Chapter 1

Novel Strategies for Solving Highly Complex NMR Spectra of Solutes in Liquid Crystals

E. Elliott Burnell[1], Cornelis A. de Lange[2], and W. Leo Meerts[3]

[1]*Chemistry Department, University of British Columbia, 2036 Main Mall, Vancouver, BC, Canada V6T 1Z1*
email: elliott.burnell@ubc.ca

[2]*Laser Centre, Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands*
email: cdelange@few.vu.nl

[3]*Molecular and Biophysics Group, Institute for Molecules and Materials, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands*
and
Physical Chemistry Department, Vrije Universiteit, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands
email: leo.meerts@science.ru.nl

The application of liquid-crystal NMR to solutes with more than eight spins has for decades been hampered by the extremely high complexity of the spectra. We describe recent experimental and novel theoretical advances to address the central problem of spectral analysis. Experimentally the application of multiple-quantum NMR is shown to be a powerful tool. Moreover, theoretical algorithms from evolutionary biology are applied to NMR of solutes in anisotropic environments. These novel spectral fitting procedures are aided by estimates of the solute degree of orientational order, based on simple phenomenological size and shape models. The apparent success of these approaches has considerably extended the realm of applications of NMR in ordered liquids.

1.1. Introduction

The Hamiltonian that describes the high-resolution NMR spectra of solutes dissolved in anisotropic solvents is well known and has proved to be an excellent predictor for such detailed experimental information. The param-

eters that determine the NMR-spectra include the chemical shieldings and indirect spin-spin couplings that are well known from NMR in isotropic liquids. Additionally, direct dipole-dipole and quadrupolar couplings are not averaged to zero in anisotropic phases and usually dominate the spectra of solutes in such solvents. The complexity of these NMR-spectra increases rapidly with the number of interacting nuclear spins. In this chapter we shall limit ourselves to ^{1}H NMR spectra.

We shall limit our discussion to the use of nematic liquid-crystal solvents that are apolar and possess uniaxial symmetry. There exists a large literature on the study of a great variety of solutes in these anisotropic phases. For simple solutes that contain a small number of nuclear spins and a high degree of symmetry, the problem posed by the spin Hamiltonian can be solved analytically.[1,2] For larger solutes with less symmetry this is no longer possible and other means have been developed. Commonly computer programmes are employed that calculate a spectrum from estimated input values for chemical shieldings, and indirect and direct dipole-dipole couplings. The calculated spectrum is then compared to the experimental one to see whether some degree of correspondence between calculated and experimental spectra can be detected. If this is the case, least-squares fitting procedures are employed for adjusting the spectral parameters in such a way that the correspondence is improved. In an iterative fashion this process hopefully converges, and good agreement between calculated and experimental spectra is obtained. If an insufficient number of transitions in the calculated spectrum can be sensibly assigned to experimental ones, the procedure is doomed to failure. In practice this approach is often notoriously time-consuming and usually limited to solutes with fewer than 8 spins.

The problems are exacerbated for molecules with low symmetry that may require as many as five independent and *a priori* unknown orientational order parameters, or for molecules that exist in several conformers that undergo fast interconversion on the NMR time scale. For such cases even crude guesses of the dipolar couplings are difficult to obtain. In addition, the anisotropic contributions to the chemical shifts make it difficult to predict values for solutes dissolved in nematic phases.

In this chapter we shall discuss both experimental and theoretical methods of considerable sophistication to address the problem of analyzing very complex spectra of large and sometimes flexible solutes in liquid-crystal solutions, including those that undergo extensive conformational change. Experimentally we shall employ the method of multiple-quantum (MQ)

NMR to develop a strategy that leads in a step-wise fashion to an analysis of the normal one-quantum NMR spectrum. Application of MQ-NMR to the highly complex spectra of solutes in nematic phases that could not be solved otherwise has shown the inherent power of this method. Several successful examples of the MQ approach will be given.

A drawback of the MQ-method is its experimental sophistication whose implementation requires detailed technical knowledge. For many years attempts have been made to develop automated spectral fitting routines, but with limited success. Recently novel and extremely powerful methods of analyzing extremely complicated NMR spectra of solutes in nematic phases have been developed. These methods are based on Evolutionary Algorithms (EA's). We use in particular Genetic Algorithms (GA's) and Evolutionary Strategies (ES's) which can be implemented into automated spectral fitting procedures that are capable of reliably locating the global minimum in an extensive search space.

It is important to realize that the novel experimental and theoretical developments that will be discussed in this chapter have led to a significant extension of the scope of liquid-crystal methods to obtain detailed information about ordered solutes that was previously thought unattainable. In the archives of many NMR laboratories there are numerous examples of highly complex NMR spectra of partially ordered solutes that never yielded to serious, painstaking and time-consuming attempts to solve them. With the present techniques the study of all kinds of interesting and rather large solutes in anisotropic liquids has now become eminently feasible, as will be shown in the following. In that sense the NMR liquid-crystal method has obtained a new lease of life.

1.2. Theory

In order to calculate high-resolution single-quantum NMR spectra of solutes in anisotropic solvents the starting point, as always in quantum mechanics, is a Hamiltonian. As is common in molecular spectroscopy we define a basis set consisting of products of nuclear spin wave functions, set up the spin Hamiltonian matrix, and perform matrix diagonalization. Transitions between matrix eigenvalues represent spectral frequencies, while intensities are obtained by calculating the squares of matrix elements of the magnetic-dipole transition operator sandwiched between pairs of eigenvectors obtained in the diagonalization procedure.[1,2]

The spin Hamiltonian (in Hz) for a collection of spin $I = \frac{1}{2}$ nuclei in a

solute orientationally ordered in a liquid-crystalline solvent is

$$\begin{aligned}\mathcal{H} = &-\frac{B_Z}{2\pi}\sum_\mu \left[\gamma_\mu(1-\sigma_\mu^{\rm iso}-\sigma_\mu^{\rm aniso})\right] I_{\mu,Z} \\ &+\sum_\mu\sum_{\nu>\mu} J_{\mu\nu}^{\rm iso}\mathbf{I}_\mu\cdot\mathbf{I}_\nu \\ &+\sum_\mu\sum_{\nu>\mu}(2D_{\mu\nu}+J_{\mu\nu}^{\rm aniso})\left[I_{\mu,Z}I_{\nu,Z}-\frac{1}{4}(I_{\mu,+}I_{\nu,-}+I_{\mu,-}I_{\nu,+}\right]\end{aligned} \tag{1.1}$$

where B_Z is the external magnetic field, γ_μ is the gyromagnetic ratio of nucleus μ, σ_μ is the chemical shielding (which has isotropic and anisotropic terms), and $J_{\mu\nu}$ is the indirect spin-spin coupling between nuclei μ and ν. For protons the anisotropic part $J^{\rm aniso}$ is to an excellent approximation zero.

The dipolar coupling between nuclei μ and ν is:

$$D_{\mu\nu} \;=\; -\frac{h\gamma_\mu\gamma_\nu}{4\pi^2}\left\langle r_{\mu\nu}^{-3}\left(\frac{3}{2}\cos^2\theta_{\mu\nu,Z}-\frac{1}{2}\right)\right\rangle \tag{1.2}$$

where $r_{\mu\nu}$ is the internuclear separation and $\theta_{\mu\nu,Z}$ is the angle between the $\mu\nu$ and magnetic field directions, and the angle brackets denote an average over solute internal and reorientational motions. With the assumption that the molecule rapidly interconverts among several 'rigid' conformers Eq. (1.2) can be written

$$D_{\mu\nu} \;=\; -\frac{h\gamma_\mu\gamma_\nu}{4\pi^2}\sum_i p_i r_{i,\mu\nu}^{-3}\sum_{k,l} S_{i,kl}\cos\theta_{\mu\nu,k}\cos\theta_{\mu\nu,l} \tag{1.3}$$

where i labels the conformer which occurs with probability p_i and $\theta_{\mu\nu,k}$ is the angle between the conformer-fixed k axis and the vector joining nuclei μ and ν which have internuclear separation $r_{i,\mu\nu}$. $S_{i,kl}$ is the kl element of the order parameter tensor for conformer i

$$S_{i,kl} \;=\; \left\langle\frac{3}{2}\cos\theta_{k,Z}\cos\theta_{l,Z}-\frac{1}{2}\delta_{kl}\right\rangle_i \tag{1.4}$$

where the angle brackets denote an average over conformer reorientational motions and $\theta_{k,Z}$ is the angle between the conformer-fixed k axis and the space-fixed Z axis which for the liquid crystals used here lies along both the director and magnetic field directions.[3] In this picture small-amplitude vibrational corrections and vibration-reorientation coupling are neglected.

It is useful to distinguish among three different categories of solutes: (i) 'Rigid' solutes with a single conformation that only undergo normal internal vibrational motion. The degree of orientational order of such solutes,

depending on solute symmetry, is determined by at most five independent order parameters S_{kl}. Examples are azulene and biphenylene (Fig. 1.1); (ii) Solutes that undergo conformational change between conformers that

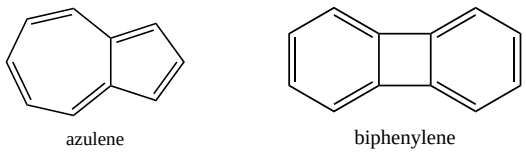

Fig. 1.1. The molecules azulene and biphenylene.

are related by symmetry. Although each conformer in principle requires up to five independent order parameters, symmetry confines this number. Examples are 2,6-dichloro-1-ethenylbenzene, biphenyl and *p*-bromo-biphenyl (Fig. 1.2); (iii) Solutes that interconvert rapidly among several symmetry-

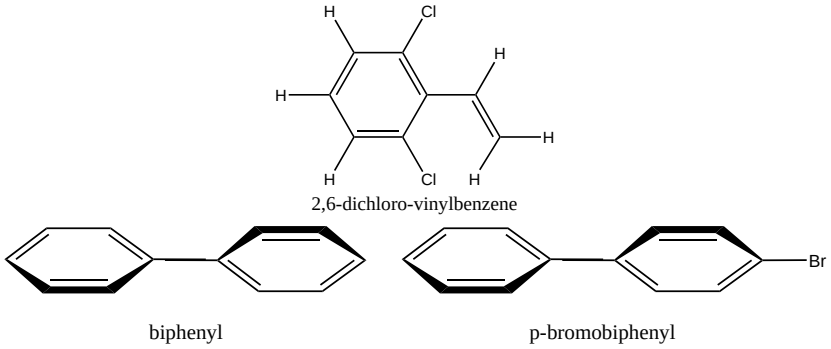

Fig. 1.2. 2,6-dichloro-1-ethenylbenzene, biphenyl and *p*-bromo-biphenyl.

unrelated conformers. Each conformer i requires its own set of independent order parameters $S_{i,kl}$. Examples are butane (C_4H_{10}, Fig. 1.3) and pentane (C_5H_{12}, Fig. 1.3).

