Journal of Biomolecular NMR, **18:** 139–144, 2000.
KLUWER/ESCOM

Selection of side-chain carbons in a high-molecular-weight, hydrophobic peptide using solid-state spectral editing methods

Kristin K. Kumashiro[a,*], Walter P. Niemczura[a], Minna S. Kim[a] & Lawrence B. Sandberg[b]
[a]*University of Hawaii at Manoa, Department of Chemistry, 2545 The Mall, Honolulu, HI 96822, U.S.A.*
[b]*Connective Tissue Laboratory, Pettis Memorial Veterans' Hospital, Loma Linda, CA 92357, U.S.A.*

Received 16 June 2000; Accepted 20 July 2000

Key words: CPMAS, elastin, solid-state NMR, spectral editing

Abstract

Solid-state spectral editing techniques have been used by others to simplify ^{13}C CPMAS spectra of small organic molecules, synthetic organic polymers, and coals. One approach utilizes experiments such as cross-polarization-with-polarization-inversion and cross-polarization-with-depolarization to generate subspectra. This work shows that this particular methodology is also applicable to natural-abundance ^{13}C CPMAS NMR studies of high-molecular-weight biopolymers. The editing experiments are demonstrated first with model peptides and then with α-elastin, a high-molecular-weight peptidyl preparation obtained from the elastic fibers in mammalian tissue. The latter has a predominance of small, nonpolar residues, which is evident in the crowded aliphatic region of typical ^{13}C CPMAS spectra. Spectral editing is particularly useful for simplifying the aliphatic region of the NMR spectrum of this elastin preparation.

Introduction

Spectral editing techniques were developed to simplify NMR spectra with many overlapping lines. In high-resolution solution NMR studies, experiments like the INEPT and DEPT sequences were used to distinguish the different types of carbons, through the J-couplings. In recent years experiments have also been developed to do the analogous task in the solid state. Zilm and co-workers used the differences in cross-polarization and depolarization due to the dipolar couplings to accomplish this task (Wu and Zilm, 1993a,b; Wu et al., 1994). Specifically, the differences in heat capacities and cross-relaxation rates were exploited. Their methodologies were utilized on samples of small organic molecules and coals (Burns, 1998). Emsley and co-workers used the scalar couplings to select carbons of differing multiplicities in solid samples with low molecular weights (Lesage et al., 1998). Specifically, compounds used in these experiments have ^{13}C peaks with narrow linewidths. In addition, this methodology is extremely insensitive; for example, ^{13}C CPMAS (cross-polarization with magic-angle spinning) spectra of isoleucine required 20 scans, and the edited spectrum with comparable signal-to-noise required 12000 scans. Therefore, the signal averaging that would be required for most large biomolecules would preclude most practical applications. For these reasons, selection via the dipolar couplings is more preferable for our high-molecular-weight system, with its typically broader features and relative sensitivity limits.

The methodology of Zilm and co-workers exploits the differences in cross-polarization and depolarization rates to select for the different carbons (Wu and Zilm, 1993a,b; Wu et al., 1994). Specifically, the directly bonded CH and CH_2 groups have markedly different cross-relaxation and cross-polarization behavior, as compared to the nonprotonated and methyl carbons. Furthermore, the methine and methylene groups are distinguishable from each other, because they are known to have different heat capacities.

*To whom correspondence should be addressed. E-mail: kristin@gold.chem.hawaii.edu

Each of the different types of carbons are resolved using one or a combination of experiments. A 'methylene-only' spectrum (also called a 'subspectrum') is generated using the cross-polarization polarization-inversion (CPPI) pulse sequence (Wu and Zilm, 1993b). This experiment utilizes a short contact time to select methine (CH) and methylene (CH_2) carbons. Then, the phases of the ^1H and ^{13}C channels are changed by 180°. This inversion exactly negates the methine intensities, but the methylene resonances are observed with negative, non-zero intensities. The nonprotonated and methyl carbons are selected using cross-polarization with depolarization (CPD) (Wu and Zilm, 1993a; Wu et al., 1994). In principle, the CPD experiment is similar to the common dipolar dephasing experiment. A long contact time is used to polarize all carbons in the sample. Then, the proton frequency is turned off for a short time. During this period the methine and methylene carbons will depolarize, due to their stronger dipolar couplings. Finally, the rf is applied to both channels. During this latter 'contact' time, the ^1H field is phase-cycled through x, −x, y, −y, while the ^{13}C channel is not. The end result is complete suppression of the ^{13}CH and $^{13}CH_2$ groups. The nonprotonated and methyl carbons are then typically 'separated' by their chemical shifts. Methyl carbons are typically found in the chemical shift range of 15–20 ppm, whereas nonprotonated carbons in most diamagnetic organic and biological solids have chemical shifts found further downfield.

