343(3)	1.342(3)
mean rC-H	1.079(6)	1.084(5)	1.082(2)	1.081(2)
r _{3,9} - r _{4,10}	0.0 (fixed)	0.0 (fixed)	-0.005(4)	-0.010(4)
C-F angle deviation*	0.93(22)	0.80(28)	0.66(11)	0.61(12)
angle C(1)-C(2)-C(3)	120.16(19)	120.19(15)	120.52(12)	120.63(12)
C(3)-H(9) deviation*	0.0 (fixed)	0.0 (fixed)	0.76(15)	0.60(16)
C(4)-H(10) deviation*	0.0 (fixed)	0.0 (fixed)	0.72(13)	0.35(10)
Orientational				
S _{VV} x 100 (ZLI1167)	-	-	-4.049(15)	-4.040(16)
S _{ZZ} x 100 (ZLI1167)	-	-	-5.991(15)	-5.994(15)
S _{VV} x 100 (ZLI1132)	-	-	7.218(20)	7.201(20)
S _{ZZ} x 100 (ZLI1132)	-	-	10.320(20)	10.322(21)
Dependent				
rC(1)-C(2)	1.388(5)	1.399(7)	1.392(3)	1.393(4)
rC(2)-C(3)	1.390(3)	1.379(8)	1.382(3)	1.380(3)
rC(3)-C(4)	1.390(3)	1.390(3)	1.390(3)	1.398(2)
rC(4)-C(5)	1.390(3)	1.390(3)	1.409(3)	1.398(2)
rC(3)-H(9)	1.079(6)	1.084(5)	1.079(3)	1.076(3)
rC(4)-H(10)	1.079(6)	1.084(5)	1.084(3)	1.086(3)
angle C(2)-C(3)-C(4)	119.73(36)	119.65(28)	119.58(19)	119.27(18)
angle C(3)-C(4)-C(5)	120.11(21)	120.16(18)	119.90(12)	120.10(11)

* positive deviations are towards the nearest fluorine atom

[†] all distances are given in Angstroms; all angles are given in degrees.

Table 7.4 - Rotation constants⁴ (MHz) and direct dipolar coupling constants¹ (Hz) used in the structural analysis of o-difluorobenzene

Constant	Observed	Corrected	Calculated	Uncertainty	Difference
B principal	2227.89	2227.58	2227.57	0.160	0.008
C principal	1323.86	1323.72	1323.74	0.068	-0.015
B 13C(1)	2226.01	2225.70	2225.64	0.160	0.059
C 13C(1)	1321.57	1321.43	1321.39	0.068	0.040
B 13C(3)	2222.37	2222.06	2222.03	0.160	0.033
C 13C(3)	1315.24	1315.10	1315.13	0.068	-0.023
B 13C(4)	2191.05	2190.74	2190.68	0.160	0.057
C 13C(4)	1309.20	1309.07	1309.06	0.068	0.012
* D2,9	140.10	143.95	144.80	0.45	-0.85
D3,9	898.90	975.32	980.95	7.70	-5.63
D4,9	139.90	143.31	140.35	0.45	2.96
D7,9	57.16	57.34	57.31	0.20	0.04
D8,9	357.34	361.57	361.54	0.43	0.03
D2,10	46.30	46.85	46.33	0.21	0.52
D3,10	173.10	177.60	179.43	0.54	-1.83
D4,10	1219.90	1328.79	1302.28	10.90	26.51
D7,10	44.19	44.31	44.29	0.20	0.02
D8,10	73.48	73.86	73.59	0.20	0.27
D9,10	415.22	422.67	424.29	0.75	-1.62
D2,11	28.70	28.94	28.93	0.21	0.01
D3,11	37.10	37.62	36.97	0.32	0.65
D4,11	130.60	134.69	133.82	0.45	0.87
D9,11	67.54	68.44	68.63	0.20	-0.20
D10,11	310.63	319.68	321.86	0.90	-2.18
D2,12	32.90	33.24	33.41	0.21	-0.17
D3,12	20.90	20.27	20.97	0.41	-0.70
D4,12	32.80	33.22	32.93	0.22	0.29
D9,12	39.60	39.95	40.00	0.20	-0.05
** D2,9	-247.60	-254.51	-254.82	0.73	0.31
D3,9	-1604.50	-1740.14	-1749.99	13.60	9.85
D4,9	-245.10	-251.04	-246.80	0.80	-4.24
D7,9	-100.78	-101.10	-100.97	0.10	-0.13
D8,9	-617.47	-624.81	-625.62	0.74	0.81
D2,10	-78.70	-79.64	-79.84	0.30	0.20
D3,10	-305.30	-313.28	-309.05	0.90	-4.23
D4,10	-2109.00	-2297.61	-2257.08	18.9	-40.53
D7,10	-76.36	-76.57	-76.77	0.10	0.20
D8,10	-126.12	-126.78	-126.73	0.10	-0.05
D9,10	-718.41	-731.36	-735.41	1.30	4.05
D2,11	-50.20	-50.63	-50.14	0.22	-0.48
D3,11	-65.20	-66.11	-64.81	0.41	1.30
D4,11	-231.20	-238.44	-236.58	0.86	-1.86
D9,11	-119.14	-120.71	-121.02	0.16	0.31
D10,11	-555.46	-571.43	-574.19	1.60	2.76
D2,12	-58.20	-58.79	-59.25	0.22	0.46
D3,12	-36.50	-36.81	-37.42	0.42	0.61
D4,12	-57.50	-58.23	-58.37	0.45	0.14
D9,12	-70.80	-71.42	-71.35	0.10	-0.07

* ZLI1167 couplings

** ZLI1132 couplings

Table 7.5 - Amplitudes of vibration of o-difluorobenzene

Number	Atom Pair	u / Å
1	C1 - C2	0.043(2)
2	C2 - C3	0.043 (tied to u1)
3	C3 - C4	0.043 (tied to u1)
4	C4 - C5	0.043 (tied to u1)
5	C1 - F7	0.042 (tied to u1)
6	C3 - H9	0.099(18)
7	C4 - H10	0.100 (tied to u6)
8	C1...C3	0.054(2)
9	C1...C4	0.063(3)
10	C1...C5	0.054 (tied to u8)
11	C1...F8	0.059 (tied to u8)
12	C1...H9	0.096 (fixed)
13	C1...H10	0.094 (fixed)
14	C1...H11	0.096 (fixed)
15	C1...H12	0.113(14)
16	C3...C5	0.055 (tied to u8)
17	C3...C6	0.064 (tied to u9)
18	C3...F7	0.064(2)
19	C3...F8	0.059 (tied to u8)
20	C3...H10	0.113 (tied to u15)
21	C3...H11	0.096 (fixed)
22	C3...H12	0.095 (fixed)
23	C4...F7	0.070(4)
24	C4...F8	0.064 (tied to u18)
25	C4...H9	0.113 (tied to u15)
26	C4...H11	0.113 (tied to u15)
27	C4...H12	0.096 (fixed)
28	F7...F8	0.098 (tied to u9)
29	F7...H9	0.108 (fixed)
30	F7...H10	0.097 (fixed)
31	F7...H11	0.108 (fixed)
32	F7...H12	0.130 (fixed)
33	H9...H10	0.155 (fixed)
34	H9...H11	0.131 (fixed)
35	H9...H12	0.119 (fixed)
36	H10...H11	0.155 (fixed)

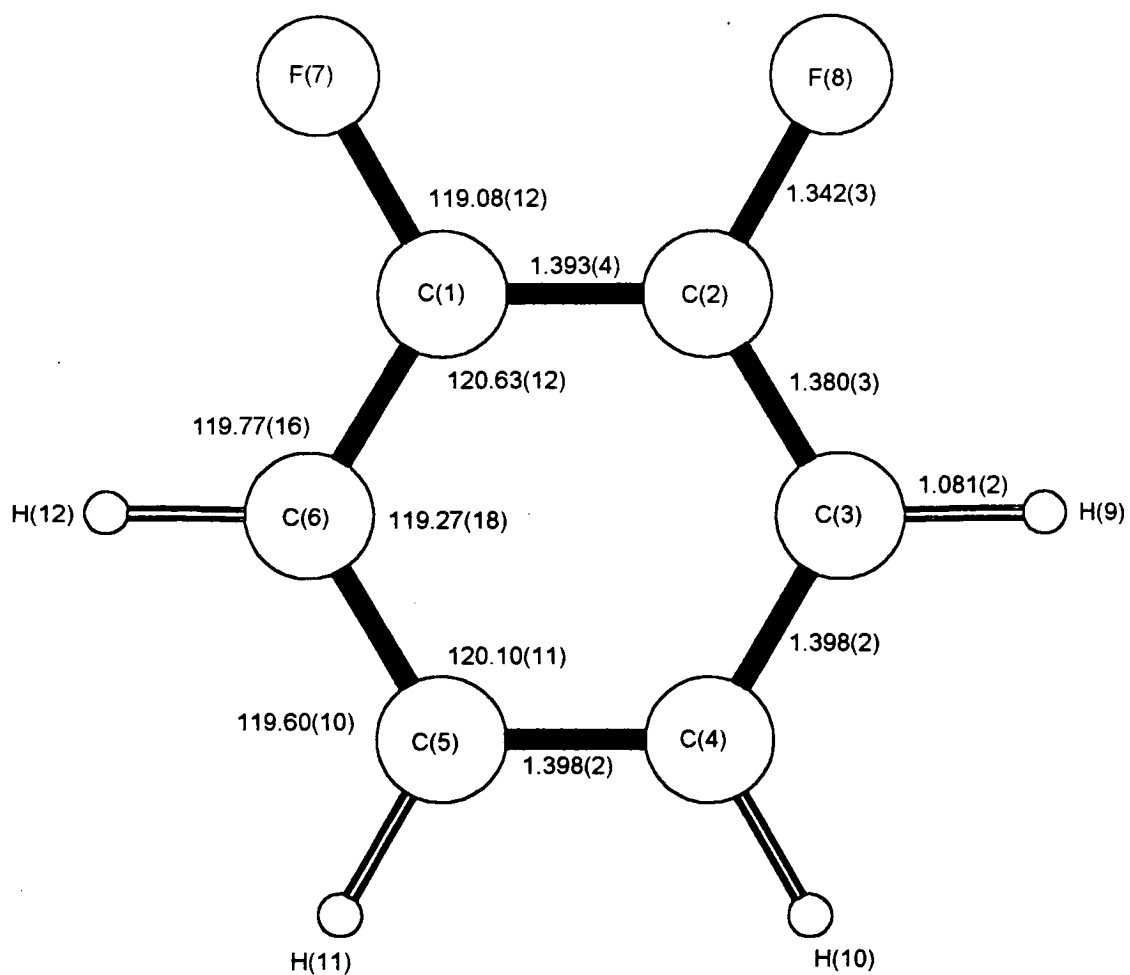
Table 7.6

Least squares correlation matrix (x100) for o-difluorobenzene. All elements with absolute value <50 have been omitted.

	P5	P7	P8	P9	P10	P11	P13	P14	P15	u1	k2
P1	-81										
P2			90	-66							
P3		-52			72						
P4						-68					
P5				57						50	
P6							-52	61	61		
P7									-55		
P8				-54							
P12								-73			
u1											57
u8											52

Parameters (P) are in the order listed in table 7.3 ; amplitudes (u) are in the order listed in table 7.5 ; scale factors (k) are in the order listed in table 7.1.

Figure 7.4 - The molecular structure of o-difluorobenzene



All distances are given in Ångströms
All angles are given in degrees

Figure 7.5

The observed and final weighted difference
molecular scattering intensity curves for o-difluorobenzene

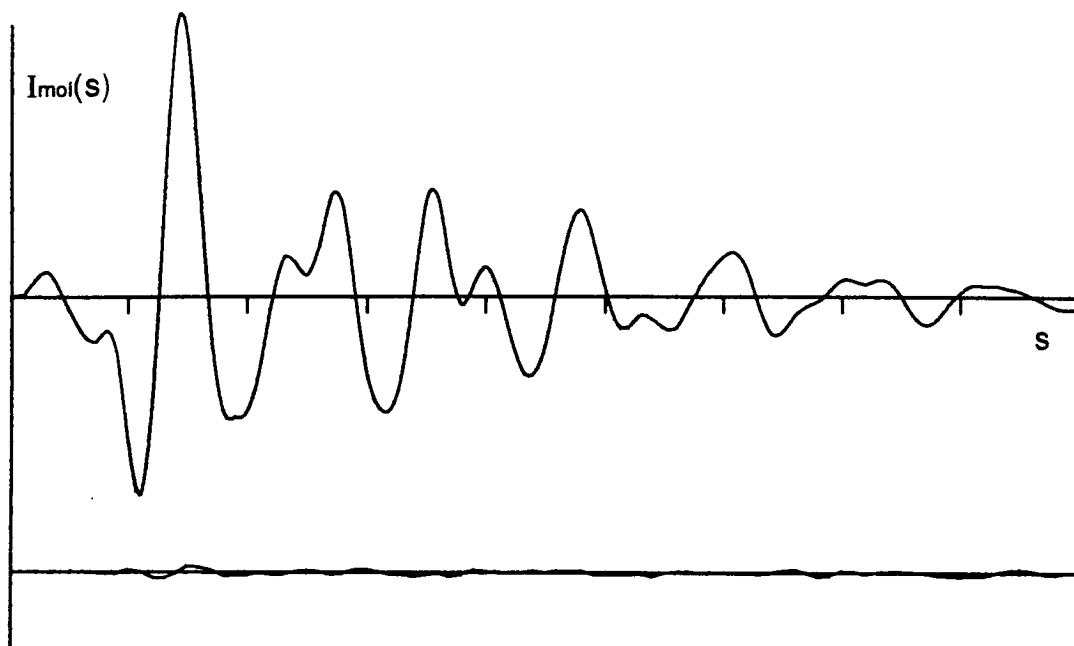
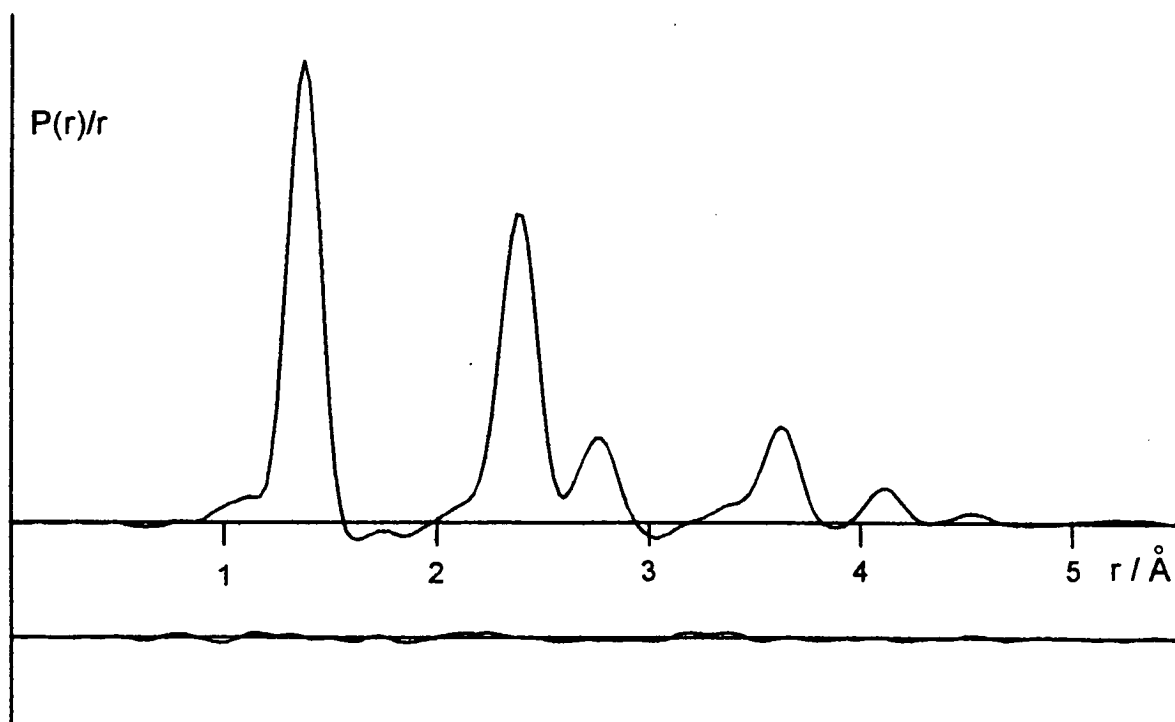


Figure 7.6

The observed and final difference
radial distribution curve, $P(r)/r$, for o-difluorobenzene



The structure of m-difluorobenzene

Although it is of the same symmetry as o-difluorobenzene (C_{2v}), m-difluorobenzene is slightly better suited to analysis by ED, as it has only three distinct C-C distances. Once more, the molecular geometry can be described by 11 independent parameters (see table 7.7) and initial values were chosen assuming a regular hexagonal structure. Vibrational amplitudes were calculated for the 38 distinct internuclear distances within the molecule.

The same refinement procedure was applied as in the case of o-difluorobenzene, starting with ED data only and introducing first MW data then LCNMR data as the refinement progressed. With ED data alone, the structure quickly settled with R_G at 7.1%. The parameter defining the difference between the mean lengths of the C-C bonds adjacent to the fluorine atoms and the remaining C-C difference refined to $-0.0104(73)$ Å, consistent with the expected shortening of adjacent bonds on fluorine substitution. After introducing several amplitudes into the refinement and refining on an r_α^o basis, an ED only structure was obtained with $R_G=6.6\%$. The difference parameter $r_{35} - r_{45}$ and the parameters determining the hydrogen atom positions could not be refined at this stage.

Eight independent rotation constants were included in the refinement using the procedure outlined in the previous section. Although this allowed the refinement of the mean C-H bond length, the value obtained, $1.097(4)$ Å, seems rather high. The remaining C-C bond length difference parameter could not be refined. The main effect of including the rotational data was

an improvement in the determination of the ring angles and the C-C-F angle. The values from both refinements can be found in table 7.7.

The LCNMR data obtained from solvent ZLI1167 were included at this stage. In general, the fit to these coupling constants was better than the fit in the case of *o*-difluorobenzene. Notable exceptions were the C-F and F-F coupling constants which were several standard deviations from their observed values. This is possibly due to the fact that indirect couplings between heavier nuclei can have a significant anisotropic component (J^{aniso}) which is inseparable from the direct coupling constant. To allow for this, the uncertainties of these couplings were increased by 1 Hz (5 Hz in the case of the directly bonded C-F).

Despite the reasonable fit to the extra data, the structure showed some peculiarities which suggested that the fit was somewhat artificial. In particular the variation in the three C-H bond lengths seemed unreasonable ($r_{2,8}=1.085(1)$ Å, $r_{4,10}=1.077(1)$ Å and $r_{5,11}=1.077$ Å). Of course, this may have been a real effect but when a refinement was carried out using data obtained using ZLI1132 the variation was reversed, with $r_{2,8}$ being significantly shorter than the other two C-H bonds. It seems that too much weight is being given to the D_{CH} coupling constants when determining these bond lengths. Again it was hoped that, by refining the structure to fit all the available data simultaneously, some of the random errors in the LCNMR data would cancel out. Unfortunately, the limitation to 50 extra data did not allow all the available LCNMR data to be used and so a selection of some of the best and worst fitting coupling constants from the ZLI1132 results were added to the ZLI1167 data for the final refinement.

The final structural refinement was based on ED data at two camera distances, 8 rotation constants, 29 direct coupling constants obtained using the solvent ZLI1167 and 13 obtained using ZLI1132. This allowed the simultaneous refinement of all 11 structural parameters as well as 4 orientation parameters and 5 amplitudes of vibration. In fact, 7 amplitudes could be refined but the current version of the program limits refinement to a total of 20 parameters. Estimated standard deviations were obtained for all 7 amplitudes by switching different amplitudes in and out of the refinement once the structure had settled.

The results of the combined structural analysis can be found in table 7.7 with the observed and calculated non-ED data listed in table 7.8. Amplitudes of vibration are shown in table 7.9 and the least squares correlation matrix calculated during the final refinement is shown in table 7.10. The molecular geometry is shown in figure 7.7 and the corresponding molecular scattering intensity and radial distribution curves are shown in figures 7.8 and 7.9 respectively.

