Iterative fitting of magic-angle-spinning NMR spectra

Groot, H.J.M. de; Smith, S.O.; Kolbert, A.C.; Courtin, J.M.L.; Winkel, C.; Lugtenburg, J.; ... ; Griffin, R.G.

JOURNAL OF MAGNETIC RESONANCE 91, 30–38 (1991)

Iterative Fitting of Magic-Angle-Spinning NMR Spectra

H. J. M. DE GROOT,*,†,‡ S. O. SMITH,*,§ A. C. KOLBERT,* J. M. L. COURTIN,† C. WINKEL,† J. LUGTENBURG,† J. HERZFELD,‡ AND R. G. GRIFFIN*

*Francis Bitter National Magnet Laboratory and Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; †Gorlaeus Laboratoria der Rijksuniversiteit te Leiden, 2300 RA Leiden, The Netherlands; and ‡Department of Chemistry, Brandeis University, Waltham, Massachusetts 02254

Received January 26, 1990

An approach for the analysis of magic-angle-spinning NMR spectra which combines nonlinear iterative and linear analytical procedures for fitting sideband spectra is presented. Parameter values, as well as statistical errors, are extracted from the experimental data. The method is particularly useful for the analysis of spectra with limited sensitivity or overlapping lines, and the fitting procedure ensures the highest possible speed. Experimental examples for the analysis of spectra with sidebands due to chemical-shift anisotropies and heteronuclear dipolar interactions are shown. The importance of statistical and systematic errors is considered. © 1991 Academic Press, Inc.

More than 30 years have passed since Andrew (*1*) and Lowe (*2*) introduced the technique of magic-angle spinning (MAS) as a means of narrowing the NMR spectra of solids. Since then MAS NMR has evolved into an indispensable technique for the investigation of a wide variety of polycrystalline and amorphous systems (*3, 4*). MAS spectra consist of centerbands at the isotropic chemical shifts flanked by rotational sidebands spaced at the spinning frequency, the number and intensity of rotational sidebands decreasing with increasing rotation speed, ω_r. When $\omega_r < \omega_0\delta$, in which $\delta = \sigma_{33} - (\frac{1}{3})\mathrm{Tr}\,\sigma$ is the anisotropy and ω_0 is the Larmor frequency, numerous sidebands are observed, and comparison of the sideband intensities with the results of numerical simulations yields information about the anisotropic interactions (chemical shift and dipolar) being averaged. The technique can only be fully exploited, however, if it is paired with sophisticated procedures for the analysis of the spectra. For example, the spectra are often noisy and it is therefore difficult to extract the relevant information, chemical-shift anisotropies and dipolar couplings. Additionally, it is extremely important to know the precision of the parameters derived from such spectra, since this affects the scientific interpretation of the data.

The purpose of the present paper is to discuss a procedure for analysis of MAS spectra involving iterative fitting of spectra with substantial noise and spectral overlap. The technique permits evaluation of the parameters describing the interactions and the associated statistical errors and recently has been successfully employed in studies of ^{13}C and ^{15}N spectra of membrane proteins (*5–9*).

§ Present address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511.

0022-2364/91 $3.00

THE FITTING PROCESS

The approach of this analysis is to incorporate all a priori information about the spectra into the fit, in order to obtain the best reliability, highest speed, and optimal efficiency. This is achieved by (i) keeping those parameters whose value can be determined in a separate experiment fixed, (ii) assuming a simple lineshape for the centerbands and sidebands in the MAS pattern, and (iii) fitting all linear parameters by means of an analytical linear least-squares fitting procedure in each iteration step. For the nonlinear fitting we use the CERN package MINUIT (*10*), which embodies a simplex as well as a gradient search procedure, in addition to a Monte Carlo optimization and a whole range of subroutines that facilitate the programming and the fitting process. In our approach we normally employ a simplex minimization followed by a gradient search for the highest accuracy. Since the theoretical simulations of the MAS spectrum are nonanalytical, the gradients are calculated from the variation of the quality of the fits associated with small changes in the parameter values.

