PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Solid State n.m.r. of Biopolymers [and Discussion]

S. J. Opella, J. G. Hexem, M. H. Frey, T. A. Cross and W. Derbyshire

Phil. Trans. R. Soc. Lond. A 1981 **299**, 665-683
doi: 10.1098/rsta.1981.0041

Phil. Trans. R. Soc. Lond. A **299**, 665–683 (1981)
Printed in Great Britain

Solid state n.m.r. of biopolymers

By S. J. Opella, J. G. Hexem, M. H. Frey and T. A. Cross
Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, *U.S.A.*

Some aspects of high-resolution ^{13}C n.m.r. of solid biopolymers are described. Spectra of amino acids, peptides, proteins and DNA are shown. Differences between solution spectra and solid state spectra of a molecule include multiple lines due to multiple crystal forms, splittings of substituted rings and asymmetric splittings of carbons bonded to nitrogen. The inability of magic angle sample spinning to remove ^{13}C–^{14}N dipolar couplings is a consequence of the ^{14}N quadrupole moment's altering the angular dependence of the interactions. Solid state n.m.r. of filamentous viruses is given as an example of new biological applications of this type of spectroscopy.

Introduction

High-resolution solid state n.m.r. can have a significant impact on the formulation and solution of many research problems in the area of biological chemistry. These methods enable comparisons to be made of molecular structure and dynamics between solutions and crystalline or amorphous solids. This is often of great importance because of the heavy reliance on crystallographic structure data for interpretation of many experimental results. Solid state n.m.r. techniques allow the study of molecules that are insoluble or are denatured when solubilized. They also extend the range of n.m.r. studies to biological supramolecular structures, such as membranes or nucleoprotein complexes, that are unapproachable with conventional high-resolution solution n.m.r. because their rotational motions are not sufficient to average out the line broadening from static anisotropic spin interactions.

Solid state n.m.r. takes advantage of the ability of radio-frequency irradiations and sample spinning to manipulate static nuclear spin interactions for the purpose of line narrowing and increasing sensitivity. The promise of dilute spin double-resonance n.m.r. for complex chemical and biological systems has been obvious since the initial descriptions (Pines *et al.* 1972*b*, 1973). In particular, the combination of magic angle sample spinning and double resonance provides the means for high-resolution n.m.r. investigations of powders (Schaefer & Stejskal 1976).

At present, solid state n.m.r. is not at the same level of applicability for biological problems as is solution n.m.r. This is in part because the initial solid state n.m.r. studies of complex molecules, including those of biological origin, do not show the simple spectral correspondences expected between solid and liquid samples. The spectral complications observed in solid state n.m.r. usually lead to a loss of resolution that makes the more complex spectra less easily analysed. In order to use solid state n.m.r. effectively, the manifestation of anisotropic spin interactions in the spectra must be understood. This requires a precise physical description of the nuclear spin system. In addition to leading to methods for increasing resolution, the solid state n.m.r. effects can be valuable sources of spectroscopic data.

Experimental

Much of the motivation for the refinement of solution n.m.r. has come from the stringent sensitivity and resolution requirements of biological studies (Opella 1977). The same situation is bound to exist for solid state n.m.r. A large part of our research effort is, by necessity, devoted to optimizing instrumentation and experimental techniques for high-resolution solid state n.m.r. The basic strategy is first to remove as much of the broadening influences as possible, and then to selectively leave one interaction for study.

The high-resolution solid state ^{13}C n.m.r. spectra most readily compared with solution ^{13}C n.m.r. spectra are obtained by decoupling ^{13}C–^{1}H dipolar interactions with radio-frequency irradiation and spinning the sample at the magic angle (54.7°) with respect to the applied magnetic field. We have published brief descriptions of the home-built double-resonance spectrometer (Cross *et al.* 1979) and probes (Opella *et al.* 1980). The spectrometer has phase

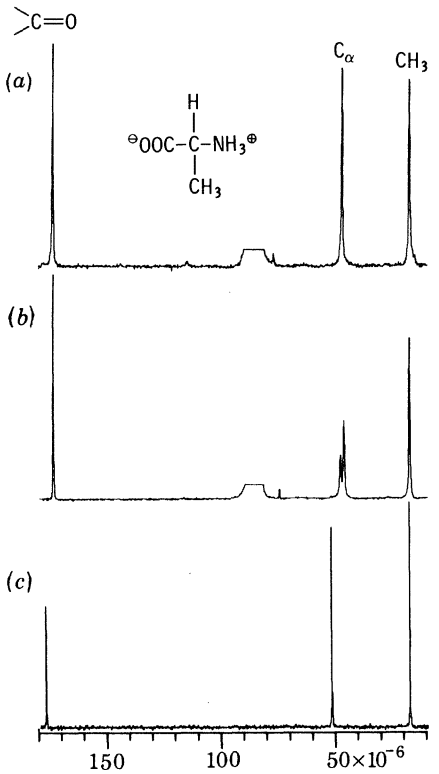

FIGURE 1. ^{13}C n.m.r. spectra of alanine. (*a*) Polycrystalline [^{15}N]Ala, obtained with 2.5 mT proton decoupling and 3.7 kHz magic angle sample spinning. (*b*) Polycrystalline [^{14}N]Ala, obtained under the same conditions. (*c*) [^{14}N]Ala in solution, obtained with modulated ^{1}H decoupling.

and frequency stability and magnet homogeneity equivalent to a high-resolution solution spectrometer. The two most critical parameters in practice for ^{13}C n.m.r. are decoupling power and spinning stability.

Figure 1 shows natural-abundance ^{13}C n.m.r. spectra of alanine. Polycrystalline alanine was packed in a conical-based rotor similar to those popularized by Andrew *et al.* (1969). The

spectra in figure 1*a*, *b* were obtained by cross-polarization of the ^{13}C magnetization from the ^{1}H spins. A ^{1}H decoupling field of 2.5 mT was applied coherently on resonance during data acquisition. Under these conditions, the downfield carbonyl line width is about 7 Hz. The simplest spectrum is for ^{15}N labelled alanine, where three well resolved resonances of equal area are seen. Even the solution spectrum (figure 1*c*) has to take into account such properties as relaxation mechanisms for different sites because of intensity differences. The solid state spectrum of alanine with natural isotopic composition (figure 1*b*) has an unusually shaped C_α resonance.

The spectra of figure 1 indicate typical resolution achievable for an amino acid, which is about 7–12 Hz line width for unsaturated carbons and 10–15 Hz for aliphatic carbons. Some other types of molecules give significantly different line widths with adamantane *ca.* 0.5 Hz, sucrose 2 Hz and amorphous polymers more than 10 Hz. Even a relatively simple molecule like alanine indicates the presence of spectral features peculiar to the solid state in its C_α resonance.

The best spectra of amino acids and peptides are obtained from samples that are carefully recrystallized. The natural-abundance samples are from Sigma Biochemicals. The ^{15}N and ^{13}C labelled compounds are from KOR Isotopes. The cyclic peptide was synthesized by Dr L. G. Pease of the University of Delaware.

Polycrystalline amino acids and peptides

We have carried out extensive studies on the ^{13}C n.m.r. properties of amino acids and small peptides in powder form (Pease *et al.* 1981; Frey & Opella 1980). This was done to develop the requisite background for protein studies as well as to take advantage of any empirical trends to help explain spectral features. Spectroscopic properties found for amino acids are also of general interest because they are a consequence of chemical functionality.