Table 7.7 - Final parameters - *meta*-difluorobenzene

Parameters [†]	ED	ED+MW	All data
Structural			
(r _{1,2} +r _{3,4} +r _{4,5})/3	1.3917(20)	1.3908(10)	1.3904(10)
(r _{1,2} +r _{3,4})/2 - r _{4,5}	-0.022(7)	-0.016(6)	-0.011(3)
r _{2,3} - r _{3,4}	0.0 (fixed)	0.0 (fixed)	0.010(3)
r _{C-F}	1.332(4)	1.336(3)	1.345(5)
mean r _{C-H}	1.085 (fixed)	1.097(4)	1.079(2)
(r _{4,10} +r _{5,11})/2 - r _{2,8}	0.0 (fixed)	0.0 (fixed)	0.000(3)
r _{4,10} - r _{5,11}	0.0 (fixed)	0.0 (fixed)	0.001(3)
angle C(1)-C(2)-C(3)	118.80(46)	118.44(37)	116.77(21)
angle C(2)-C(3)-C(4)	121.74(25)	121.90(21)	122.80(19)
angle C(2)-C(3)-F(9)	118.41(44)	118.58(15)	117.87(12)
angle C(3)-C(4)-H(10)	120.0 (fixed)	120.0 (fixed)	119.93(21)
Orientalional			
S _{YY} x 100 (ZLI1167)	-	-	-6.969(13)
S _{ZZ} x 100 (ZLI1167)	-	-	-5.306(13)
S _{YY} x 100 (ZLI1132)	-	-	9.357(23)
S _{ZZ} x 100 (ZLI1132)	-	-	6.791(17)
Dependent			
r _{C(1)-C(2)}	1.384(2)	1.386(2)	1.392(2)
r _{C(3)-C(4)}	1.384(2)	1.386(2)	1.382(2)
r _{C(4)-C(5)}	1.406(6)	1.401(5)	1.398(2)
r _{C(4)-H(10)}	1.085 (fixed)	1.097(4)	1.080(2)
r _{C(5)-H(11)}	1.085 (fixed)	1.097(4)	1.079(3)
r _{C(2)-H(8)}	1.085 (fixed)	1.097(4)	1.079(2)
angle C(3)-C(4)-C(5)	118.74(37)	118.56(28)	118.53(18)
C(4)-H(10) deviation*	0.63 (fixed)	0.72 (fixed)	0.81(20)

* a positive deviation is towards the adjacent fluorine atom

[†] all distances are given in Angstroms; all angles are given in degrees.

Table 7.8 - Rotation constants⁴ (MHz) and direct dipolar coupling constants² (Hz) used in the structural analysis of m-difluorobenzene

Constant	Observed	Corrected	Calculated	Uncertainty	Difference
B principal	1760.53	1760.14	1760.18	0.200	-0.039
C principal	1197.35	1197.22	1197.25	0.065	-0.036
B 13C(1)	1752.14	1751.74	1751.65	0.200	0.086
C 13C(1)	1193.29	1193.16	1193.13	0.065	0.036
C 13C(2)	1194.68	1194.55	1194.55	0.200	-0.003
B 13C(4)	1751.52	1751.12	1751.17	0.065	-0.049
C 13C(4)	1189.63	1189.50	1189.54	0.200	-0.032
C 13C(5)	1188.09	1187.96	1187.94	0.065	0.022
* D2,8	1170.97	1258.14	1276.44	9.00	-18.30
D3,8	169.90	173.82	173.53	0.50	0.29
D4,8	41.00	41.42	41.91	0.30	-0.49
D5,8	28.60	28.77	27.38	0.30	1.39
D2,10	45.00	45.58	45.77	0.22	-0.19
D3,10	171.40	176.21	175.09	0.60	1.12
D4,10	1463.77	1587.57	1580.36	12.40	7.21
D5,10	202.30	207.50	206.45	0.54	1.05
D8,9	424.03	428.33	428.34	0.50	-0.01
D8,10	85.82	86.81	86.97	0.11	-0.16
D9,10	342.19	345.77	345.82	0.36	-0.05
D2,11	27.30	27.52	27.39	0.20	0.13
D3,11	43.40	43.84	43.57	0.22	0.27
D4,11	173.60	178.19	178.11	0.55	0.08
D5,11	1181.75	1287.99	1277.74	10.70	10.25
D8,11	51.82	52.19	52.18	0.13	0.01
D9,11	71.58	71.81	71.68	0.13	0.13
D10,11	494.75	504.31	505.35	0.96	-1.04
D3,12	33.50	33.78	34.80	0.31	-1.02
D4,12	52.00	52.70	52.62	0.30	0.08
D9,12	52.70	52.85	52.96	0.13	-0.11
D10,12	102.50	103.84	103.76	0.15	0.08
D3,7	40.70	40.74	42.32	1.40	-1.58
D4,7	26.80	26.80	26.77	1.50	0.03
D2,9	154.10	155.24	153.88	1.32	1.36
D3,9	743.60	757.51	765.18	6.40	-7.67
D4,9	124.40	124.97	124.14	1.70	0.83
D5,9	34.90	34.90	36.03	1.30	-1.13
D7,9	70.30	70.24	72.04	1.16	-1.80
** D2,8	-1560.10	-1677.23	-1633.68	11.70	-43.55
D3,8	-221.00	-226.16	-226.03	0.55	-0.13
D2,10	-59.90	-60.68	-59.94	0.31	-0.74
D3,10	-219.60	-225.77	-226.93	0.68	1.16
D4,10	-1942.70	-2106.43	-2103.59	16.40	-2.84
D5,10	-270.60	-277.49	-277.16	0.73	-0.33
D9,10	-438.00	-442.71	-442.74	0.47	0.03
D2,11	-35.30	-35.58	-35.05	0.40	-0.53
D5,11	-1507.60	-1644.52	-1635.34	13.70	-9.18
D8,11	-66.18	-66.66	-66.79	0.13	0.13
D10,11	-657.67	-670.27	-672.11	1.30	1.84
D3,12	-47.20	-47.60	-46.30	0.31	-1.30
D10,12	-137.48	-139.26	-139.33	0.22	0.07

* ZLI1167 couplings

** ZLI1132 couplings

Table 7.9 - Amplitudes of vibration of m-difluorobenzene

Number	Atom Pair	u / Å
1	C1 - C2	0.040 (tied to u3)
2	C3 - C4	0.040 (tied to u3)
3	C4 - C5	0.040(2)
4	C1 - F7	0.039 (tied to u3)
5	C2 - H8	0.077 (tied to u7)
6	C4 - H10	0.077 (tied to u7)
7	C5 - H11	0.077(10)
8	C1...C3	0.052(2)
9	C1...C4	0.065(4)
10	C1...C5	0.053 (tied to u8)
11	C1...H8	0.098 (fixed)
12	C1...F9	0.065(2)
13	C1...H10	0.094 (fixed)
14	C1...H11	0.096 (fixed)
15	C1...H12	0.098 (fixed)
16	C2...C4	0.053 (tied to u8)
17	C2...C5	0.067 (tied to u9)
18	C2...F7	0.058 (tied to u8)
19	C2...H10	0.096 (fixed)
20	C2...H11	0.095 (fixed)
21	C4...C6	0.053 (tied to u8)
22	C4...F7	0.070(5)
23	C4...H8	0.096 (fixed)
24	C4...F9	0.058 (tied to u8)
25	C4...H11	0.098 (fixed)
26	C4...H12	0.096 (fixed)
27	C5...F7	0.066 (tied to u12)
28	C5...H8	0.095 (fixed)
29	C5...H10	0.098 (fixed)
30	F7...H8	0.130 (fixed)
31	F7...F9	0.095(8)
32	F7...H10	0.097 (fixed)
33	F7...H11	0.108 (fixed)
34	F7...H12	0.130 (fixed)
35	H8...H10	0.131 (fixed)
36	H8...H11	0.119 (fixed)
37	H10...H11	0.155 (fixed)
38	H10...H12	0.131 (fixed)

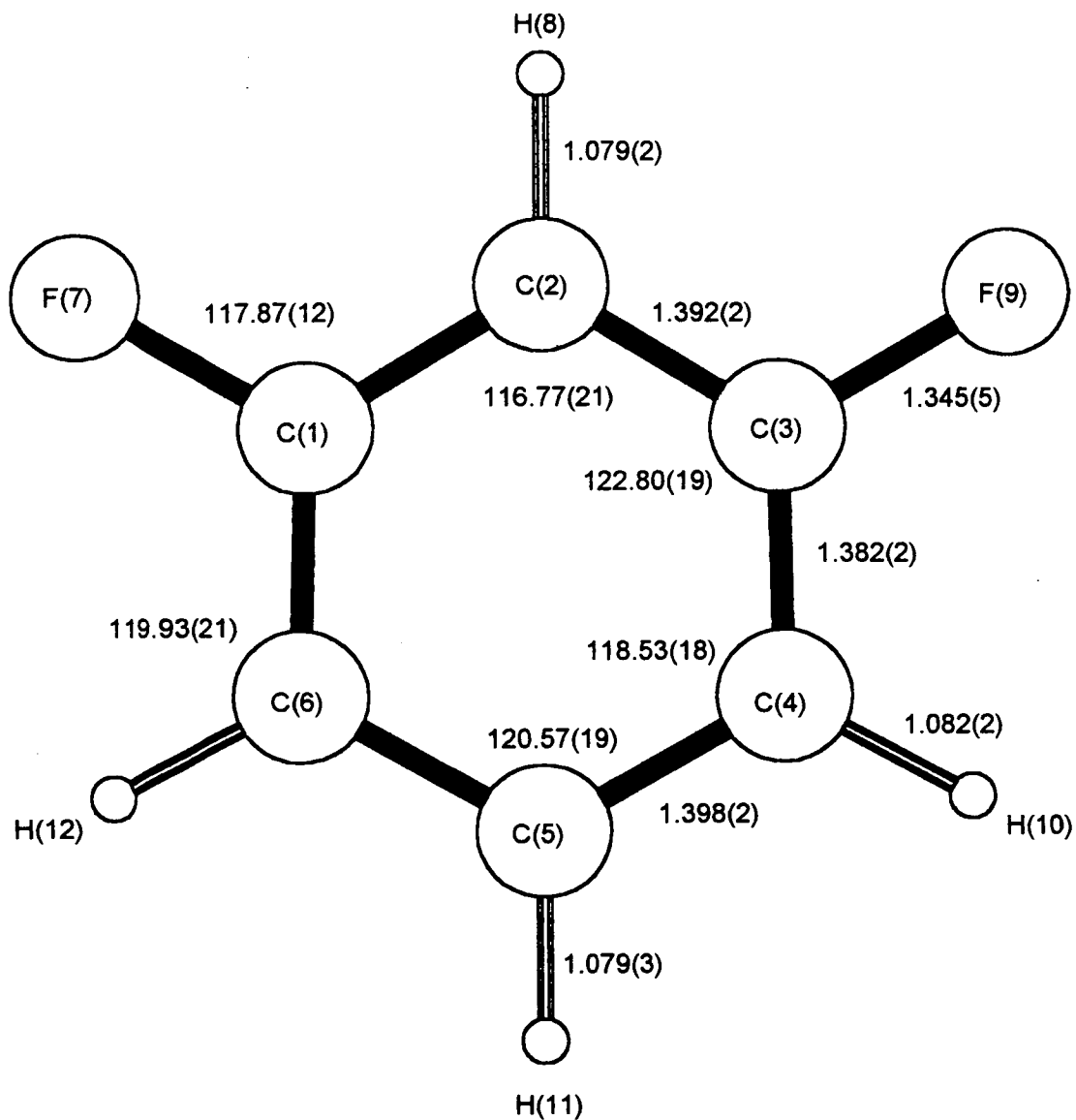
Table 7.10

Least squares correlation matrix (x100) for m-difluorobenzene.
All elements with absolute value <50 have been omitted.

	P4	P8	P9	P10	P11	P12	P14	P15	k2
P1	-84								
P2					-80				
P3				-70					
P4		-67	72						
P5						-71	71		
P6		-53							
P8			-83	81					
P9				-61	57				
P11						-55	52		
P13								-68	
u3									61
u9									58

Parameters (P) are in the order listed in table 7.7 ; amplitudes (u) are in the order listed in table 7.9 ; scale factors (k) are in the order listed in table 7.1.

Figure 7.7 - The molecular structure of m-difluorobenzene



All distances are given in Ångströms
All angles are given in degrees

Figure 7.8

The observed and final weighted difference
molecular scattering intensity curves for m-difluorobenzene

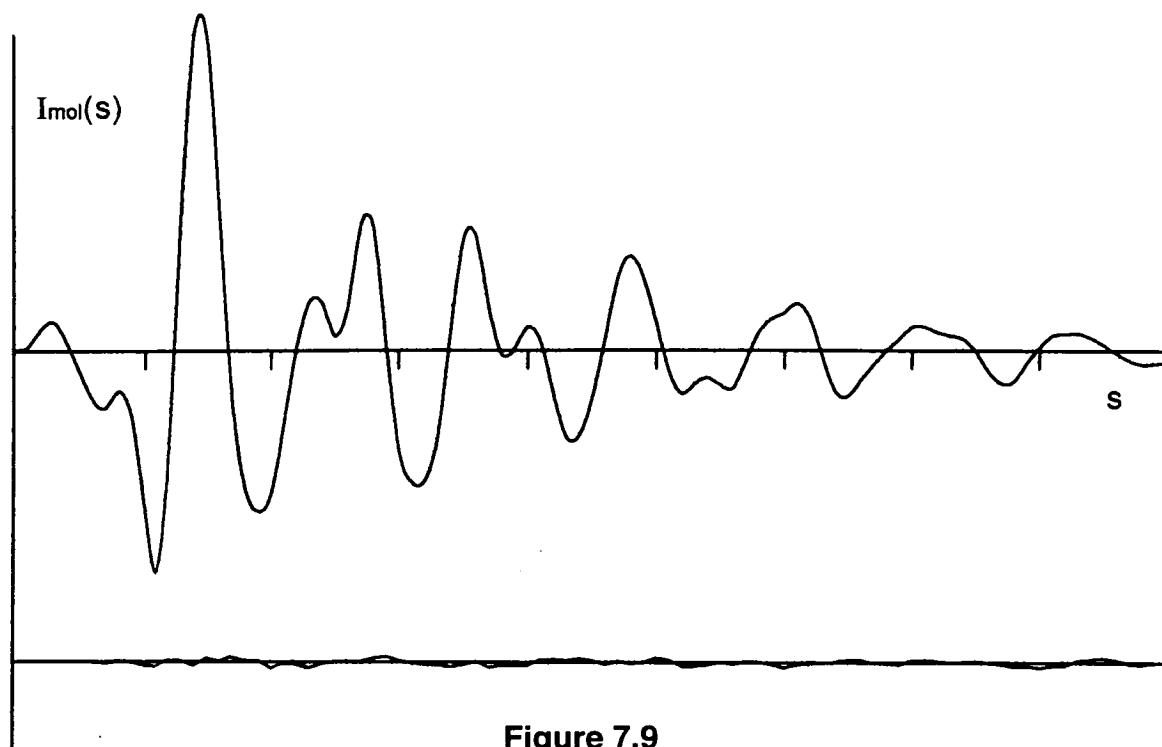
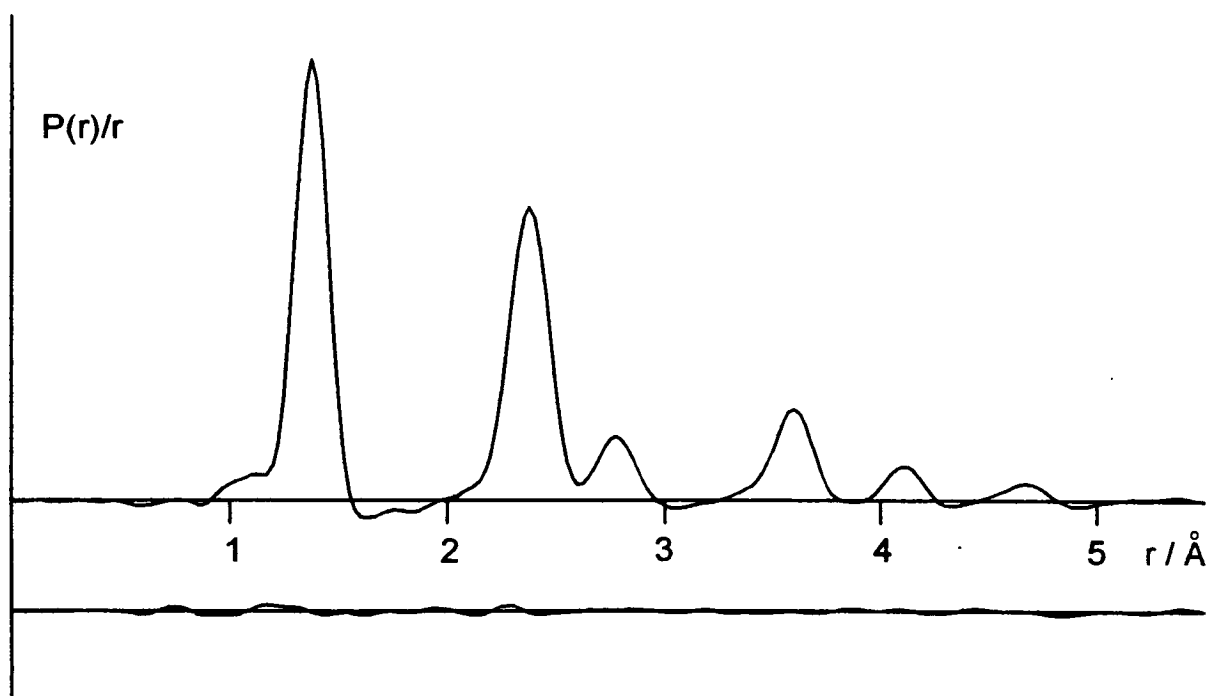


Figure 7.9

The observed and final difference
radial distribution curve, $P(r)/r$, for m-difluorobenzene



The structure of p-difluorobenzene

Of the three isomeric difluorobenzenes, p-difluorobenzene is the best suited to structural analysis by electron diffraction. The molecule can be assumed to have D_{2h} symmetry and only 6 independent parameters are required to fully describe the molecular geometry (see table 7.11). Furthermore, of the 22 distinct internuclear distances, there are only two C-C bonded distances, one C-F bonded distance and one C-H bonded distance.

Using ED data alone, R_G quickly dropped to 6.6% with only the four parameters describing the heavy atom positions refining (including the C-C difference parameter). Introducing amplitudes of vibration into the refinement and switching to an r_α^0 basis led to a slightly improved fit, with $R_G=6.0\%$. An attempt was made to refine the remaining two structural parameters (r_{C-H} and the C-H angle deviation) but only the bond length would refine with a reasonable e.s.d., 1.086(7) Å. The angle parameter was therefore fixed at zero for the final refinement.

No MW data are available for p-difluorobenzene as the molecule has no permanent dipole moment. However, this is not a major problem as the ring structure is already fairly well determined by the ED data. At this stage the first LCNMR data set (obtained using ZLI1167) was introduced. Initially this allowed simultaneous refinement of all six structural parameters but as the weight given to the LCNMR data was increased the ring geometry changed significantly. In particular, the parameter $r_{12} - r_{23}$ changed from -0.0135(46) Å, obtained in the ED analysis, to +0.0115(46) Å. The effect of fluorine substitution is usually to shorten the adjacent C-C bonds⁷ and there

is no apparent reason why p-difluorobenzene should differ in this respect (particularly as the ED structure is in agreement with this observation). Further experimentation showed that by excluding the C-F and F-F coupling constants, this parameter once more refined to a negative value. As with similar couplings in the previous two studies, this can be attributed to an anisotropic component of one or more of the indirect coupling constants.

The structure obtained using the ZLI1167 data was consistent with data obtained using two further solvents ZLI1132 and ZLI1695. The introduction of these data caused no significant change to the molecular geometry and the final refinement was carried out using ED data recorded at two camera distances and 33 direct dipolar coupling constants (see table 7.12). The final parameters can be found in table 7.11 and the vibrational amplitudes in table 7.13. Table 7.14 shows the least squares correlation matrix from the final refinement. The molecular geometry is also shown in figure 7.10 and the corresponding molecular scattering intensity and radial distribution curves are shown in figures 7.11 and 7.12 respectively.

Table 7.11 - Final parameters - *para*-difluorobenzene

Parameters [†]	ED	ED + LCNMR
Structural		
(r _{1,2} +r _{2,3})/2	1.395(2)	1.391(1)
r _{1,2} - r _{2,3}	-0.014(5)	-0.008(2)
r _{C-F}	1.334(4)	1.344(2)
r _{C-H}	1.086(7)	1.078(2)
angle C(6)-C(1)-C(2)	121.98(20)	122.24(16)
C-H angle deviation*	0.0 (fixed)	0.81(8)
Orientalional		
S _{VV} x 100 (ZLI1167)	-	-2.883(10)
S _{ZZ} x 100 (ZLI1167)	-	-4.904(8)
S _{VV} x 100 (ZLI1132)	-	6.127(21)
S _{ZZ} x 100 (ZLI1132)	-	11.464(17)
S _{VV} x 100 (ZLI1695)	-	-3.811(13)
S _{ZZ} x 100 (ZLI1695)	-	-6.046(10)
Dependent		
r _{C(1)-C(2)}	1.388(1)	1.387(1)
r _{C(2)-C(3)}	1.402(4)	1.395(2)
angle C(1)-C(2)-C(3)	119.03(10)	118.88(8)

* a positive deviation is towards the adjacent fluorine atom

[†] all distances are given in Angstroms; all angles are given in degrees.