Since conformational probabilities and order parameters for each conformer only occur as products in the expressions for the dipolar couplings, a separation can only be achieved if we have some source of independent information about either property. From a large body of research it is now accepted wisdom that a key contribution to the average degree of orientational order is the short-range interaction between the solute and the solvent

Fig. 1.3. The molecules butane (C_4H_{10}, left) and pentane (C_5H_{12}, right).

environment. An important breakthrough in understanding the physical processes underlying partial orientation in anisotropic phases was achieved by the introduction of so-called 'magic mixtures' in which the short-range contribution is deemed to be by far the dominant one. Moreover, simple phenomenological models based on the size and shape of the solute have been developed that can predict order parameters at approximately the 10% level. For an overview of this highly significant development we refer to the literature.[4,5] Clearly, by employing one of the simple size and shape models for estimating order parameters developed by several research groups, a very useful tool for restricting the possible values of the dipolar couplings is available to us. Employing such models has been a crucial factor in the successes of our 1Q high-resolution and multiple-quantum NMR investigations, as well as to the automated fitting procedures.

1.3. Multiple-quantum NMR

The normal single-quantum NMR spectrum consists of transitions between spin energy states that involve the flip of a single spin. For high-spin systems there are many such transitions. The idea of MQ-NMR is to use transitions that involve the mutual flip of many spins. When the number of spins flipped approaches the number of spins in the molecule, the number of possible transitions is greatly reduced. Weitekamp wrote an excellent review that covers a wide range of ideas and experiments.[6] Here we concentrate on the application to the analysis of the high-resolution NMR spectra of solutes orientationally ordered in liquid-crystal solvents, where the technique is used for spectral analysis. We use examples mainly from our own work to demonstrate the use and limitations.

The idea is exemplified by the example of benzene,[6] Fig. 1.4. The normal 1Q spectrum consists of 74 lines which show no regular pattern. The 6Q spectrum is a single line, involving the flip of all six spins. The 5Q spectrum is a doublet because all spins are equivalent: there is one transition involving $|\alpha\alpha\alpha\alpha\alpha\alpha\rangle \rightarrow |\alpha\beta\beta\beta\beta\beta\rangle$ and another involving $|\alpha\alpha\alpha\alpha\alpha\beta\rangle \rightarrow |\beta\beta\beta\beta\beta\beta\rangle$.

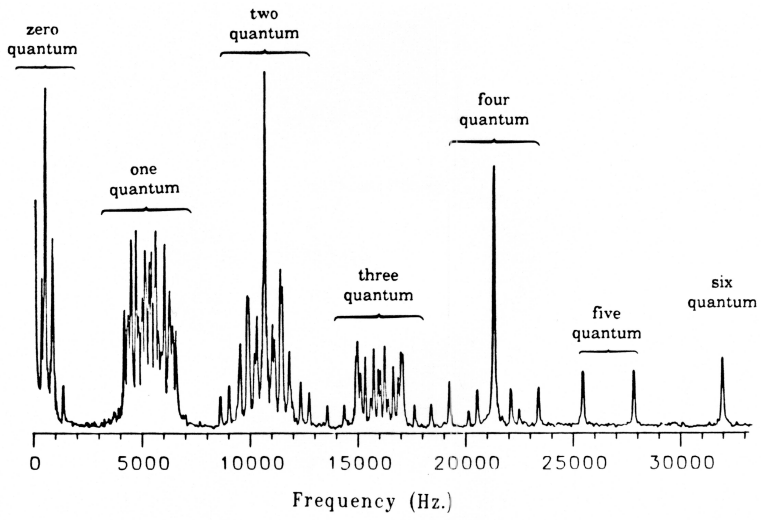

Fig. 1.4. The multiple-quantum spectrum of benzene dissolved in N-(4-ethoxybenzylidene)-4'-*n*-butylaniline (EBBA) obtained using pulse sequence A of Table 1.1. Reproduced with permission from Ref. [7].

The 4Q spectrum consists of three triplets because there are three ways of choosing the two spins that are not flipped, being those that are either *ortho*, *meta* or *para* to each other. The central line of this triplet involves 4Q coherences between spin states such as $|\beta\alpha\alpha\alpha\alpha\alpha\rangle$ and $|\alpha\beta\beta\beta\beta\beta\rangle$. It can be shown that all spectral parameters can be obtained by analysis of the three highest order and relatively simple MQ spectra.[6]

Of course in the case of benzene the 1Q spectrum is readily analyzable with the aid of a computer — in this case computer analysis is 'trivial' because of the high D_{6h} symmetry and hexagonal structure which means there is essentially a single unknown (the order parameter S) needed to predict the three independent dipolar couplings and a single chemical shift. As we shall discuss below, for more complicated spin systems, the MQ-NMR spectra are an invaluable tool to aid spectral analysis.

1.3.1. *The MQ experiment and some variants*

Since the selection rule for NMR spectroscopy is, in general, that only 1Q transitions are directly observable, the MQ spectra must be obtained indirectly. In general, any three-pulse sequence (see Fig. 1.5) can be used

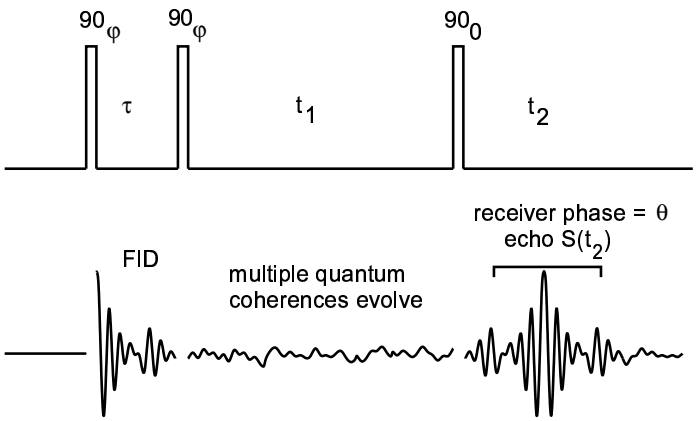

Fig. 1.5. The pulse sequences discussed are all variants of the sequence shown, with phase cycling as described in Table 1.1. Sequence A is a simple example of a generic three-pulse sequence, any of which can in general produce a MQ-NMR spectrum. In general, some of the MQ-NMR coherences that are generated by the second pulse, and that evolve during the variable time t_1, are transferred back into observable coherences by the third pulse. The MQ-NMR spectrum is obtained by Fourier transforming as a function of t_1 the echo signal that follows the third pulse. In practice this is accomplished via a 2D transform of the signals acquired as a function of t_1 and t_2, and projecting the absolute value of the resultant spectrum onto the F_1 axis to give the MQ-NMR spectrum, as shown in Fig. 1.4. (Reproduced from Ref. [8] with permission from Academic Press.)

to generate the MQ spectrum. The first pulse generates observable X and Y magnetizations from equilibrium Z magnetization. During a waiting time τ the X and Y magnetizations evolve under the spin Hamiltonian, and other unobservable 1Q coherences are generated by the dipolar and spin-spin parts of this Hamiltonian. The second pulse converts these other coherences into MQ coherences of all possible orders. Between the second and third pulses the MQ coherences evolve under the spin Hamiltonian. An important feature of this evolution is that the effect of chemical shift offset (in the rotating-frame) is enhanced by the MQ order such that NQ coherences evolve proportional to N times this offset frequency. The third pulse serves to rotate some of the MQ coherences back to unobservable 1Q coherences, which then further evolve under the Hamiltonian to form an observable echo at a time τ following the third pulse. A Fourier transform of the echo amplitude as a function of t_1, the time between pulses 2 and 3, gives in general a spectrum containing all MQ orders as shown in Fig. 1.4. In practice, a two-dimensional experiment is performed in which an echo

Table 1.1. Phase cycles for the three pulse sequences[a].

sequence	phase	phase list
A[b]	ϕ	0
	θ	0
B[c]	ϕ	$0,\ 90/N,\ 2\times 90/N,\ 3\times 90/N$ .... $(4N-1)\times 90/N$
	θ	0, 90, 180, 270
C[d]	ϕ	$0,\ 360/n,\ 2\times 360/n,\ 3\times 360/n$ .... $(n-1)\times 360/n$
	θ	0

[a] As defined in Fig. 1.5, ϕ refers to the phase of pulses 1 and 2, and θ refers to the receiver phase. (Reproduced from Ref. [8] with permission from Academic Press.)
[b] Sequence A gives all MQ orders in a single, wide spectrum (see *e.g.* Fig. 1.4). For this experiment the phases ϕ and θ can be any arbitrary constant angle. Zero degrees was chosen for convenience.
[c] Sequence B is used for the selective detection of the NQ (and $2NQ$, $3NQ$...) spectrum. N is the quantum order being selectively detected.
[d] Sequence C is used for the 3D experiment, where n is an arbitrary integer that should be at least $2N_{\max}+1$. A separate signal is acquired and stored for each value of n, giving a three-dimensional interferogram that is a function of t_1, ϕ, and t_2.

signal is collected for each increment of t_1, and the absolute value of the double Fourier transform with respect to t_1 and t_2 is projected onto the F_1 axis to give the MQ spectrum. Collection of the entire echo in this manner both improves the signal-to-noise ratio and sometimes reveals more transitions in the MQ spectrum than are observable with simply using the echo amplitude.

Historically, as shown in Fig. 1.4, sequence A of Table 1.1 was used with a large frequency offset to spread out the various MQ orders which precess during time t_1 (between pulses 2 and 3) at N times this offset. The large spectral width required to separate the various orders leads to a very wide spectral width with poor digital resolution, and even to problems in exciting the entire spectrum with the pulse powers and widths normally used.[8]

Several schemes have been employed to improve the digital resolution by the selective detection of one order at a time. In one method, field gradient pulses are applied before and after the third pulse.[9] The relative integrated amplitude of these pulses is chosen such that before the third pulse, the effective Z gradient adds $N\nu$ to the offset, which is recovered after pulse 3 by a gradient N times as large. Only the NQ coherences are coherently

added to the echo signal, the others being destroyed by the gradients which are set to refocus only the N^{th} quantum order.