One of the most important features of ^{13}C n.m.r. spectra of biological molecules in the solid state is the unusual line shapes observed for carbons bonded to nitrogen (Lippmaa *et al.* 1979; Hexem *et al.* 1981; Groombridge *et al.* 1980; Frey & Opella 1980). This is illustrated in figure 1*b*. The comparison of figure 1*a* and *b* shows that the asymmetric doublet of the alanine α carbon resonance is due to the spin 1 ^{14}N nucleus, which has a quadrupole moment, unlike the ^{15}N spin $\frac{1}{2}$ nucleus. The effect of ^{14}N on ^{13}C spectra can also be substantial for non-bonded carbons. The ^{13}C n.m.r. spectra of polycrystalline arginine (figure 2*a*) demonstrate the devastating effect of three bonded nitrogens on the line width of the guanido carbon (C_ζ) as well as more generalized line broadening for other carbons not bonded to nitrogens. The upfield aliphatic carbon resonances of arginine are significantly broader than those of most other amino acids in the solid state. The C_ζ resonance of arginine is generally useful for solution studies (figure 2*b*) because of its distinguishing chemical shift and narrow line width due to lack of nearby protons (Allerhand 1979); however, the broadening from the ^{14}N is so great that it is unlikely that such sites will be observed in solid proteins.

Splitting of carbon resonances arises from heteronuclear dipolar coupling between ^{13}C and ^{14}N spins which survive magic angle sample spinning at rotation rates that are expected to effectively suppress these interactions (Hexem *et al.* 1981). The quadrupole moment of ^{14}N modifies the angular dependence of its dipolar coupling with other spins. In this section, we present the theory showing how ^{14}N–^{13}C dipolar coupling modified by the ^{14}N quadrupolar interaction leads to asymmetric line shapes when powder samples are subjected to magic angle

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

sample spinning. Since the ^{13}C and ^{14}N spins are in different environments with respect to nuclear spin states and energies, the spin states of each nucleus are described separately.

The interactions that determine the ^{14}N spin states are the Zeeman coupling with the applied magnetic field, the electronic shielding leading to chemical shift, and the quadrupole interaction. Most nitrogen scalar couplings are small enough to be neglected in solid state

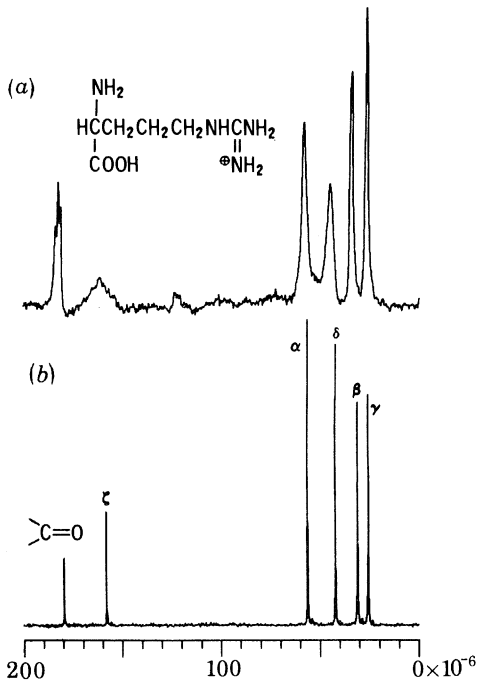

FIGURE 2. ^{13}C n.m.r. spectra of arginine. (*a*) Polycrystalline sample; (*b*) aqueous solution.

n.m.r. Relaxation effects are ignored, since if the nuclear spin states were not long lived with respect to the dipolar coupling, then 'self-decoupling' as described by Spiess *et al.* (1977) would occur, eliminating ^{13}C–^{14}N dipolar interactions and the doublet structure of the ^{13}C resonances.

The Hamiltonians of these ^{14}N interactions commute, allowing the use of eigenfunctions of the Zeeman interaction as the basis for computing the total Hamiltonian. Thus, the ^{14}N spin states are described as a linear combination of the Zeeman spin states

$$\phi_n = \sum_m C_{nm}|m\rangle; \quad n = 1, 2, 3; \quad m = +1, 0, -1. \tag{1}$$

Expectation values of the ^{14}N vector spin operators determine the ^{14}N magnetization, and these must be evaluated in terms of these new spin functions when the influence of ^{14}N spins on ^{13}C spins is described. Irreducible spherical tensor operators are appropriate for examining the effects of rotations on nuclear spin states, and in this formalism the vector spin operators are

$$\tau_{10} \equiv S_0 \equiv S_z, \tag{2}$$

$$\tau_{1,\pm1} \equiv S_{\pm1} = \mp\tfrac{1}{\sqrt{2}}S_{\pm}; \tag{3}$$

τ_{lm} are the irreducible spherical tensor operators of rank $l = 1$; S_z, S_+ and S_- are the usual vector spin operators in Cartesian coordinate systems.

In the presence of a large applied magnetic field these operators are best considered in the interaction representation. This is done by using Wigner rotation matrices and transforming to a coordinate system rotating at the ^{14}N Larmor frequency, ω_N:

$$T_{lm} = \sum_{m'} D^{l}_{m',m}(0, 0, \omega_N t)\, \tau^{N}_{1m} \tag{4}$$

$$= \tau_{lm} e^{-im\omega_N t}. \tag{5}$$

T_{lm} are the lth-rank irreducible spherical tensor components of the spin operators and $D^{l}_{m'm}$ (α, β, γ) are the lth-order Wigner rotation matrices, with arguments of the Euler angles (α, β, γ) describing the relation of the old and new coordinate systems.

Tensor coupling of the vector spin operators with the applied magnetic field, electric field gradient or electronic shielding then lead to manifestation of those interactions. Thus we write the total ^{14}N Hamiltonian

$$\mathcal{H}^{tot}_N = \sum_{\lambda} \sum_{l} \sum_{m=-l}^{+l} C^{\lambda} (-1)^{m} R^{\lambda}_{l,-m} T^{\lambda}_{lm}, \tag{6}$$

where R^{λ}_{lm} are irreducible spherical tensor components of the spatial dependence of interaction type λ (Zeeman, chemical shift and quadrupole); T^{λ}_{lm} are irreducible tensor components of the spin dependence of interaction of type λ, both written in a coordinate system in which the z axis lies along the applied magnetic field. C^{λ} is a rotationally invariant constant.

The effect of magic angle sample spinning on the anisotropic chemical shielding can simply be considered as averaging it to its isotropic value, and this can be included in the definition of ω_N or ω_C. In contrast, the effect of sample rotation on the quadrupole interaction is not straightforward. Noting that terms with $l = 0, 1$ vanish for the quadrupole interaction, the total ^{14}N spin Hamiltonian can be written

$$\mathcal{H}^{tot}_N = -\omega_N T^{Z}_{10} + C^{q} \sum_{m=-2}^{2} (-1)^{m} R^{q}_{2,-m} T^{q}_{2m}. \tag{7}$$

C^{q} is a rotationally invariant constant of the quadrupole interaction with value $eQ/2S(2S-1)\,\hbar$, eQ being the electric quadrupole moment of the ^{14}N nucleus and S its spin quantum number.

The spatial components must be transformed from the principal axis system (p.a.s.) of the quadrupole interaction to the laboratory system, which is spinning about an axis inclined at the magic angle (54.7°) with respect to the applied magnetic field. This is done by using the transformation properties of irreducible spherical tensors under rotations by

$$R^{\lambda}_{2,m} = \sum_{m'} D^{2}_{m',m}(0, \theta_m, \omega_r t) \sum_{m''} D^{2}_{m'',m}(\alpha^{q}, \beta^{q}, \gamma^{q})\, \rho^{q}_{2m''}, \tag{8}$$

where θ_m is the magic angle, $\omega_r t$ is the azimuthal angle through which the sample has rotated. in time t while spinning at a rotation rate ω_r, α^{q}, β^{q}, γ^{q} are the Euler angles orientating the p.a.s. of the quadrupole interaction in the laboratory frame, and $\rho^{q}_{2m''}$ are tensor components of the quadrupole interaction in its p.a.s. For the quadrupole interaction, these are given by:

$$\rho^{q}_{20} = \tfrac{\sqrt{3}}{2} e q_{zz}, \tag{9}$$

$$\rho^{q}_{2,\pm 2} = \tfrac{1}{2}\eta e q_{zz}; \quad \rho^{q}_{2,\pm 1} = 0, \tag{10}$$

where eq_{zz} is the electric field gradient along the z axis of the p.a.s. and

$$\eta = (q_{yy} - q_{xx})/q_{zz} \tag{11}$$

is the asymmetry parameter of the quadrupole interaction.