Table 7.12
Direct dipolar coupling constants³ (Hz) used
in the structural analysis of p-difluorobenzene

Constant	Observed	Corrected	Calculated	Uncertainty	Difference
* D1,8	87.48	89.59	89.34	0.22	-0.17
D2,8	766.30	827.89	827.46	6.20	0.43
D7,8	228.58	231.22	231.37	0.27	-0.15
D1,9	31.81	32.10	31.96	0.10	0.14
D2,9	132.81	135.80	135.74	0.30	0.06
D8,9	363.61	369.08	369.43	0.55	-0.35
D8,10	54.73	54.99	54.96	0.10	0.03
D2,11	17.75	17.84	17.59	0.20	0.25
D8,11	33.47	33.64	33.59	0.10	0.06
D2,12	22.40	22.57	22.67	0.20	-0.10
D8,12	44.21	44.56	44.50	0.10	0.06
** D1,8	-185.73	-190.35	-189.96	0.60	-0.38
D2,8	-1689.53	-1826.16	-1826.75	13.80	0.59
D7,8	-506.69	-512.68	-513.41	0.60	0.73
D1,9	-72.38	-73.04	-72.82	0.30	-0.22
D2,9	-307.00	-313.91	-314.02	0.80	0.11
D8,9	-850.37	-862.92	-863.70	1.30	0.78
D8,10	-126.26	-126.86	-126.81	0.10	-0.05
D2,11	-38.47	-38.66	-38.76	0.30	0.10
D8,11	-73.71	-74.11	-74.04	0.10	-0.07
D2,12	-47.98	-48.35	-48.39	0.36	0.04
D8,12	-93.80	-94.56	-94.57	0.10	0.01
*** D1,8	115.33	118.07	118.08	0.37	-0.01
D2,8	987.32	1066.32	1065.35	7.90	0.97
D7,8	293.16	296.49	296.80	0.34	-0.31
D1,9	40.15	40.52	40.20	0.20	0.32
D2,9	165.06	168.77	168.75	0.40	0.02
D8,9	448.28	455.11	455.48	0.68	-0.37
D8,10	68.22	68.54	68.48	0.10	0.06
D2,11	22.65	22.76	22.68	0.28	0.08
D8,11	43.12	43.35	43.29	0.10	0.06
D2,12	29.84	30.07	29.87	0.28	0.20
D8,12	58.37	58.83	58.82	0.10	0.01

* ZLI1167 couplings

** ZLI1132 couplings

*** ZLI1695 couplings

Table 7.13 - Amplitudes of vibration of p-difluorobenzene

Number	Atom Pair	u / Å
1	C1 - C2	0.043(1)
2	C2 - C3	0.043 (tied to u1)
3	C1 - F7	0.043 (tied to u1)
4	C2 - H8	0.084(10)
5	C1...C3	0.053(1)
6	C1...C4	0.067(5)
7	C1...H8	0.098 (fixed)
8	C1...H9	0.096 (fixed)
9	C1...F10	0.070(5)
10	C2...C5	0.070 (tied to u6)
11	C2...C6	0.054 (tied to u5)
12	C2...F7	0.059 (tied to u5)
13	C2...H9	0.098 (fixed)
14	C2...F10	0.065(2)
15	C2...H11	0.095 (fixed)
16	C2...H12	0.096 (fixed)
17	F7...H8	0.130 (fixed)
18	F7...H9	0.108 (fixed)
19	F7...F10	0.076(7)
20	H8...H9	0.154 (fixed)
21	H8...H11	0.119 (fixed)
22	H8...H12	0.131 (fixed)

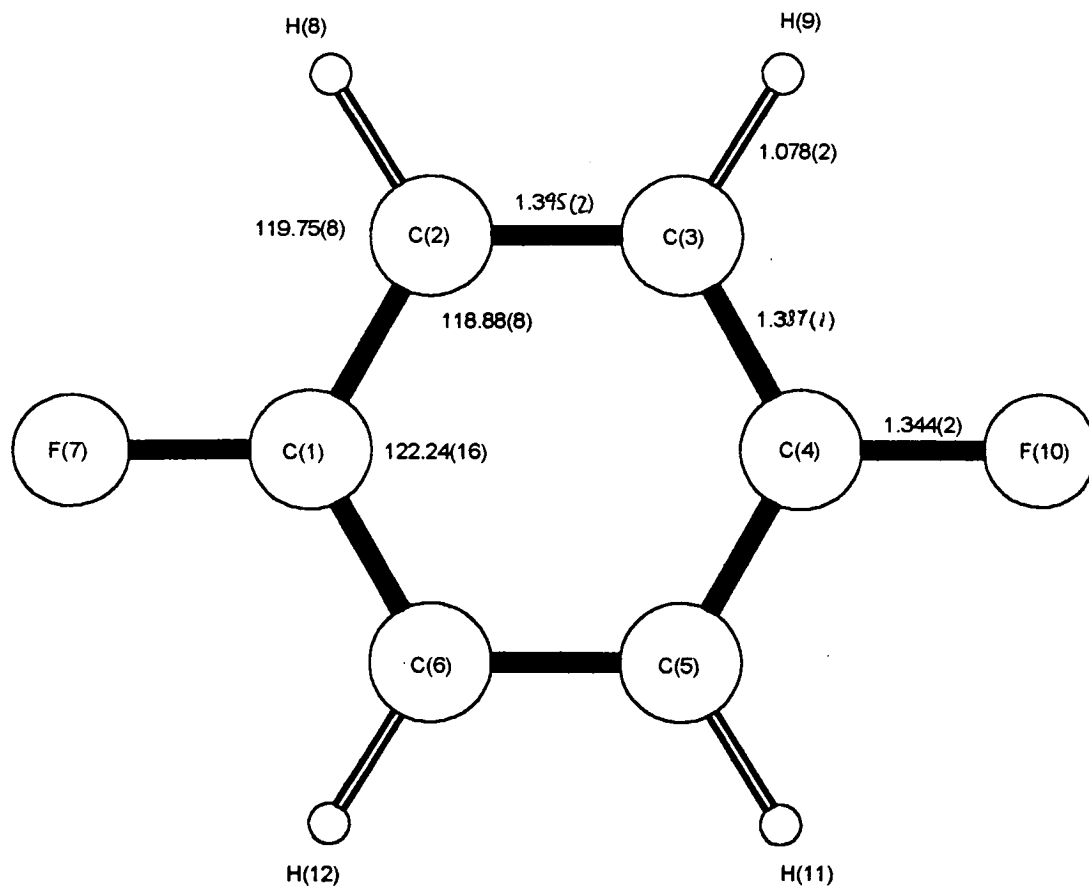
Table 7.14

Least squares correlation matrix (x100) for p-difluorobenzene.
All elements with absolute value <50 have been omitted.

	P3	P5	P6	P7	P8	P9	P10	P11	P12	u1	u5	k2
P1	-62											
P2			51									
P3		55								54		
P4				-72	-59	75	68	-73	-58			
P5			-68									
P7						-91		89				
P8							-72		68			
P9								-92				
P10									-73			
u1											52	60
u5												56

Parameters (P) are in the order listed in table 7.11 ; amplitudes (u) are in the order listed in table 7.13 ; scale factors (k) are in the order listed in table 7.1.

Figure 7.10 - The molecular structure of p-difluorobenzene



All distances are given in Ångströms
All angles are given in degrees

Figure 7.11

The observed and final weighted difference
molecular scattering intensity curves for p-difluorobenzene

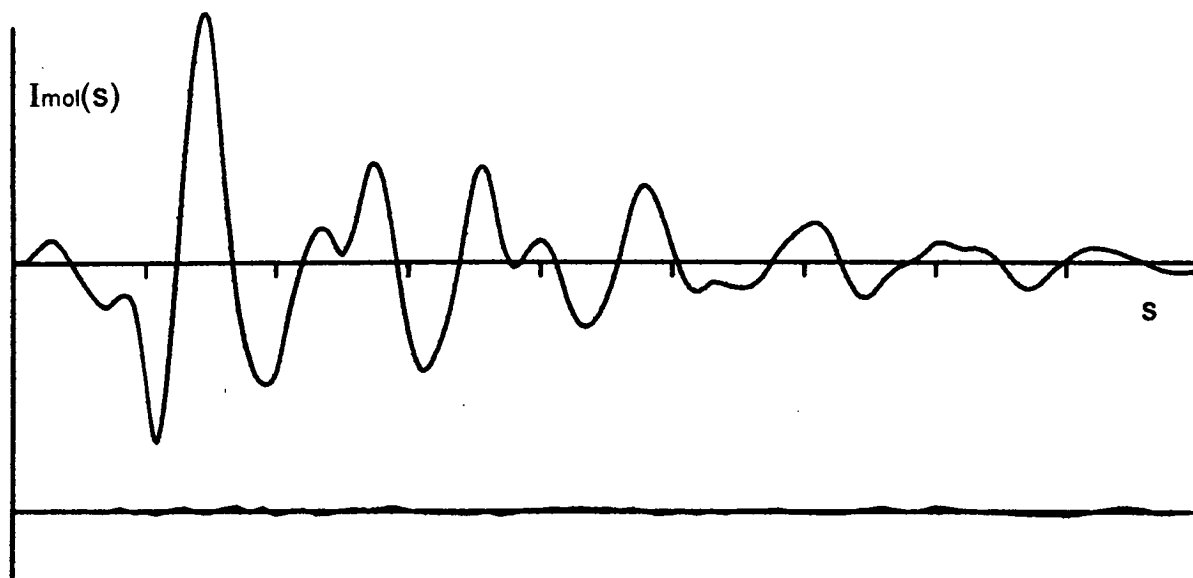
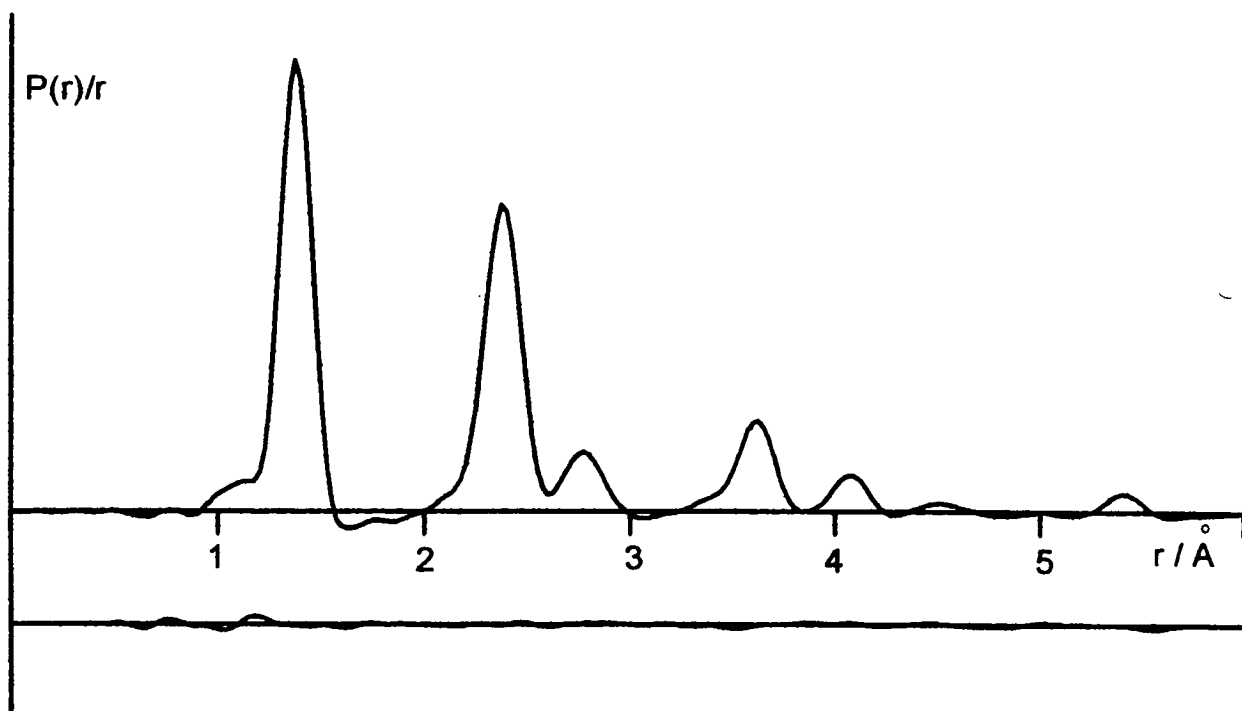


Figure 7.12

The observed and final difference
radial distribution curve, $P(r)/r$, for p-difluorobenzene



Conclusions

Combined analysis

The above studies serve to highlight some important points about the technique of combined analysis. In the first instance, it is clear that the use of ED data alone was insufficient to determine any of the structures completely. In general, ED data supplied information on the ring geometry and the positions of the fluorine atoms but, in most cases, C-C bond length difference parameters could not be refined. A notable exception to this is the difference parameter in p-difluorobenzene but even this parameter had a fairly high e.s.d. in the ED analysis. The addition of MW data, where available, allowed some of these difference parameters to be refined but still little could be said about the positions of the hydrogen atoms. Only on introducing the LCNMR data could any of the C-H angles be refined successfully. Furthermore, the results were generally improved when two or more sets of direct coupling constants were included. This helped to reduce the effects of random errors in the LCNMR data (see below).

The case of m-difluorobenzene, in particular, shows how powerful the technique of combined analysis can be. In the ED analysis, only 6 of the 11 structural parameters could be refined whereas in the final combined analysis all 11 structural parameters, as well as 4 orientational parameters, were successfully refined.

Structures

The structures obtained for the three difluorobenzenes are largely consistent with expectations arising from previous studies^{7,8,9,10}. In particular the effects of fluorine substitution on the internal ring angles are easily predicted; the *ipso* angle opens out while the adjacent angles are reduced by approximately half the amount. The remaining ring angles are also affected but to a lesser degree. This is a well determined and widely studied effect^{11,12} and applies to a whole range of electron withdrawing substituents. In the case of the difluorobenzenes, the ring geometry can be considered as a superposition of the effects arising from the individual fluorine atoms.

In *o*-difluorobenzene these effects are in competition, angle C(6)-C(1)-C(2) is increased by the *ipso* fluorine but is decreased by the *ortho* fluorine, although to a lesser degree. For this reason, the ring geometry of *o*-difluorobenzene is remarkably close to that of a regular hexagon with the largest deviation of an internal angle being 0.73(18) degrees. Not surprisingly this is the angle at atoms C(3) and C(6) where the effects are not in direct competition.

m-difluorobenzene is very different in that the effects of the two fluorine atoms are combined constructively. This produces some very large angular distortions within the ring, particularly at angle C(1)-C(2)-C(3) which, in the final combined analysis, refined to 116.77(21) degrees. In fact, this angle seems much more distorted than might be expected which is perhaps an indication that the final value for the parameter $r_{2,3} - r_{3,4}$ is unrealistic.

The angles produced in the ED+MW refinement, with this parameter fixed at zero, seem more in line with expectations. It is hoped that this discrepancy will be resolved when **ed92** is modified to allow the use of more than 50 non-ED data.

The fluorine atoms in p-difluorobenzene are sufficiently far apart to have essentially independent effects on the ring angles. For this reason, the ring angles are similar to the *ipso* and *ortho* angles observed in fluorobenzene⁷.

In general, the increase of the ring angle at a carbon atom substituted with an electron withdrawing group is associated with a shortening of the two adjacent bonds. This has been rationalised in terms of hybridisation effects¹² in which there is an increase in the p character of the sp² hybrid orbital of the substituted carbon which points towards the substituent. This effectively leads to a decrease in the p character of the remaining sp² orbitals and hence a shortening of the C-C bonds.

Similar effects are observed for the difluorobenzenes. In general, the bonds adjacent to the fluorine substituted carbon are shortened by between 0.01 and 0.02 Å. However, care should be taken when assessing these results as the bond length differences are not well determined in some cases. A notable exception to this rule is the C(1)-C(2) bond in o-difluorobenzene. Being adjacent to both fluorine substituents, it is not unreasonable to expect this bond to be extremely short. In fact it turns out to be the longer than the C(2)-C(3) bond which is only adjacent to one fluorine. This could be rationalised in terms of steric repulsion between the fluorine atoms but this explanation is inconsistent with the observation that

the fluorine atoms are bent towards each other, by 0.61(12) degrees. It seems likely, therefore, that the lengthening of this bond is due to electronic, rather than steric, factors. It can be imagined that the electron withdrawing effect of the fluorine atoms leaves the carbon atoms with a slight positive charge. The carbon atoms would then experience a degree of electrostatic repulsion which counteracts the bond shortening effect described above. It is satisfying to note that a similar effect is observed in *o*-difluorobenzene⁸ although the uncertainties involved make it impossible to draw any firm conclusions.

The use of combined analysis, in particular the inclusion of LCNMR data, allows more subtle structural effects to be studied. For example, in all three molecules the C-H bonds adjacent to the fluorine substituents are bent towards the fluorine atom by between half and one degree. In fact, the consistency of this result is a good indication of the accuracy of the technique. In the case of *m*-difluorobenzene, the angle of the bond C(2)-H(8) is fixed by the symmetry of the molecule and instead it is the C-F bonds which deviate towards the hydrogen atom by 0.7(2) degrees. Unfortunately, the C-H bond lengths are not determined with sufficient accuracy to allow an investigation into how they are affected by fluorine substitution in the ring. *Ab initio* calculations have suggested that C-H bond lengths are likely to change by only a few thousandths of an Ångström¹³.

Uncertainties

On the whole, the use of **MM3** to calculate vibrational corrections for data to be used in combined analysis has proved successful. In particular, the large uncertainties associated with the corrections to rotation constants do not seem to prevent useful structural information from being obtained. The calculation of corrections to direct dipolar coupling constants also seems to be satisfactory, especially in view of the consistency of the results regarding the C-H angle deviations. Nonetheless, several points must be raised concerning the estimated uncertainties of such corrections. In general, assuming an uncertainty of 10% in the correction term d^h seems to result in data which are consistent with a single structure. However, it seems that extra caution should be taken when employing D_{CF} or D_{FF} coupling constants. The possibility of a significant anisotropy of the associated indirect coupling constants should be taken into account, either by excluding such couplings from the refinement or by increasing their uncertainties.

It was noted above that, in several cases, the calculated direct coupling constants do not agree with the observed values. In particular, some of the coupling constants of the directly bonded C-H atom pairs are different by three or four standard deviations. This can perhaps be attributed to an underestimation of the uncertainties of the vibrational corrections. However, discrepancies are also found for a number of coupling constants in which the correction term is very small. There are a number of possible explanations for such inconsistencies, some of which are outlined below.

The uncertainties used in the combined analysis are a combination of uncertainties in the original observed data and uncertainties in the vibrational correction terms. For couplings in which the vibrational correction is extremely small, the uncertainty in the original observation becomes dominant. If such a coupling does not fit well in the combined analysis then it is possible that this uncertainty has been underestimated. In most cases, these uncertainties are produced by the programs used in the analysis of the LCNMR spectra. However, for the most part, these programs do not take into account such factors as uncertainties in the line positions from the NMR experiment or uncertainties in the indirect coupling constants (whether due to a significant anisotropic contribution or to solvent or temperature dependency of J). The possibility of misassignment of one or more lines is also ignored.

During the combined analysis, the molecular geometry is refined on an r_{α}° basis which should be consistent with the LCNMR data. However, when calculating the theoretical molecular scattering intensities the r_{α}° distances are first transformed to r_g distances using perpendicular amplitudes (K), calculated from the vibrational analysis. If these amplitudes were changed then the ED data would still produce the same r_g values but the corresponding r_{α}° values (and hence the calculated direct couplings) would differ. At present no account is taken of the uncertainties of the calculated perpendicular amplitudes (perhaps arising from the use of rectilinear rather than curvilinear normal co-ordinates) and so this presents another possible source for discrepancies in the calculated direct couplings.

Finally, it should not be forgotten that the sample in the LCNMR experiment is in a different phase to the samples in the ED and MW experiments. The

possibility of structural variations between the molecules in these phases should not be ruled out. Although the molecules and solvents used in this work were chosen to minimise the possibility of such distortions, it can not be assumed that the LCNMR data and gas-phase data are completely compatible. However, the fact that the couplings fit as well as they do would suggest that any such differences must be small and the use of LCNMR data in the combined analysis seems justified.