The fitting process is represented schematically in Fig. 1. First, the theoretical MAS line intensities are calculated for all the components in the spectrum based upon the initial values of the parameters. These intensities are then convolved with Lorentzian or Gaussian lines and superposed to generate the MAS spectrum. The simulation is then scaled, and a polynomial background, up to fifth order, obtained by a linear least-squares fit of the experimental spectrum, is added. The χ^2 between the experimental and the theoretical spectra, which is used as the criterion in the optimization procedure, is then calculated. MINUIT then adjusts the parameter set, and the process continues. Only the parameters that determine the sideband intensities are optimized by iterating the MAS simulation. For instance, if only the linewidths, isotropic shifts, or relative contributions of the components are different from the parameters set in the previous iteration, calculation of a new set of line intensities is not necessary and is therefore bypassed. When the χ^2 is at the noise level, estimated from the baseline of the experimental spectrum, the iteration ends, and the spectrum and fit are plotted.

Central to the analysis procedure is the MAS iteration loop, which calculates a new set of line intensities whenever any relevant parameters are changed. This loop is called many times, and is therefore optimized for speed. For the evaluation of the line intensities of the MAS pattern various approaches may be chosen depending on the application. By far the fastest procedure to generate the line intensities is the Herzfeld–Berger method (*11*), which is used mainly for the analysis of spectra with sideband intensities due to chemical-shift anisotropies. This procedure involves obtaining the line intensities via a previously calculated look-up table. For the 1D dipolar-chemical-shift spectrum, a free induction decay for spins evolving under the appropriate Hamiltonian is evaluated, using procedures described by Munowitz and Griffin (*12*), in a slightly modified form. The time-domain simulation contains a minimum number of points n_p, equal to the number of peaks in the spectrum. The signal is calculated over one rotor period and, utilizing the periodicity of the Hamiltonian, is replicated outside the powder average loop. Off-resonance effects are incorporated after the replication of the FID. Numerical Fourier transformation of this time-domain signal yields the set of line intensities that describes the MAS pattern.

In particular, when the Herzfeld–Berger method is employed, the whole fitting pro-

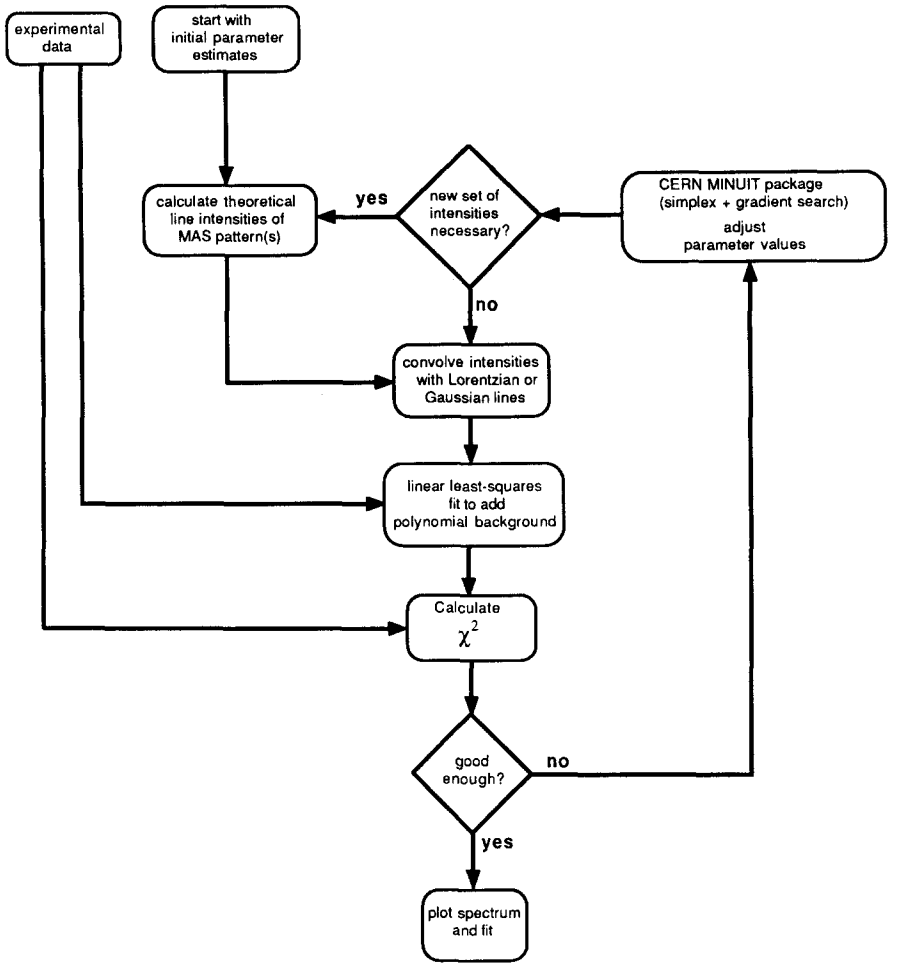

FIG. 1. Schematic diagram of the fitting analysis procedure. For highest efficiency inside the iteration loop a linear least-squares fitting procedure optimizes the linear parameters for each set of nonlinear parameters.

cess, simplex minimization plus gradient search, is very efficient. Typical CPU times on a VAXstation II may vary from three minutes, for a spectrum with a single MAS pattern, to one hour, for a spectrum with four independent overlapping resonances. For problems in which it is necessary to simulate an FID, such as the dipolar-chemical-shift experiment, analysis of a simple spectrum may take one hour and for more complicated cases the program runs overnight.

After completion of the fit, estimates for the errors are calculated. The MINUIT program evaluates parabolic statistical error estimates for each nonlinear parameter. With appropriate scaling these errors represent a good estimate of the statistical errors for the values of the parameters determined in the fits (*13*). The errors (±) given throughout this paper represent a 95% confidence interval.

EXPERIMENTAL EXAMPLES

Among the important fields of application of MAS NMR is the study of large biomolecules, where this technique is one of the few methods to obtain detailed information about their structure and function on the atomic level (*14*). In particular, the combination of specific isotope labeling and MAS NMR difference spectroscopy is very powerful and to date has allowed the study of systems with molecular weight up to 100 kDa. The analysis procedures described here are often essential for the interpretation of the MAS NMR spectra from the heavier systems (>10 kDa). They form a good example to demonstrate the usefulness of iterative fitting analysis.

In Fig. 2a we show a fit of a spectrum, taken of the protein bacteriorhodopsin (bR), the light-driven proton pump of *Halobacterium halobium* (MW ~27 kDa). The molecule contains a retinylidene chromophore in the active site of a protein with 248 amino acid residues that is embedded in a membrane. For the experiment of Fig. 2a the molecule has been labeled (99% ^{13}C) at C-5 of the retinylidene moiety. The data were acquired with a spinning speed $\omega_r/2\pi$ = 2.5 kHz, chosen in order to produce the signal from the label exactly degenerate with the carbonyl resonance.

Even for this highly unfavorable case a reasonably accurate anisotropy parameter of $\omega_0\delta/2\pi$ = 9.5(0.8) kHz could be obtained by performing a serial analysis of two spectra at the same spinning speed. First, a natural-abundance spectrum was analyzed, with Gaussian lines, to characterize the carbonyl and aromatic resonances (at 176

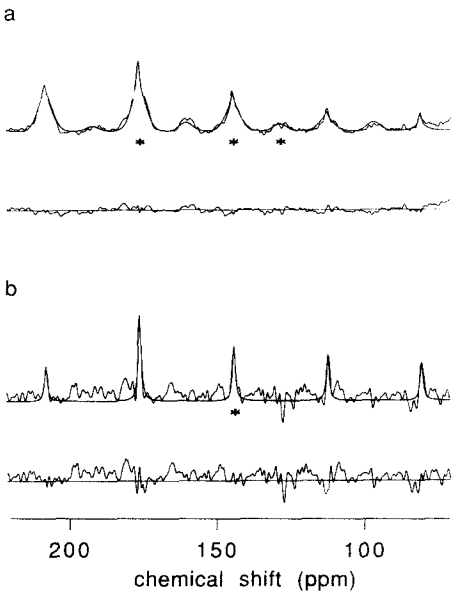

FIG. 2. ^{13}C MAS spectrum (a) and difference spectrum (b) for [5-^{13}C]retinal-labeled bacteriorhodopsin under heteronuclear (CW) decoupling at spinning speed $\omega_r/2\pi$ = 2.5 kHz. The centerbands of the label (144.3 ppm), the natural-abundance carbonyl (176 ppm), and the natural-abundance aromatic carbons (129 ppm) are marked with an asterisk. The results of the analyses are represented by the smooth line in the spectra. Below each spectrum the residues, data minus fit, are shown.