The 3×3 matrix of the total Hamiltonian for an arbitrary orientation of the applied magnetic field may be found by substituting (8) into (7). The coefficients C_{nm} in (1) determining the ^{14}N spin states for that orientation are obtained by diagonalizing this matrix.

Since it is assumed that self-decoupling does not occur, meaning that there are no time-dependent interactions leading to transitions between the ^{14}N spin states, the density matrix for the spin states in the interaction representation is

$$\rho_{nm} = C_{nm} e^{-i(n-m)\omega_N t} \tag{12}$$

The expectation values of the ^{14}N vector spin operators for the nth spin state are given by

$$\langle T_{1m} \rangle = \mathrm{tr}\, \{\rho_{nm} \tau_{1m} e^{-im\omega_N t}\}_{m=+1,0,-1} \tag{13}$$

and are found to be

$$\left.\begin{aligned} \langle S_0 \rangle_n &= C^*_{n1} C_{n1} - C^*_{n3} C_{n3}; \\ \langle S_{+1} \rangle_n &= C^*_{n2} C_{n1} + C^*_{n2} C_{n3}; \\ \langle S_{-1} \rangle_n &= C^*_{n1} C_{n2} + C^*_{n3} C_{n2}. \end{aligned}\right\} \tag{14}$$

The ^{13}C spin properties are described by a Hamiltonian consisting of the Zeeman interaction and the heteronuclear dipolar coupling with the nitrogen spins. This is written in terms of irreducible spherical tensor components as

$$\mathscr{H}_C^{\mathrm{tot}} = -\omega_C \tau^Z_{10} + C^{\mathrm{d}} \sum_{m=-2}^{2} (-1)^m D^2_{m',m}(0, \theta_m, \omega_r t) \sum_{m''} D^2_{m'',m'}(\alpha^{\mathrm{d}}, \beta^{\mathrm{d}}, \gamma^{\mathrm{d}})\, \rho^{\mathrm{d}}_{2m''} \tau^{\mathrm{d}}_{2m} \tag{15}$$

where the transformation from the laboratory frame to the principal axis system of the dipolar interaction has been included.

The Euler angles α^{d}, β^{d}, γ^{d} define the ^{13}C–^{14}N internuclear vector in the laboratory coordinate system, the x axis of which is determined once the orientation of the quadrupolar p.a.s. with respect to the applied magnetic field has been established. By setting $\alpha^{\mathrm{q}} = 0$, the x axis of the laboratory frame is coplanar with the laboratory z axis and the quadrupole coupling tensor z axis. In the case of an axially symmetric quadrupole coupling tensor, the z axis of the quadrupole p.a.s. and the internuclear vector are then collinear and $\alpha^{\mathrm{q}} = \alpha^{\mathrm{d}} = 0$.

The quantities $\rho^{\mathrm{d}}_{2m''}$ are tensor components in the principal axis system of the dipolar interaction; these are given by

$$\left.\begin{aligned} \rho^{\mathrm{d}}_{20} &= \tfrac{\sqrt{3}}{2} R_{\mathrm{CN}}{}^{-3}; \\ \rho^{\mathrm{d}}_{2,\pm1} &= \rho^{\mathrm{d}}_{2,\pm2} = 0. \end{aligned}\right\} \tag{16}$$

The tensor components τ^{d}_{2m} represent the spin dependence of the dipolar interaction and are constructed from the vector spin operators of the ^{13}C and ^{14}N spins by using the Wigner–Eckart theorem. After transformation to the ^{13}C rotating frame, zero-order average Hamiltonian theory is applied to the dipolar spin operators to obtain secular terms of the dipolar Hamiltonian to use in a first-order perturbation treatment:

$$\langle T^{\mathrm{d}}_{2,m} \rangle_{\mathrm{av}} = \frac{\omega_C}{2\pi} \int_0^{2\pi/\omega_C} C(l_1 l_2 L; m_1 m_2 M)\, \tau^C_{1,m} e^{-im\omega_C t} \langle T^N_{1,m} \rangle \, \mathrm{d}t, \tag{17}$$

where $C(l_1, l_2, L; m_1, m_2, M)$ are the Clebsch–Gordon coefficients.

When this is done, the spin operators contributing secular terms to the dipolar Hamiltonian are

$$\left.\begin{aligned} T^{\mathrm{d}}_{20} &= \tfrac{2}{\sqrt{6}} I_0 \langle S_0 \rangle; \\ T^{\mathrm{d}}_{2,\pm1} &= \tfrac{1}{\sqrt{2}} I_0 \langle S_{\pm1} \rangle; \\ T^{\mathrm{d}}_{2,\pm2} &= 0; \end{aligned}\right\} \tag{18}$$

where $\langle S_0\rangle$, $\langle S_{+1}\rangle$ and $\langle S_{-1}\rangle$ are given in (14). This procedure is equivalent to the van Vleck truncation of internal Hamiltonians, but in this case additional terms, diagonal in the I spin, are retained. These terms correspond to the C and D terms of the familiar expression for the dipolar interaction (Abragam 1961) and physically they represent the fact that expectation values of the ^{14}N raising and lowering operators are stationary when viewed in a frame rotating at the carbon Larmor frequency, and thus contribute to secular splitting of the carbon resonances.

The dipolar Hamiltonian is evaluated in a basis set consisting of the product functions of the ^{13}C and ^{14}N spin states:

$$\left.\begin{aligned}\psi_n^+ &= |\alpha^{\mathrm{c}} > \cdot |\,\phi_n^{\mathrm{N}}\rangle;\\ \psi_n^- &= |\beta^{\mathrm{c}} \cdot |\,\phi_n^{\mathrm{N}}\rangle.\end{aligned}\right\} \tag{19}$$

The selection rule $\Delta m = \pm 1$ remains valid for the ^{13}C absorptions, which are given by

$$\begin{aligned}\omega_{^{13}\mathrm{C}}^n &= \langle\psi_n^+|\,\mathscr{H}_{\mathrm{C}}^{\mathrm{tot}}|\,\psi_n^+\rangle - \langle\psi_n^-|\,\mathscr{H}_{\mathrm{C}}^{\mathrm{tot}}|\psi_n^-\rangle\\ &= \omega_{\mathrm{c}} + \sqrt{\tfrac{1}{6}}C^{\mathrm{d}}R^{\mathrm{d}}_{2,0}\langle S_0\rangle_n + \tfrac{1}{4}C^{\mathrm{d}}\{R^{\mathrm{d}}_{2,1}\langle S_{+1}\rangle_n - R^{\mathrm{d}}_{2,1}\langle S_{-1}\rangle_n\}.\end{aligned} \tag{20}$$

Thus for each orientation of the quadrupole coupling tensor and corresponding orientations of the internuclear vector, evaluated for the sample inclined at the magic angle, and having rotated through the angle $\omega_{\mathrm{r}}t$, the ^{13}C absorption is split into three components with respect to its isotropic chemical shift position.

Sample rotation rates cannot approach the 1–5 MHz values of the quadrupole coupling constants of ^{14}N. A zero-order average over the rotation period is therefore not valid, and higher orders of the average Hamiltonian treatment must be included to account for sample rotation; because of the mathematical difficulty of this we treat sample rotation by numerically integrating the dipolar Hamiltonian over the azimuthal angle $\omega_{\mathrm{r}}t$. The ^{14}N spin states and dipolar coupling are determined in a time that is small compared with the rotation period. The effect of sample rotation is treated as a many site exchange problem, and an average chemical shift is seen. This is valid when the rotation rate is fast compared with the range of shifts experienced by the ^{13}C nucleus. This range is determined by $\omega_{\mathrm{d}}(\gamma_{\mathrm{c}}\gamma\hbar/r^3)$ and is small compared with the 3–4 kHz rotation rates employed. For the ^{15}N case the truncated dipolar Hamiltonian depends only on $P_2\cos\beta$, which is rotationally averaged to zero with magic angle spinning. This is also true for the second term on the right-hand side of (20). However, for ^{14}N the presence of additional terms in the secular dipolar Hamiltonian, dependent on spherical harmonics with $m = \pm 1$, results in non-zero average dipolar shifts.