References

- ¹T.Väänänen, J.Jokisaari & J.Lounila, *J.Magn.Reson.*, 49, (1982), 73
- ²T.Väänänen, J.Jokisaari & J.Lounila, *J.Magn.Reson.*, 54, (1983), 436
- ³A.Pulkkinen, J.Jokisaari & T.Väänänen, *J.Mol.Struct.*, 144, (1986), 359
- ⁴S.Doraiswamy & S.D.Sharma, *J.Mol.Struct.*, 110, (1983), 81
- ⁵O.P.Singh, J.S.Yadav & R.A.Yadav, *Proc.Indian Acad.Sci.*, 99(3), (1987), 159
- ⁶L.Ernst, D.N.Lincoln & V.Wray, *J.Magn.Reson.*, 21, (1976), 115
- ⁷G.Portalone, A.Domenicano & I.Hargittai, *J.Mol.Struct.*, 118, (1984), 53
- ⁸P.B.Liescheski, D.W.H.Rankin & H.E.Robertson, *Acta Chemica Scand.*, A42, (1988), 338
- ⁹D.G.Anderson, S.Cradock, P.B.Liescheski & D.W.H.Rankin, *J.Mol.Struct.*, 216, (1990), 181
- ¹⁰S.Cradock, P.B.Liescheski & D.W.H.Rankin, *J.Magn.Reson.*, 91, (1991), 316

¹¹A.Domenicano, "*Accurate Molecular Structures: Their Determination and Importance*", Chapter 18, eds. A.Domenicano & I.Hargittai, Oxford University Press, (1992), 437

¹²L.Nygaard, I.Bojesen, T.Pederson & J.Rastrup-Anderson, *J.Mol.Struct.*, 2, (1968), 209

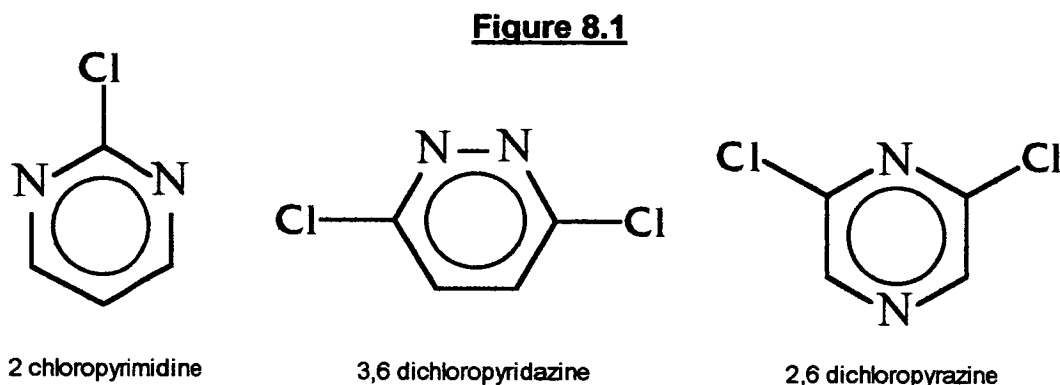
¹³C.W.Bock, M.Trachtman & P.George, *J.Mol.Struct. (Theochem)*, 122, (1985), 155

Chapter 8

The Structures of Some Chlorine Substituted Heteroaromatic Compounds

Introduction

The difluorobenzenes, described in the previous chapter, are ideally suited to a combined analysis by ED, LCNMR and MW data. This has also proved to be the case for a number of heteroaromatic compounds^{1,2} and other substituted benzenes^{3,4}. A logical extension to this work is to study substituted heteroaromatic compounds, three of which are presented here. The molecules chosen are 2-chloropyrimidine, 3,6-dichloropyridazine and 2,6-dichloropyrazine, and are shown in figure 8.1.



These particular molecules were chosen for a number of reasons. They are all of C_{2v} symmetry and are therefore better suited to ED analysis than asymmetrically substituted aromatics. However, the small number of abundant spin- $\frac{1}{2}$ nuclei (2 or 3 hydrogen atoms) is approaching the limit at which useful LCNMR data can be obtained. Even if couplings due to natural abundance ^{13}C and ^{15}N can be found, the maximum number of data possible is much less than in the case of the difluorobenzenes. Furthermore, the presence of quadrupolar ^{14}N can lead to line broadening which make it difficult to obtain useful spectra (cf. the analysis of 2-chloropyridine in chapter 5). The relatively heavy chlorine atoms have the greatest effect on both the ED and MW data, making the determination

of the positions of the ring atoms more difficult than for the difluorobenzenes. In many ways, therefore, these molecules can be considered to be testing the limit of the applicability of the combined analysis techniques which have proved so successful in previous studies.

Experimental

Samples of the three compounds were obtained from the Aldrich Chemical Company and used without further purification. ED data were recorded using the Edinburgh apparatus using procedures describe in chapter 1. Details of the experimental temperatures and the ranges of the data obtained can be found in table 8.1.

An LCNMR spectrum was obtained of 2-chloropyrimidine in the solvent EBBA (see figure 2.5), at 25°C. The spectrum of 3,6-dichloropyridazine was recorded in the solvent E5, also at 25°C. The analyses of these spectra is described below. Attempts to obtain a useful LCNMR spectrum of 2,6 dichloropyrazine were unsuccessful. Rotation constants of the principle isotopic species and of the ^{37}Cl substituted isotopomers of 2-chloropyrimidine and 3,6-dichloropyridazine were recorded at the University of Connecticut, by Robert K. Bohn⁵.

N.B. The ED data and LCNMR spectra of these three compounds were originally recorded as part of an undergraduate project⁶. Unfortunately, the unavailability of reliable force fields for these molecules meant that the analysis was incomplete. In the present work a complete reanalysis of the LCNMR and ED data has been undertaken and the new MW data have been incorporated where available.

Vibrational corrections

Data from the three techniques, ED, LCNMR and MW spectroscopy, were reduced to a common basis using the corrections described in chapter 4. The vibrational analysis was carried out using the modified version of the molecular mechanics program, **MM3**, as described in chapter 6.

Table 8.1

Experimental details and weighting parameters for the ED analysis
of 2-chloropyrimidine, 3,6-dichloropyridazine and 2,6-dichloropyrazine

	2-chloropyrimidine		3,6-dichloropyridazine		2,6-dichloropyrazine		
Camera dist. / mm	256.99	198.30	94.66	257.98	97.41	257.98	97.41
Nozzle temp. / K	386	386	386	443	443	443	443
$\Delta s / \text{\AA}^{-1}$	0.2	0.4	0.4	0.2	0.4	0.2	0.4
$s_{\text{min}} / \text{\AA}^{-1}$	2.0	4.0	10.0	2.0	14.0	2.0	12.0
$s_{w1} / \text{\AA}^{-1}$	4.0	6.0	11.0	4.0	15.0	4.0	14.0
$s_{w2} / \text{\AA}^{-1}$	14.4	19.6	30.0	14.0	30.0	14.0	30.4
$s_{\text{max}} / \text{\AA}^{-1}$	16.8	22.4	35.2	16.4	35.2	16.4	35.2
Correl ^l parameter	0.4630	-0.0017	0.3935	0.4900	0.4207	0.4835	0.1628
Scale factor	0.941	0.903	0.775	0.836	0.750	0.859	0.801
Wavelength / \AA	0.0567	0.0567	0.0566	0.0567	0.0568	0.0567	0.0567

2-chloropyrimidine

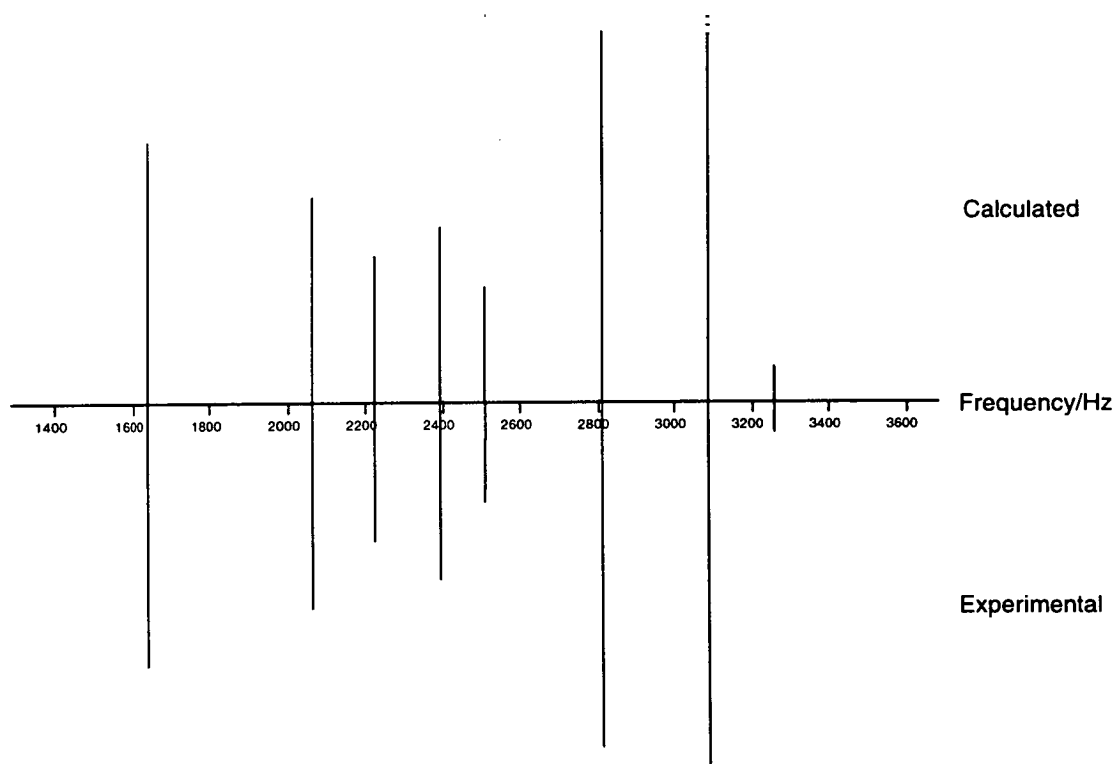
Analysis of the LCNMR spectrum

The ^1H spectrum of 2-chloropyrimidine, in the liquid crystal solvent EBBA, was recorded at room temperature. This was not the first choice of solvent but solubility problems prevented the use of the preferred solvent E5. The resulting second order spectrum was analysed using the programs **lcsim** and **sliquor**, as described above. Values for the indirect coupling constants were obtained from a sample run using D_2O as a solvent. The signs of these couplings were deduced by comparison with the equivalent couplings in unsubstituted pyrimidine⁷. The parent spectrum yields only two direct dipolar coupling constants (the two D_{HH} couplings) and can not be used to provide any structural information, as two independent orientation parameters must also be determined. For this reason, the assignment some of the ^{13}C and ^{15}N satellite peaks is essential if the LCNMR data are to be used in a combined structural analysis. A simulation of the parent spectrum is shown in figure 8.2. Unfortunately, the quality of the spectrum was not particularly good with some of the lines being quite broad due to the presence of the quadrupolar ^{14}N nuclei. This made the identification of satellite peaks extremely difficult. Not enough ^{13}C satellite peaks could be assigned with sufficient confidence to allow the refinement of any of the couplings involving atoms C(2) or C(5) and no ^{15}N satellites were found. However, from the $^{13}\text{C}(4)$ subspectrum a total of eleven satellites were identified allowing the refinement of the three couplings, $D_{4,8}$, $D_{4,9}$ and $D_{4,10}$. Although this does not provide a great deal of structural information, these three couplings and the two D_{HH} couplings were included

in the combined analysis (see below). When the structural analysis was complete the structure obtained was used in a further attempt to identify some of the $^{13}\text{C}(2)$, $^{13}\text{C}(5)$ and ^{15}N satellite peaks but this proved unsuccessful. The $^{13}\text{C}(2)$ couplings are small and the predicted satellites lie underneath the broad lines of the parent spectrum. None of the ^{15}N satellites were observed, presumably due to the low natural abundance of the isotope. The final values obtained for the direct coupling constants which could be determined can be found in the combined analysis section, below.

Figure 8.2

Simulation of the ^1H LCNMR spectrum of
2-chloropyrimidine in the solvent EBBA



Structure refinement of 2-chloropyrimidine

The molecular geometry can be defined in terms of 9 independent structural parameters and these are listed in table 8.2. Initial values were chosen to produce reasonable bond lengths and all angles were initially set at 120°. Parallel and perpendicular vibrational amplitudes were calculated for each of the 27 distinct internuclear distances using the program **MM3**, as described above.

Using ED data only the principal heavy atom parameters were introduced into the refinement one at a time until R_G settled at 6.3%. It is interesting to note that both of the ring bond difference parameters could be refined, even at this early stage. Including some of the more important vibrational amplitudes in the refinement led to a gradual reduction in R_G . In particular, the refinement of the amplitudes associated with the three bond ring distances and the N...Cl amplitude seemed to improve the fit to the ED data significantly. A final ED-only structure was obtained with an R_G factor of just 5.3% which is probably due to the fact that ED data were recorded at three camera distances for this molecule. Only the parameters involving the positions of the hydrogen atoms could not be refined, although the parameter defining the difference between the two C-N bond lengths seemed quite large in the final refinement.

The three independent rotation constants were included as extra data, initially with reduced weight. As the weight was gradually increased, the C-N difference parameter decreased until reaching a final value of 0.004(3) Å, which seems more realistic than the value obtained in the ED analysis. An attempt was made to refine the mean C-H bond length but

although the e.s.d. for this parameter was reasonably small the bond length itself seems quite large 1.098(6) Å.

Finally, the five direct dipolar coupling constants obtained from the LCNMR analysis were included. Perhaps not surprisingly, the C-H bond length difference could not be refined with the limited amount of data available and so was fixed at zero. However, the addition of the LCNMR data did allow the C(5)-C(4)-H(8) angle to be refined as well as the mean C-H bond length, which decreased to a more reasonable value of 1.090(4) Å. With the exception of $D_{4,10}$, the calculated values for the extra data fitted the observed values to within one or two standard deviations (see table 8.3).

The final refinement is based on ED data recorded at three camera distances, three rotation constants and just five direct coupling constants. The parameters obtained at the various stages of the refinement can be found in table 8.2 and the final values of the vibrational amplitudes are listed in table 8.4. Table 8.5 shows the least squares correlation matrix calculated during the final refinement. The molecular geometry obtained from the combined analysis of data from all three sources is shown in figure 8.3 and the corresponding molecular scattering intensity and radial distribution curves are shown in figures 8.4 and 8.5 respectively.

Table 8.2 - Final parameters: 2-chloropyrimidine

Parameters [†]	ED	ED+MW	All data
Structural			
(r _{1,2} +r _{3,4} +r _{4,5})/3	1.3497(10)	1.3524(4)	1.3525(4)
mean r _{CN} - r _{CC}	-0.041(8)	-0.073(5)	-0.075(4)
r _{1,2} - r _{3,4}	-0.018(4)	-0.004(3)	-0.008(3)
r _{C-Cl}	1.728(2)	1.729(2)	1.729(2)
(2xr _{4,8} +r _{5,9})/3	1.080 (fixed)	1.098(6)	1.090(4)
r _{4,8} - r _{5,9}	0.0 (fixed)	0.0 (fixed)	0.0 (fixed)
angle N(1)-C(2)-N(3)	128.53(21)	128.17(20)	127.97(19)
angle C(2)-N(3)-C(4)	114.43(36)	115.77(29)	116.07(24)
angle C(5)-C(4)-H(8)	118.2 (fixed)	119.1 (fixed)	121.02(37)
Orientalional			
S _{VV} x 100 (EBBA)	-	-	2.439(22)
S _{ZZ} x 100 (EBBA)	-	-	7.129(47)
Dependent			
r _{N(1)-C(2)}	1.327(2)	1.326(2)	1.324(2)
r _{N(3)-C(4)}	1.345(3)	1.330(2)	1.332(2)
r _{C(4)-C(5)}	1.377(6)	1.401(3)	1.402(3)
r _{C(4)-H(8)}	1.08 (fixed)	1.098(6)	1.090(4)
r _{C(5)-H(9)}	1.08 (fixed)	1.098(6)	1.090(4)
angle N(3)-C(4)-C(5)	123.52(33)	122.04(28)	121.78(23)
angle C(4)-C(5)-C(6)	115.57(31)	116.20(23)	116.31(24)

[†] all distances are given in Ångströms; all angles are given in degrees.

Table 8.3 - Rotation constants⁵ (MHz) and direct dipolar coupling constants (Hz) used in the structural analysis of 2-chloropyrimidine

Constant	Observed	Corrected	Calculated	Uncertainty	Difference
B principal	1705.71	1705.35	1705.33	0.18	0.02
C principal	1331.95	1331.75	1331.78	0.10	-0.03
C 37Cl(7)	1302.32	1302.12	1302.12	0.10	-0.00
D8,9	-267.12	-273.19	-273.81	0.70	0.62
D8,10	-36.32	-36.59	-36.28	0.20	-0.31
D4,9	-162.78	-166.73	-166.40	0.55	-0.33
D4,10	-17.98	-18.12	-19.71	0.40	1.59
D4,8	-720.24	-801.30	-796.66	6.8	4.64

Table 8.4 - Amplitudes of vibration of 2-chloropyrimidine

Number	Atom Pair	u / Å
1	C2 - CL7	0.048(3)
2	C2 - N3	0.033(3)
3	N3 - C4	0.033 (tied to u2)
4	C4 - C5	0.034 (tied to u2)
5	C4 - H8	0.077 (fixed)
6	C5 - H9	0.078 (fixed)
7	N1...N3	0.060(3)
8	N1...C4	0.070(2)
9	N1...C5	0.062 (tied to u7)
10	N1...CL7	0.069 (tied to u8)
11	N1...H8	0.094 (fixed)
12	N1...H9	0.095 (fixed)
13	N1...H10	0.098 (fixed)
14	C2...C4	0.061 (tied to u7)
15	C2...C5	0.067 (tied to u8)
16	C2...H8	0.096 (fixed)
17	C2...H9	0.092 (fixed)
18	C4...C6	0.062 (tied to u7)
19	C4...CL7	0.075(2)
20	C4...H9	0.099 (fixed)
21	C4...H10	0.095 (fixed)
22	C5...CL7	0.086(5)
23	C5...H8	0.100 (fixed)
24	CL7...H8	0.112 (fixed)
25	CL7...H9	0.096 (fixed)
26	H8...H9	0.157 (fixed)
27	H8...H10	0.130 (fixed)

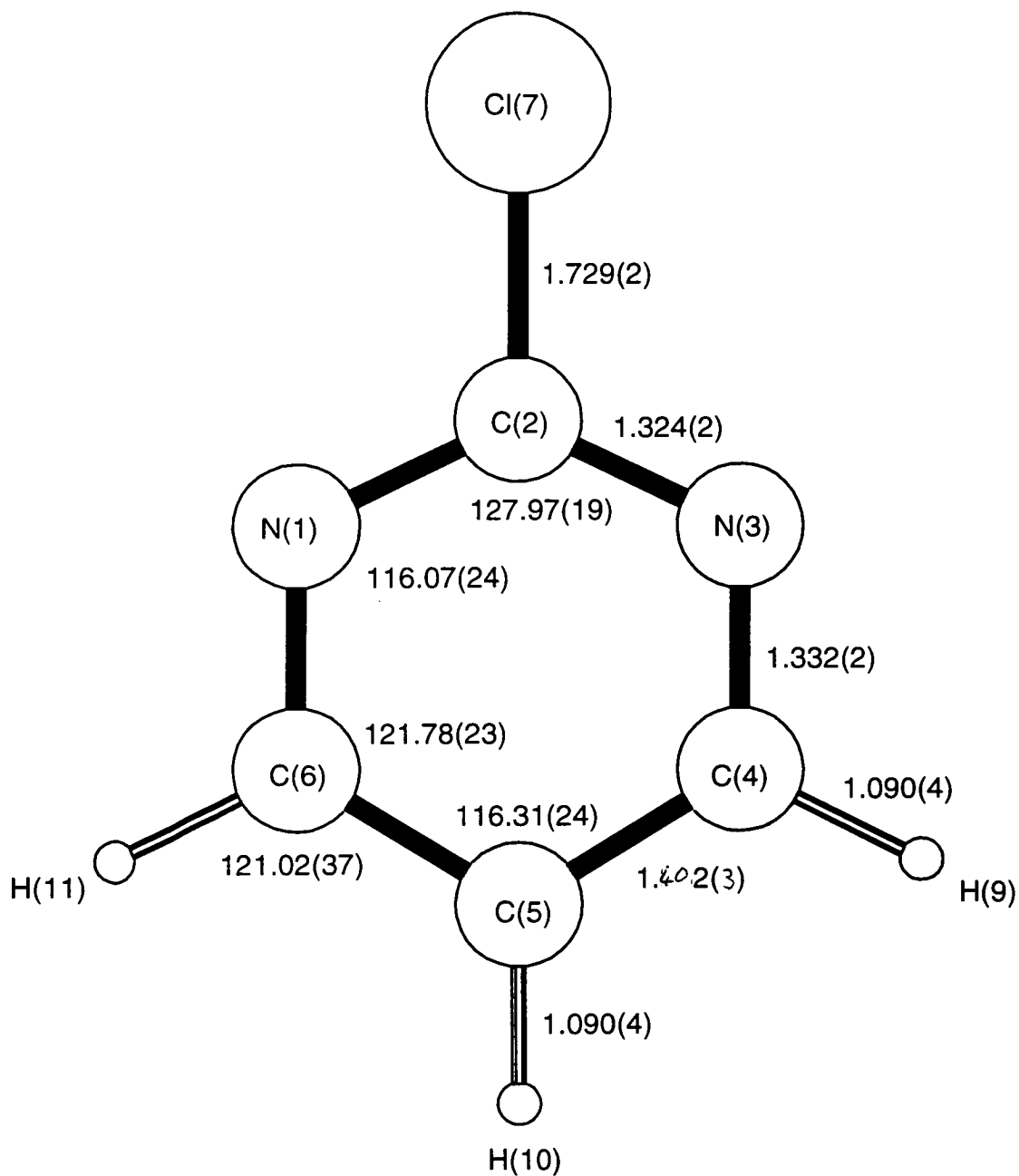
Table 8.5

Least squares correlation matrix (x100) for 2-chloropyrimidine.
All elements with absolute value <50 have been omitted.