This work shows that the family of spectral editing techniques is extremely useful to the study of elastin, an amorphous, insoluble, high-molecular-weight (MW) protein with significant hydrophobic character. Elastin is assembled from a soluble monomeric form, called tropoelastin (∼70 kDa), and it comprises mostly amino acids with small, nonpolar side chains (Rosenbloom, 1996). The amino acid analysis of a typical elastin preparation shows that glycine (Gly, G), valine (Val, V), proline (Pro, P), and alanine (Ala, A) make up over 80% of the protein; i.e., these four amino acids account for >800 per 1000 residues. Interestingly, these four amino acids are typically found in repeating polypeptide motifs, such as $(VPGVG)_n$ and $(APGVGV)_n$. Alanines are also found in the crosslinking regions. As for the remaining 20% of the amino acids, leucine (Leu) and phenylalanine (Phe) together account for another 10%, and the last 10% comes from Hyp, Asp, Thr, Ser, Glu, Ile, Tyr, Lys, Arg, and the desmosine and isodesmosine crosslinks, which are derived from lysines. Models (Gray et al., 1973) which attempted to explain the structure function relationships present in elastin and the elastic fiber draw on the hydrophobic nature of the four major amino acids. α-Elastin is a soluble form of insoluble elastin, obtained by treating the latter with oxalic acid. The peptide 'fragments' isolated in a typical α-elastin sample have high molecular weight (≫10 kDa). These types of soluble preparations have been used in some of the pioneering biochemical studies of elastin (Partridge and Davis, 1955; Partridge et al., 1955; Whiting et al., 1974).

At present, elastin is isolated from tissue samples, and site-specific and uniform-labeling strategies are not straightforward. Therefore, strategies such as the spectral editing methodology, which target the natural-abundance ^{13}C populations in these types of biological macromolecules, must be implemented.

To demonstrate the selectivity of the CPPI and CPD experiments, the spectral editing methodology is first applied to model peptides. Specifically, the dipeptides glycylalanine (Gly-Ala), valylglycine (Val-Gly), and valylalanine (Val-Ala) are chosen because they contain all or most of the different types of carbons, i.e., methyl, methylene, methine, and nonprotonated carbon. Moreover, the choice of dipeptides is closely confined to those which contained the major amino acids of elastin. The spectra presented for all model compounds clearly show the effectiveness of the editing methodology. The editing experiments are then successfully applied to the elastin preparation to produce the appropriate subspectra.

Materials and methods

The dipeptides Gly-Ala, Val-Gly, Val-Ala were obtained from Sigma and used without further purification. The α-elastin was prepared and purified as described by others (Partridge and Davis, 1955; Partridge et al., 1955). For this study, a lyophilized α-elastin preparation was used.

All experiments were performed on a Varian Unity Inova with a proton resonance frequency of 400 MHz. A Varian double-resonance magic-angle-spinning (MAS) probe was used with 7 mm silicon nitride rotors and Torlon caps. Typical sample spinning speeds were 5–6 kHz. Typical B_1 field strengths used in the cross-polarization were $\gamma B_1/2\pi = 48$–50 kHz. TPPM decoupling was used during acquisition, with $\gamma B_1/2\pi = 55$–65 kHz. Chemical shifts are externally referenced to the TMS scale.