and 129 ppm, respectively), and second, the spectrum of the sample containing the ^{13}C-5 label was fitted by adding a set of Lorentzian lines for the contribution from the label, with the parameters for the natural-abundance resonances held constant. The asymmetry parameter $\eta = (\sigma_{22} - \sigma_{11})/\delta$, however, could only be determined when the resonance from the ^{13}C-5 label was separated from the natural-abundance background by taking the difference of the spectrum from the labeled sample and that from the unlabeled sample (*15*). This is demonstrated in Fig. 2b. The fit of this rather noisy spectrum yields $\omega_0\delta/2\pi = -8.8(0.9)$ kHz with $\eta = 0.6(0.2)$, in good agreement with earlier results on high-quality spectra, $\omega_0\delta/2\pi = -9.4(0.2)$ kHz and $\eta = 0.6\ (0.1)$ (*16*).

An important aspect is the sensitivity of the method for systematic errors, i.e., errors other than those from the noise in the spectrum. In Figs. 3a–3c we show a set of three

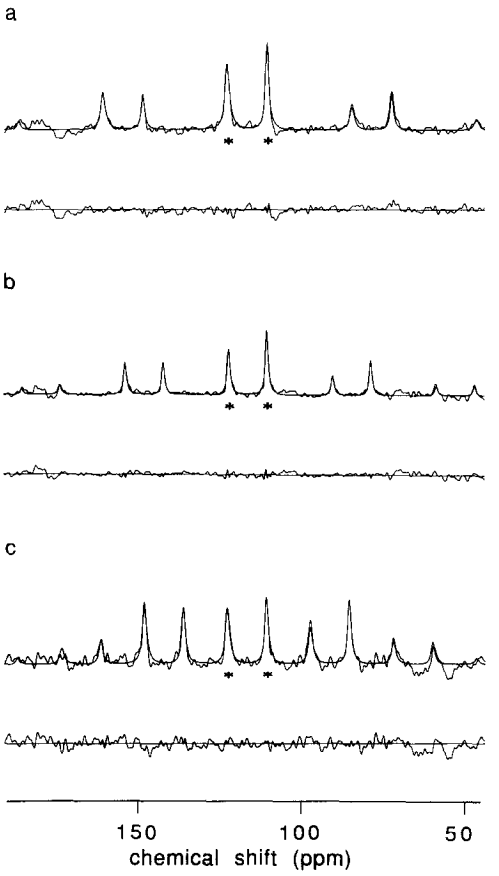

FIG. 3. ^{13}C MAS difference spectra for [14-^{13}C]retinal-labeled bacteriorhodopsin under heteronuclear (CW) decoupling at spinning speeds $\omega_r/2\pi = 3.0$ kHz (a), 2.5 kHz (b), and 2.0 kHz (c). The centerbands are marked with an asterisk. The results of the analyses are represented by the smooth line in the spectra. Below each spectrum the residues are shown.

bR difference spectra, this time labeled with ^{13}C at the 14 position of the retinylidene moiety and acquired with $\omega_r/2\pi$ = 3.0, 2.5, and 2.0 kHz, respectively. The results of the analysis are listed in Table 1. It is seen that at least two sets of rotational sidebands are needed in order to obtain accurate results. For the upfield component (111 ppm) good results are obtained already for $\omega_r/2\pi$ = 2.5 kHz, while for the downfield resonance lower spinning speeds are necessary. The reason is a difference in the asymmetry parameter η. For the upfield resonance $\omega_0\delta/2\pi$ = 5.5(0.3) kHz with η = 0.9(0.1), and for the downfield component $\omega_0\delta/2\pi$ = −5.7(0.3) kHz with η = 0.7(0.1) (*15*). It is seen that at the higher speed η is underestimated, and this causes problems when η is small. At the same time the best fit to δ is also changing with ω_r. The reason for this behavior has been known for some time (*11*). The lower η, the more information moved into one of the wings of a spin-$\frac{1}{2}$ powder pattern, and it is exactly this information that is suppressed at the highest speeds with MAS NMR. It is important to realize, however, that in such cases a high-speed spectrum yields spurious results, not simply a larger statistical error that includes the correct result, as may be concluded from the analysis of the downfield resonance in the spectra of Fig. 3. Fortunately, asymmetry parameters tend to be rather large ($\eta > 0.5$) for nonsaturated carbons in biological systems, and therefore spinning speeds $\omega_r < \omega_0\delta/2$ should be used in order to obtain reliable results.