	P7	P8	P10	P11	u2	u7	u8	k2
P2		-72		-54	77			
P4	76							
P5				50				
P7		-77				-52		
P8					-52			
P9			61					
P11					-56			
u2								55
u7							59	

Parameters (P) are in the order listed in table 8.2 ; amplitudes (u) are in the order listed in table 8.4 ; scale factors (k) are in the order listed in table 8.1.

Figure 8.3 - The molecular structure of 2-chloropyrimidine



All distances are given in Ångströms
All angles are given in degrees

Figure 8.4

The observed and final weighted difference
molecular scattering intensity curves for 2-chloropyrimidine

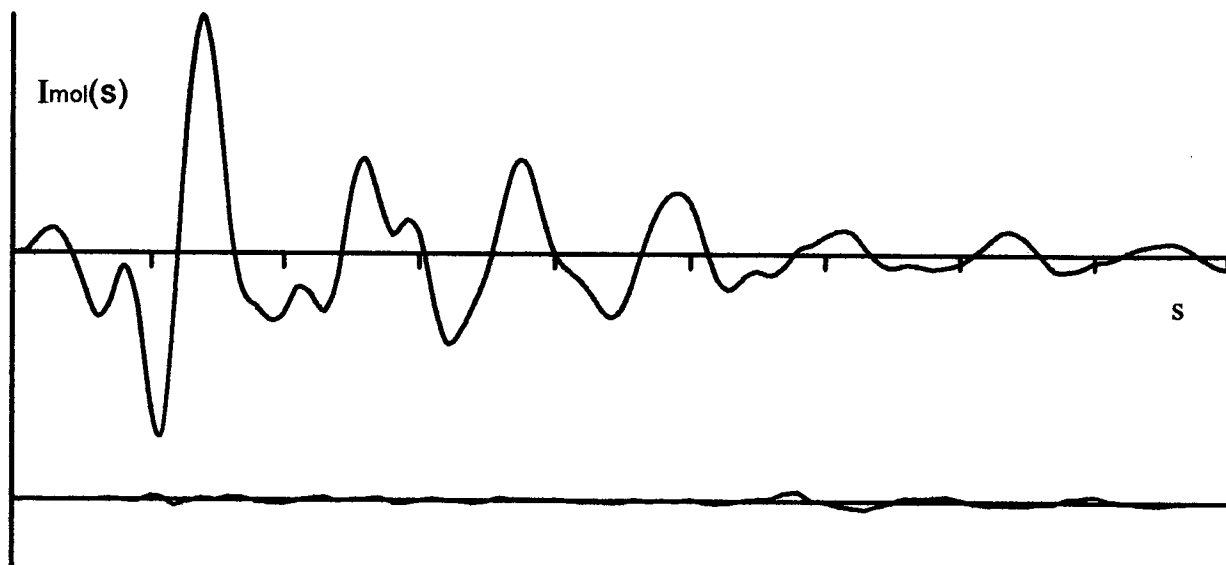
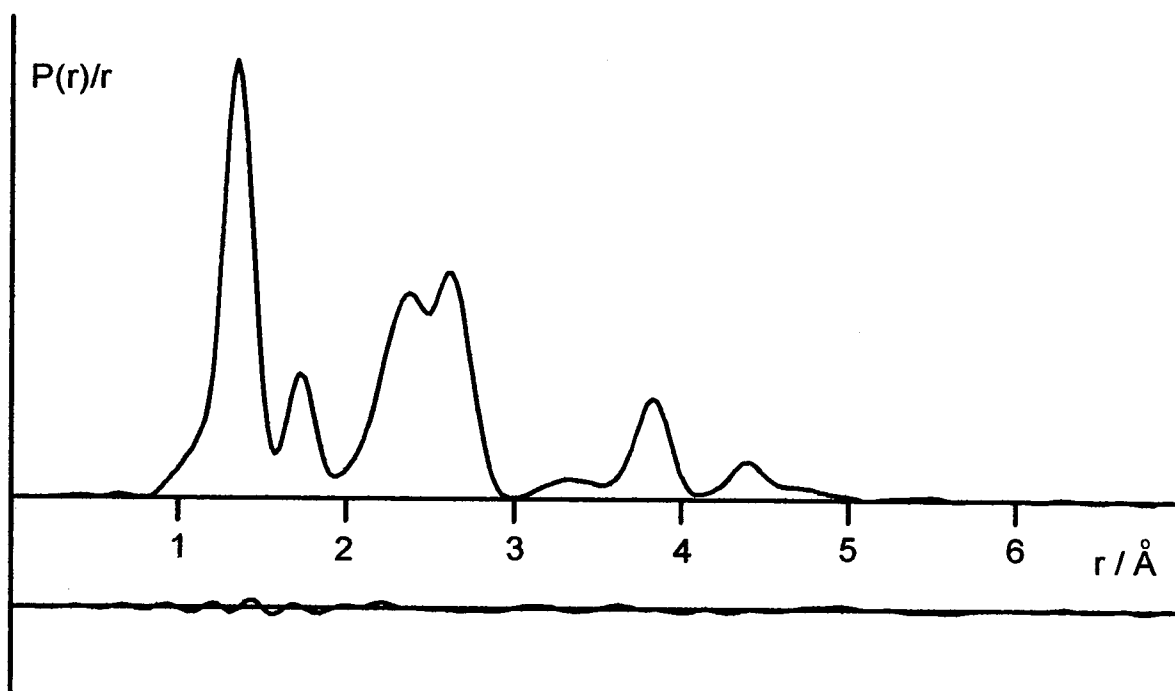


Figure 8.5

The observed and final difference
radial distribution curve, $P(r)/r$, for 2-chloropyrimidine



3,6-dichloropyridazine

Analysis of the LCNMR spectrum

The ^1H spectrum of 3,6-dichloropyridazine was recorded at room temperature using the liquid crystal solvent E5. Indirect coupling constants were obtained from a sample dissolved in CD_2Cl_2 . Because the molecule has just two hydrogen atoms (which are magnetically equivalent) the resulting spectrum consists of a doublet of splitting $3D_{\text{HH}}$. This immediately allows one of the orientation parameters to be determined (as the internuclear vector lies parallel to the y-axis). These two peaks were very broad indeed due to the quadrupolar ^{14}N nuclei, which are close to the hydrogen nuclei. On first examination it seemed as if the satellite subspectra were first-order but it soon became apparent that there were fewer peaks than expected. In fact for each ^{13}C nucleus only two satellites were observed rather than the doublet of doublets that might be predicted. On using **lcsim** and **sliquor** to analyse the spectrum it was found that the satellite subspectra were in fact second-order and that two of the expected four lines for each ^{13}C were of such low intensity that they could not be detected. Such spectra are often known as "deceptively simple" and are quite common for spin systems of this kind⁸. As a consequence of this effect it is impossible to obtain values for both D_{CH} coupling constants with any degree of accuracy. Instead it is the sum of the two couplings which is well determined. A similar effect would be expected in the ^{15}N satellite subspectrum but, unfortunately, the ^{15}N satellites are lost under the broad peaks of the parent spectrum. Therefore, from a total of seven possible direct coupling constants, only three LCNMR data have been obtained

(D_{HH} , $D_{4,8}+D_{4,9}$ and $D_{3,8}+D_{3,9}$). This effectively amounts to only one piece of structural information as there are two independent orientation parameters which must also be determined.

Structure refinement of 3,6-dichloropyridazine

The molecular geometry can be described in terms of nine independent structural parameters (see table 8.6). As in the previous examples, initial values were chosen to give bond lengths typical of their types and angles were assumed to be 120° . Parallel and perpendicular amplitudes were calculated for each of the 25 distinct internuclear distances within the molecule.

From the start of the refinement it was apparent that the ED data was not of a particularly high quality. The initial refinements converged with an R_G of 13.6%, much higher than that achieved for any of the other molecules studied. The most noteworthy structural change was a decrease in the N(2)-C(3)-Cl(7) angle to $114.27(41)^\circ$ which corresponds to a deviation of almost four degrees towards the nitrogen atoms. Attempts to refine some of the vibrational amplitudes were not entirely successful and many of the amplitudes had to be fixed for the remainder of the refinement. It seems likely that this is a consequence of the poor quality of the ED data.

Including the four independent rotation constants in the refinement led to a unrealistic change in the geometry. A reasonable fit to the MW data was only achieved when the ring bond difference parameters swapped signs so that the C-N bonds became longer than the C-C bonds. This was an

important clue as to where the refinement was going wrong. If the geometry of the ring is considered to be fixed it can be seen that variation of the N-C-Cl angle can produce two structures which are almost equivalent from the point of view of the electron diffraction experiment. In one case the C-Cl bond is bent towards the nitrogen atoms and in the other it is bent towards the C-H bonds. By chance, the initial refinements had produced the former structure but the MW data showed this to be incorrect. To investigate the other possible structure, a second ED refinement was made. This time, however, the N-C-Cl angle was given a starting value of 123° and, as predicted, the structure refined equally well but with a deviation of the chlorine atoms towards the adjacent hydrogen atoms. Once more the MW data were introduced and on this attempt they proved to be compatible with the new structure (see table 8.7).

This is a good example of additional data being used to overcome the limitations of ED analysis. In the final refinement using ED and MW data all the heavy atom parameters were refined, apart from the difference parameter $r_{\text{CN}} - r_{\text{NN}}$ which was fixed at zero. Attempts to refine this parameter resulted in a C-N bond length which was far too short, $1.308(7) \text{ \AA}$, and a correspondingly lengthened N-N bond length.

Although only three LCNMR data were available, these were included in the refinement and the two independent orientation parameters were allowed to refine. Because only one piece of structural information is attainable from the LCNMR data, the C-H bond length was fixed at 1.085 \AA and only the angle was refined. Perhaps surprisingly, this refinement was successful, giving a C-C-H angle of $122.63(35)^\circ$. A further refinement was carried out

with rC-H fixed this time at 1.080 Å. This produced no significant change in the ring structure or in the C-C-H angle which now refined to 122.64(35)°.

The final refinement is based on ED data recorded at two camera distances, 4 rotation constants and 3 LCNMR observations. Table 8.8 lists the final values obtained for the amplitudes of vibration, although most of these were fixed throughout the refinement. Table 8.9 shows the least squares correlation matrix calculated during the final refinement. The final molecular geometry is shown in figure 8.6 and the corresponding molecular scattering intensity and radial distribution curves are shown in tables 8.7 and 8.8 respectively.

Table 8.6 - Final parameters: 3,6-dichloropyridazine

Parameters [†]	ED	ED+MW	All data
Structural			
(r _{1,2} +r _{2,3} +r _{3,4} +r _{4,5})/4	1.363(2)	1.365(2)	1.365(2)
(r _{3,4} +r _{4,5})/2 - (r _{1,2} +r _{2,3})/2	0.055(9)	0.051(11)	0.048(12)
r _{3,4} - r _{4,5}	0.0 (fixed)	-0.025(13)	-0.026(14)
r _{2,3} - r _{1,2}	0.0 (fixed)	0.0 (fixed)	0.0 (fixed)
r _{C-Cl}	1.725(2)	1.723(2)	1.723(2)
r _{C-H}	1.085 (fixed)	1.085 (fixed)	1.085 (fixed)
angle N(1)-N(2)-C(3)	119.09(28)	118.76(32)	118.70(34)
angle N(2)-C(3)-Cl(7)	123.57(57)	122.96(70)	122.84(72)
angle C(3)-C(4)-H(8)	122.0 (fixed)	122.0 (fixed)	122.63(35)
Orientational			
S _{VV} × 100 (EBBA)	-	-	-33.48(34)
S _{ZZ} × 100 (EBBA)	-	-	9.48(10)
Dependent			
r _{N-N}	1.336(5)	1.339(6)	1.341(6)
r _{C-N}	1.336(5)	1.339(6)	1.341(6)
r _{C(3)-C(4)}	1.391(5)	1.378(8)	1.375(8)
r _{C(4)-C(5)}	1.391(5)	1.403(9)	1.401(10)
angle N(2)-C(3)-C(4)	124.36(31)	124.85(33)	124.82(34)
angle C(3)-C(4)-C(5)	116.55(23)	116.39(28)	116.49(30)

[†] all distances are given in Angstroms; all angles are given in degrees.

Table 8.7 - Rotation constants⁵ (MHz) and direct dipolar coupling constants (Hz) used in the structural analysis of 3,6-dichloropyridazine

Constant	Observed	Corrected	Calculated	Uncertainty	Difference
B principal	709.72	709.640	709.641	0.040	-0.001
C principal	634.02	633.977	633.983	0.022	-0.006
B 37Cl(7)	692.30	692.220	692.033	0.440	0.187
C 37Cl(7)	620.60	620.560	619.887	0.220	0.673
D _{8,9}	2497.28	2527.20	2527.17	3.00	0.03
D _{4,8} + D _{4,9}	1315.28	1191.30	1192.28	16.00	-0.98
D _{3,8} + D _{3,9}	-156.46	-157.24	-157.27	1.20	0.03

Table 8.8 - Amplitudes of vibration of 3,6-dichloropyridazine

Number	Atom Pair	u / Å
1	N1 - N2	0.044 (fixed)
2	N2 - C3	0.045 (fixed)
3	C3 - C4	0.045 (fixed)
4	C4 - C5	0.045 (fixed)
5	C3 - CL7	0.033(5)
6	C4 - H8	0.077 (fixed)
7	N1...C3	0.070 (fixed)
8	N1...C4	0.082(19)
9	N1...C5	0.063 (fixed)
10	N1...CL7	0.082 (fixed)
11	N1...H8	0.105 (fixed)
12	N1...H9	0.099 (fixed)
13	N1...CL10	0.079 (fixed)
14	C3...C5	0.060 (fixed)
15	C3...C6	0.074 (tied to u8)
16	C3...H8	0.099 (fixed)
17	C3...H9	0.098 (fixed)
18	C3...CL10	0.080 (fixed)
19	C4...CL7	0.078 (fixed)
20	C4...H9	0.099 (fixed)
21	C4...CL10	0.077 (fixed)
22	CL7...H8	0.148 (fixed)
23	CL7...H9	0.120 (fixed)
24	CL7...CL10	0.094(6)
25	H8...H9	0.157 (fixed)

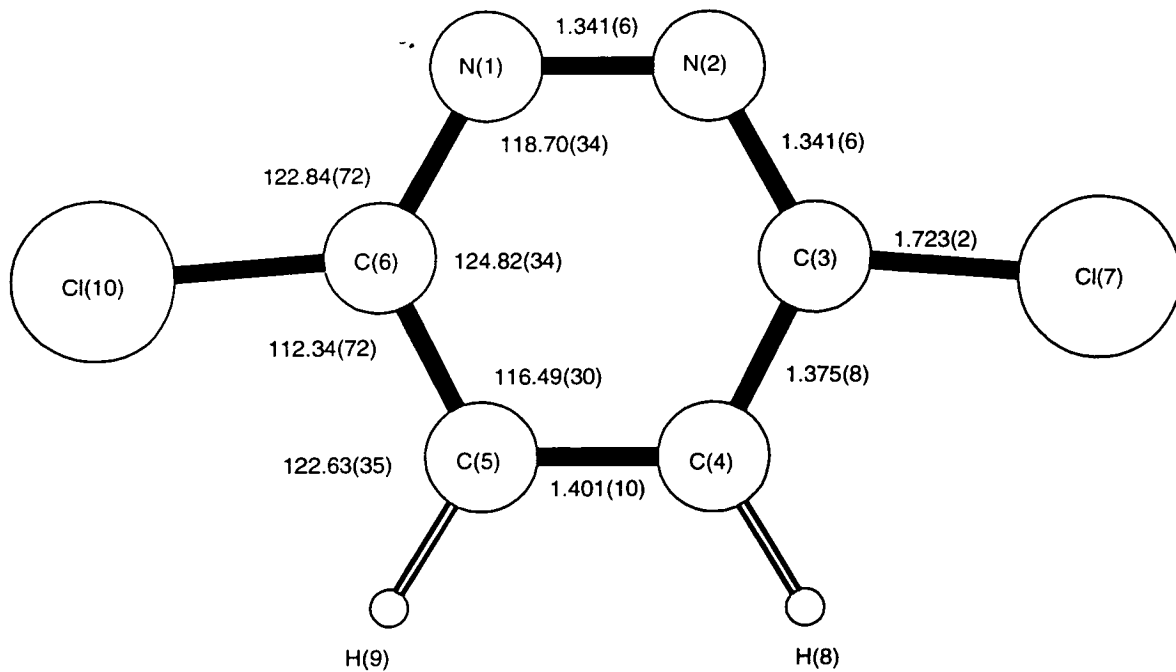
Table 8.9

Least squares correlation matrix (x100) for 3,6-dichloropyridazine.
All elements with absolute value <50 have been omitted.

	P3	P7	P8	P9	P10	P11	u8	k2
P1	-91		-52	59	-69		-53	
P2		87	64	68	-64	90		59
P3		51	57	-60	64		55	
P7			87			73		55
P8								51
P9					-87	67		
P10						-73		
P11								52
u5								60

Parameters (P) are in the order listed in table 8.6 ; amplitudes (u) are in the order listed in table 8.8 ; scale factors (k) are in the order listed in table 8.1.

Figure 8.6 - The molecular structure of 3,6-dichloropyridazine



All distances are given in Ångströms
All angles are given in degrees

Figure 8.7

The observed and final weighted difference
molecular scattering intensity curves for 3,6-dichloropyridazine

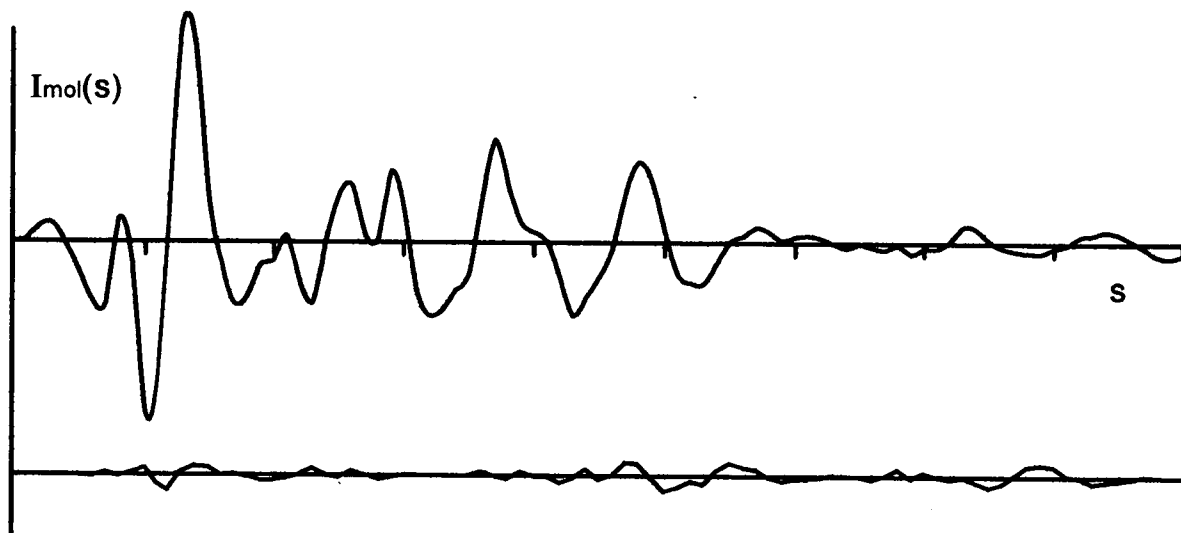
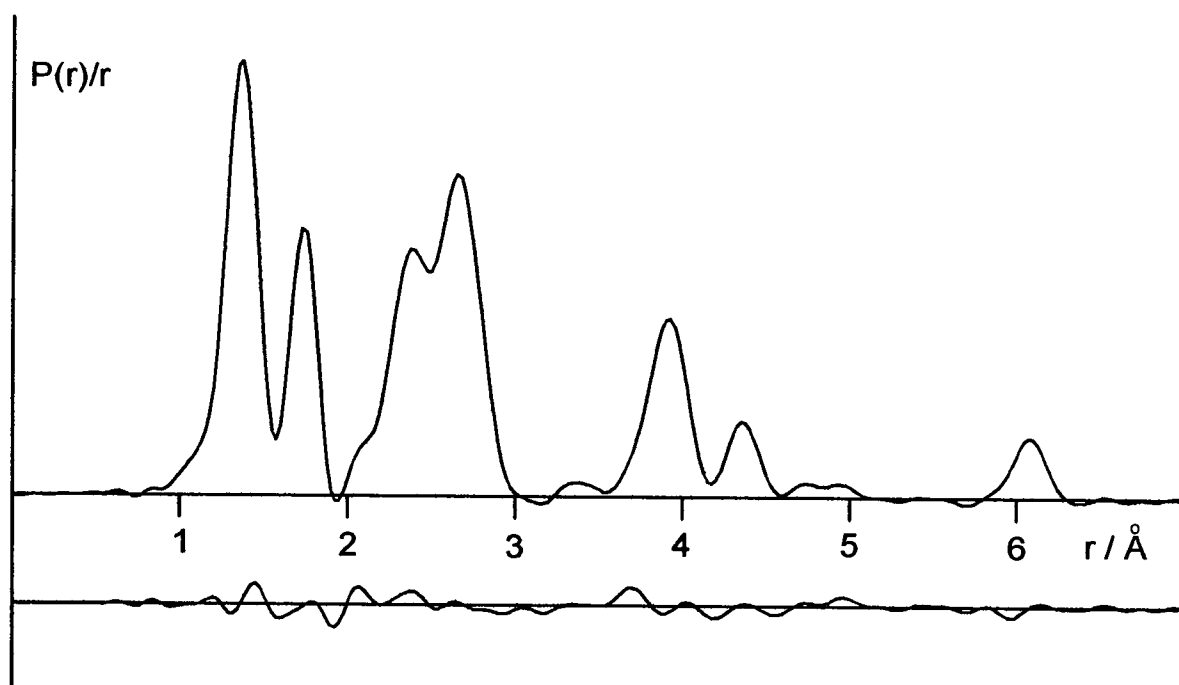


Figure 8.8

The observed and final difference
radial distribution curve, $P(r)/r$, for 3,6-dichloropyridazine



2,6-dichloropyrazine

An LCNMR spectrum of 2,6-dichloropyrazine was obtained using the liquid crystal solvent E5. As in the case of 3,6-dichloropyridazine, this consisted of a doublet parent spectrum with very broad lines. Unfortunately, no satellites could be found in this spectrum, which had a very high signal to noise ratio. In any case, it seems likely that some of the satellite peaks would be lost under the broad parent peaks; only the particularly large orientation parameters of the 3,6-dichloropyridazine experiment allowed some of the satellites to be found. As only one coupling constant could be measured from the spectrum (D_{HH}) and two orientation parameters are required for a molecule of C_{2v} symmetry, no structural information could be obtained. It is unfortunate that this is also the only molecule in this study for which no MW data were available. Nevertheless, it was decided to carry out a structural refinement using the ED data only.