To calculate a spectrum for an axially symmetric quadrupole coupling tensor, we calculate the shifts averaged over the rotation period for all the orientations of β^{q}, and give an intensity for each shift corresponding to $\sin\beta^{\mathrm{q}}$. This is shown in figure 3 where the calculation was performed for 10^3 orientations of β^{q} in the interval 0 to π for $e^2Qq/h = -1.20$ MHz and $R_{\mathrm{CN}} = 1.49$ Å† at a magnetic field strength of 3.54 T. The quadrupole coupling constant is typical for the α-amino group of amino acids at 77 K, neglecting asymmetry in the quadrupole coupling tensor.

The calculated ^{13}C spectrum consists of three constituent powder patterns, corresponding to the three absorptions into which the carbon absorption is split. The upfield two of these are nearly superimposed and correspond to the shifts caused by the highest and lowest energy states of ^{14}N. The expectation values of the ^{14}N spin raising and lowering operations for $n = 1$

† 1 Å = 10^{-10} m = 10^{-1} nm.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

and $n = 3$ have the same sign, and shift the carbon resonance approximately the same amount in the same direction. The shift due to the $n = 2$ spin state of ^{14}N is twice as large and occurs in the opposite direction. Thus the downfield peak in figure 3 is twice as broad with approximately twice the range of shifts as the upfield peaks. The downfield area most clearly shows the powder pattern nature of the ^{13}C absorption. It is the same as a first-order quadrupole

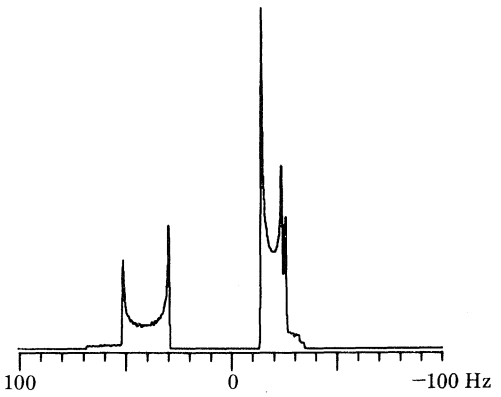

FIGURE 3. Theoretical spectrum of the alanine α carbon, calculated for a C–N bond length of 1.49 Å in a magnetic field of 3.54 T; $e^2qQ/h = -1.11$ MHz; $\eta = 0$; β^d and β^q are collinear; $\alpha^q = \alpha^g = 0$. Dipolar shifts are calculated to ± 0.1 Hz for 10^3 data points.

powder pattern for a spin 1 nucleus spinning at the magic angle where zero-order but not higher orders of the quadrupole effect are suppressed (Maricq & Waugh 1979).

Figure 5 compares the experimental C_α resonance of figure 1*b* of alanine with the broadened theoretical spectrum from the data of figure 3. Clearly, the theory described above predicts the experimental ^{13}C resonance line shape. Amino groups split carbons with the more intense peak upfield. The ^{13}C n.m.r. spectrum of 2-methyl-4-nitroaniline in figure 4 shows both this

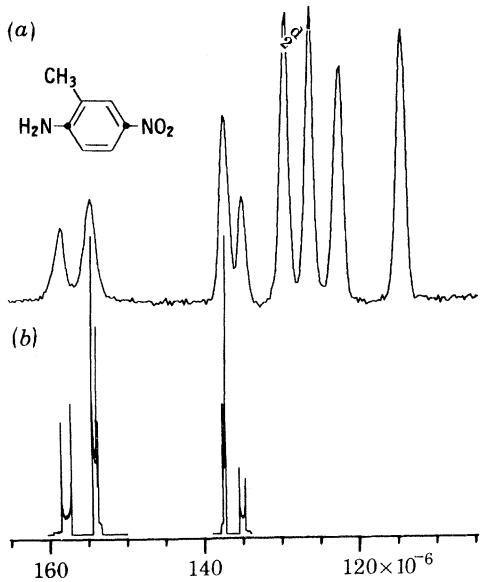

FIGURE 4. (*a*) ^{13}C n.m.r. of spectrum solid 2-methyl-4-nitroalanine. Theoretical spectra calculated as described in the text are shown in (*b*).

amino split and a reverse split from the nitro group. This is a consequence of the opposite signs of the quadrupole coupling for the two types of nitrogen. The spectra were calculated with a C–N bond length of 1.49 Å and quadrupole coupling constants of −2.65 MHz for the amino group and 1.6 MHz for the nitro group. In this case, the use of axial symmetry for the quadrupole coupling tensor is an approximation, used here because the asymmetry parameter and the orientation of the z axis of the electric field gradient have not been determined for this compound at room temperature. The theoretical spectra must then be regarded as giving a qualitatively correct explanation of the experiment.

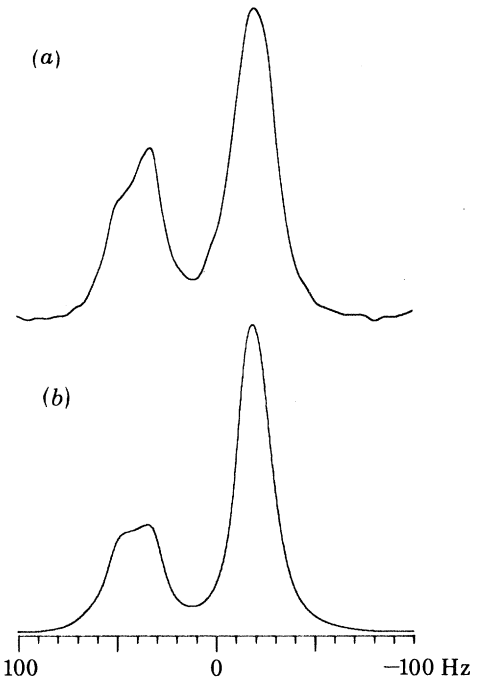

FIGURE 5. ^{13}C n.m.r. spectra of the α carbon resonance of polycrystalline alanine. (*a*) Experimental data as in figure 1; (*b*) theoretical data of figure 3 with 12 Hz line broadening added.

When the z axis of the electric field gradient and the ^{13}C–^{14}N bond are not collinear, which generally happens when the quadrupole coupling tensor is not axially symmetric, the dipolar shifts of the ^{13}C resonance are dependent on both the asymmetry parameter η and the angle between the internuclear vector and the z axis of the quadrupole coupling tensor. It is possible to calculate the dependence of the split between the components of the asymmetric doublet on this angle, which then is a source of geometrical information about the molecule, and which may prove useful in determining peptide group conformations of polycrystalline or amorphous proteins, since splits of both carbonyl and α carbon resonance are induced by the ^{14}N of the peptide bond. An example of such data is shown in figure 6, which contains the carbonyl and α carbon lines of the dipeptide tyrosylleucine. The peptide bond nitrogen splits its α carbon resonance about 53 Hz compared with the 144 Hz of its carbonyl resonance (Frey & Opella 1980). This large difference cannot be explained solely on difference in bond lengths and must be due to the angles between the bonds and the z axis of the quadrupole coupling tensor with the same ^{14}N of the peptide bond.

High-resolution ^{13}C spectra of powders are often confusing because of a multiplicity of lines. More lines than numbers of carbons can result from a number of factors. It is usually easy to recognize the doublets due to ^{13}C–^{14}N interactions because of their broadness and asymmetry. Other common sources of multiple lines include the presence of different crystal forms and the effect of subsequent groups on aromatic rings that have an asymmetric conformation in the solid state, causing adjacent carbons to have different environments (Schaefer *et al.* 1977; Lippmaa *et al.* 1978; Maricq & Waugh 1977).