For spectra with very good signal-to-noise, however, deviations from the Lorentzian or Gaussian lineshape may lead to a small systematic error. In that case slightly better results may be obtained by performing a direct comparison between the line intensities calculated in the computer simulations and the experimental intensities obtained from integrating the spectra.

As the final example, we show the natural-abundance ^{15}N 1D dipolar-chemical-shift spectrum of bR in Fig. 4. In the 1D dipolar-chemical-shift spectrum, which is obtained with acquisition under homonuclear proton decoupling, the sideband intensities are due to the combination of ^{15}N chemical-shift anisotropy and ^{15}N–^{1}H heteronuclear dipolar coupling. Serial analysis of two spectra, the first acquired under continuous heteronuclear decoupling to determine the chemical-shift anisotropy and the second, the dipolar-chemical-shift spectrum, is a good procedure for extracting N–H bond lengths from two one-dimensional spectra. This method is of particular

TABLE 1

Anisotropy Parameters of the Chemical-Shift Tensors (δ, η) and Isotropic Shifts (σ_i) for the Two Resonances of [14-^{13}C]Retinal in Dark-Adapted Bacteriorhodopsin at Various Spinning Speeds

$\omega_r/2\pi$ (kHz)	σ_i (ppm)	$\omega_0\delta/2\pi$ (kHz)	η	σ_i (ppm)	$\omega_0\delta/2\pi$ (kHz)	η
2.0	110.8 (0.3)	5.4 (0.5)	0.9 (0.2)	122.6 (0.3)	−5.6 (0.5)	0.7 (0.2)
2.5	110.9 (0.3)	5.5 (0.5)	0.9 (0.2)	122.2 (0.3)	−6.0 (0.5)	0.6 (0.2)
3.0	110.5 (0.3)	6.1 (0.7)	0.8 (0.2)	122.3 (0.3)	−6.7 (0.8)	0.5 (0.2)

Note. The errors (±) are given in parentheses. They represent a 95% confidence interval and are statistically determined. The errors for the isotropic shifts are determined by the accuracy of the calibration.

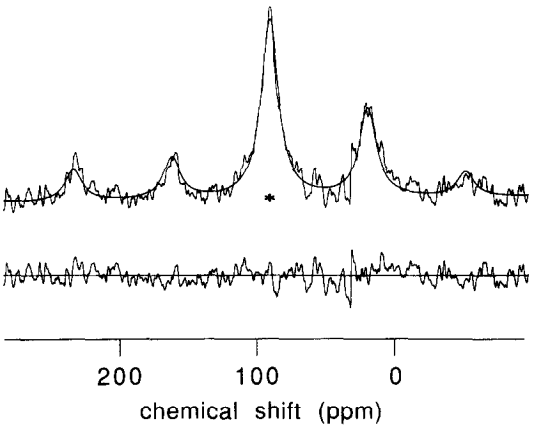

FIG. 4. 1D dipolar-chemical-shift ^{15}N MAS spectrum of natural-abundance bR under homonuclear (MREV-8) decoupling at spinning speed $\omega_r/2\pi = 2.24$ kHz. The centerband is marked with an asterisk. The results of the analysis are represented by the smooth line. The residues, data minus fit, are shown below the spectrum.

interest for large biomolecules, where acquisition of a full 2D dipolar-chemical-shift spectrum is often not feasible, due to sensitivity requirements.

The results of the fit of this spectrum are listed in Table 2. For the chemical-shift parameters, δ and η, previously determined values were employed as input to the fit (*8*). The protein contains many peptide bonds of various kinds. Nevertheless, note the striking similarity between the natural abundance protein parameters and those for [^{15}N]acetylvaline, also listed in Table 2 (*17*). The N–H bond length is approximately the same, and this appears to be the predominant factor which determines the difference in sideband intensities compared to the chemical-shift spectrum. The polar angle of the N–H bond in the principal-axis system of the ^{15}N chemical-shift tensor, χ, is also the same within experimental errors, whereas the azimuthal angle (to σ_{11}), ψ, does not have any discernible influence, for small values of χ and η (*17*), and is therefore not included in Table 2.