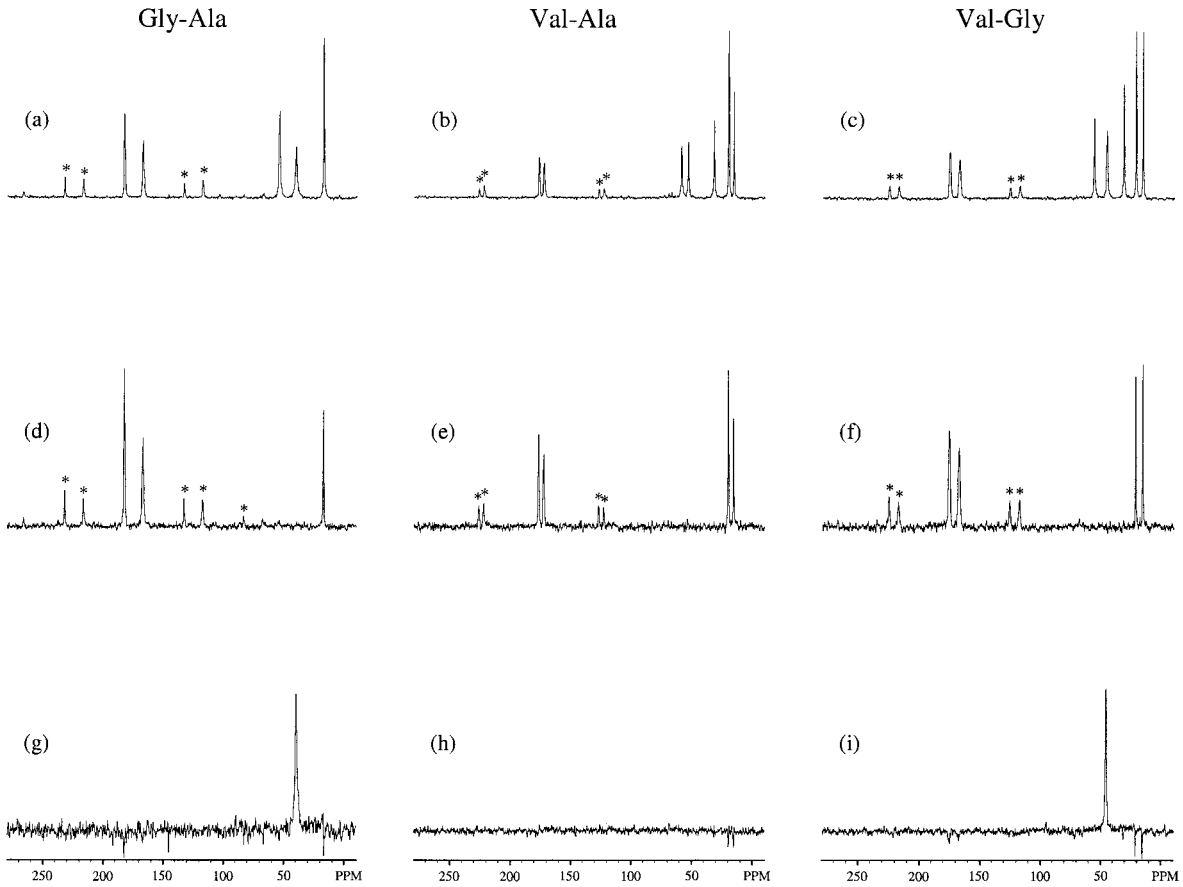

Figure 1. Spectral editing of dipeptides. Full CPMAS spectra of (a) Gly-Ala (32 scans), (b) Val-Ala (64 scans), and (c) Val-Gly (64 scans) are illustrated. The CPD pulse sequence was used to obtain subspectra with nonprotonated carbons and methyls, as shown for (d) Gly-Ala (32 scans), (e) Val-Ala (64 scans), and (f) Val-Gly (64 scans). Methylene-only subspectra were acquired with the CPPI sequence, for (g) Gly-Ala (128 scans), (h) Val-Ala (1024 scans), and (i) Val-Gly (1024 scans). All data were acquired with TPPM decoupling.

Pulse sequences for spectral editing are included in work by Zilm and co-workers (Wu et al., 1994). Optimal spectral editing parameters were determined as 140 μs (CPD and CP) or 1600 μs (CPPI) spin-locking time on ^{1}H, a maximum of 450 μs for spin-locking on the ^{13}C, cross-polarization times of 40 μs ('short contact time') or 1.5 ms ('long contact time'), a depolarization time of 200 μs, and a polarization inversion time of 32 μs.