An alternative method is to phase-cycle the first and second pulses and the receiver phase in the manner described by sequence B of Table 1.1. In this way only the desired quantum order (and multiples of it) survive the signal addition and contribute to the MQ spectrum.[10] In addition, only the $+N$Q or $-N$Q spectrum is obtained (see Fig. 1.6), and one can work on-resonance. The spectral width is then much narrower, and thus the digital resolution can be much improved. One problem with this and the previous gradient technique is that the experiment measures a single quantum order, and must be repeated many times in order to collect each MQ order independently; the collection of several orders requires a lot of time.

A solution to the problem of needing a separate experiment to collect each order is to collect all orders at the same time in a 3D experiment, as suggested by Weitekamp[6] and as demonstrated by Syvitski *et al.*[8] The idea is (for each time t_1) to collect an echo for each value of the phase ϕ of pulses 1 and 2 which are varied between 0 and 360° in n increments, sequence C of Table 1.1. Fourier transformation with respect to the phase ϕ gives planes of signal F_1, F_2 where the various MQ spectra are the projections onto the F_1 axis, as demonstrated in Fig. 1.7. As can be seen in Fig. 1.6, the 4Q and 6Q spectra are virtually identical to those collected using the selective detection method of sequence B in Table 1.1: the important difference is that it took the same time using the 3D sequence to collect spectra for *all* orders as it did to collect a single order in the selective experiment.

1.3.2. *Spectral analysis with MQ*

As pointed out in the introduction, with fewer than about 8 spins $I = \frac{1}{2}$, especially for solutes with high enough symmetry (for which there are only one or two independent order parameters) the spectrum can usually be solved using the high-resolution 1Q spectrum and line assignment techniques. More complicated systems are difficult, and the difficulty increases with increasing the number of spins, with lowering the symmetry, and with solute flexibility. Several examples where analysis of the MQ solute spectra proved invaluable, and for which it is not clear that the 1Q spectrum could have been analyzed, are the following.

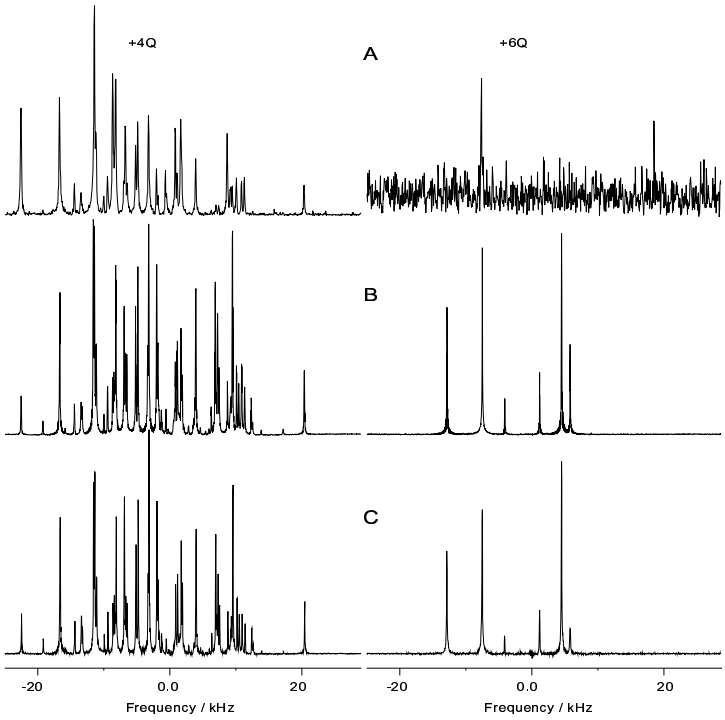

Fig. 1.6. Comparison of the $+4Q$ and $+6Q$ MQ-NMR spectra of *para*-chlorotoluene in ZLI-1132 obtained with different techniques. (A) Two sections of the spectrum collected in 14 h 5 m using sequence A of Table 1.1 with τ set to 20 ms, four scans/t_1 increment, 8192 t_1 increments, F_2 spectral width of 83333 Hz, and F_1 spectral width of 1MHz. The absolute value of the projection of the complex F_2 spectra onto the F_1 axis gives the F_1 spectrum, which is the entire MQ-NMR spectrum. The spectrometer frequency was offset by 60000 Hz from the center of the spectrum in order to separate the various orders. Under these conditions, the higher order coherences are poorly detected, presumably due to the combination of the large offset with the long duration (10 μs) of the 90° pulses. The line at ≈18 kHz in the 6 Q spectrum is an artifact. (B) The absolute value of the spectrum obtained from a selective detection of the +4 Q complex spectrum requiring 14 h 10 m (and the +6 Q complex spectrum requiring another 10 h 40 m) obtained using sequence B of Table 1.1 with N equal 4 (and 6), τ set to 20 ms, two (and one) scans for each of 16 (and 24) phase increments per t_1 increment, 1024 t_1 increments, F_3 spectral width of 33333 Hz, and F_1 spectral width of 70 kHz. (C) Two slices of the absolute value of the 3D experiment (displayed in Fig. 1.7) obtained using sequence C of Table 1.1. The intensity differences between the 6 Q spectra obtained using sequences B and C are most likely due to spectrometer fluctuations. For comparison, the ratios of vertical scales for the 6 Q to the 4 Q spectra are 17 × for (A), 1.35 × for (B), and 1.42 × for (C). The t_1 signals were zero filled once before the Fourier transform. The digital resolution is then 61 Hz per point in (A) and 34 Hz per point in (B) and (C). The peaks of very strong lines have been chopped off for clarity. (Reproduced from Ref. [8] with permission from Academic Press.)

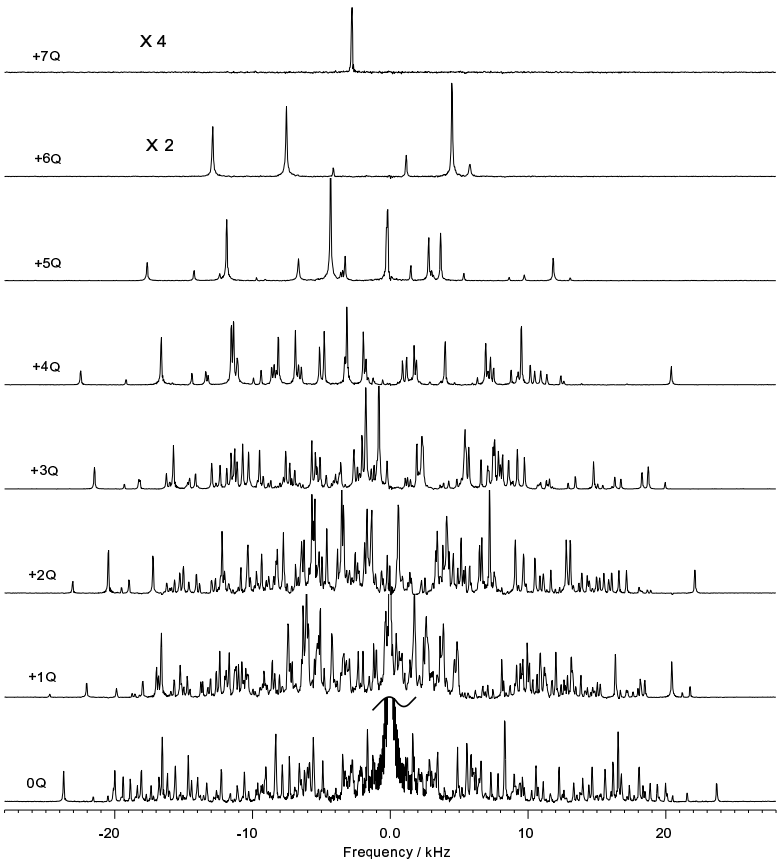

Fig. 1.7. Absolute values of multiple-quantum spectra of *para*-chlorotoluene in ZLI-1132 obtained in 14 h 25 m from the 3D experiment using sequence C of Table 1.1 with τ set to 20 ms, two scans for each of n=16 phase increments per t_1 increment, 1024 t_1 increments, F_3 spectral width of 33333 Hz, and F_1 spectral width of 70 kHz. The Fourier transforms F_3 of the echo signals $S(t_2)$ have been summed onto the F_1 dimension. The result is a two-dimensional spectrum where F_2 are the pseudo ϕ spectra, and F_1 the desired multiple quantum spectra. The positive MQ-NMR orders (F_1 slices 9 to 16) are shown, which correspond to the 0 to $+7Q$ spectra. The peaks of very strong lines have been chopped off for clarity. The Bruker software multiplies the first point of $S(t_2)$ by $\frac{1}{2}$ to correct the rolling baseline caused by the staggered acquisition of real and imaginary signals. While this is important for normal high-resolution NMR experiments, for the 3D MQ NMR experiment it results in small amounts of all MQ orders showing up in the spectrum for each MQ order: it is necessary to set FCOR = 1 in the Bruker software[11] in order to correct this problem and to obtain MQ spectra that contain only one order. In the spectra displayed, the extra signal from all MQ orders was removed by subtracting the signal obtained from an order higher than $N_{\max}$. (Reproduced from Ref. [8] with permission from Academic Press.)

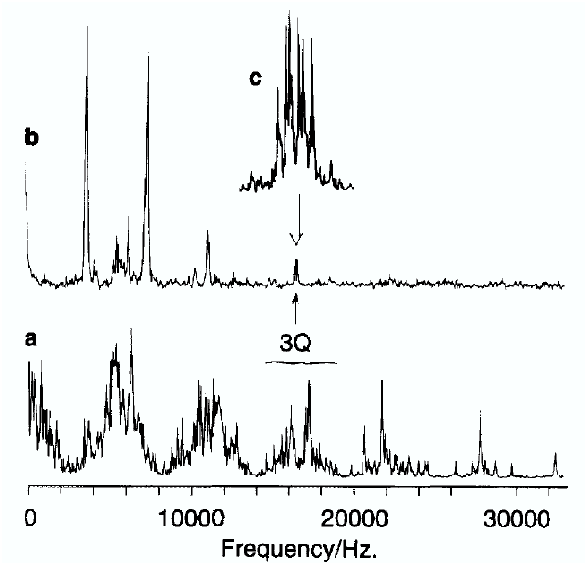

Fig. 1.8. (a) The 400 MHz nonselective MQ-NMR spectrum of 1,3-dichloro-2-ethenylbenzene dissolved in EBBA acquired using pulse sequence A of Table 1.1. (b) The 400 MHz selective excitation MQ-NMR spectrum of 1,3-dichloro-2-ethenylbenzene dissolved in EBBA using pulse sequence A of Table 1.1 with the first two pulses being frequency selective for the aromatic protons (see text). (c) An expanded view of the 3Q region noted in (b). (Reproduced from Ref. [12] with permission from Academic Press.)