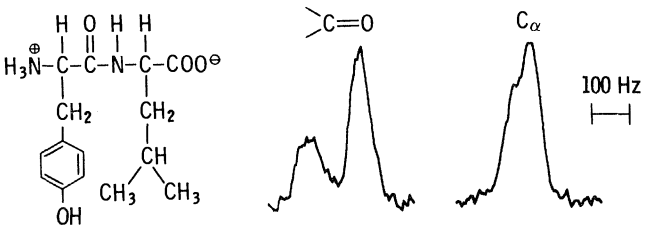

FIGURE 6. ^{13}C n.m.r. spectra of the peptide carbonyl and α carbon of the dipeptide Tyr-Leu.

Many molecules can crystallize in several forms. If these different crystals are present in the same powder, then the n.m.r. spectrum often has multiple lines for each site. A particularly clear case of this is shown in figure 7 with leucine. Since this molecule is isotopically labelled with ^{15}N replacing ^{14}N, no confusion results from the quadrupole moment of ^{14}N. The spectrum

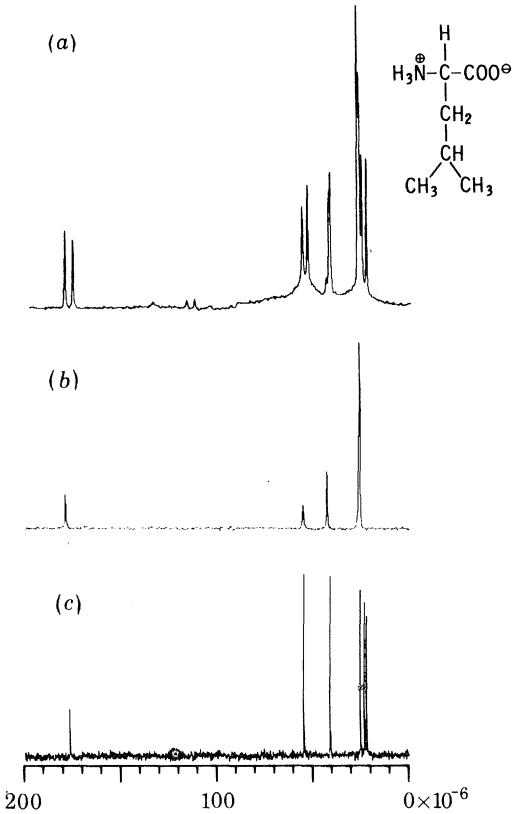

FIGURE 7. ^{13}C n.m.r. spectra of leucine. (*a*) Polycrystalline [^{15}N]Leu sample; (*b*) recrystallized [^{15}N]Leu; (*c*) Leu in aqueous solution.

in figure 7*a* shows multiple aliphatic peaks and, particularly clearly, two well resolved carbonyl peaks. After recrystallization, the spectrum in figure 7*b* results, which is much simpler. Obviously one of the most important experimental manipulations of solid state n.m.r. is sample preparation. Leucine has a long aliphatic side chain and the spectrum of figure 7*b* demonstrates an additional feature of ^{13}C n.m.r. of some powders; large changes in isotropic chemical shift are apparent in comparing spectra from aqueous solution with those from the solid state. For leucine these shifts among the γ and δ carbons are as large as 5×10^{-6} and significantly change the appearance of methylene and methyl resonances as shown in figure 7. Such large shifts with change of phase are surprising for saturated carbons, which have relatively small chemical shift anisotropy (Pines *et al.* 1972*a*).

The aromatic region of the ^{13}C n.m.r. spectrum of tyrosine is shown in figure 8. Both the ε and δ carbon resonances are split because of the ring substitutions. For comparison the solution spectrum of tyrosine only has four lines in the aromatic region, compared with the nearly resolved six lines of tyrosine in the solid state.

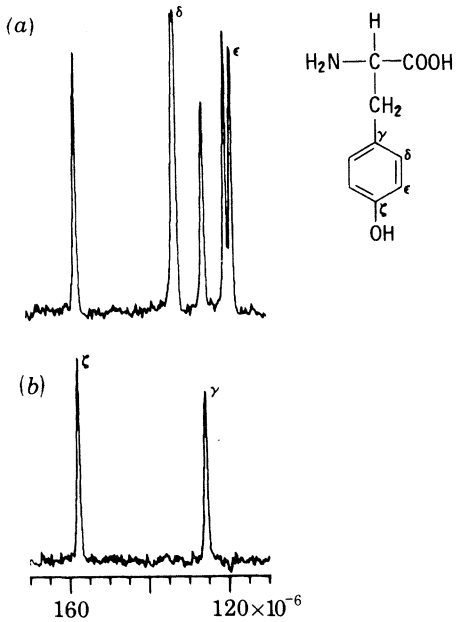

FIGURE 8. ^{13}C n.m.r. spectra of polycrystalline tyrosine. (*a*) Aromatic region; (*b*) non-protonated aromatic carbons. Obtained as described in the text.

In addition to the aromatic spectral features, figure 8 illustrates the use of a resonance selection technique (Opella & Frey 1979). As molecules increase in size and complexity it is necessary to increase practical resolution; procedures that simplify n.m.r. spectra are therefore important. This is true for both solution and solid state n.m.r. It is often desirable to observe only the signals from the non-protonated carbons in a sample in natural-abundance ^{13}C n.m.r. spectroscopy. Resonance assignments and spectral simplifications result from the selection of this limited class of carbons.

Because of the somewhat greater line widths and the same effective chemical shift dispersion resulting in a loss of basic resolution, selection procedures are even more important in the solid state than in solution. For solids, static heteronuclear dipolar coupling can be used as a basis

for selection of non-protonated carbons. A brief delay without ^{1}H irradiation is inserted between the development of ^{13}C magnetization with the mix pulse and data acquisition with full decoupling; this causes a significant loss in intensity for carbons with strong dipolar couplings compared with those with weak couplings.

We have observed large spectral changes between solution and powder samples of amino acids and linear small peptides in most cases. While the changes in line shape of carbons bonded to nitrogen are well understood and have the potential for providing new molecular information, other sources of splittings or shifts are not as interesting at present. Since our understanding of the chemical shift is so much less than of dipolar or quadrupolar interactions, little physical insight is available from noting these chemical shifts or splittings. Certainly the asymmetric structures of substituted aromatic rings can be explained reasonably, but the physical basis of the size of the chemical shifts and their variability from compound to compound are not readily described.

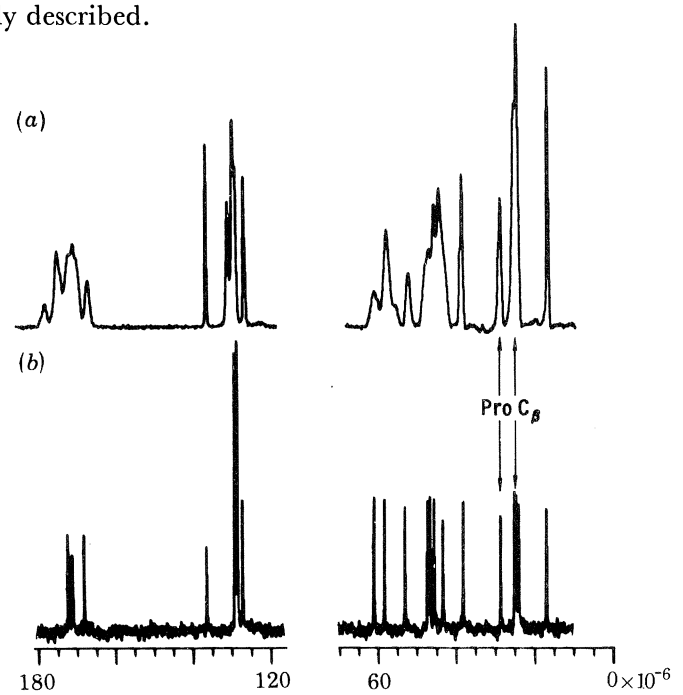

FIGURE 9. ^{13}C n.m.r. spectra of cyclo(D-Phe-Pro-Gly-D-Ala-Pro). (*a*) Polycrystalline sample; (*b*) sample dissolved in $CDCl_3$.