Results and discussion

Spectral editing data for the dipeptides are shown in Figure 1. Spectra in Figures 1a–c show the full ^{13}C CPMAS spectra obtained for Gly-Ala, Val-Gly, and Val-Ala. The two peaks corresponding to the isotropic chemical shifts of the carbonyl carbons are clearly resolved. Spinning sidebands are indicated by asterisks (*). Peaks from 40–60 ppm correspond to the Cα carbons. The methine peak of valine (in Val-Gly and Val-Ala) has an isotropic chemical shift of 30–31 ppm, and the methyls are found the furthest upfield, at 15–21 ppm.

The subspectra in Figures 1d–f were obtained using the CPD pulse sequence, and they clearly show only the nonprotonated and methyl carbons. In each of the three subspectra, the isotropic peaks of the backbone carbonyls and their respective spinning sidebands are visible. In addition, the methyl peaks are clearly resolved.

Subspectra illustrated in Figure 1g–i were obtained using CPPI, to select for the methylenes. In these samples, the Cα of Gly is observed. This methylene peak is phased positively for presentation. In the Val-Ala sample, there are no signals observed, as expected.

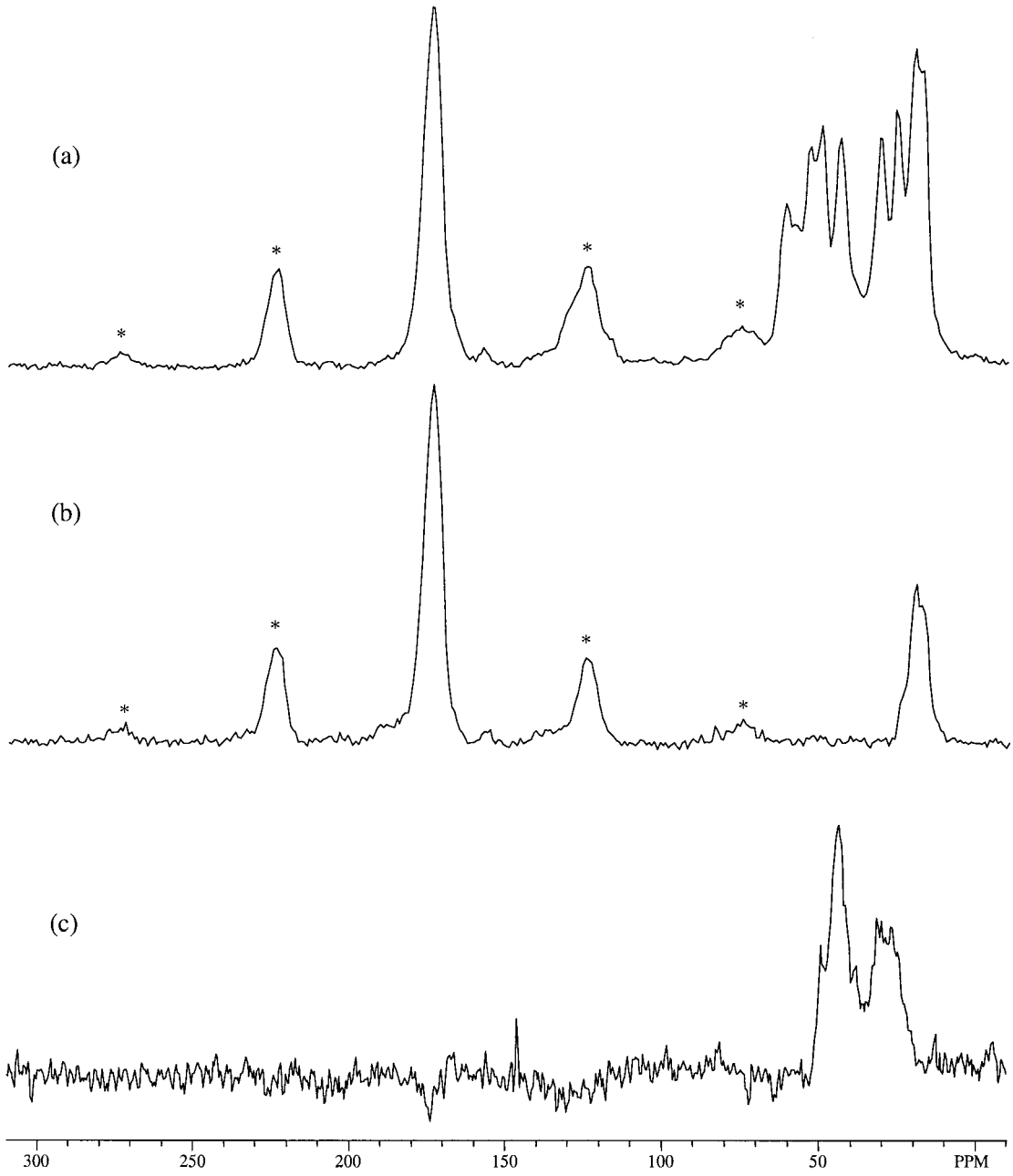

Figure 2. Spectral editing data acquired on α-elastin. All data were acquired with TPPM decoupling. (a) Full CPMAS spectrum; (b) CPD spectrum, with selection of nonprotonated and methyl carbons; (c) CPPI data, with selection of methylenes.

Residual, negative or dispersive peak intensities correspond to only the most mobile nonprotonated, methine, and methyl carbons. As shown in these data, as well as in those published by Zilm and co-workers (Wu et al., 1994), the methylene selection is definitive; i.e., the methylenes will be observed with much greater relative intensity and opposite phase as compared to those of the other types of carbons. Finally, it was also shown (Wu et al., 1994) that the methylene peak(s) will have a much lower intensity than that observed in the full CPMAS spectrum, due to the short contact times used in the selection. As expected, the signal-to-noise figures of the subspectra in Figure 1g–

i are much lower than those observed for the CPMAS or even the CPD sequences.

The ^{13}C CPMAS spectrum for the α-elastin preparation is shown in Figure 2a. In addition to the backbone carbonyl peak, there are spinning sidebands and features which correspond to the aromatic carbons. There are at least eight features in the crowded aliphatic region. Due to the natural linewidths in amorphous, heterogenous systems such as the elastin peptides, there is a great deal of overlap and baseline resolution is not obtained.