1.3.2.1. *2,6-dichloro-1-ethenylbenzene*

This solute (see Fig. 1.2) has only six protons and hence it might be expected to produce an orientationally ordered six-spin NMR spectrum that is easy to analyze. That is not the case. One problem is that the ethenyl group makes an unknown dihedral angle with the benzene ring, and the molecule has no symmetry when this group is out of plane. Hence, there are five unknown order parameters. In addition, the anisotropies in chemical shifts of the ethenyl protons lead to problems in assigning initial values to these shifts. It turns out that assigning the wings of the spectrum to the aromatic protons is easy, and all the problems lie in the cluttered central clump of transitions. This central part from the ethenyl protons was only assigned once the MQ spectrum was obtained and solved.[12,13]

This molecule was also used to investigate the possibility of selectively exciting a subset of spins and so leading to a spectrum from which long-

range intergroup couplings between aromatic and ethenyl protons could more easily be assigned. Here this is possible because there is only weak coupling between the distinct groups of spins. This weak coupling exists because the dipolar couplings between groups of spins are much weaker than those within a group, and there is negligible mixing of basis states between different spin groups. The same non-mixing of basis states occurs in high-resolution NMR when the chemical-shift difference between two spins is much greater than their J coupling. Hence, it is possible to irradiate transitions that involve only one spin group.

It turns out that the outer lines in the normal 1Q NMR spectrum involve only aromatic protons and do not depend on intra-ethenyl couplings. Thus it is possible to excite MQ-coherences among the three aromatic protons by making the first two pulses of sequence A of Table 1.1 frequency selective for outer lines of the spectrum, and hence for the aromatic protons only (a so-called DANTE[14] pulse is used for this purpose). The highest order spectrum then observed is the three-quantum spectrum (from the three aromatic protons) that then consists of eight transitions (see Fig. 1.8) which all depend on the long-range dipolar couplings between aromatic and ethenyl protons. In this way information about these long-range couplings is obtained without needing knowledge of the much stronger short-range couplings and the anisotropies in the chemical shifts. These small long-range couplings are difficult to obtain otherwise.

1.3.2.2. *Biphenyl*

Biphenyl (see Fig. 1.2) is interesting because it forms the core of many liquid crystals. It has 10 spins $I = \frac{1}{2}$ and its orientational order is described by two independent order parameters. Its interest lies in determining the dihedral angle which might be affected by the anisotropic environment. The 1Q spectrum is very complicated. The eight-quantum spectrum was analyzed and provided excellent starting parameters for a fit to the 1Q high-resolution spectrum.[15]

1.3.2.3. *Butane*

Butane (see Fig. 1.3) is arguably the most complicated solute whose spectrum has been analyzed using MQ-NMR. There are 10 spins $I = \frac{1}{2}$, three conformers (only two of which are related by symmetry), and three unknown independent order parameters for each of the two symmetry-unrelated conformers. The conformational problem further complicates

matters because the relative populations of conformers is the subject of debate, and may well be affected by the interaction between the solute and the surrounding molecules in the condensed phase as well as the anisotropic part of the ordering potential. Of course the investigation of these issues is what makes the study of butane interesting. The horribly complicated 1Q spectrum which has few distinguishing features (Fig. 1.9 left) was analyzed by first fitting the selectively detected 7Q and 8Q spectra (Fig. 1.9 right) which gave sufficiently accurate NMR parameters that almost all transitions in the 1Q spectrum could be assigned immediately.[16]

1.3.2.4. *Some problems*

Although the MQ technique is quite successful in extending the number of spins that can be accommodated, there remain problems. It is still necessary to generate trial MQ spectra and to match transitions with the experimental spectrum. Without a lot of effort (*i.e.* following the evolution of the density matrix through the three-pulse sequence), the intensities of the calculated spectrum are meaningless, and one must deal with matching frequencies only and ignoring intensities. The advantage is that there are far fewer lines than in the 1Q spectrum. In general, all spectral parameters are determinable from analysis of a combination of the $N_{\max}$, $N_{\max}-1$ and $N_{\max}-2$ spectra.[6] For simple systems for which there is a reasonable estimate of the structure, a crude starting point can be an initial fit with the order parameters and chemical shifts — in this case, one may be able to get away with fitting the relatively sparse $N_{\max}$ and $N_{\max}-1$ spectra. In the case of butane (see above), the $N_{\max}-2$ and $N_{\max}-3$ spectra were used for the analysis of the dipolar couplings and chemical shifts.

Studies in our laboratory demonstrate that analysis of relatively poor-resolution MQ spectra give NMR parameters that are quite accurate. This accuracy may be sufficient for many purposes, but is lower than that obtainable from the high-resolution 1Q spectrum. For more accurate determination of the spectral parameters, those obtained from analysis of the MQ-NMR spectra provide initial estimates for fitting using 1Q high-resolution line assignment techniques — in every case in our laboratory, the MQ parameters predict a spectrum in which one can immediately assign, with no ambiguity, almost all transitions, and hence obtain spectral parameters to very high accuracy.

While MQ-NMR techniques have proven to be quite successful in the analysis of complicated NMR spectra of orientationally ordered molecules,

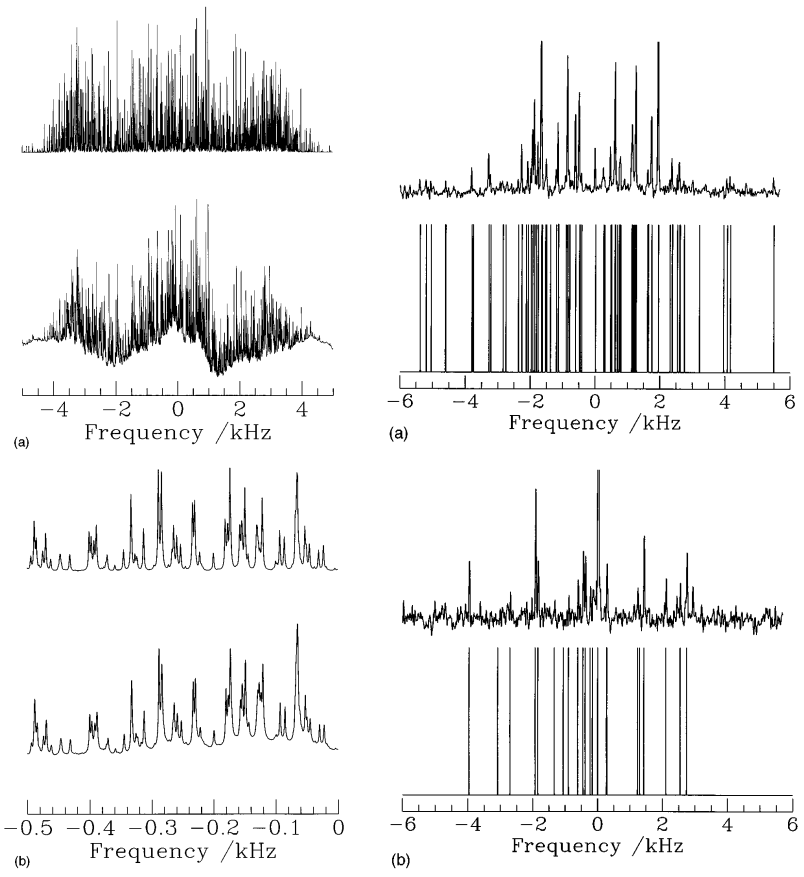

Fig. 1.9. Left: (a) Experimental (bottom) and simulated (top) spectra of partially oriented butane in 55 wt% ZLI 1132 / EBBA. (b) Expanded region of the spectra in (a). Right: (a) Experimental (top) and simulated (bottom) seven-quantum spectra of partially oriented butane. (b) Experimental (top) and simulated (bottom) eight-quantum spectra of partially oriented butane. For the simulated spectra of (a) and (b), the line intensities have been arbitrarily set equal since the intensity of each MQ transition is a complicated function of the preparation time τ and the parameters in the spin Hamiltonian, and does not provide any information for the present study. (Reproduced from Ref. [16] with permission from the American Institute of Physics.)

the experiments are time consuming, and the analysis of complicated spin systems requires considerable effort and NMR-expertise. While it is very powerful, it is certainly not an automated spectral analysis tool.

1.4. Automated analysis using evolutionary strategies

Attempts to develop automated computer routines to efficiently search parameter space in order to find unique values for the spectral parameters that reproduce an observed high-resolution NMR spectrum were made at an early stage, initially for isotropic liquids. Such an alternative approach to the more conventional methods of spectral analysis was first suggested by Diehl *et al.*[17,18] In a similar spirit Stephenson and Binsch developed a basic algorithm which relies on a matrix method derived from a general formulation of the least-squares problem.[19] This approach was implemented in the computer programme DAVINS (Direct Analysis of Very Intricate NMR Spectra),[20] and was reasonably successful. Subsequently the more challenging task of applying the method to the NMR spectra of solutes in anisotropic liquids was undertaken. The programme DANSOM (Direct Analysis of NMR Spectra of Oriented Molecules) was first used in the analysis of the spectra of a number of allyl halides[21] and cyclopentene.[22] However, the limitations of the method in these applications became apparent, because background corrections proved troublesome and for the more complicated cases the intervention of skilled operators during the fitting procedure was usually required. Later the Cosenza group in Italy achieved some notable successes, despite the need for significant operator intervention to avoid the problem of trapping in local minima.[23,24]

In a separate development the use of a Genetic Algorithm (GA) for the analysis of NMR spectra in integrated form of solutes in liquid crystals was reported in a short paper that gave little detail.[25] It appears that trapping in local minima during the convergence process was problematic. In a follow-up paper[26] it was made clear that GA fitting methods were employed to obtain a first approximation to the experimental NMR spectrum. The best GA fit obtained was then used as a starting point for the assignment of individual lines and subsequent least-squares adjustment and refinement. Apart from these two papers we have found no further applications of this method in the literature.

In this section we discuss the implementation of a sophisticated and robust Genetic Algorithm to fit highly complex anisotropic NMR spectra that avoids problems commonly encountered in the automated analysis of NMR spectra of solutes in liquid-crystalline solvents. In particular, the GA-method is inherently capable of circumventing situations where the convergence procedure gets trapped in local minima, without ever reaching the desired global minimum. Our routine is extremely efficient, requires

minimal operator interference, and constitutes essentially an intelligent parameter search method. In the absence of reasonable initial guesses for order parameters or dipole-dipole couplings and chemical shielding parameters convergence is still reached and excellent fits are obtained. Sometimes additional information about order or spectral parameters is available which can then be used to narrow the search range and accelerate the convergence process. Owing to the enormous advances in computer technology and parallel processing the analysis of complex anisotropic NMR spectra now becomes eminently feasible. We shall deal with several convincing examples, *viz.* the solutes azulene, biphenylene, *p*-bromo-biphenyl and butane.

Despite the successes that we obtained by employing the GA-method, a weakness of the procedure is that it represents a global optimizer without any strategy. This becomes a limitation when the parameter search space gets very extensive. Currently the development of Evolutionary Algorithms (EA's) is an active and fruitful area of research. From the literature[27] it is known that novel Evolutionary Strategies (ES's) can be very successful in improving the search for an optimal solution. Hence, we decided to explore and implement different EA's as well. We shall demonstrate how successful this approach is when studying *n*-pentane as a solute in different nematic phases.

1.4.1. *Limiting parameter search space*

With a large number of spectral parameters it is important to limit the parameter search space in the fitting procedure as much as one can. Any knowledge available from independent sources about possible parameter values and their uncertainties is highly desirable. Since in practice the experimental spectra are dominated by the direct dipolar couplings, we shall focus on restricting their possible values by any means available.

1.4.2. *The GA-method*

A fit using genetic algorithms (GA) mimics the concepts of natural reproduction and selection processes. For a detailed description of the GA the reader is referred to the original literature on evolutionary or genetic algorithms.[27–29] A detailed description of the GA used in this investigation can be found in Refs. 30–32.

The molecular parameters are encoded in binary or real type, each parameter to be optimized representing a gene. The vector of all genes, which contains all molecular parameters, is called a chromosome. In an initial

step the values of all parameters are set to random values between lower and upper limits which are chosen by the user. The quality of the solutions then are evaluated by a fitness function.A proper choice of this fitness function is of vital importance for the success of the GA convergence. In Ref. 30 and 31 the fitness function F_{fg} has been defined as:

$$F_{fg} = \frac{(\boldsymbol{f}, \boldsymbol{g})}{\|\boldsymbol{f}\| \, \|\boldsymbol{g}\|}. \tag{1.5}$$

Here $\boldsymbol{f}$ and $\boldsymbol{g}$ are the vector representations of the experimental and calculated spectrum, respectively. The inner product $(\boldsymbol{f}, \boldsymbol{g})$ is defined with the metric $\boldsymbol{W}$ which has the matrix elements $W_{ij} = w(|j - i|) = w(r)$ as:

$$(\boldsymbol{f}, \boldsymbol{g}) = \boldsymbol{f}^T \boldsymbol{W} \boldsymbol{g}, \tag{1.6}$$

and the norm of $\boldsymbol{f}$ as $\|\boldsymbol{f}\| = \sqrt{(\boldsymbol{f}, \boldsymbol{f})}$; similar for $\boldsymbol{g}$. For $w(r)$ we used a triangle function[30] with a width of the base of Δw:

$$w(r) = \begin{cases} 1 - |r| \, / \left(\frac{1}{2}\Delta w\right) & \text{for } |r| < \frac{1}{2}\Delta w \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

The above defined fitness function is able to smooth in a controlled way the error landscape and therefore allows the GA to locate the global minimum. The width of the function $w(r)$ critically determines the ability of the GA to converge to the global minimum and also the speed of convergence. The smoothing of the error landscape allows to sense regions far from the minimum. The GA convergence is obtained in a well-defined procedure. At first the function $w(r)$ should be chosen relatively broad; $\Delta w \approx 15 - 20$ times the line width of an individual NMR transitions in the spectrum. In this way, a first set of parameters is obtained, which still has to be refined. This is done by decreasing Δw and narrowing the limits of the parameter space to be searched in the fit. Decreasing Δw improves the accuracy of the molecular parameters obtained from the fit, while narrowing the parameter space leads to an improved sampling in the region of the minimum. This of course is critical in the procedure, but can in most cases be done automatically. In a final calculation Δw is set to zero. Usually full GA convergence to the best set of parameters is achieved by narrowing Δw to zero in one or two steps.

One optimization cycle, including evaluation of the fitness of all solutions, is called a generation. Pairs of chromosomes are selected for reproduction and their information is combined via a crossover process. Since crossover combines information from the parent generations, it basically explores the error landscape. The values of a small number of bits in the

binary encoding of the molecular parameters is changed randomly by a mutation operator. Mutation can be viewed as exploration of the fitness surface. The best solutions within a generation are excluded from mutation. This elitism prevents already good solutions from being degraded. Mutation prevents the calculation from being trapped in local minima, as is often the case with more conventional fitting routines.

The GA described above together with the library PGAPack version 1.0[33] are very suited for massive parallel computation. The larger the starting population for a given problem, the faster is the convergence in terms of generations. On the other hand a large starting population tends to slow the algorithm down due to computational demands. A large number of fast processors will circumvent this problem and will lead to fast and straightforward assignments of the spectra. The computational speed is inversely proportional to the number of processors. Therefore this kind of algorithm is perfectly applicable for parallel processing. Modern computer cluster systems make these calculations very feasible, even for complex spectra.

The automatic GA fitting of anisotropic NMR spectra represents one step up in complexity compared to the fitting of high-resolution Laser Induced Fluorescence (LIF) spectra[30–32] that was performed with the same GA procedure before. The reason is that the LIF spectra often show some kind of regularity. This is in general completely absent in anisotropic NMR spectra because of the complex nature of the spin Hamiltonian (Eq. (1.1)).

1.5. Results obtained with the GA-method

In order to test the validity of the GA-method for very complicated ^{1}H NMR spectra of solutes dissolved in nematic phases, we decided to consider molecules of increasing size and flexibility requiring strategies for spectral analysis of increasing sophistication. We started with the relatively simple case of the planar 'rigid' molecules azulene ($C_{10}H_8$) and biphenylene ($C_{12}H_8$), see Fig. 1.1, each possessing C_{2v} symmetry. Both solutes were dissolved in a 'magic mixture' composed of 55 wt% Merck ZLI 1132 / EBBA. The orientational order of each solute can be described by two independent order parameters. Both solutes had been studied before and their eight-spin ^{1}H NMR spectra analyzed. Azulene had been studied in the nematic liquid crystal EBBA and its ^{1}H spectrum analyzed by conventional means.[34] Biphenylene had been studied in the 55 wt% Merck ZLI 1132 / EBBA 'magic mixture' by employing MQ methods.[35] The results of these previous studies can be used as guidance.

In our GA-fits on azulene and biphenylene a two-step procedure was implemented. To start with, a relatively crude geometrical structure of the solute was assumed. In the first step a GA-fitting procedure was performed in which two independent order parameters and the chemical shielding parameters were varied until the best possible correspondence with the experimental spectrum was obtained. Owing to the approximations about the structure made in this first step, complete agreement could of course not be achieved. However, in practice the choice of geometry does not have to be perfect in order to obtain sufficient common features between calculated and experimental spectra to proceed from there. After the best but far from perfect fit of step one was obtained, approximate dipolar couplings were calculated. In a second GA-procedure, these approximate dipolar couplings were then used as starting values for a fit in which all dipole-dipole couplings and chemical shielding parameters were varied until convergence was reached. For both solutes exceptionally good fits to the experimental spectra were obtained after step two, because no constraints were placed on the dipolar coupling and chemical shielding parameters. When we compare the dipolar couplings to those obtained in earlier work which was performed under slightly different experimental conditions, the correspondence is pleasing, giving us confidence that the GA-method can be of great value for the analysis of very complicated high-resolution NMR spectra.

Next, we turned to a solute that possesses more than one configuration, with rapid interconversion between them on the NMR time scale. As an illustrative example we focused on the nine-spin system *p*-bromo-biphenyl (see Fig. 1.2) dissolved in three different nematic phases, *viz.* Merck ZLI 1132, EBBA, and the 55 wt% Merck ZLI 1132 / EBBA 'magic mixture'. This solute interconverts rapidly between two symmetry-related conformers and had not been studied before. The orientational order of one conformer can be described by the three S-parameters S_{xx}, S_{zz} and S_{yz} where the molecular axes are defined with the x-axis along the inter-ring CC-bond and the y and z-axes in planes bisecting the dihedral angles between the rings. Apart from a change of sign in S_{yz}, the orientational order of the second conformer can be described by the same parameters. Since a description of the orientational order of the related biphenyl only requires two order parameters by symmetry, S_{yz} is taken to be close to zero. Inter-ring couplings are averaged by rotation about the central CC bond and thus the molecule has 16 independent dipolar couplings and 5 chemical shielding parameters.

Again a two-step procedure was performed. In the first step an estimated geometrical structure of the solute was taken and two independent order parameters, chemical shielding parameters and the dihedral angle between the two rings were varied in a GA-fitting procedure until the best correspondence with the experimental spectrum was obtained. With the approximations made, complete agreement could not be expected. However, the correspondence was still good enough for approximate dipolar couplings to be calculated. These then served as starting values in a second GA-calculation in which all dipole-dipole couplings and chemical shielding parameters were varied until convergence was reached. In all three liquid-crystal solvents employed exceptionally good agreement between experimental and fitted ^{1}H spectra was obtained.[36]

For this problem a modern cluster of PC's was used. The CPU time required amounted to $\sim$ 64 minutes per GA-run, which with 16 processors came to 4 minutes wall clock time per run. A typical result is presented in Fig. 1.10.

The next level in complexity consists of a solute that undergoes conformational change between symmetry-unrelated conformers. In order to see whether GA-fitting would be successful under these challenging conditions, we addressed the example of butane (C_4H_{10}, see Fig. 1.3) dissolved in the 55 wt% Merck ZLI 1132 / EBBA 'magic mixture'. As in the case of biphenylene, butane had been studied before employing MQ methods.[16] In fact, for our GA-analysis we used the good-quality single-quantum spectrum of butane in 'magic mixture' (Fig. 1.9) measured in this earlier work.

The conformers of butane can be approximated with the rotational isomeric state (RIS) model.[37] In this model each C–C torsion bond is assumed to exist in three local arrangements, *trans* (t), *gauche* plus (g_+) and *gauche* minus (g_-), with dihedral angles of 0°, and roughly +112° and −112°, respectively, corresponding to the angles at the minima of the rotational potential profile. These angles are close to 120°. In principle, butane can exist at other dihedral angles, with the extremes being the eclipsed and staggered forms. However, since steric hindrance is larger in the eclipsed configuration, the staggered forms are strongly favoured. In butane the possible staggered conformations are anti or t (C_{2h} symmetry), g_+ (C_2 symmetry) and g_- (C_2 symmetry), the latter two being related by symmetry. Each independent conformer requires three independent order parameters, leading to a total of six. There is an energy difference E_{tg} between *trans* and *gauche* conformers. The rapid interconversion between these conformers on the NMR time scale leads to a total of seven observable direct dipolar cou-

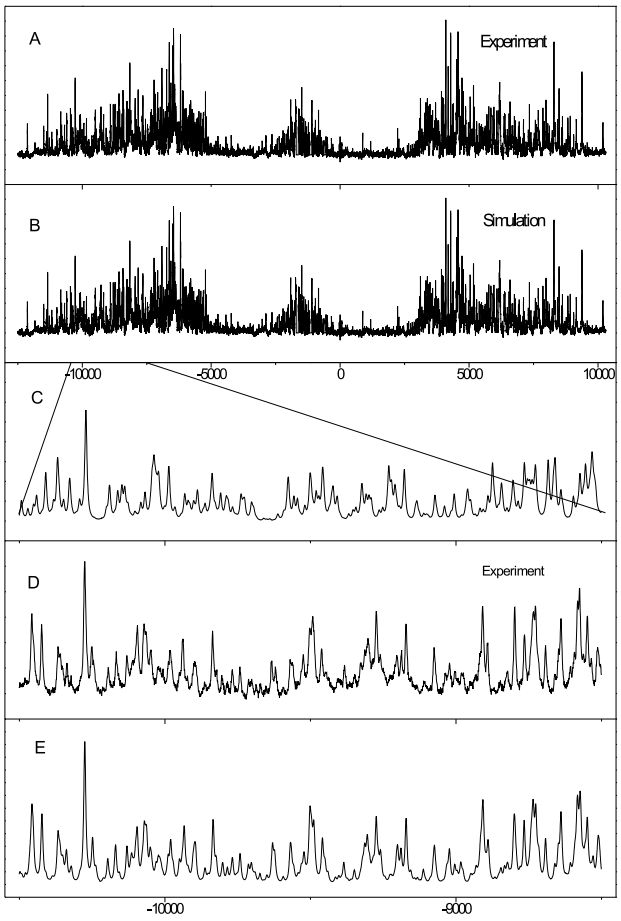

Fig. 1.10. ^{1}H NMR spectrum of *p*-bromo-biphenyl dissolved in 'magic mixture'. The experimental spectrum (A) is compared with the final fit obtained by varying 16 dipolar couplings and 5 chemical shieldings (B); in the next frames the spectrum is enlarged to show the fit obtained by varying 2 S-parameters, 5 chemical shieldings and the dihedral angle (C), the experimental spectrum (D), and the fit obtained by varying 16 dipolar couplings and 5 chemical shieldings (E). All horizontal scales are in Hz. (Reproduced from Ref. [36] with permission from Elsevier Science Publishers B.V.)

plings that are averaged over the internal motions, as well as two chemical shielding parameters.

Since conformational probabilities and order parameters only occur in Eq. (1.3) in product form, we need independent information to achieve

a separation between these quantities. For solutes dissolved in a 'magic mixture' the degree of orientational order depends primarily on the size and shape of the solute molecule. The choice of size and shape model is not essential, and we employ the so-called Circumference-Integral (CI) approach that has been used successfully in a study of 46 unrelated solutes in the magic mixture. We fix the ratio of the two interaction parameters of the model to the value 23.529 found in an earlier study of 46 solutes in the magic mixture.[38] The value chosen for the one remaining parameter (a force constant in the CI-model) k is the one that leads to a calculated butane spectrum that possesses the same width as the observed one. In order to proceed, we make a reasonable estimate of the geometries of the different butane conformers and we use the CI-model to obtain relatively crude starting values for the order parameters associated with the different butane configurations.

With a two-step approach similar to that discussed above, with order tensors obtained with the CI-model for every conformer, and with an E_{tg} energy difference estimated from other sources, a set of trial dipolar couplings was established for the first step of the fitting procedure. First, a GA-fit was performed on chemical shieldings and order parameters until some resemblance between experimental and calculated spectra was obtained. From this tentative fit for the order tensors, dipolar couplings were estimated that served as starting parameters for the next step. In step two in a new GA-fitting procedure the chemical shieldings and all the dipolar couplings were varied until excellent agreement between experimental and calculated spectra was obtained (see Fig. 1.11). The dipolar couplings and chemical shieldings obtained from our GA-fitting procedure proved to be virtually identical to those derived previously from the same experimental spectrum following an extensive MQ-approach.[16] This pleasing result lends additional credence to the power of GA-fitting.

Considering the ease with which spectra of butane dissolved in a liquid-crystal solvent can now be reliably analyzed by employing the GA-method, we have recently performed an extensive study of butane in an anisotropic solvent as a function of temperature. In a detailed computer simulation effort in which the solute-solvent interaction is carefully taken into account the degree of orientational order of the various solute conformers can be predicted and the results compared to our experiments. It can be expected that such studies will lead to a better understanding of conformational change in the condensed phase than is currently available.[39]

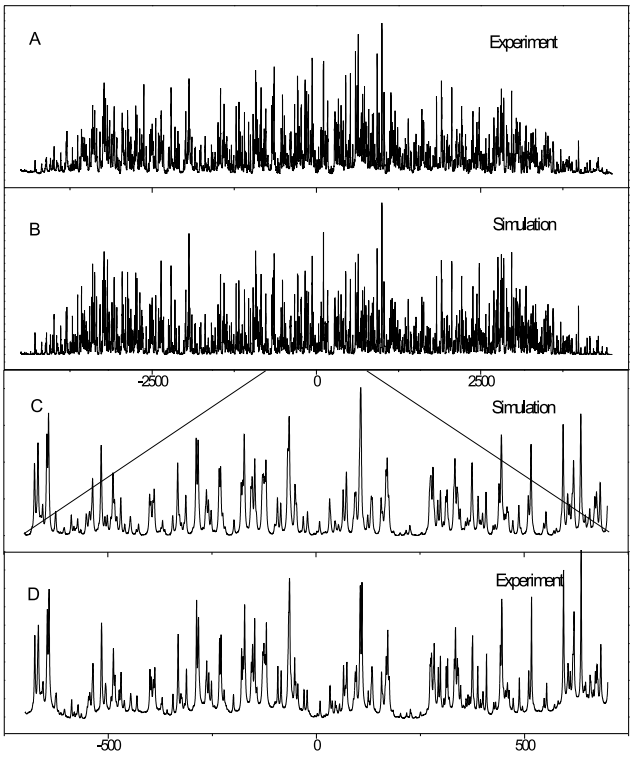

Fig. 1.11. ^{1}H NMR spectrum of butane dissolved in 'magic mixture'. The experimental spectrum (A) is compared with the final fit obtained by varying 7 dipolar couplings and 2 chemical shieldings (B); in the next frames the spectrum is enlarged to show the experimental spectrum (D), and the fit (C). All horizontal scales are in Hz.

In summary, the robustness of the GA-fitting method has been demonstrated for solutes that belong to various categories of increasing complexity:[40] rigid solutes with a single conformation (*e.g.* the eight-spin systems azulene and biphenylene), solutes interconverting between a number of symmetry-related conformers (*e.g. p*-bromo-biphenyl), and solutes undergoing conformational change between non-symmetry-related conformers (*e.g.* the ten-spin system butane). In the following section the even more challenging example of *n*-pentane dissolved in different nematic phases will be treated. This will require the use of Evolutionary Algorithms with a higher degree of sophistication than the relatively simple GA-method.

1.6. Spectral analysis using Evolutionary Algorithms

The ^{1}H NMR spectrum of pentane (see Fig. 1.3) dissolved in Merck ZLI 1132 (shown in Fig. 1.12) has no easily recognizable patterns that could aid the spectral analysis. A visual inspection of the spectrum gives no clue whatsoever of the magnitudes of the dipolar couplings. When applying an EA to such a problem it should be realized that the computational time increases very steeply with the search ranges used for every parameter. It is therefore prudent to choose initial values and ranges for each parameter that are as narrow as possible, but still wide enough to include the correct value. In the following we shall describe a strategy to do just that.

High-resolution NMR spectral values were used for indirect spin-spin couplings which were initially assumed to remain unchanged in the spectrum of the orientationally ordered solute. High-resolution NMR values were also used as initial estimates of chemical shifts and the search ranges were chosen to be consistent with centering the calculated with the experimental spectrum.

Pentane exists in several conformers that interchange rapidly on the NMR time scale. As with butane, initial estimates of the eleven independent dipolar couplings were quite problematic because they depend on conformer structures, order tensors and conformer probabilities. Again, we assumed that the conformer problem can be described adequately by the rotational isomeric state (RIS) approximation.[37] For n-pentane there are then nine RIS conformers, some being related by symmetry. The four independent conformers are tt (C_{2v}, singly degenerate), tg (C_1, four-fold degenerate), g_+g_+ (C_2, two-fold degenerate) and g_+g_- (C_s, two-fold degenerate). It is necessary to assign values for conformer probabilities. The extra energy involved in a *gauche* conformer E_{tg} was set to the value 650 cal, with the weight of the g_+g_- conformer set to zero because of the additional energy involved with the steric interaction of the two methyl groups in this configuration (the 'pentane' effect).[37] For the structure of each conformer we used the same bond angles and bond lengths that were used for a previous analysis of butane.[16]

Finally, we required estimates of the pentane conformer order tensors in the two liquid-crystal solvents 'magic mixture' and 1132. Solute orientational order in 'magic mixture' is dominated by a single orientational mechanism that depends on solute size and shape alone.[4,5] Hence, phenomenological size and shape models are appropriate for the prediction of conformer order parameters in this case. For 1132 the situation is some-

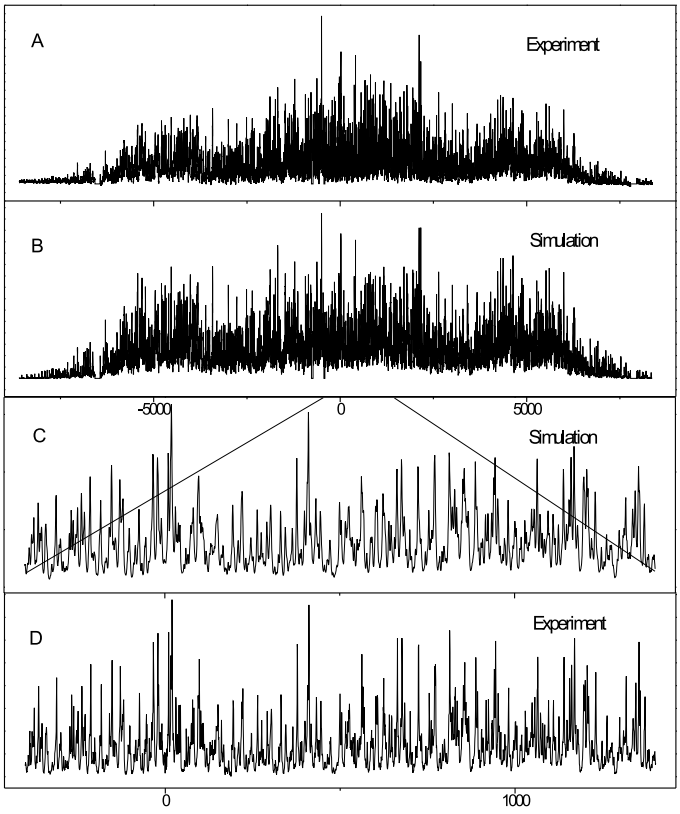

Fig. 1.12. ^{1}H NMR spectrum of pentane dissolved in 1132. The experimental spectrum (A) is compared with the final fit obtained by varying 11 dipolar couplings, 9 indirect spin-spin couplings and 3 chemical shieldings (B); in the next frames the spectrum is enlarged to show the experimental spectrum (D), and the fit obtained by varying 11 dipolar couplings, 9 indirect spin-spin couplings and 3 chemical shieldings (C). All horizontal scales are in Hz. The 11 indirect spin-spin couplings were varied only in the final fitting procedure; in earlier iterations they were given their isotropic values. (Reproduced from Ref. [41] with permission from the American Institute of Physics.)

what different, because at least one additional mechanism that depends on the anisotropic electrostatic properties of the solute contributes to the orientational potential. However, in the case of alkanes these additional mechanisms are considered to be of minor importance and can be neglected.[42] Thus the moniker 'magic solute' is often used for alkanes. Hence, size and shape models (with the interaction constant adjusted to account for changes

in liquid-crystal solvent and temperature) are probably also appropriate to describe the orientational order of pentane in 1132. These considerations are supported by literature data on propane[43] for which the ratios of the two propane order parameters are approximately equal in the liquid crystals 'magic mixture' and 1132. As in the case of butane, we employ the so-called CI-model[38] to obtain order parameters for the various pentane conformers to probably better than 10%.[4,5] Again, we fix the ratio of the two interaction parameters of the model to the value 23.529.[38] The value chosen for the one remaining parameter k is the one that leads to a calculated pentane spectrum that possesses the same width as the observed one.

Spectra calculated using the order parameters estimated above bore resemblance to the experimental ones in the sense that differences were appreciable, but the overall underlying features were reasonable. An initial, unsuccessful attempt with the GA fitting procedure was performed using (and varying within reasonable limits) the above chemical shifts, order parameters for each conformer, and conformer probabilities. Such parameters cannot be expected to give an exact solution because they are based on models and guesses, and hence may have significant uncertainties. The search was abandoned in favour of one that varied the eleven independent dipolar couplings and three chemical shifts that could, in principle, yield an exact match to the experimental spectrum. Indeed there are fewer independent dipolar couplings than conformer order parameters, giving further justification for basing the second attempt on the dipolar couplings.

In order to estimate reasonable initial values and ranges for the dipolar couplings, spectra were calculated for various values of order parameters ($\pm 10\%$ of the values above), E_{tg} (between 400 and 1000 cal/mol) and the dihedral angle for a *gauche* rotation (between 112° and 120°). The highest and lowest values of each dipolar coupling resulting from these calculations provided the search ranges for the dipolar couplings. These ranges seemed consistent with spectra that reproduce reasonably the overall features and width of the experimental one.

We then started a fit using the GA-method which was successful in unraveling the spectrum of the nine-spin system *p*-bromo-biphenyl[36] and the ten-spin system butane.[40] Initially this was unsuccessful for the spectrum of pentane in 1132 for a number of reasons. The uncertainties in some of the larger dipolar constants were still quite large. As a consequence the GA procedure was not able to converge. Probably the main reason for this failure comes from the fact that the GA procedure is a global optimizer

without any strategy. If the search ranges become too large there is insufficient coverage of the parameter space to locate the global minimum. We solved this problem initially by breaking up the uncertainty regions for a number of parameters into regions so small that we were certain the GA would converge. The size of the smaller search regions was determined with test spectra. This was done for the dipolar couplings with the largest uncertainties. Now a grid search was set up in which the GA was applied in the different subregions, until a subregion was found for which the GA succeeded in reproducing the experimental spectrum. In this way we obtained a calculated spectrum that is essentially a perfect fit to the experimental one.[41]

It is clear that this grid search process rapidly will become very expensive. If for example n uncertainty regions are divided into m subregions the size of the grid will be m^n, a number that increases steeply with n and m. We therefore decided to implement different Evolutionary Algorithms. From the literature[27] it is known that Evolutionary Strategies (ES's) can be very successful in improving the search for an optimal solution.

The ES algorithm starts with one or more parents. A parent is a trial solution that corresponds to a vector of parameters like in the genetic algorithm, with the difference that each vector has an extra set of parameters corresponding to the type of strategy used. This parent generates offspring. The quality or performance of these children is checked and, depending on the strategy, the next parent is generated. There are several different strategies for the generation of the offspring as well as the generation of the next parent. For a more detailed description see the literature.[27]

The offspring is generated from the parents in a mutative step-size fashion. The magnitude and direction of the steps are dependent on previous history and on the quality of offspring generated. A drawback of the standard ES is the fact that the mutations of the decision and the strategy parameters are subject to independent random processes. If for example an individual with a large step size undergoes only a very small change in the decision parameters and this small change turns out to yield a high fitness, the large step size will be inherited by the next generation. As a result, the fitness in the next mutations may worsen. This problem is resolved in derandomized (DR) algorithms which make the random mutations in decision and strategy parameters dependent on each other. This idea was implemented first as DR1 and soon improved by the concept of accumulated information,[44] which is called DR2. The history of the optimization is recorded and the evolution of the mutation ellipsoid is partially governed by past successful mutations.

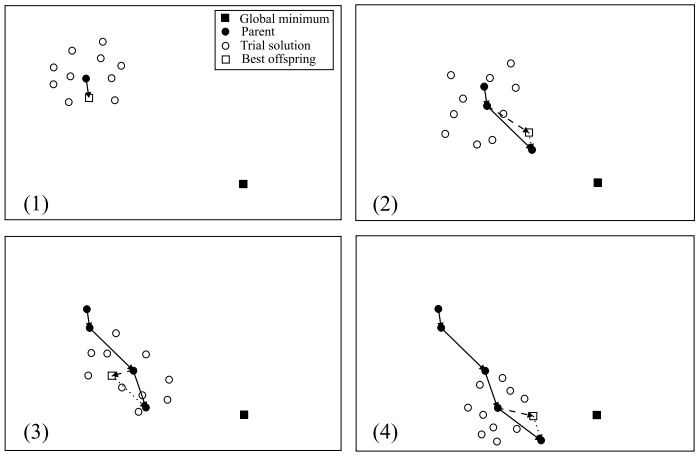

Fig. 1.13. The first four generations of an Evolutionary Strategy: (1) An initial population is generated, and the best offspring is used as the next parent. (2) The offspring is spread over a larger area in the second generation due to the relatively large step made in the previous generation. The vector from the parent to the best offspring (dashed line) is combined with the (shortened) mutation vector of the last generation (dotted line) to generate the new parent (solid line). (3) Due to the correlation between the past two mutations the search range has been extended again in the general direction of both mutations while it has been limited in the perpendicular direction. The best offspring is now a local minimum. The memory effect of the evolutionary algorithm, which incorporates past mutation vectors into the calculation of the next parent, helps to overcome the local minimum and the next parent is still closer to the global minimum. (4) The barrier between the local and global minima has been overcome, and the optimization is progressing towards the global minimum. (Reproduced from Ref. [41] with permission from the American Institute of Physics.)

A further improvement was achieved by Hansen and Ostenmeier[45] with the Covariance Matrix Adaptation Evolution Strategy (CMA-ES). It turns out to be a particularly reliable and highly competitive evolutionary algorithm for local optimization and, surprisingly at first sight, also for global optimization.[46] The CMA-ES does not leave the choice of strategy parameters open to the user — only the population size can be set. Finding good strategy parameters is considered to be part of the algorithm design.

Figure 1.13 depicts the first four generations of an evolutionary strategy and demonstrates the effect of the chosen strategy. In general the Evolution Strategies converge faster and are more robust than the genetic algorithm.

Next, we implemented this algorithm into our programme, selected the estimated uncertainty regions discussed above for the dipolar couplings,

and performed a fit of the spectrum of pentane in 'magic mixture'. This spectrum was even harder to deal with than that of pentane in 1132, because the density of lines is larger (smaller dipolar couplings) and the signal-to-noise ratio was smaller, while at the same time the background removal was more difficult. This background is due to an abundance of unresolved dipolar splittings from the liquid-crystal molecules themselves that typically contain of order 20 ^{1}H spins. It turned out that the CMA-ES procedure was immediately successful in obtaining an essentially perfect fit of the observed spectrum and determining the parameters unambiguously. At that point there was only one small glitch left. Due to the large density of lines it was relatively difficult to remove completely the background from the observed spectrum. In spite of this the CMA-ES succeeded in fitting all parameters and transitions. Improvement of the accuracy in the final values of the parameters can be realized by a better way of dealing with the background.[41]

With our initial fit we were in the position to carry out an almost perfect removal of the background from the experimental spectrum. This was performed in the following way. The calculated spectrum resulting from the CMA-ES was subtracted from the experimental one. To remove the high-frequency noise in this difference a degree of smoothing was applied. Finally this smoothed difference was subtracted from the original experimental spectrum. In this way an experimental spectrum was obtained which is almost background-free. The result for part of the spectrum is shown in Fig. 1.14. Using this approach we were able to improve the parameters in a final CMA-ES fit. In addition, it was possible to determine accurately the small indirect spin-spin couplings $J_{\mu\nu}$ between nuclei μ and ν.[41]

The background removal procedure as described in the previous paragraph was repeated for the spectrum of pentane in 1132 (Fig. 1.12) and again we were able to fit the indirect spin-spin coupling constants. Within estimated error the results for the $J_{\mu\nu}$ for pentane in 1132 and in magic mixture were identical, as expected.