The molecular geometry can be described in terms of nine independent structural parameters (see table 8.10). Initial values were chosen to give bond lengths typical of their type and all angles were assumed to be 120° . Parallel and perpendicular amplitudes were calculated for each of the 25 distinct internuclear distances within the molecule.

The data was clearly of a better quality than that obtained for 3,6-dichloropyridazine and the structure quickly settled producing a fit of $R_G=9.4\%$. The internal ring angles increased at the chlorine substituted carbon atoms and decreased at the nitrogen atoms. As might be expected with only ED data available, the difference between the two C-N bond lengths could not be refined, nor could the hydrogen atom positions.

The principal vibrational amplitudes were then allowed to refine which improved the fit to the ED data slightly. However, problems were experienced when attempting to refine the amplitudes associated with the N...N distance as it decreased dramatically to an unrealistic value of 0.029 Å. It was therefore fixed at its initial value of 0.066 Å for the remainder of the refinement.

The final refinement was carried out on an r_{α}° basis using ED data recorded at two camera distances. The values obtained for the amplitudes of vibration can be found in table 8.11 and the least squares correlation matrix is shown in table 8.12. The final molecular geometry is shown in figure 8.9 and the corresponding molecular scattering intensity and radial distribution curves are shown in figures 8.10 and 8.11 respectively.

Table 8.10 - Final parameters: 2,6-dichloropyrazine

Parameters [†]	Final values (ED)
Independent	
(r _{1,2} +r _{2,3} +r _{3,4})/3	1.356(1)
(mean r _{CN}) - r _{CC}	-0.076(5)
r _{1,2} - r _{3,4}	0.0 (fixed)
r _{C-Cl}	1.731(2)
r _{C-H}	1.080 (fixed)
angle C(6)-N(1)-C(2)	114.32(30)
angle N(1)-C(2)-C(3)	124.10(27)
angle N(1)-C(2)-Cl(7)	116.81(24)
angle C(2)-C(3)-H(8)	120.0 (fixed)
Dependent	
r _{C-C}	1.407(4)
r _{C(1)-N(2)}	1.331(1)
r _{C(3)-N(4)}	1.331(1)
angle C(2)-C(3)-N(4)	119.01(37)
angle C(3)-N(4)-C(5)	119.44(47)

[†] all distances are given in Angstroms; all angles are given in degrees.

Table 8.11 - Amplitudes of vibration of 2,6-dichloropyrazine

Number	Atom Pair	u / Å
1	N1 - C2	0.037 (tied to u2)
2	C2 - C3	0.039(3)
3	C3 - N4	0.037 (tied to u2)
4	C2 - CL7	0.047(2)
5	C3 - H8	0.077 (fixed)
6	N1...C3	0.039(6)
7	N1...N4	0.066 (fixed)
8	N1...CL7	0.073(4)
9	N1...H8	0.096 (fixed)
10	C2...N4	0.038 (tied to u6)
11	C2...C5	0.065 (fixed)
12	C2...C6	0.039 (tied to u6)
13	C2...H8	0.098 (fixed)
14	C2...H9	0.097 (fixed)
15	C2...CL10	0.082(4)
16	C3...C5	0.040 (tied to u6)
17	C3...CL7	0.074 (tied to u8)
18	C3...H9	0.098 (fixed)
19	C3...CL10	0.079(5)
20	N4...CL7	0.081 (tied to u15)
21	N4...H8	0.099 (fixed)
22	CL7...H8	0.140 (fixed)
23	CL7...H9	0.102 (fixed)
24	CL7...CL10	0.112(6)
25	H8...H9	0.133 (fixed)

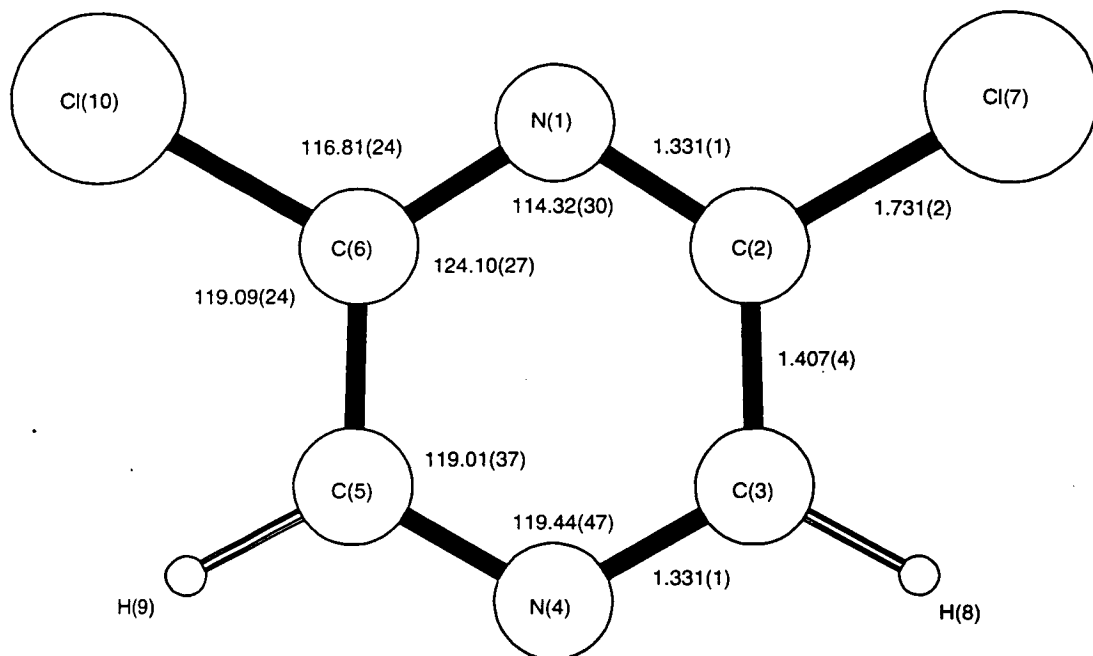
Table 8.12

Least squares correlation matrix (x100) for 2,6-dichloropyrazine.
All elements with absolute value <50 have been omitted.

	P2	P6	P7	u2	u6	k1
P1	-57					
P2				68		
P4		-52				
P6			-91		52	
P7					-57	
u2						58
u4						68

Parameters (P) are in the order listed in table 8.10 ; amplitudes (u) are in the order listed in table 8.11 ; scale factors (k) are in the order listed in table 8.1.

Figure 8.9 - The molecular structure of 2,6-dichloropyrazine



All distances are given in Ångströms

All angles are given in degrees

Figure 8.10

The observed and final weighted difference
molecular scattering intensity curves for 2,6-dichloropyrazine

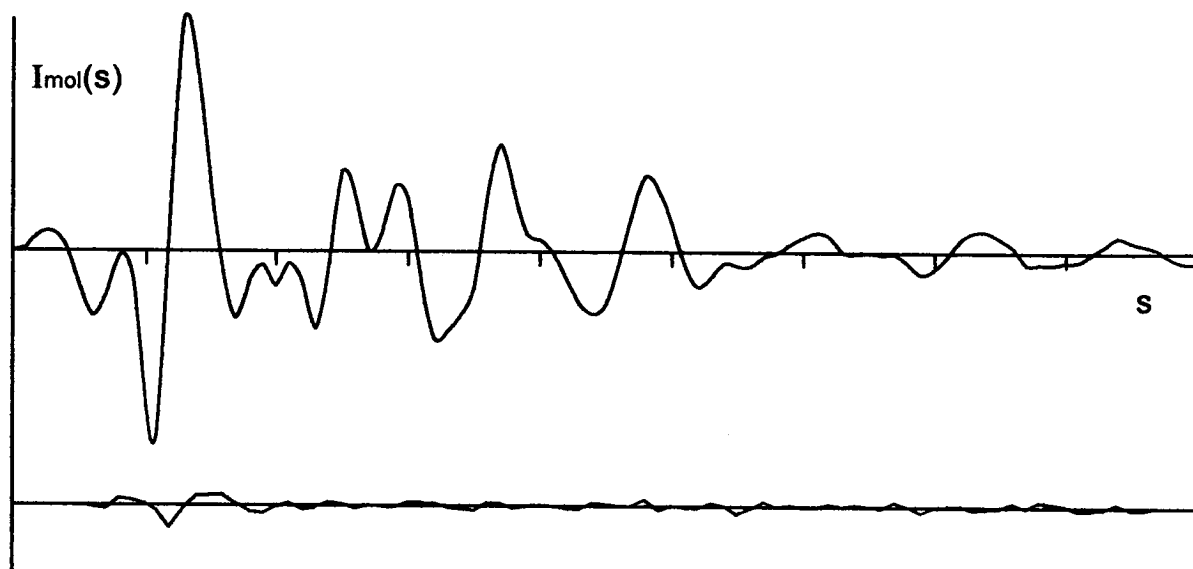
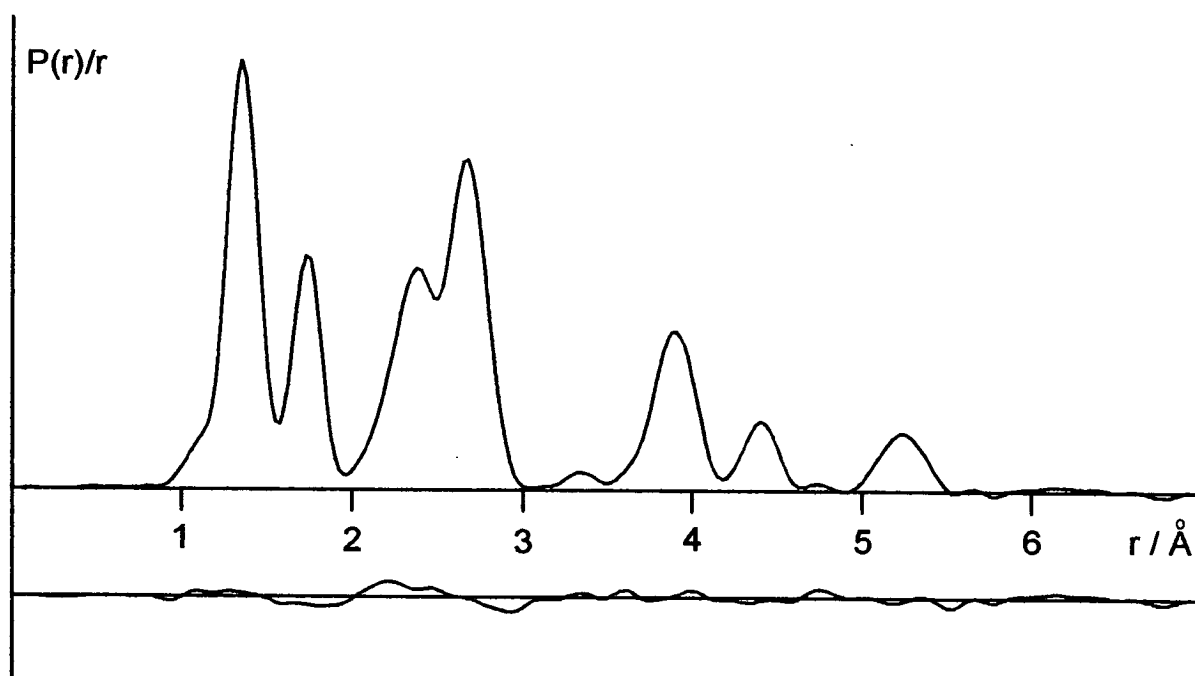


Figure 8.11

The observed and final difference
radial distribution curve, $P(r)/r$, for 2,6-dichloropyrazine



Conclusions

Initial doubts as to the suitability of these molecules for structural determination by combined analysis of ED, MW and LCNMR data, were to some extent warranted. In particular it proved difficult to obtain sufficient LCNMR data to allow the complete determination of the hydrogen atom positions. This limitation is largely due to the presence of quadrupolar ^{14}N nuclei which broadens the peaks in the spectrum and makes the assignment of satellites particularly difficult. Nonetheless, it has been shown that, even with a minimal amount of LCNMR data, some extra structural information can be obtained. It should be remembered that when few LCNMR data are used the hydrogen atom positions obtained are largely uncorroborated and any errors in the LCNMR analysis are likely to go undetected. Once more, the use of data obtained from a number of different solvents would go some way to overcoming this uncertainty but none of these compounds were readily soluble in most of the liquid crystal solvents available.

As with the fluorobenzenes, the distortions of the internal ring angles of these molecules can be predicted by the superposition of effects from the substituent atoms and heteroatoms. The effect of the chlorine substituent is similar to that of fluorine; the *ipso* angle is increased by approximately 1.7 degrees and the adjacent angles are reduced by about half that amount (these values are taken from the structure of chlorobenzene³). The presence of a nitrogen heteroatom in the ring has a similar but opposite effect, with the angle at the nitrogen decreasing by close to 3 degrees and the adjacent angles increasing by about 4 degrees. In this case however,

the effect extends to the *meta* and *para* angles which both decrease by about 1.5 degrees (these values are taken from the structure of pyridine⁹).

Table 8.13 shows the predicted ring angles for each of the three molecules studied, as well as the actual values obtained from the structure refinements. It can be seen immediately that the agreement is particularly good for the first two molecules, but that the angle at N(4) of 2,6-dichloropyrazine is quite different from the predicted value. This suggests that either the observed structure is incorrect or the superposition principle is breaking down; it not easy to say which without further data.

Table 8.13 - Prediction of internal ring angles by superposition of effects

Molecule	Angle	Predicted	Observed
2-chloropyrimidine	N(1)-C(2)-N(3)	129.3	127.97(19)
	C(2)-N(3)-C(4)	114.5	116.07(24)
	N(3)-C(4)-C(5)	122.4	121.78(23)
	C(4)-C(5)-C(6)	116.8	116.31(24)
3,6-dichloropyridazine	N(1)-N(2)-C(3)	120.0	118.70(34)
	N(2)-C(3)-C(4)	123.8	124.82(34)
	C(3)-C(4)-C(5)	116.2	116.49(30)
2,6-dichloropyrazine	C(6)-N(1)-C(2)	113.5	114.32(30)
	N(1)-C(2)-C(3)	124.2	124.10(27)
	C(2)-C(3)-N(4)	121.2	119.01(37)
	C(3)-N(4)-C(5)	115.7	119.44(47)

It is not possible to draw any firm conclusions about the effects of substitution on the ring bond lengths because, in all but the first case, some of the difference parameters were fixed during the refinement. On the whole, however, the results are generally consistent with a shortening of the adjacent bonds on chlorine substitution. The origin of this effect is explained in the preceding chapter with regard to fluorine substitution.

Another structural feature worth noting is the extraordinary deviation of the chlorine atoms in 3,6-dichloropyrazine away from the adjacent nitrogen atoms (by more than 5 degrees). This can easily be explained in terms of repulsion between the lone pairs of the chlorine and nitrogen atoms and is consistent with the deviation of the C-Cl bond in 2-chloropyridine¹⁰ which moves by 3.9 degrees away from the nitrogen. However, the structure of 2,6-dichloropyrazine shows quite the opposite effect which can not easily be explained. The fact that the geometry of 2,6-dichloropyrazine is inconsistent both in the deviation of the C-Cl bond and the predicted internal ring angles is perhaps an indication that the structure has been incorrectly refined. However, an attempt was made to refine the structure with the N-C-Cl angle starting at 124° but it quickly returned to the original value of 116.8°. It is likely that MW data, should they become available, would help to either verify or disprove this structure.

References

¹S.Cradock, C.Purves & D.W.H.Rankin, *J.Mol.Struct.*, 220, (1990), 193

²P.B.Liescheski & D.W.H.Rankin, *J.Mol.Struct.*, 196, (1989), 1

³S.Cradock, J.M.Muir & D.W.H.Rankin, *J.Mol.Struct.*, 220, (1990), 205

⁴S.Cradock, P.B.Liescheski & D.W.H.Rankin, *J.Magn.Reson.*, 91, (1991), 316

⁵R.K.Bohn, University of Connecticut, Storres, private communication

⁶E.M.Brown, Honours project, University of Edinburgh, 1990

⁷P.Diehl, T.Bjorholm & H.Bösiger, *J.Magn.Reson.*, 42, (1981), 390

⁸J.W.Emsley & J.C.Lindon, "NMR Spectroscopy using Liquid Crystal Solvents",
Permagon Press, Oxford, (1975), 43

⁹W.Pyckhout, N.Horemans, C. van Alsenoy, H.J.Geise & D.W.H.Rankin, *J.Mol.Struct.*,
156, (1987), 315

¹⁰C.A.Brookman, Ph.D.Thesis, University of Edinburgh, (1993)

Chapter 9

Final Conclusions and Further Work

Two of the main objectives of this work were to improve the programs used in the analysis of LCNMR spectra and to investigate the use of the molecular mechanics program **MM3** to calculate the vibrational corrections necessary if spectroscopic data are to be used to supplement electron diffraction data. The success of the former aim has been demonstrated in that several spectra which had previously remained unsolved have since been reanalysed and successfully assigned using the modified programs.

The progress made with **MM3** is more difficult to assess and can only be truly tested by its continued application to a number of different molecules. In particular, the methods used to estimate the uncertainties associated with direct dipolar coupling constants may well be refined as more molecules are studied. The use of LCNMR data obtained for one molecule using a number of different liquid crystal solvents can also provide valuable information about the compatibility of such data with gas-phase data. It has been seen that often the results obtained using data from one solvent can be quite different to those obtained using data from another (see chapter 7). The importance of using as many uncorrelated data as are available can not be over stressed. This should be made easier when the limitation of the current version of **ed92** to a maximum of 50 non-ED data is removed. At that time it would be useful to reanalyse the difluorobenzenes using all the available data.

A considerable advantage of **MM3** over conventional normal co-ordinate analysis programs is that it is just as well suited to molecules with low symmetry as it is to highly symmetric molecules. The lack of suitable force fields proved to be a major obstacle in the structural analysis of

asymmetrically substituted pyridines and diazines, recently carried out at Edinburgh¹. It is hoped that **MM3** could be used in their reanalysis, particularly now that some of the problems involved in obtaining LCNMR data have been overcome (see chapter 5 for the analysis of the LCNMR spectrum of 2-chloropyridine).

A major limiting factor in the LCNMR analysis of nitrogen containing heteroaromatic compounds seems to be the line broadening effect of the quadrupolar ¹⁴N nuclei. Although in principle this might be overcome by using decoupling experiments, this is not possible with the instrumentation available at Edinburgh. An alternative approach would be the use of ¹⁵N labelled compounds. This would have the added advantage of removing the need to search for ¹⁵N satellites which, at a natural abundance of just 0.37%, are often extremely difficult to locate.

The structural determination of the difluorobenzenes shows how powerful the technique of combined analysis can be in the best of circumstances. However, it is perhaps even more encouraging to note that even the very limited data available in the analysis of 2-chloropyrimidine and 3,6-dichloropyridazine significantly improved the geometry determination when compared to the analysis using electron diffraction data alone.