The spectrum in Figure 2b was generated using the CPD pulse sequence. CPD spectra of α-elastin contain the backbone carbonyl peak at 173 ppm. The broad feature of relatively low intensity in the range of 120–160 ppm correspond to unsaturated, non-protonated sp^2 carbons, such as those found in aromatic or heteroaromatic structures. The low intensities observed in this region are consistent with the amino acid analysis data.

The simplification with editing methods is clearly shown in the upfield region of the CPD spectrum. Herein, the majority of the aliphatic peaks have been suppressed, as expected. The upfield region of the CPD spectrum clearly indicates the presence of methyl groups, which, again, is expected with the predominance of alanines and valines in the elastin preparation. The methyl lineshape may be deconvolved into four peaks. The predominant contributions to the methyl lineshape come from peaks with isotropic chemical shifts of 16.0 ppm (33% of overall methyl intensity) and 19.3 ppm (50% of overall methyl intensity). The shoulders observed on the methyl feature may be attributed to peaks at 23.6 ppm (9%) and 13.1 ppm (4%). Clearly, this type of information cannot be directly observed in a more routine natural-abundance ^{13}C CPMAS spectrum of elastin.

Figure 2c shows the methylene-only subspectrum. Most of the CH_2 intensity is due to glycine and proline. In this sample, 335.20 glycines are found per thousand residues, in a standard amino acid analysis. Similarly, there are 109.72 prolines. Assignment of the tallest peak (43 ppm) to the Cα of glycine is based on relative intensities, and it is consistent with solution and solid-state NMR data, both as free amino acids (Pretsch et al., 1983), as well as those found in proteins and peptides (Saito and Yokoi, 1992; Wang et al., 1996). The other major features are identified on the downfield shoulder of the Cα-Gly peak and in the region of 24–31 ppm, which are tentatively assigned to the methylene carbons of proline. Again, the assignment is based on the relative amounts of proline found in this sample, as well as data in the literature. Minor contributions to the methylene lineshape are also observed, consistent with amino acid analyses which indicate that smaller amounts of leucine and isoleucine residues are also present. As expected, all other peaks in the aliphatic and aromatic region have been suppressed. There is a negative peak intensity observed for the backbone carbonyl peak. As noted earlier, non-protonated carbons with some mobility, such as in elastin, may appear in the CPPI spectrum, but the feature will have a low intensity (compared to the full CPMAS) and, more importantly, phase opposite to that observed for *any* methylene.

It is noteworthy to address the issue of mobility with regard to spectral editing. The inventors of this methodology concede that fast, large-amplitude motions, such as those present in some proteins, may compromise the effectiveness of this editing strategy. (Wu and Zilm, 1993a,b; Wu et al., 1994). However, this elastin preparation appears not to fall into this category. For example, direct polarization data (not shown) indicate that there is no significant portion of the protein which has enough mobility to preclude the use of these experiments.

Conclusions

We have demonstrated that the spectral editing techniques based on dipolar interactions can be successfully applied to a large biomacromolecule such as this elastin preparation. It is not feasible to make complete peak assignments, due to the nature of this protein. However, this methodology has been shown to be extremely useful for the spectral simplification of this high-molecular weight biopolymer. The spectral editing techniques allow one to focus on the predominant amino acids in elastin preparations and similar proteins. Specifically, their roles in the repeating polypeptide motifs or crosslinking domains may be examined by focusing on carbons in one or two of the predominant amino acids. For example, most glycines in elastin are found in the hydrophobic repeats. Using the approaches demonstrated herein, the Cα of Gly, clearly resolved in CPPI data, are ideal subjects in an experimental design. Similarly, observation or ‘targeting’ of the methyl carbons is more straightforward with solid-state editing; the CPD data will select the methyl carbons, most of which is due to the valines and alanines. And, although it is not pos-

sible to observe a single valine site, it is reasonable, feasible, and relevant (to the study of elastin) to observe the population of valine methyls and to ascertain structural and dynamical parameters of this amino acid in this biopolymer.

In future work, this family of techniques will be used in tandem with other methods in high-resolution solid-state NMR to identify structural characteristics in elastin and its related peptides.

Acknowledgements

K.K.K. acknowledges financial support from the CAREER program of the National Science Foundation (MCB-9733035). W.P.N. acknowledges helpful discussions with K.W. Zilm and S.T. Burns.

References

Burns, S.T. (1998) Ph.D. Thesis, Yale University, New Haven, CT.
Gray, W.R., Sandberg, L.B. and Foster, J.A. (1973) *Nature*, **246**, 461–466.
Lesage, A., Steuernagel, S. and Emsley, L. (1998) *J. Am. Chem. Soc.*, **120**, 7095–7100.
Partridge, S.M. and Davis, H.F. (1955) *Biochem. J.*, **61**, 21–30.
Partridge, S.M., Davis, H.F. and Adair, G.S. (1955) *Biochem. J.*, **61**, 11–20.
Pretsch, E., Seibl, J., Simon, W. and Clerc, T. (1983) *Tables of Spectral Data for Structure Determination of Organic Compounds*, English translation of the second German edition. Springer-Verlag, Berlin.
Rosenbloom, J. (1996) In *The Elastic Fiber in Health and Disease* (Ed., Moshell, A.N.), Evanston, IL, Vol. 30, pp. 1–15.
Saito, H. and Yokoi, M. (1992) *J. Biochem.*, **111**, 376–382.
Wang, J., Parkhe, A.D., Tirrell, D.A. and Thompson, L.K. (1996) *Macromolecules*, **29**, 1548–1553.
Whiting, A.H., Sykes, B.C. and Partridge, S.M. (1974) *Biochem. J.*, **141**, 573–575.
Wu, X., Burns, S.T. and Zilm, K.W. (1994) *J. Magn. Reson.*, **A111**, 29–36.
Wu, X. and Zilm, K.W. (1993a) *J. Magn. Reson.*, **A102**, 205–213.
Wu, X. and Zilm, K.W. (1993b) *J. Magn. Reson.*, **A104**, 119–122.