TABLE 2

Chemical-Shift Anisotropy (σ_i, δ, η) and Dipolar Interaction (χ, r_{N-H}) Parameters for a Small Molecule and for the Natural-Abundance Resonance of Bacteriorhodopsin

Sample	σ_i (ppm)	σ_{11} (ppm)	σ_{22} (ppm)	σ_{33} (ppm)	χ (degrees)	r_{N-H} (Å)
Natural-abundance bR	93.6 (0.4)	20 (4)	56 (4)	206 (4)	24	1.06 (0.02)
[^{15}N]Acetylvaline	96.7 (0.2)	32 (3)	57 (3)	201 (3)	22	1.049 (0.007)

Note. Only a slight dependence of the fit on ψ was observed, which is therefore not listed in the table. The errors (±) are given in parentheses. They represent a 95% confidence interval and are statistically determined. The errors for the isotropic shifts are determined by the accuracy of the calibration.

EXPERIMENTAL PROCEDURES

All NMR experiments were performed with a homebuilt spectrometer equipped with a 7.4 T magnet. The probe was also homebuilt and incorporated MAS rotors and stafors purchased from Doty Scientific, Inc. (South Carolina). The ^{1}H, ^{13}C, and ^{15}N $\pi/2$ pulse lengths were $\sim$3, $\sim$6, and $\sim$10 μs, respectively. Recycle times were 2 s and CP mixing times were 2 ms. For dipolar-chemical-shift experiments MREV-8 pulsed decoupling (*18*) was used. ^{13}C shifts are referenced to external TMS and ^{15}N shifts relative to 5.6 M $^{15}NH_4Cl$ in H_2O.

Purple membrane was obtained from a culture of the JW-3 strain using the purification procedure of Oesterhelt and Stoeckenius (*19*). The synthesis of [5-^{13}C]- and [14-^{13}C]retinal is described elsewhere (*20, 21*). The retinals were incorporated into bR following a procedure similar to that described by Smith *et al.* (*9*) and pelleted in a centrifuge before packing into rotors. For each protein spectrum, $\sim$40,000 transients were acquired with a recycle delay of 2 s. The spinning speeds were kept constant within $\sim$2 Hz with a spinning speed controller previously described (*15*).

Anisotropies are defined following Haeberlen (*22*). In the principal axis system of the chemical-shift tensor the principal components σ_{33}, σ_{22}, and σ_{11} are arranged so that

$$|\sigma_{33} - \sigma_i| \geq |\sigma_{11} - \sigma_i| \geq |\sigma_{22} - \sigma_i|$$

with the isotropic shift $\sigma_i = (\frac{1}{3})\mathrm{Tr}\sigma$. The anisotropy parameters are calculated according to

$$\delta = \sigma_{33} - \sigma_i$$

$$\eta = (\sigma_{22} - \sigma_{11})/\delta.$$

For the dipolar-chemical-shift spectrum, the parameters needed to characterize the ^{15}N–^{1}H dipolar interaction, r_{NH}, plus two polar angles χ, ψ that describe the relative orientation of the dipolar and chemical-shift tensors, are defined following Munowitz and Griffin (*12*).

CONCLUSIONS

Iterative computer fitting of MAS NMR spectra offers promising possibilities for the analysis of data with poor signal-to-noise or overlapping lines. Provided that sufficient sideband intensity is present ($\omega_r < \omega_0\delta/2$), chemical-shift anisotropies and bond lengths can be extracted from (1D) spectra with reasonable precision. The method yields a null signal, produced by subtracting the experimental and calculated data, which provides a visual assessment of the errors. In addition the statistical errors are evaluated. Knowledge of these errors is crucial to interpreting changes in chemical-shift or dipolar tensors. Together with MAS NMR difference spectroscopy, these fitting methods are particularly useful for the study of large biomolecules.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health (GM-23403, GM-23289, GM-36810, and RR-00995). H.J.M.d.G is a recipient of a research career development fellowship (Akademie-Onderzoeker) from the Royal Dutch Academy of Sciences.