Reference

¹C.A.Brookman, Ph.D. Thesis, The University of Edinburgh, 1993

Appendix A

FORTRAN77 source code

A.I Example ED92 Model Subroutine

```
SUBROUTINE COORD(X,Y,Z)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
REAL*4 RM
COMMON/M80/PAR(30),R(100),RM(100),V(100)
DIMENSION X(100),Y(100),Z(100)
C Model for 2 chloropyrimidine (C4H3CLN2)
C Atoms 1 and 3 N ; 2,4,5 and 6 C ; 7 Cl ; 8-10 H
C Molecule in YZ plane
C Z-axis is C2 axis ; Y-axis through both Nitrogens
C R1 means N1-C2 OR C2-N3
C R2 means N3-C4 OR N1-C6
C R3 means C4-C5 OR C6-C5
C R4 means C4-H8 OR C6-H10
C R5 means C5-H9
C PAR(1)=(R1+R2+R3)/3
C PAR(2)=(R1+R2)/2-R3
C PAR(3)=R1-R2
C PAR(4) means C2-CL7
C PAR(5)=(2*R4+R5)/3
C PAR(6)=R4-R5
C PAR(7) means N1C2N3
C PAR(8) means C2N3C4 or C6N1C2
C PAR(9) means C5C4H8 or C5C6H10
RAD=3.141592/180.0
A1=PAR(7)*RAD
A2=PAR(8)*RAD
A3=PAR(9)*RAD
R1=(6.0*PAR(1)+3.0*PAR(3)+2.0*PAR(2))/6.0
R2=(6.0*PAR(1)+2.0*PAR(2)-3.0*PAR(3))/6.0
R3=(3.0*PAR(1)-2.0*PAR(2))/3.0
R4=PAR(6)/3.0+PAR(5)
R5=(3.0*PAR(5)-2.0*PAR(6))/3.0
Y(3)=R1*DSIN(A1/2.0)
Z(3)=0.0
Y(1)=-Y(3)
Z(1)=0.0
Y(2)=0.0
Z(2)=R1*DCOS(A1/2.0)
Y(7)=0.0
Z(7)=Z(2)+PAR(4)
A4=A2-(90.0*RAD+DASIN(Z(2)/R1))
Y(4)=Y(3)+R2*DSIN(A4)
Z(4)=-R2*DCOS(A4)
Y(6)=-Y(4)
Z(6)=Z(4)
Y(5)=0.0
Z(5)=Z(4)-DSQRT(R3**2.0-Y(4)**2.0)
Y(9)=0.0
Z(9)=Z(5)-R5
A5=A3+DARCOS(Y(4)/R3)-90.0*RAD
Y(8)=Y(4)+R4*DSIN(A5)
Z(8)=Z(4)-R4*DCOS(A5)
Y(10)=-Y(8)
Z(10)=Z(8)
```

```
c      *** Determine dependent parameters ***
      CALL ED92QC(X,Y,Z,3,4,5,0,P12)
      CALL ED92QC(X,Y,Z,4,5,6,0,P13)
      PAR(12)=P12
      PAR(13)=P13
      PAR(14)=R1
      PAR(15)=R2
      PAR(16)=R3
      PAR(17)=R4
      PAR(18)=R5
      RETURN
      END
```

A.II Example ED92 Extra Data Subroutine

```
SUBROUTINE EXTRA(X,Y,Z,E)
  IMPLICIT DOUBLE PRECISION (A-H,O-Z)
  COMMON/M80/ PAR(30),RT(250)
  DIMENSION E(100),X(100),Y(100),Z(100),S(6),G(50),AM(50),RR(3)

c   *** Calculation of D couplings ***
c   :
c   *** Array G() contains magic numbers ***
  GC=87.1659
  GH=346.574
  G(1)=0.0
  G(2)=GC
  G(3)=0.0
  G(4)=GC
  G(5)=GC
  G(6)=GC
  G(8)=GH
  G(9)=GH
  G(10)=GH
c   *** Array S() contains orientation tensor elements ***
  S(4)=PAR(10)/100.
  S(6)=PAR(11)/100.
  S(1)=- (S(4)+S(6))
  S(2)=0
  S(5)=0
  S(3)=0
c   *** Subroutine ED92XN calculates D couplings ***
  E(6)=ED92XN(X,Y,Z,G,S,4,8)
  E(7)=ED92XN(X,Y,Z,G,S,2,8)
  E(3)=ED92XN(X,Y,Z,G,S,2,9)
  E(4)=ED92XN(X,Y,Z,G,S,4,9)
  E(5)=ED92XN(X,Y,Z,G,S,4,10)
  E(1)=ED92XN(X,Y,Z,G,S,8,9)
  E(2)=ED92XN(X,Y,Z,G,S,8,10)

c   *** Calculation of rotation constants ***
c   *** Array AM() contains atomic masses ***
  NA=10
  AMH=1.007825
  AMC=12.
  AMN=14.0067
  AMCL=34.968853

  DO 10 I=2,6
10  AM(I)=AMC
    AM(1)=AMN
    AM(3)=AMN
    AM(7)=AMCL

  DO 20 I=8,10
20  AM(I)=AMH
```

```
c    *** Subroutine ED92XM returns rotation constants in RR() ***  
call ED92XM(AM,X,Y,Z,NA,RR)  
E(8)=RR(1)  
E(9)=RR(2)  
E(10)=RR(3)
```

```
c    *** Repeat for 37Cl isotope ***  
AM(7)=36.9658  
call ED92XM(AM,X,Y,Z,NA,RR)  
E(11)=RR(1)  
E(12)=RR(2)  
E(13)=RR(3)
```

```
RETURN  
END
```

A.III Modifications to LCSIM

INPUT - makelcsim source code

```
dimension j(7,7),w(7),type(7),itype(7),kh(7),mult(7),index(7)
character*30 filename
character*40 gampfile
real nn,x,y,z,type,j,w,maxf,minf,mini
integer itype,kh,mult,ixyz
c Multiplicities of nuclei with 0 for dummy entries
data (mult(i),i=1,7)/0,2,2,3,2,2,2/
data (kh(i),i=1,7)/0,0,0,0,0,0,0/
print*,'makelcsim by E.Brown 4/9/92'
print*,'-----'
print*,'This program allows you to interactively create an input'
print*,'file for the LCNMR interpretation program - LCSIM.'
print*,' '
print*,'Input is in free format unless specified.'
print*,' '
print*,' '
write(6,100)
100 format(' Name of file to create: ',)$)
read(5,200)filename
200 format(a30)
open(1,file=filename)
call space
15 write(6,300)
300 format(' Number of atoms (spinning nuclei only): ',)$)
read(5,*)NN
if (NN.lt.2.or.NN.gt.7.or.NN.ne.INT(NN)) then
  print*,'Sorry, integers between 2 and 7 only.'
  go to 15
endif
write(1,400)NN
400 format(F7.4)
call space
print*,'Input co-ordinates ,types and index numbers of atoms.'
print*,'Types are as follows:-'
print*,' '
print*,'H ..... 1          15N .... 4'
print*,'C ..... 2          19F .... 5'
print*,'14N .... 3          29Si ... 6'
print*,' '
print*,'Order is X Y Z Type Index'
do 10 i=1,nn
20 write(6,1000)
read(5,*)x,y,z,type(i),index(i)
if (type(i).lt.1.or.type(i).gt.6) then
  print*,'ERROR - retype line'
  goto 20
else
  write(1,500)z,x,y,type(i),index(i)
  itype(i)=type(i)
endif
```

```

10  continue
500  format(4F7.3,I2)
     call space
600  format(F10.3)
700  format(7I1)
     print*,'Extent of calculations.'
     print*,' '
     write(6,800)
800  format(' Minimum frequency: ',)$
     read(5,*)minf
     write(6,810)
810  format(' Maximum frequency: ',)$
     read(5,*)maxf
     write(6,820)
820  format(' Minimum intensity: ',)$
     read(5,*)mini
     write(1,600)minf
     write(1,600)maxf
     write(1,600)mini
     call space
     write(1,700)(mult(itype(i)+1),i=1,7)
     do 345 i=1,nn
345  kh(i)=1
     write(1,700)(kh(i),i=1,nn)
     call space
     print*,'Input indirect (J) coupling constants.'
     do 30 i=1,nn-1
     do 40 ii=i+1,nn
     write(6,900)index(i),index(ii)
     read(5,*)j(i,ii)
40  continue
     write(1,950)(j(i,ii),ii=i+1,nn)
30  continue
900  format(' J(',I2,',',I2,') = ',)$
950  format(F10.3)
     call space
     print*,'Input chemical shifts (Hz)'
     do 50 i=1,nn
     write(6,960)index(i)
     read(5,*)w(i)
50  continue
     write(1,970)(w(i),i=1,nn)
960  format(' Atom',i2,' :',)$
970  format(F15.4)
     call space
1000 format(':',)$
     WRITE(6,6220)
6220 FORMAT(' The principal rotation axes a,b,c may be related to the',
# ' axes x,y,z used for',/, ' orientation in the direct coupling',
# ' experiment in one of the following six ways;',/, ' specify which'
# , ' by typing in the integer (1-6).')
     if (iflg70.eq.0) print*,'N.B. 0 implies no vibrational corrections
# to be made.'
     WRITE(6,6221)
6221 FORMAT('      1 2 3 4 5 6',/, ' A  x  x  y  y  z  z',/,
# ' B  y  z  x  z  x  y',/, ' C  z  y  z  x  y  x')
     write(6,6329)
6329 format('No.: ',)$

```

```

        READ*,xyzabc
        ixyz=xyzabc
        write(1,6330)ixyz
6330  format(il)
        print*,' '
        if (xyzabc.ne.0) then
            write(6,6222)
6222  format('File containing vibrational info. (e.g. dcXYZ) : ', $)
            read(5,6223)gampfile
            write(1,6223)gampfile
6223  format(a40)
        endif
        print*, 'End of input'
        print*, 'Written to file ', filename
        end

subroutine space
print*,' '
print*,' '
return
end

```

OUTPUT - spectrum plotting routines

```

subroutine xdraw(fmin, fmax, name, stnsor, iflgop, intmin, intmax, inob)
character*20 name
character*1 ans
integer iflgop
c **** inob=1 :don't call gclear in xplot ie no blank page.
inob=1
iflgop=1
print*, 'Please select device to be used and then exit'
print*, 'NB. To use gpchem select either php7550a4 or php7550a3'
print*, '    To use pschem (WHICH COSTS MONEY !!) select hposta4'
print *,' '
print*, 'Eg. type select php7550a3;exit'
print*,' '
print*, 'For further information type help.'
call groute(' ') ! lets you choose where output shall go
c ***** choose to have orientation par. on graph *****
5 print*, 'Do you wish to have the orientation tensor on your graph'
write(6,10)
10 format(' Please enter y/n :', $)
read(5,20) ans
20 format(a1)
if(ans.eq.'N'.or.ans.eq.'n') iflgop=0
if(ans.ne.'Y'.and.ans.ne.'y'.and.iflgop.ne.0) goto 5
if (iflgop.eq.0) then
    print*, 'Orientation parameters will NOT be printed'
else
    print*, 'Orientation parameters WILL be printed'
endif
call gopen ! opens window on selected device
call grpsiz(xs,ys) ! inquire size of device page (mm)
call gvport(xs*0.1,ys*0.1,xs*0.8,ys*0.8) ! defines viewport size

```



```

call gwbox(xs,ys,0.) ! maps page to device page
call xplot(frmin,frmax,name,stnsor,iflgop,intmin,intmax,inob)
call gclose ! close window on selected device
print*,'Your output should be in a file called '
print*,'HPGL for gpchem (use lpr -b -Pgpchem HPGL)'
print*,'POST for pschem (use lpr -Ppschem POST)'
return
end

subroutine xplot(frmin,frmax,name,stnsor,iflgop,intmin,intmax,inob)
real stnsor(5),intmax,intmin,zero,exp,calc
character*1 shorter(20),shorten(30)
character*20 name
character*60 titles
integer iflgop,hardplot
logical expd
if(inob.eq.1) then
  hardplot=1
else
  call gclear
  inob=0
  hardplot=0
endif
call glimit(frmin,frmax,intmin,intmax,0.,0.) ! defines user coord.
limits
  call gscale ! maps user units to mm coords
  open(10,file='lcsim.calc')
  inquire(file='lcsim.exp',exist=expd)
  if (expd.ne.0) open(11,file='lcsim.exp')
c Both file formats should be free format
c each line containing one frequency and one intensity
c Draw axis
  call raxlas(2)
  call raxis(1,0.,2.,1)
c Plot calculated spectrum
  call gwicol(-2.,1) ! gwicol sets width to 2 pixel & colour to
white
  110 read(10,*,end=200)freq,amp
  if (freq.lt.frmin.or.freq.gt.frmax) goto 110
  call gvect(freq,0,0) ! gvect(x,y,0) moves to x,y
c gdash sets line style (0=solid,1-3=dotted,4-7=dashed,8-10=dotdash)
  if(amp.le.(0.9*intmax)) goto 199
  call gdash(0)
  call gvect(freq,(0.85*intmax),1) ! gvect(x,y,1) draws to x,y
  call gvect(freq,(0.85*intmax),0)
  call gdash(2)
  call gvect(freq,(0.9*intmax),1)
  goto 110
  199 call gdash(0)
  call gvect(freq,amp,1)
  goto 110
  200 continue
c Plot experimental spectrum
  if (expd.eq.0) goto 220
  210 read(11,*,end=220)freq,amp
  if (freq.lt.frmin.or.freq.gt.frmax) goto 210
  call gdash(4)

```

```

call gvect(freq,0.,0)
if(amp.le.(-0.9*intmin)) goto 209
  call gvect(freq,(0.85*intmin),1)
  call gdash(2)
  call gvect(freq,(0.9*intmin),1)
  goto 210
209 call gvect(freq,-amp,1)
  goto 210
220 continue
close(10)
if (expd.ne.0) close(11)
c **** remove blank spaces at end of string name *****
shorter(20)=name
j=0
do 208 i=1,20
if(shorter(i).ne.' ')shorten(i)=shorter(i)
if(shorter(i).eq.' ')j=i+1
208 continue
titles='LCSIM - '//name
k=.5*(30-(9+j))
do 207 i=1,k
titles=' '//titles
207 continue
l=30-k
c Annotate plot
call rtxhei(3.) ! height 3 mm
call rtxjus(1,1) ! coords indicate bottom middle
call rtx(-1,titles,(frmax+frmin)/2.,.95*intmax)
do 2000 i=21,1
call rtxc(-1,shorten(i))
2000 continue
zero=(-intmin/(intmax-intmin))
calc=zero+(intmax/(2*(intmax-intmin)))
exp=zero-(-intmin/(2*(intmax-intmin)))
call gscamm
call grpsiz(xs,ys)
call rtxjus(0,2) ! left centre
if (hardplot.eq.0) then
  call rtx(-1,'Frequency/Hz',.9*xs,(ys*zero*0.75)+0.1*ys)
  call rtx(-1,'Calculated',.9*xs,ys*calc*0.75+0.1*ys)
  call rtx(-1,'Experimental',.9*xs,ys*exp*0.75+0.1*ys)
else
  call rtx(-1,'Frequency/Hz',.905*xs,(ys*zero*0.75)+0.125*ys)
  call rtx(-1,'Calculated',.905*xs,ys*calc*0.75+0.125*ys)
  call rtx(-1,'Experimental',.905*xs,ys*exp*0.75+0.125*ys)
endif
c ***** write orientation tensor to plot *****
if(iflgop.eq.0) goto 10
call rtxhei(3.) ! height 3 mm
call rtxjus(0,1) ! Coordinates indicate bottom left
syy=-(stnsor(1)+stnsor(2))
call rtx(-1,'Sxx = ',.01*xs,.24*ys)
call rtxnc(stnsor(2),5)
call rtx(-1,'Syy = ',.01*xs,.20*ys)
c rtxnc writes number continuing from last write statement
call rtxnc(syy,5)
call rtx(-1,'Szz = ',.01*xs,.16*ys)
call rtxnc(stnsor(1),5)

```

```

call rtx(-1,'Sxz = ',.01*xs,.12*ys)
call rtxnc(stnsor(3),5)
call rtx(-1,'Syz = ',.01*xs,.08*ys)
call rtxnc(stnsor(4),5)
call rtx(-1,'Sxy = ',.01*xs,.04*ys)
call rtxnc(stnsor(5),5)
10  iflgop=1
    call gempty
    return
    end

subroutine xopen(isit)      ! opens an X-window for plot
print*, 'Please Wait - opening window....'
call groute('select lx11;exit') ! select xterminal for output
call gopen      ! open graphics device
call grpsiz(xs,ys) ! inquire size of device page (mm)
call gvport(xs*0.1,ys*0.1,xs*0.8,ys*0.75) ! defines viewport size
call gwbox(xs,ys,0.) ! maps page to device page
isit=1
return
end

subroutine xclose(isit)    ! closes X-window
print*, 'Closing window...'
call gclose
isit=0
return
end

```

CALCULATION - make vibrational corrections

```

FUNCTION DDCNMR(I,J,COOR,S,xyzabc,gampfile,ltu,utl)
REAL GAMMA(6),COOR(4,50),S(5),Dalpha,correction
integer xyzabc,ltu(50),utl(50)
character*40 gampfile
DATA GAMMA/245.017,61.605,17.702,-24.832,230.509,-48.714/
C   * Determine Nuclear Distance
R=(COOR(1,I)-COOR(1,J))**2+(COOR(2,I)-COOR(2,J))**2
& +(COOR(3,I)-COOR(3,J))**2
R3=R**1.5
R=R**0.5
C   * Determine Direction Cosines
ZZ=(COOR(3,J)-COOR(3,I))/R
ZY=(COOR(2,J)-COOR(2,I))/R
ZX=(COOR(1,J)-COOR(1,I))/R
C   * Determine TERMK
NUCI=COOR(4,I)
NUCJ=COOR(4,J)
TERMK=-2.*GAMMA(NUCI)*GAMMA(NUCJ)
C   * Determine Sxx & Syy
SXX=S(2)
SYY=-(S(1)+S(2))
C   * Determine the Direct Dipolar Coupling Term
Dalpha=TERMK/R3*(S(1)*ZZ**2+SYY*ZY**2+SXX*ZX**2

```

```

& +2.*S(3)*ZX*ZZ+2.*S(4)*ZY*ZZ+2.*S(5)*ZX*ZY)
  if (xyzabc.eq.0) goto 998
c   Make vibrational corrections to obtain Do
  open(8,file=gampfile)
1   goto (2,3,4,5,6,7) xyzabc
2   READ(8,100,end=999)N1,N2,CXX,CYY,CZZ,CXY,CYZ,CXZ,DELX,DELY,DELZ
  GO TO 8
3   READ(8,100,end=999)N1,N2,CXX,CZZ,CYY,CXZ,CYZ,CXY,DELX,DELZ,DELY
  GO TO 8
4   READ(8,100,end=999)N1,N2,CYY,CXX,CZZ,CXY,CXZ,CYZ,DELY,DELX,DELZ
  GO TO 8
5   READ(8,100,end=999)N1,N2,CYY,CZZ,CXX,CYZ,CXZ,CXY,DELY,DELZ,DELX
  GO TO 8
6   READ(8,100,end=999)N1,N2,CZZ,CXX,CYY,CXZ,CXY,CYZ,DELZ,DELX,DELY
  GO TO 8
7   READ(8,100,end=999)N1,N2,CZZ,CYY,CXX,CYZ,CXY,CXZ,DELZ,DELY,DELX
8   continue
  maxij=max(ltu(i),ltu(j))
  minij=min(ltu(i),ltu(j))
100  format(1x,2i2,9f10.6)
  if(max(n1,n2).eq.maxij.and.min(n1,n2).eq.minij) then
    R=SQRT(DELX**2+DELY**2+DELZ**2)
    ZX=DELX/R
    ZY=DELY/R
    ZZ=DELZ/R
    g1=gamma(coor(4,i))*sqrt(2.)
    g2=gamma(coor(4,j))*sqrt(2.)
    BRAC=7.0*(CXX*ZX**2+CYY*ZY**2+CZZ*ZZ**2+2.0*(CXY*ZX*ZY+CXZ*ZX*ZZ+C
#YZ*ZY*ZZ))- (CXX+CYY+CZZ)
    PHIXX=CXX-10.0*(CXX*ZX**2+CXY*ZX*ZY+CXZ*ZX*ZZ)+2.5*ZX**2*BRAC
    PHIYY=CYY-10.0*(CYY*ZY**2+CXY*ZX*ZY+CYZ*ZY*ZZ)+2.5*ZY**2*BRAC
    PHIZZ=CZZ-10.0*(CZZ*ZZ**2+CXZ*ZX*ZZ+CYZ*ZY*ZZ)+2.5*ZZ**2*BRAC
    PHIXY=CXY-5.0*((CYY+CXX)*ZX*ZY+CXY*(ZX**2+ZY**2)+CXZ*ZY*ZZ+CYZ*ZX*
#ZZ)+2.5*ZX*ZY*BRAC
    PHIXZ=CXZ-5.0*((CZZ+CXX)*ZX*ZZ+CXZ*(ZX**2+ZZ**2)+CXY*ZY*ZZ+CYZ*ZX*
#ZY)+2.5*ZX*ZZ*BRAC
    PHIYZ=CYZ-5.0*((CZZ+CYY)*ZY*ZZ+CYZ*(ZY**2+ZZ**2)+CXY*ZX*ZZ+CXZ*ZY*
#ZX)+2.5*ZY*ZZ*BRAC
    RD=SXX*(DELX**2-DELZ**2)+SYY*(DELY**2-DELZ**2) +
& 2*(SXY*DELX*DELY+SXZ*DELX*DELZ+SYZ*DELY*DELZ) ! missing term
    RPERP=SXX*PHIXX
    RPAR=-SXX*PHIZZ+SYY*(PHIYY-PHIZZ)+2.0*(S(3)*PHIXY+S(5)*PHIXZ+S(4)*
#PHIYZ)
    RC=RPAR+RPERP
  else
    goto 1
  endif
  close(8)
  correction=1.
c   Safeguard against division by zero error
  if (RC.ne.0.and.RD.ne.0) correction=(1-(RC/(RD+RC)))
  Dnought=Dalpha/correction
998  if (xyzabc.eq.0) DDCNMR=Dalpha
  if (xyzabc.ne.0) DDCNMR=Dnought
  return
999  stop 'End of gampfile reached too soon'
  END