One conclusion based on the background of amino acid and peptide spectroscopy is that changes in isotropic chemical shifts between solid and solution reflect the conformational flexibility of a molecule. The actual shifts may result from intermolecular interactions, but the accessibility to different environments indicates the capability of changing structure. This is also consistent with the largest changes in chemical shift occurring in long-chain aliphatic amino acids where rotations of methyl and methylene groups are readily allowed.

Cyclic peptides of defined conformation have provided additional information on the sensitivity of ^{13}C isotropic chemical shift to the change from solution to the solid state. The cyclic pentapeptide (D-Phe-Pro-Gly-D-Ala-Pro), which contains a β and a γ turn, forms excellent crystals suitable for X-ray diffraction and solid state n.m.r. (Pease 1979). Figure 9 compares the ^{13}C n.m.r. spectra of the pentapeptide in aqueous solution and in the solid state. The most striking feature of this comparison is that there are essentially no changes in isotropic chemical

shift between the two states. This is in contrast to the large changes observed in linear peptides or amino acids, such as shown for leucine in figure 7. The overall appearance of the spectra in figure 9*a* and *b* are somewhat different, but this is almost exclusively due to the ^{14}N–^{13}C interactions which split and broaden carbonyl and α carbon resonances. Since this is a cyclic peptide, there is no terminal carboxyl group, so all of the carbonyl carbons ($168–173 \times 10^{-6}$) are broadened and split into asymmetric doublets in the solid state spectrum. Clearly some resolution is apparent in this region in the solid state, and with deconvolution and spectral fitting this resonance region could be fully characterized. Similar comments hold for the α carbons ($45–62 \times 10^{-6}$), which are also split into asymmetric broadened doublets. The conformation of the peptide backbone can be deduced from these data sets as explained for the dipeptide with figure 6.

The rest of the aliphatic region shows strong similarities between solution and solid state chemical shifts. The solid state spectrum is only slightly broader than the solution spectrum, giving a level of resolution that is quite manageable. One of the most distinguishing features of this type of cyclic peptide is the difference in conformation (Karle 1980) and C_β chemical shift of the two proline residues (Pease & Watson 1978). One of the proline β carbon resonances is shifted upfield to about 25×10^{-6} while the other remains near 29×10^{-6}. A comparison of this spectral region for the solution and solid samples shows that this characteristic feature of peptide conformation is retained in the solid state (Pease *et al.* 1981). This is a strong indication of the peptide's having the same conformation in solution and the solid state; it is also the first example of resolution of secondary or tertiary structure induced chemical shift dispersion.

Solid DNA

There are disproportionately few n.m.r. studies of high molecular mass native DNA because of the very broad line widths of nucleotide resonances that are a consequence of the motions of the polymer being too slow to average out static nuclear spin line broadening mechanisms. Solution studies have been carried out on low molecular mass or single-stranded materials, because they have narrow resonances that can be characterized by using conventional high-resolution spectrometers. Most work has been done by observing ^{31}P signals. This includes solution studies of model systems (Hogan & Jardetzky 1979; Bolton & James 1980; Shindo 1980) as well as some recent solid state n.m.r. work on a variety of samples including high molecular mass DNA in solution (Opella *et al.* 1981).

^{13}C n.m.r. studies have been extremely limited by the unfavourable dynamical properties of DNA. Recently Rill *et al.* (1980) have reported natural-abundance ^{13}C n.m.r. spectra of 140 base pair fragments of duplex DNA in solution. Even, in this favourable case, the results are limited by sensitivity and line widths of the polymer.

To date, reliable ^{13}C n.m.r. spectra of high molecular mass DNA have not been reported. Part of the difficulty comes from the truly inconvenient motional properties of the DNA double helix. In fact, if the sugars and bases have the same motions as found for the phosphodiester backbone, then the ^{13}C resonances will be very broad from efficient ^{13}C–^{1}H dipolar relaxation. This type of broadening cannot be removed by radio-frequency decoupling or magic angle sample spinning.

To obtain high-resolution ^{13}C spectra we resort to the prescription of Haeberlen & Waugh (1969), where the sample is immobilized to eliminate motions effective in inducing relaxation, even though we replace that line broadening by what at first glance is even worse: line broaden-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

ing due to static chemical shift anisotropy and dipolar interactions. The benefit, of course, is that available experimental procedures can effectively remove these effects but not those due to relaxation. Figure 10 shows the natural abundance ^{13}C n.m.r. spectrum of solid calf thymus DNA, obtained by solid state n.m.r. All carbon resonances are present in the full DNA spectrum. The aromatic spectral regions shown as insets contain only non-protonated carbon resonances.

In the complete ^{13}C spectrum of DNA the only assignable resonances are from the thymine methyl at 12×10^{-6} and the thymine C-5 and cytosine C-5 resonances. The dissection of the base resonances is complicated by the presence of so many nitrogen atoms. The C-5 positions of the pyrimidines are the only carbons not directly bonded to at least one ^{14}N. The thymine and cytosine C-5 resonances can be distinguished on the basis of their isotropic chemical shift and from selection procedure for non-protonated carbons.

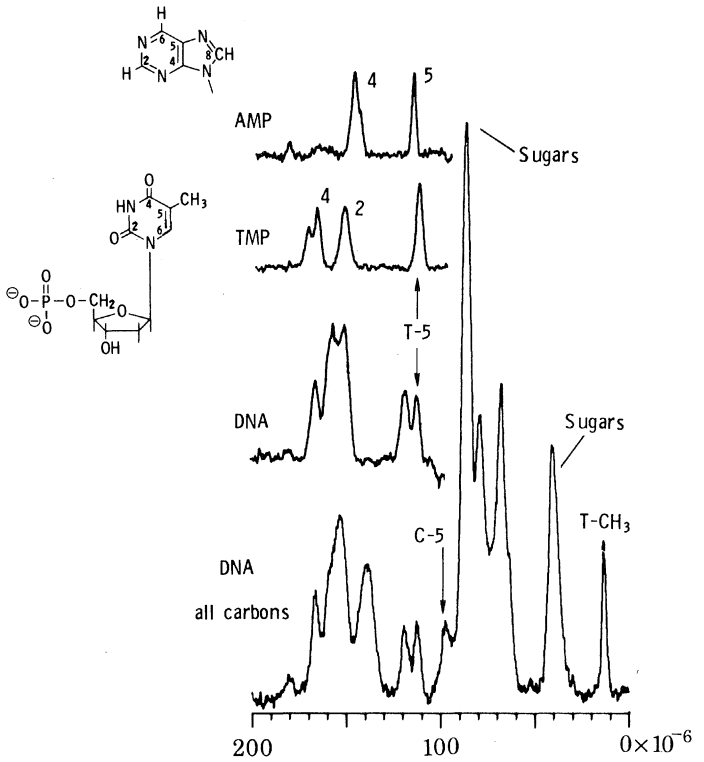

FIGURE 10. ^{13}C n.m.r. of solid DNA and nucleotides.

The aromatic spectral regions of DNA, AMP and TMP shown in figure 10 are much less complicated and broadened than expected on the basis of the amino acid results for carbons bonded to multiple nitrogens, especially the arginine C_ζ resonance. The reason for this may be in the angular dependence of the ^{13}C–^{14}N coupling, with apparent single-line spectra resulting if the z axis of the electric field gradient and the C–N bond form an angle of 90° to each other, as could be true for these planar rings (J. G. Hexem & S. J. Opella, unpublished results).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Application of solid state n.m.r. to the study of filamentous viruses

The filamentous bacteriophages provide several opportunities for using solid state n.m.r to study otherwise unapproachable biochemical problems. Two of these that use ^{13}C n.m.r. will be discussed here. The major coat protein of the viruses is very hydrophobic and consequently it is completely insoluble in water without the presence of detergents or lipids; the pure protein can therefore only be studied by solid state n.m.r. The viruses in solution behave as rigid rods of very high particle mass, and the resulting lack of efficient motional averaging makes solution n.m.r. inapplicable.