REFERENCES

1. E. R. ANDREW, A. BRADBURY, AND R. G. EADES, *Nature (London)* **182,** 1659 (1958).
2. I. J. LOWE, *Phys. Rev. Lett.* **2,** 285 (1959).
3. M. MEHRING, "Principles of High-Resolution NMR in Solids," Springer-Verlag, Berlin, 1983.
4. R. G. GRIFFIN, W. P. AUE, R. A. HABERKORN, G. S. HARBISON, J. HERZFELD, E. M. MENGER, M. G. MUNOWITZ, E. T. OLEJNICZAK, D. P. RALEIGH, J. E. ROBERTS, D. J. RUBEN, A. SCHMIDT, S. O. SMITH, AND S. VEGA, "Magic Angle Spinning," Lecture Notes for the Enrico Fermi School of Physics, 8–18 July 1986, Elsevier Science, Amsterdam, 1986.
5. J. HERZFELD, S. K. DAS GUPTA, M. R. FARRAR, G. S. HARBISON, A. MCDERMOTT, S. L. PELLETIER, D. P. RALEIGH, S. O. SMITH, C. WINKEL, J. LUGTENBURG, AND R. G. GRIFFIN, *Biochemistry,* **29,** 5567 (1990).
6. S. O. SMITH, I. PALINGS, M. E. MILEY, J. COURTIN, H. J. M. DEGROOT, J. LUGTENBURG, R. A. MATHIES, AND R. G. GRIFFIN, *Biochemistry,* **29,** 8158 (1990).
7. H. J. M. DEGROOT, S. O. SMITH, J. COURTIN, E. VANDERBERG, C. WINKEL, J. LUGTENBURG, R. G. GRIFFIN, AND J. HERZFELD, *Biochemistry,* **29,** 6873 (1990).
8. H. J. M. DE GROOT, G. S. HARBISON, J. HERZFELD, AND R. G. GRIFFIN, *Biochemistry* **28,** 3346 (1989).
9. S. O. SMITH, H. J. M. DEGROOT, R. GEBHARD, J. M. L. COURTIN, J. LUGTENBURG, J. HERZFELD, AND R. G. GRIFFIN, *Biochemistry* **28,** 8897 (1989).
10. F. E. JAMES AND M. ROOS, Minuit computer code, CERN, Geneva, Program D-506.
11. J. HERZFELD AND A. E. BERGER, *J. Chem. Phys.* **73,** 6021 (1980).
12. M. G. MUNOWITZ AND R. G. GRIFFIN, *J. Chem. Phys.* **76,** 2848 (1982); **77,** 2217 (1982).
13. W. T. EADIE, D. DRIJARD, F. E. JAMES, M. ROOS, AND B. SADOULET, "STATISTICAL METHODS IN EXPERIMENTAL PHYSICS," NORTH-HOLLAND, AMSTERDAM, 1971.
14. S. O. SMITH AND R. G. GRIFFIN, *Annu. Rev. Phys. Chem.* **39,** 511 (1988).
15. H. J. M. DE GROOT, V. COPIÉ, S. O. SMITH, P. J. ALLEN, C. WINKEL, J. LUGTENBURG, J. HERZFELD, AND R. G. GRIFFIN, *J. Magn. Reson.* **77,** 251 (1988).
16. G. S. HARBISON, S. O. SMITH, J. A. PARDOEN, J. M. L. COURTIN, J. LUGTENBURG, J. HERZFELD, R. A. MATHIES, AND R. G. GRIFFIN, *Biochemistry* **24,** 6955 (1985).
17. J. E. ROBERTS, G. S. HARBISON, M. G. MUNOWITZ, J. HERZFELD, AND R. G. GRIFFIN, *J. Am. Chem. Soc.* **109,** 4163 (1987).
18. P. MANSFIELD, *J. Phys.* **C4,** 1444 (1971); W.-K. RHIM, D. D. ELLEMAN, AND R. VAUGHAN, *J. Chem. Phys.* **58,** 1772 (1973).
19. D. OESTERHELT AND W. STOECKENIUS, *in* "Methods in Enzymology" (L. Grossman and K. Moldave, Eds.), Vol. 31, p. 667, Academic Press, New York, 1973.
20. J. M. L. COURTIN, G. K. 'T LAM, A. J. M. PETERS, AND J. LUGTENBURG, *Recl. Trav. Chim. Pays-Bas* **104,** 284 (1985).
21. J. A. PARDOEN, C. WINKEL, P. P. J. MULDER, AND J. LUGTENBURG, *Recl. Trav. Chim. Pays-Bas* **103,** 135 (1984).
22. U. HAEBERLEN, "Advances in Magnetic Resonance, Suppl. 1, High Resolution NMR in Solid: Selective Averaging," Academic Press, New York, 1976.