```

A.IV Modifications to MM3

```
subroutine cyvin (lun,lisamp,temp,dxi,dyi,dzi)
include 'COMMON.PAR'
common/atoms/natom,x(maxatom),y(maxatom),z(maxatom),
$      itype(maxatom),name(maxtype),wt(maxtype)
.
.
.
REAL COVAR(3,3),RMASSI,RMASSJ,UIA,UIB,UJA,UJB,FWT
REAL MYCONST,TEMP1,TEMP2,BRACKETS
INTEGER ALPHA,BETA
LOGICAL EXISTEXP ! true if file containing exp. freq. exists
REAL FEXP(300) ! contains experimental frequencies
INQUIRE(FILE='freq.exp',EXIST=EXISTEXP)
.
.
.
      natoms = 3 * natom
c do-loop over normal vibrational modes
      do 50 L=nnrcrd,natoms
.
.
.
c two do-loops over atoms to calculate amplitudes for each atom pair
c
      OPEN(97,FILE='dcmm3')
      OPEN(99,FILE='covar.mm3')
      WRITE(99,116)
116  FORMAT(' i j      Cxx      Cyy      Czz      Cxy      Cyz
$      Cxz')

      iamp = 0
      do 30 j = 2, natom
        i2 = j - 1
        do 40 i = 1, i2
.
.
.
          DO 112 ALPHA=1,3
          DO 112 BETA=1,3
          COVAR(ALPHA,BETA)=0.0
112  CONTINUE
          IF (EXISTEXP.NE.0) THEN
            w=FEXP(1) ! use experimental frequencies if available
          ELSE
            w = sqrt (abs (slam(1))) * fcon
          ENDIF
.
.
.
```

```

MYCONST=16.858
TEMP1=2.0*0.71942*W/T
TEMP2=(1.0 + exp (-TEMP1)) / (1.0 - exp (-TEMP1))
FWT=(MYCONST/W)*TEMP2
DO 111 BETA=1,3
DO 111 ALPHA=1,3
    UIA=EIGVEC(3*(I-1)+ALPHA,L)
    UIB=EIGVEC(3*(I-1)+BETA,L)
    UJA=EIGVEC(3*(J-1)+ALPHA,L)
    UJB=EIGVEC(3*(J-1)+BETA,L)
    BRACKETS=(UIA-UJA)*(UIB-UJB)
    COVAR(ALPHA,BETA)=COVAR(ALPHA,BETA)+BRACKETS*FWT
111 CONTINUE
50 continue

.
.
.
WRITE(97,115) i,j,COVAR(1,1),COVAR(2,2),COVAR(3,3),COVAR(1,2),
$ COVAR(2,3),COVAR(1,3),(x(i)-x(j)),(y(i)-y(j)),(z(i)-z(j))
WRITE(99,114) i,j,COVAR(1,1),COVAR(2,2),COVAR(3,3),COVAR(1,2),
$ COVAR(2,3),COVAR(1,3)
114 FORMAT(2i3,6F10.6)
115 FORMAT(1x,2i2,9F10.6)

40 continue
30 continue

CLOSE(99)
CLOSE(97)

.
.
.
return
end

```

Appendix B

Programs for the analysis of LCNMR spectra

Programs

The various programs required to obtain D^α couplings from LCNMR spectra can be found in the directory `/u/user1/chem/dwhr06/bin` and can therefore be accessed most easily if the user's **PATH** is set to include this directory. Details of how to do this can be found in the **Unix** documentation.

The source files can also be found on the process `/u/user1/chem/dwhr06`, in their respective subdirectories. The main programs are:

lcsim - formerly **lequor** now modified to include vibrational corrections. Used to determine starting values for D° couplings, approximate orientation parameters and to assign experimental spectra.

sliquor - used to refine D° couplings and other spectral parameters to fit an experimental spectrum.

bmgv - used to correct refined D° couplings (determined above) to D^α couplings which yield r_α structural information.

Other useful programs are:-

makelcsim - used to make input files for **lcsim**

dcexpand - used to convert **dcXYZ** file produced by vibrational analysis program **GAMP** (described elsewhere) for use with **lcsim**.

scale - scales, or sets threshold for, intensities of experimental lines in files **lcsim.exp** and **sliq.exp**.

rotax - used to transform a set of cartesian coordinates to a principal inertial axis system.

Overview

In solving LCNMR spectra we start with an approximate structure, known isotropic spectral parameters and an experimental spectrum. The program **lcsim** is used to experiment with values for the orientation parameters until the simulated spectrum is close enough to the experimental spectrum to allow the assignment of some (if not all) of the peaks. This assignment along with initial values for the D^0 couplings are passed into **sliquor** for refinement resulting in a set of D^0 couplings which fit the spectrum "exactly". These couplings are then vibrationally corrected to D^α couplings using **bmgv** and are ready for use in structural determination.

N.B. in simple cases where the spectrum is first order, D^0 values may be obtained directly from the experimental spectrum thereby eliminating the need for **lcsim** and **sliquor**. However, **bmgv** must still be used, as D^α values are required to obtain r_α structural parameters.

Using LCSIM

The program is run by typing **lcsim**, assuming the user's **PATH** has been set up as described above (otherwise type **~dwhr06/bin/lcsim**). Wherever possible, an X-terminal or Sun running X-windows should be used; that

way the simulated spectrum will be of far greater resolution than the 'asterisk plot' obtained using a text only terminal.

N.B. if the user wishes to send simulated spectra to an X-window or to a hardcopy device, they must edit their shell startup file (**.login**, **.bashrc** etc.) to include one of the following lines.

```
. /usr/local/uniras/base/uni.profile    (using bash)
```

```
source /usr/local/uniras/base/uni.profile    (using cs)
```

Input

Input may be via a file or directly from the terminal. In the former case the program **makelcsim** is available which asks for information in a user-friendly manner (I hope) and creates an input file for **lcsim**. This method is recommended as it is least likely to result in error (some error checking is included in **makelcsim**).

Input from the terminal, from within **lcsim**, is similar but the program is more fussy about formats etc. You will find the program asking you for what may seem to be obvious information (e.g. the spin multiplicity of a ^1H nucleus - **makelcsim** knows this for itself).

Conclusion - use makelcsim.

On running **makelcsim** the user is first prompted for the name of the input file to create - any **unix** permissible filename is allowed. The next prompt is for the number of spinning nuclei in the system. It should be noted that low

abundance nuclei should not be included at this stage - "satellite only" spectra should be simulated for each low abundance nucleus in turn. For each nucleus, the user must now input cartesian co-ordinates, nuclear type and index number. The numbers corresponding to different nuclear types are given with the prompt. The index numbers allow the programs to communicate using the user's own numbering scheme (only numbers from 1-99 are currently allowed). The structural input is now complete.

The next three values to be input define the minimum frequency, maximum frequency and minimum intensity of the calculated spectral lines. There is little disadvantage in choosing as wide a range as the complete experimental spectrum. For the purposes of plotting, the axes may be changed from within **lcsim**. It is best to start with a low value for the minimum intensity (~0.1). This may be increased later if too many lines are calculated.

Prompts follow for the indirect (J) coupling constants and chemical shifts (Hz) using the user's numbering scheme as input above. Finally, information is input concerning vibrational corrections to convert D^α couplings (calculated from an r_α structure) to D^o couplings (as observed experimentally). This involves first inputting an integer from 1 to 6 which defines how the axes used in the vibrational analysis program (A,B and C) relate to the axes used for the input of coordinates (X,Y and Z). The former axis system can be determined by examining the **gamp** output file (**dumpXYZ**), the **mm3** output file (**TAPE4.MM3**) or the **asym20** input file. If a zero is input at this stage then the program will make no vibrational corrections, hence the simulation will be more crude. Assuming vibrational corrections are to be made, then the next input is the name of the file

containing the information necessary to do so. If either **mm3** or **asym20** have been used for the vibrational analysis, the filename should point to the file **dcmm3** or **dcasym** produced by the relevant program (including a pathname if necessary). If **gamp** has been used for the vibrational analysis the program **dcexpand** must first be used to expand the file. This stems from the fact that **dcXYZ** usually contains information on only one of each set of related atom pairs. The filename of the output file of **dcexpand** should then be used in the input to **lcsim**.

Using dcexpand

This program first prompts for input and output filenames (e.g. input **dcXYZ**, output **dcXYZ.expanded**) and also a file in which information about the conversion process is stored. This means that the expansion can be repeated, if the force field is altered, without the need to re-input all the information. On the first run, however, the user is given the index numbers of two atoms and is asked to input any symmetry related pairs. This is repeated for all pairs of atoms given in the input to **gamp**.

Simulating a Spectrum

Assuming an input file has been made and its name given to **lcsim** when prompted, the program prints out a list of commands and gives the prompt '**lcsim Commands:**'. If a simulated spectrum is required (almost always the case) the command '**B**' must be carried out first, which will lead to prompting for some further information (most of which is contained in the input file and can be read from there).

Simulating a spectrum can be laborious, often relying on trial and error in varying the orientation tensor. It is simplified if the symmetry is high, when certain orientation parameters are known to be zero. Commands used to vary the orientation parameters are:-

sab=r Set a certain parameter to a specified value
 e.g. **sxx=0.032**

i=r Set increment to a specified value e.g. **i=0.0001**

sab+ or **sab-** Increment or decrement **S_{ab}** by **i** e.g. (using the above values) **sxx+** (**S_{xx}** becomes 0.0321) then **sxx---** (**S_{xx}** becomes 0.0318)

vs View current values of orientation parameters

Note: **S_{yy}** may not be varied directly as it is implied by **S_{xx}+S_{yy}+S_{zz}=0**

To view the effect of the orientation parameters on the spectrum one of the following commands is used:-

G Usually used on a non X-windows terminal - produces a crude plot composed of asterisks (there is a suspected bug in this routine which causes lines which should be "off-screen" to appear elsewhere on the plot).

- X** If the program is run under X-windows then a window will be opened displaying the simulated spectrum. If there is a file in the current directory called **lcsim.exp** then the program will assume that this contains experimental frequencies and intensities, and will plot these below the calculated spectrum. This file should contain frequencies and intensities (in free format), one pair to a line. The intensities should range from 0 (?) to around 3 to be similar to the calculated values but may take any value (the program **scale** can be used to scale or set a threshold to the intensities in **lcsim.exp**).
- P** This may be used to redirect the simulated spectrum to any suitable device, e.g. **gpchem** or **pschem**. Onscreen instructions tell you how to select these. This does not require the program to be run under X-windows.
- Y & F** Used to change the limits of the axes of the simulated spectrum. **F** changes the frequency axis, **Y** changes the intensity axis. **Y** has no effect on the output of **G**.

Viewing Direct Couplings

To view the effect of the orientation parameters on the **D** couplings the command **vd** is used e.g. **vd8,10** (shows coupling $D_{8,10}$). In addition, **vd** with no parameters lists all the couplings.

Information obtained from lcsim

After an approximate fit has been made to the experimental spectrum, the following information should be noted:

1. Approximate values for the D couplings
2. Approximate orientation parameters
3. The peak assignment

The peak assignment is shown by the integer preceding each calculated frequency in the table generated by any of the plot commands. At present this cannot be sent directly to a file other than by redirecting the standard output or by using the **unix** command, *script* (type *man script* for help with this). The standard X-windows cut and paste features can also be useful in this regard.

Useful Tips

- For molecules with low abundance nuclei such as ^{13}C or ^{29}Si , it is best to first simulate the spectrum due to high abundance nuclei only. Using orientation parameters obtained from this, the expected positions of the satellite peaks may be calculated by introducing the low abundance nuclei one at a time to the input. The results for each low abundance nucleus should be compared together. Clearly it makes little sense to assign a single experimental satellite to two different spin systems. Any peaks which cannot be assigned with a reasonable degree of confidence should not be used in the refinement stage (see below).

- There are no hard and fast rules when it comes to finding orientation parameters. Often it is simply a case of systematically varying the unknown tensor elements and noting the values at which the simulated spectrum starts to resemble the actual spectrum. This essentially leads to an n-dimensional grid (where n is the number of unknown orientation parameters) in which similarity to the experimental spectrum is recorded. Sections of the grid where the fit is reasonable can then be examined in more detail and the process repeated until the stage where assignment of lines becomes possible.
- The above process can often be simplified if one or more couplings can be measured directly from the spectrum. For example, the program may only be needed to identify satellite peaks.
- A direct coupling constant arising from a pair of atoms lying parallel to an axis will depend only on the orientation parameter associated with that axis and the interatomic distance. For example, for atoms lying parallel to the x-axis, the direct coupling will be independent of all orientation parameters other than S_{xx} .

Using SLIQUOR

After approximate D^0 couplings have been found using **lcsim**, these can be refined to fit the experimental spectrum more precisely. This is the purpose of the program **sliquor**.

Input to **sliquor** is similar to the input to **lcsim**, the major differences being that no structure is required and that no information concerning vibrational

corrections is needed. Information entered is automatically saved to a named file which allows the calculation to be re-run. This file may be edited if any changes are required.

The program is run by typing *sliquor*. This assumes that the user's **PATH** has been set up as described for **lcsim**. The name of an input file must then be entered. If this file exists, it is used for input, otherwise a new input file is created. The next few inputs are self-explanatory and are very similar to those for **lcsim** (or **makelcsim**). The prompt **Group/Single Nuclei**, however, does not appear in **makelcsim**. If you are unsure about its meaning then simply type 1 for each spinning nucleus in the system you are studying.

Note: The **Group/Single Nuclei** flags (known elsewhere as **kh** flags) are used to simplify problems with, for example, a freely rotating methyl group. This can effectively be treated as one nucleus of multiplicity 4. In such cases it would also be necessary to define a diagonal D coupling (i.e. $D_{n,n}$) and this may be done later on in the input.

The final numerical input is the peak assignment, as determined using **lcsim**. If this is a straightforward simulation, with no refinement required, simply type 0 for the number of peaks. If peaks have been assigned then the identification numbers and experimental frequencies are entered as **I4,F10.3**.

At this stage it is possible to change the output options (selecting whether various tables and plots are produced).

If the option to refine spectral parameters has been selected, the number of parameter sets must now be entered. The parameter sets themselves are entered according to the formatting guide displayed as a row of exclamation marks. This is best explained by example:

Input	Meaning
! !!!!! !!!!!	
:d 8 9 910	refine $D_{8,9}$ and $D_{9,10}$ to the same value
:d 810	refine $D_{8,10}$
:w 8 8 9 9	refine chemical shifts of atoms 8 and 9 to the same value
:j 810	refine $J_{8,10}$

The above example might be used in a system where atoms 8 and 10 are equivalent and so couple equally to atom 9. It is a bad idea to refine more spectral parameters than there are assigned peaks. In such cases it is best to refine parameters a few at a time until more peaks can be identified. In most cases it can be assumed initially that chemical shifts and J couplings do not change significantly from their values in isotropic solvents.

Output

Sliquor can output a simulated spectrum in much the same way as **lcsim**. The main difference is that there are no orientation parameters printed (**sliquor** refines spectral parameters independently of orientation). Refined spectral parameters are printed along with standard deviations.

Useful Hints

- Any changes made to the initial values of the spectral parameters may result in a change of the peak assignment numbers. The program will ignore any lines after the first one it fails to match which can cause a "crash". For this reason, it is advisable to first calculate the spectrum with no parameters refining, in order to check that the assignment is still valid. To help with this, the number of lines matched and the highest line number matched are displayed after the list of calculated frequencies.
- Missassignment of even a single line can prevent the refinement from fitting the rest of the spectrum. Resist the temptation to assign as many lines as possible at the beginning of the refinement. Start with only the lines that can be assigned with a high degree of confidence.
- Intensities can be as important as frequencies when assigning lines. They can also be misleading! Broad lines in the experimental spectrum can be much more intense than they appear (peak areas, rather than peak heights, should be used if possible). Beware also for overlapping lines which can make experimental peaks seem more intense than expected.

Using BMGV

If the direct dipolar coupling constants, obtained above, are to be used to calculate r_{α} structural parameters, they must first be vibrationally corrected from D^0 to D^{α} . This is done using the program **bmgv**. Upon running the

program (by typing **bmgv**) the user is prompted for the names of the following files.

1. The file containing the covariance matrices and internuclear vectors. This will usually be called **dcXYZ**, **dcmm3** or **dcasym**, depending on the program used to carry out the vibrational analysis (i.e. **GAMP**, **MM3** or **ASYM20** respectively).
2. A file which will store the remainder of the input so that the calculation can easily be repeated (for example, if changes are made to the force field). Inputting the name of an existing file will cause the program to attempt to use that file for subsequent input.
3. The name of the output file (a warning will be given if an existing file is chosen).

The remainder of the input is quite straightforward. The index numbers of atoms of each of the available types are entered first (ending each list with a zero) followed by the values of the orientation parameters. The relationship between the axis system used by the vibrational analysis program and that used in the LCNMR analysis is then defined, using the same method as that described for **lcsim**. When this has been done, the values of the D° couplings are entered. Zero (or simply RETURN) can be entered for any couplings which are unknown.

The input is now complete and the results are written to the specified file, in the form of a table. This includes the correction terms (**dh**) and the corrected coupling constants (**Dcorr**) as well as values calculated from the internuclear vectors (**dhcalc** and **Dcalc**) which can be used to verify that the axis systems have been correctly defined. If all is well then the **Dcorr**

values (i.e. D^α) are ready to be used in the structural analysis. Uncertainties can be determined by combining the errors from the **sliquot** refinement with the errors associated with the vibrational correction (these can be taken to be approximately 10% of **dh**).

Appendix C

Notes on the use of MM3

These notes are intended as a supplement to the full **MM3** manual and serve only to highlight aspects specific to the calculation of vibrational corrections to electron diffraction, microwave and liquid crystal NMR data.

MM3 input files are most conveniently produced using the program **MINP** (as described in the **MM3** manual). Once created, **MM3** can be started and the line number of the first input line must be entered (usually 1). In order to calculate vibrational corrections option 4 must be selected at the main menu (i.e. Block diagonal followed by Full Matrix Method with various temperature and printout options). The next input is the temperature for which the calculations are to be made. N.B. inputting 0. at this stage results in calculations being made for room temperature and so a suitably low value (e.g. 0.01) must be used if calculations for absolute zero are required.

The next two values to be input are flags which determine the various output options required and it is simplest to enter 3 at both prompts. The program will then perform the required calculations and the bulk of the output is written to a text file called **TAPE4.MM3**. This contains parallel and perpendicular amplitudes of vibration as well as corrections to rotation constants. If it is required that experimentally determined frequencies should be used in the calculations then a file called **freq.exp** must be created which contains the frequencies, one to a line, in the order calculated by **MM3**. The correct ordering of the lines is best achieved by examining the symmetry assignments of the calculated frequencies, after an initial run of the program. In addition, the program **vibplt** can be used to produce animated pictures of the normal modes (if running under X-windows). It is important to remember that **MM3** will always use the

values contained in **freq.exp** if it is present in the working directory. The same applies to a file called **struct.exp** which can be used to force **MM3** to use a geometry based on the co-ordinates contained within the file (in free format).

The covariance matrices required to calculate vibrational corrections to direct coupling constants are written to a file called **dcmm3**, in a suitable format to be read directly by the program **bmgv** (described elsewhere).

Appendix D

Lecture Courses and Conferences Attended

Lecture Courses

X-ray Crystallography (1 unit)
Dr.R.O.Gould & Dr.A.J.Blake

NMR Spectroscopy (1 unit)
Dr.I.H.Sadler & Dr.D.Reed

Inorganic Cluster Chemistry (1 unit)
Prof.B.F.G.Johnson

Industrial Chemistry (1 unit)
Unilever Research & I.C.I.

UNIX I (1 unit)
Edinburgh Regional Computing Centre

UNIX II (1 unit)
Edinburgh Regional Computing Centre

Shell Programming (1 unit)
Edinburgh Regional Computing Centre

Conferences

Fourth European Symposium on Gas Electron Diffraction
Firbush, Scotland, 1991

Sixth Austin Conference on Molecular Structure
Austin, Texas, 1992

University of Strathclyde Inorganic Club Conference (3 units)
1990-1993

Departmental Meetings and Seminars, 1990-1994