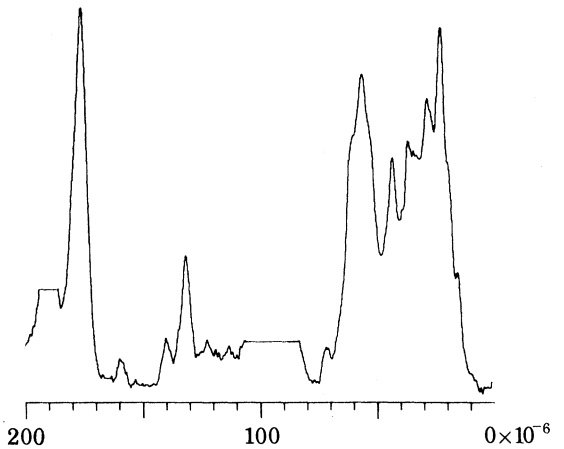

Figure 11. ^{13}C n.m.r. spectrum of lyophilized powder of fd coat protein.

The virus fd is a filamentous virus that infects *Escherichia coli* (Marvin & Hohn 1969) and is a long flexible filament with dimensions 9 nm × 900 nm in solution (Newman *et al.* 1977). It is constructed from 2700 copies of the 5000 molecular mass major coat protein, which surround the single-stranded circular DNA of 6400 nucleotides. Several copies of a minor coat protein are located at one end of the filament. The total particle mass of the virion is 16.4×10^6, of which 88% is the major coat protein. Fibres of fd give X-ray diffraction patterns that have been used for structural studies (Marvin *et al.* 1974). The picture that emerges is that the DNA is extended lengthways in a closed tube of coat proteins, which are in an overlapping helical array. We chose fd for n.m.r. studies because it is relatively small, it has simple organization and small coat protein subunits, it is experimentally tractable, and extensive prior work has been done.

One of the most important experimental procedures that can be carried out is the isolation of the coat protein subunits. The virus can be disrupted in the presence of detergents at high temperatures. Sephadex chromatography separates the coat proteins from the DNA. Removal of the detergent from the coat proteins causes the protein to aggregate and come out of solution. This sample is intractable because of its insolubility. The amorphous coat protein is an acceptable form of sample for solid state n.m.r.

Figure 11 is the natural-abundance ^{13}C n.m.r. spectrum of the isolated fd coat protein (Opella *et al.* 1979). It contains signals from all of the carbons in the protein. The aliphatic carbons are in the region $10–70 \times 10^{-6}$. All aromatic carbons occur in the region $110–165 \times 10^{-6}$

with some separation of peaks apparent. The carbonyl carbons make up the bond centred at 175×10^{-6}. In this spectrum the only identifiable resonance is that of the two Tyr C_ζ at 160×10^{-6}.

Figure 12 compares the aromatic carbon region with all resonances to the spectrum obtained under conditions selecting for non-protonated carbon resonances. The protein has 50 amino acids with a total of 10 non-protonated aromatic resonances from 3 Phe, 2 Tyr and 1 Trp (Nakashima & Konigsberg 1974). The resonance at 160×10^{-6} is from the two Tyr_ζ carbons. The resonances at 138×10^{-6} are from the 3 Phe C_γ and the Trp ϵ_2 carbon; the contribution

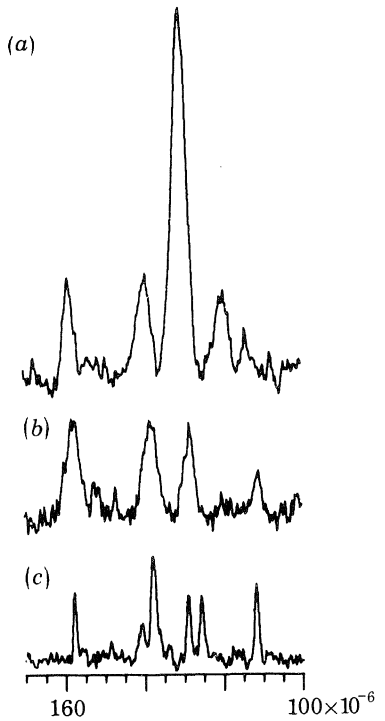

FIGURE 12. Aromatic region of ^{13}C n.m.r. spectrum of lyophilized fd and constituent amino acids. (*a*) Complete spectrum of fd; (*b*) non-protonated aromatic carbon resonances of fd; (*c*) non-protonated aromatic carbon resonances of a powder mixture of the constituent amino acids.

from Trp carbon is probably reduced in the solid spectra owing to ^{14}N-induced splitting. Intensity at 130×10^{-6} is due to the 2 Tyr C_γ and Trp δ_2 carbon; these signals are completely overwhelmed by other ring carbons in the conventional proton-enhanced spectrum of figure 2. Near 110×10^{-6} is the resonance from the single γ carbon of Trp-26. Its assignment is unambiguous because there is only one Trp in the protein.

The line widths observed for the amorphous protein are significantly broader than those of the crystalline amino acids and peptides. It is not clear what this line width is due to, although chemical shift-like terms resulting from crystal packing or magnetic anisotropies may be responsible.

The study of biological supramolecular structures in solution by n.m.r. is a tremendously challenging area of research. The experimental obstacles are formidable but if they can be overcome, truly unique new information will be available. Figure 13*a* shows the starting point spectroscopically for ^{13}C n.m.r. of fd in solution. In spite of long signal averaging on a very

concentrated and large sample of fd in solution, very little resonance intensity is observable by conventional high-resolution n.m.r. High-power proton decoupling and magic angle sample spinning experiments show that the severe broadening is due to ^{1}H–^{13}C dipolar couplings and ^{13}C chemical shift anisotropy. These static line broadening mechanisms reflect the lack of motional averaging in such a large particle as a filamentous virus. The lack of resolved aromatic ($110–160 \times 10^{-6}$) and carbonyl ($170–180 \times 10^{-6}$) carbon resonance intensity is particularly clear when the protein spectrum of figure 12 is compared with the virus spectrum in figure 13 and indicates the difficulties associated with conventional high-resolution ^{13}C n.m.r. of large structures.

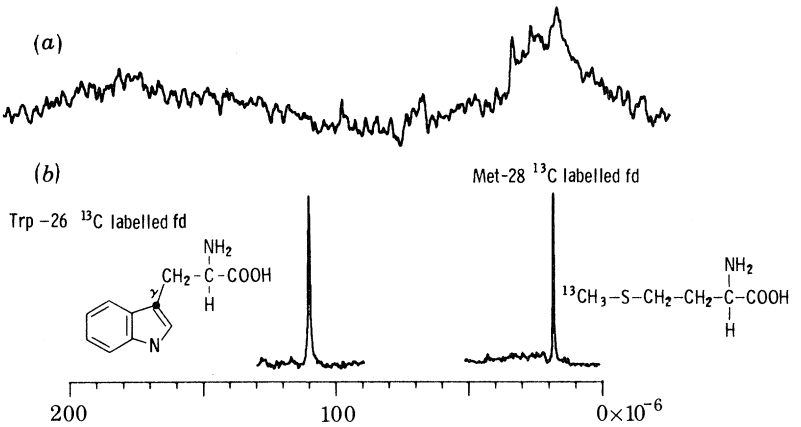

FIGURE 13. ^{13}C n.m.r. spectra of fd virus in 'solution'. (*a*) Spectrum obtained with conventional solution state techniques; (*b*) spectrum of selectively enriched virus, prepared as described in the text, obtained by using solid n.m.r. techniques.