What was left to calculate were the statistical errors in the parameters. This is in general a difficult task for EA's. However, in our case it was relatively simple. The converged result of the EA not only produces the best values for the fit constants, but at the same time the quantum numbers of the individual transitions are assigned. This allows a classical least-squares fit using the assigned frequencies. In Meerts and Schmitt[32] this is called

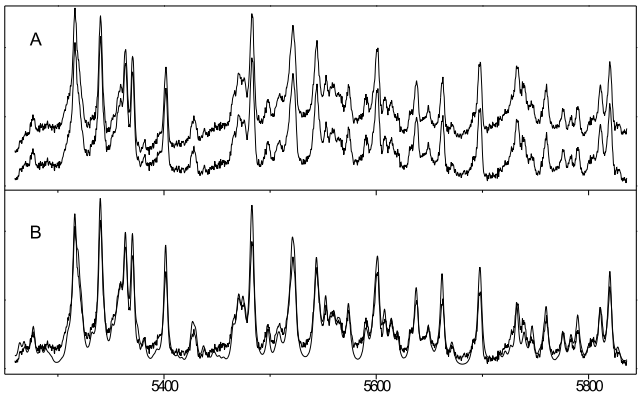

Fig. 1.14. A small part of the ^{1}H NMR spectrum of pentane dissolved in a magic mixture. This shows the effect of the background removal method discussed in the text. (A) The upper and lower traces show the experimental spectra without and with background removal, respectively. (B) The noise free trace is the calculated spectrum and the other the experimental spectrum after removal of background. All horizontal scales are in Hz. (Reproduced from Ref. [41] with permission from the American Institute of Physics.)

an 'assigned fit' and the definitions of the statistical errors and correlation coefficients are discussed in Appendix B of that paper.

1.7. Conclusions

NMR-studies of solutes dissolved in anisotropic solvents provide a rich source of information about molecular structure in the condensed phase. In particular, the direct dipolar couplings in liquid-crystal environments are not averaged to zero, and usually dominate the NMR-spectra. Since these direct dipolar couplings depend on distances between nuclear spins, obtaining these quantities from experiment is an essential step. Application of the liquid-crystal technique to solutes with low symmetry and with more than approximately 8 spins is seriously hampered by the fact that the single-quantum NMR-spectrum shows a multitude of often overlapping transitions without any apparent structure or pattern. Analyzing such spectra and extracting the dipolar couplings from them is a challenging task that for many years appeared to present insurmountable problems. The difficulties encountered were exacerbated in the case of solutes that interconvert among several symmetry-unrelated conformers. For many years

the extension of the liquid-crystal method to larger solutes was therefore considered 'impossible'.

In this chapter we focus both on recent experimental and novel theoretical methods to address the central problem of spectral analysis. In recent years the development of multiple-quantum NMR has significantly increased the possibilities of liquid-crystal NMR. High-order multiple-quantum spectra are much simpler than single-quantum ones, and their analysis provides an approximate determination of the direct dipolar couplings that contain most of the relevant information. These estimates can then be employed as a starting point for the detailed analysis of the high-resolution single-quantum spectrum. Various sophisticated versions of multiple-quantum NMR have been developed and successfully applied to the study of relatively large and complicated solutes dissolved in liquid-crystals. All these advanced techniques have the drawback that they require considerable effort and NMR-expertise.

Recently a novel approach has been developed in which mathematical algorithms from evolutionary biology (Evolutionary Algorithms, EA's) were applied to NMR of solutes in anisotropic environments. These methods comprise applications of a relatively simple Genetic Algorithm (GA), as well as the use of more sophisticated Evolutionary Strategies (ES's). These techniques have been described in some detail and have been applied to a number of representative solutes, ranging from 'rigid' solutes with eight spins to the flexible butane and even the pentane molecules. The availability of enormous computer power, and the fact that dealing with the Hamiltonian problem in liquid-crystal NMR lends itself to extensive parallel programming, are the key elements why these spectral fitting techniques are eminently successful. The convergence progress can be accelerated considerably if the search ranges for the spectral parameters can be limited in a sensible fashion. *A priori* estimates of the degree of orientational order of solutes in liquid crystals to a level of better than 10%, based on simple phenomenological size and shape models, are very effective in this context.

In summary, through the application of modern experimental multiple-quantum methods and novel spectral analysis techniques based on evolutionary strategies, the liquid-crystal NMR-method can now be extended to much larger and more complicated solutes than previously possible. Introduction of these methods has therefore given liquid-crystal NMR a new lease of life.

Acknowledgments

E.E.B. acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada. W.L.M. thanks the National Computer Facilities of the Netherlands Organization of Scientific Research for a Grant on the Dutch supercomputing facility SARA.

References

1. P. Diehl and C. L. Khetrapal, *NMR Basic Principles and Progress*. vol. *1*, (Springer-Verlag, Berlin, 1969, p. 1).
2. J. W. Emsley and J. C. Lindon, *NMR Spectroscopy using Liquid Crystal Solvents*. (Pergamon Press, Oxford, 1975).
3. E. E. Burnell and C. A. de Lange, *Chem. Phys. Letters*. **76**, 268, (1980).
4. E. E. Burnell and C. A. de Lange, Prediction from molecular shape of solute orientational order in liquid crystals, *Chemical Reviews*. **98**, 2359, (1998).
5. E. E. Burnell and C. A. de Lange (Eds.), *NMR of Ordered Liquids*. (Kluwer Academic, Dordrecht, The Netherlands, 2003).
6. D. P. Weitekamp, *Adv. Magn. Reson.* **11**, 111, (1983).
7. J. C. T. Rendell, Ph. D. thesis, University of British Columbia, 1987.
8. R. T. Syvitski, N. Burlinson, E. E. Burnell, and J. Jeener, *J. Magn. Reson.* **155**, 251, (2002).
9. D. N. Shykind, J. Baum, S. B. Liu, and A. Pines, *J. Magn. Reson.* **76**, 149, (1988).
10. L. W. Field and M. L. Terry, *J. Magn. Reson.* **69**, 176, (1986).
11. XWIN-NMR Software Manual, Bruker Analytik GmbH, 2000.
12. J. C. T. Rendell and E. E. Burnell, *J. Magn. Reson.* **A 112**, 1, (1995).
13. J. C. T. Rendell and E. E. Burnell, *Mol. Phys.* **90**, 541, (1997).
14. G. A. Morris and R. Freeman, *J. Magn. Reson.* **29**, 433, (1978).
15. T. Chandrakumar, J. M. Polson, and E. E. Burnell, *J. Magn. Reson.* **A 118**, 264, (1996).
16. J. M. Polson and E. E. Burnell, *J. Chem. Phys.* **103**, 6891, (1995).
17. P. Diehl, S. Sýkora, and J. Vogt, *J. Magn. Reson.* **19**, 67, (1975).
18. P. Diehl and J. Vogt, *Org. Magn. Reson.* **8**, 638, (1976).
19. D. S. Stephenson and G. Binsch, *J. Magn. Reson.* **37**, 395, (1980).
20. D. S. Stephenson and G. Binsch, *J. Magn. Reson.* **37**, 409, (1980).
21. D. S. Stephenson and G. Binsch, *Org. Magn. Reson.* **14**, 226, (1980).
22. D. S. Stephenson and G. Binsch, *Mol. Phys.* **43**, 697, (1980).
23. F. Castiglione, M. Carravetta, G. Celebre, and M. Longeri, *J. Magn. Reson.* **132**, 1, (1998).
24. F. Castiglione, G. Celebre, G. De Luca, and M. Longeri, *J. Magn. Reson.* **142**, 216, (2000).
25. H. Takeuchi, K. Inoue, Y. Ando, and S. Konaka, *Chem. Lett.* **11**, 1300, (2000).
26. K. Inoue, H. Takeuchi, and S. Konaka, *J. Phys. Chem. A*. **105**, 6711, (2001).

27. I. Rechenberg, *Evolutionsstrategie - Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. (Frommann-Holzboog, Stuttgart, 1973).
28. J. H. Holland, *Adaption in Natural and Artificial Systems*. (MI: The University of Michigan Press, Ann-Arbor, 1975).
29. D. E. Goldberg, *Genetic Algorithms in search, optimisation and machine learning*. (Addison-Wesley, Reading Massachusetts, 1989).
30. J. A. Hageman, R. Wehrens, R. de Gelder, W. L. Meerts, and L. M. C. Buydens, *J. Chem. Phys.* **113**, 7955, (2000).
31. W. L. Meerts, M. Schmitt, and G. Groenenboom, *Can. J. Chem.* **82**, 804, (2004).
32. W. L. Meerts and M. Schmitt, *Int. Rev. Phys. Chem.* **25**, 253, (2006).
33. D. Levine. PGAPack V1. 0, PgaPack can be obtained via anonymous ftp from: ftp://ftp.mcs.anl.gov/pub/pgapack/pgapack.tar.z, (1996).
34. A. D. Hunter, E. E. Burnell, and T. C. Wong, *J. Molec. Structure.* **99**, 159, (1983).
35. J. M. Polson and E. E. Burnell, *J. Magn. Reson.* **A 106**, 223, (1994).
36. W. L. Meerts, C. A. de Lange, A. C. J. Weber, and E. E. Burnell, *Chem. Phys. Lett.* **441**, 342, (2007).
37. P. J. Flory, *Statistical Mechanics of Chain Molecules*, Wiley-Interscience, New York, 1969.
38. D. S. Zimmerman and E. E. Burnell, *Mol. Phys.* **78**, 687, (1993).
39. R. Berardi, E. E. Burnell, C. A. de Lange, W. L. Meerts, A. C. J. Weber, and C. Zannoni. Work in progress.
40. A. C. J. Weber, E. E. Burnell, W. L. Meerts, and C. A. de Lange. To be published.
41. W. L. Meerts, C. A. de Lange, A. C. J. Weber, and E. E. Burnell, *J. Chem. Phys.* **130**, 044504, (2009).
42. A. F. Terzis, C. D. Poon, E. T. Samulski, Z. Luz, R. Poupko, H. Zimmermann, K. Müller, H. Toriumi, and D. J. Photinos, *J. Am. Chem. Soc.* **118**, 2226, (1996).
43. E. E. Burnell, L. C. ter Beek, and Z. Sun, *J. Chem. Phys.* **128**, 164901, (2008).
44. A. Ostermeier, A. Gawelczyk, and N. Hansen, *Step-Size Adaptation Based on Non-Local Use of Selection Information. In Parallel Problem Solving from Nature (PPSN3)*. (Springer, 1994).
45. N. Hansen and A. Ostermeier, Completely derandomized self-adaptation in evolution strategies, *Evolutionary Computation.* **9**(2), 159–195, (2001).
46. N. Hansen and S. Kern. Evaluating the CMA evolution strategy on multimodal test functions. In ed. X. Yao, *et al., Parallel Problem Solving from Nature PPSN VIII*, vol. 3242, *LNCS*, pp. 282–291. Springer, (2004).