One of the initial goals of the research on filamentous virus was to obtain the equivalent of a high-resolution spectrum for single carbon sites that could be assigned. An advantage of the filamentous bacteriophage system is that fd infects *E. coli*, which can be grown on a chemically defined medium. Specifically labelled amino acids can therefore be incorporated into the fd coat protein. Because the protein only has 50 amino acids there are some residues present in very low quantity, in particular the single Trp-26 and the single Met-28. Therefore, when these amino acids are labelled in a single site with ^{13}C the observed resonance is automatically assigned.

Figure 13 demonstrates the success of the entire procedure. Two different growths of *E. coli* infected with fd were performed, one with C_{ε}-labelled Met added to a mixture of natural-abundance amino acids and the other with C_{γ}-labelled Trp. The virus was then isolated from these growths and purified on caesium chloride gradients.

The observed ^{13}C signals were obtained by cross-polarization and observed in the presence of high-power proton decoupling while spinning the sample at the magic angle. Narrow single-line spectra result. Clearly, the strategy of selecting for individual sites of the small subunits by ^{13}C labelling and the application of solid state n.m.r. procedures yields high-resolution spectra. The information in the spectra in figure 13 is limited to the isotropic chemical shift and the relaxation parameters because the n.m.r. procedures applied to narrow lines remove chemical shift anisotropy or dipolar couplings as observable parameters. However, experiments

can be designed to put back selectively one interaction for study while retaining the single-site resolution.

We thank C. M. Gall for help with the Met-28 labelled fd virus experiment. This work is being supported by grants from the National Institutes of Health (GM-24266) and the American Cancer Society (NP-225A). T.A.C. is supported by a Cell and Molecular Biology Training Grant. S.J.O. is a fellow of the A. P. Sloan Foundation (1980–2).

References (Opella *et al.*)

Abragam, A. 1961 *The principles of nuclear magnetism.* Oxford University Press.
Allerhand, A. 1979 *Acct. chem. Res.* **11**, 469–474.
Andrew, E. R., Farnell, L. F., Firth, M., Gledhill, T. D. & Roberts, I. 1969 *J. magn. Reson.* **1**, 27–34.
Bolton, P. H. & James, T. L. 1980 *J. Am. chem. Soc.* **102**, 25–31.
Cross, T. A., DiVerdi, J. A., Wise, W. B. & Opella, S. J. 1979 In *NMR and biochemistry*, pp. 67–74. New York: Dekker.
Frey, M. H. & Opella, S. J. 1980 *J. chem. Soc. chem. Commun.*, pp. 474–475.
Groombridge, C. J., Harris, R. K., Packer, K. J., Say, B. J. & Tanner, S. F. 1980 *J. chem. Soc. chem. Commun.*, pp. 174–175.
Haeberlen, U. & Waugh, J. S. 1969 *Phys. Rev.* **185**, 420–429.
Hexem, J. G., Frey, M. H. & Opella, S. J. 1981 *J. Am. chem. Soc.* (In the press.)
Hogan, M. E. & Jardetzky, O. 1979 *Proc. natn. Acad. Sci. U.S.A.* **76**, 6341–6345.
Karle, I. L. 1980 In *Perspectives in peptide symmetry* (ed. T. Wieland, R. Geiger & A. Eberle). Basel: S. Korger. (In the press.)
Lippmaa, E., Alla, M., Roude, H., Teeaar, R., Heinmaa, I. & Kundla, E. 1979 In *Proc. 20th Congress Ampere*, p. 87.
Lippmaa, E., Alla, M. A., Pehk, T. J. & Engelhardt, G. 1978 *J. Am. chem. Soc.* **100**, 1929–1931.
Maricq, M. M. & Waugh, J. S. 1979 *J. chem. Phys.* **70**, 3300–3316.
Marvin, D. A. & Hohn, B. 1969 *Bact. Rev.* **33**, 172–209.
Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J. & Marvin, F. J. 1974 *J. molec. Biol.* **88**, 581–600.
Nakashima, Y. & Konigsberg, W. 1974 *J. molec. Biol.* **88**, 598–600.
Newman, J., Swinney, H. L. & Day, L. A. 1977 *J. molec. Biol.* **116**, 593–606.
Opella, S. J. 1977 *Science, N.Y.* **198**, 158–165.
Opella, S. J. & Frey, M. H. 1979 *J. Am. chem. Soc.* **101**, 5854–5856.
Opella, S. J., Frey, M. H. & Cross, T. A. 1979 *J. Am. chem. Soc.* **101**, 5856–5857.
Opella, S. J., Frey, M. H. & DiVerdi, J. A. 1980 *J. magn. Reson.* **37**, 165–169.
Opella, S. J., Wise, W. B. & DiVerdi, J. A. 1981 *Biochemistry, Wash.* (In the press.)
Pease, L. G. 1979 In *Peptides: structure and biological function* (*Proceedings of the Sixth American Peptide Symposium*) (ed. E. Gross & J. Meienhofer), pp. 197–200. Rockford: Pierce Chemical Co.
Pease, L. G. & Watson, C. 1978 *J. Am. chem. Soc.* **100**, 1279–1286.
Pease, L. G., Frey, M. H. & Opella, S. J. 1981 *J. Am. chem. Soc.* (In the press.)
Pines, A., Gibby, M. G. & Waugh, J. S. 1972*a* *Chem. Phys. Lett.* **15**, 373–376.
Pines, A., Gibby, M. G. & Waugh, J. S. 1972*b* *J. chem. Phys.* **56**, 1176–1777.
Pines, A., Gibby, M. G. & Waugh, J. S. 1973 *J. chem. Phys.* **59**, 569–590.
Rill, R. L., Hilliard, P. R., Bailey, J. T. & Levy, G. C. 1980 *J. Am. chem. Soc.* **102**, 418–420.
Schaefer, J. & Stejskal, E. O. 1976 *J. Am. chem. Soc.* **98**, 1031–1032.
Schaefer, J., Stejskal, E. O. & Buchdahl, R. 1977 *Macromolecules* **10**, 384–405.
Shindo, H. 1980 *Biopolymers* **19**, 509–522.
Spiess, H. W., Haeberlen, U. & Zimmerman, H. 1977 *J. magn. Reson.* **25**, 55–66.

Discussion

W. Derbyshire (*Department of Physics, University of Nottingham, U.K.*). My question is related to the 'peculiar' structure obtained in the proton enhanced spectra of carbon atoms directly bonded to nitrogen. We are told that the observed splitting decreases with increasing frequency. First, is this splitting expressed in hertz or millionths, and secondly, is the splitting inversely proportional to frequency, to the square of frequency or to something else? This information would help us to decide if first- or second-order electric quadrupolar effects are involved.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

S. J. Opella. Figure 14 contains a plot of the splitting of the 'centre of mass' of the upfield and downfield shifted absorptions as a function of field strength. This calculation was done for the same parameters as the theoretical spectrum of alanine shown in figure 3, and the detailed behaviour in each case will depend on the values of the parameters.

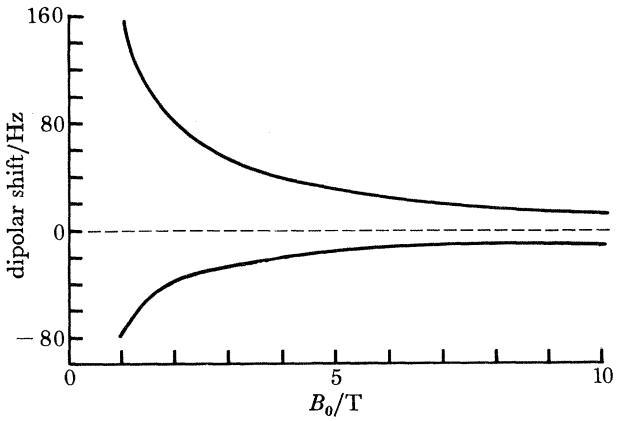

Figure 14. Plot of splitting of the C_α resonance of alanine induced by ^{14}N as a function of magnetic field strength.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES