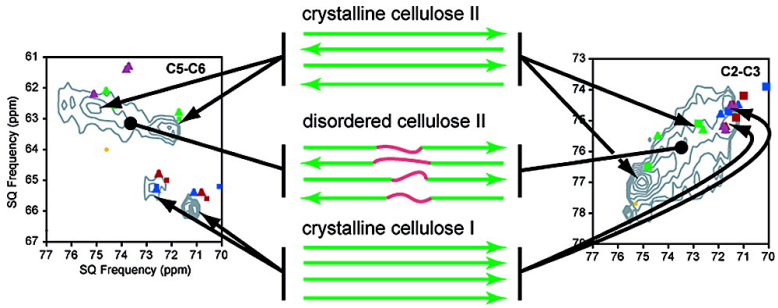
Article

Chemical Shift Correlations in Disordered Solids

Sylvian Cadars, Anne Lesage, and Lyndon Emsley

J. Am. Chem. Soc., **2005**, 127 (12), 4466-4476• DOI: 10.1021/ja043698f • Publication Date (Web): 05 March 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 10 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 03/05/2005

Chemical Shift Correlations in Disordered Solids

Sylvian Cadars, Anne Lesage, and Lyndon Emsley*

Contribution from the Laboratoire de Chimie (UMR-5182 CNRS-ENS), Laboratoire de Recherche Correspondant du CEA (23 V), Ecole Normale Supérieure de Lyon, 69364 Lyon, France

Received October 16, 2004; E-mail: Lyndon.Emsley@ens-lyon.fr

Abstract: We show how two-dimensional chemical shift conditional probability distributions can be extracted from experimental NMR correlation spectra of disordered solids. We show that transverse dephasing times are of central importance in determining the resolution (and sensitivity) of these probability distributions. These conditional probability distributions provide a new source of structural information characteristic of disordered solids, which is much more sensitive to structure than the individual-atom chemical shift distributions. The structural information contained in these distributions is clearly a potentially extremely rich source for understanding disorder. This is illustrated with examples of a phosphorus-containing organic compound and with a sample of disordered cellulose where the different structural allomorphs present are identified from the distributions.

1. Introduction

Liquid-state NMR has become an extremely powerful method for characterizing complex molecular systems, including the determination of complete protein structures. This success is largely due to the very high resolution obtained in spectra of liquids. Each different atom gives rise to a narrow resonance that can be identified in the spectrum using sophisticated multidimensional correlation techniques. The source of these narrow lines is the rapid molecular tumbling present in solution that leads to an averaging of any conformational disorder, leading to a single average chemical shift. In solids, molecular tumbling is largely absent, so solid-state NMR spectroscopy requires more sophisticated techniques to remove various sources of anisotropic broadening. Today, anisotropic broadening can be selectively removed by fast magic angle spinning (MAS),[1,2] while powerful heteronuclear decoupling techniques[3−7] have been implemented to remove the strong dipolar interactions of dilute spins such as carbon-13 with protons. The result is that high-resolution spectra can be obtained for highly ordered crystalline solids, yielding line widths of less than 50 Hz.[8] This recently enabled the complete assignment and structure determination of solid microcrystalline proteins containing up to almost 100 residues.[9−12]

Nevertheless, most solids contain some level of disorder. The result is that a given atom in a molecule will find itself in slightly different structural environments in going from one molecule (or subunit) to another. The result is a corresponding variation in isotropic chemical shift from one molecule to another, resulting in an inhomogeneous broadening of the NMR resonance. A similar broadening can arise from differences in magnetic susceptibility in different parts of the sample. Since this broadening is due to *variations in the isotropic frequency*, it cannot be removed by the averaging techniques described above. Indeed, even under fast MAS and strong heteronuclear decoupling conditions, disordered solids often yield carbon-13 line widths of several hundred hertz. This broadening is not always a disadvantage, as many authors have shown in the past that the chemical shift distribution (if it can be measured) can in favorable cases be interpreted in terms of structure.

One consequence of spectral broadening is that resonances corresponding to different chemical sites are often overlapped, preventing characterization. We have recently shown that this limitation can be overcome by recording pairwise correlation spectra.[13] Since the chemical shift variations induced by disorder are usually correlated between neighboring sites, two-dimensional (2D) correlation spectra yield peaks that can have a resolution much higher than those of the one-dimensional (1D) spectrum. We illustrated this principle using two-dimensional refocused INADEQUATE experiments.[14] This correlation effect

(1) Andrew, E. R.; Bradbury, A.; Eades, R. G. *Nature* **1958**, *182*, 1659.
(2) Lowe, I. J. *Phys. Rev. Lett.* **1959**, *2*, 285−287.
(3) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951−6958.
(4) Fung, B. M.; Khitrin, A. K.; Ermolaev, K. *J. Magn. Reson.* **2000**, *142*, 97−101.
(5) Detken, A.; Hardy, E. H.; Ernst, M.; Meier, B. H. *Chem. Phys. Lett.* **2002**, *356*, 298−304.
(6) De Paepe, G.; Hodgkinson, P.; Emsley, L. *Chem. Phys. Lett.* **2003**, *376*, 259−267.
(7) De Paepe, G.; Elena, B.; Emsley, L. *J. Chem. Phys.* **2004**, *121*, 3165−3180.
(8) Laws, D. D.; Bitter, H. M. L.; Jerschow, A. *Angew. Chem., Int. Ed.* **2002**, *41*, 3096−3129.
(9) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H. *Chembiochem* **2001**, *2*, 272−281.
(10) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98−102.
(11) Bockmann, A.; Lange, A.; Galinier, A.; Luca, S.; Giraud, N.; Juy, M.; Heise, H.; Montserret, R.; Penin, F.; Baldus, M. *J. Biomol. NMR* **2003**, *27*, 323−339.
(12) Igumenova, T. I.; Wand, A. J.; McDermott, A. E. *J. Am. Chem. Soc.* **2004**, *126*, 5323−5331.
(13) Sakellariou, D.; Brown, S. P.; Lesage, A.; Hediger, S.; Bardet, M.; Meriles, C. A.; Pines, A.; Emsley, L. *J. Am. Chem. Soc.* **2003**, *125*, 4376−4380.
(14) Lesage, A.; Bardet, M.; Emsley, L. *J. Am. Chem. Soc.* **1999**, *121*, 10987−10993.

10.1021/ja043698f CCC: $30.25

is actually widespread and is likely to be observed in most glassy[15] or polymer samples,[16] for example.

In this article, we explore the information content of the correlation peaks observed in these experiments and show that they can be used to determine structural features that cannot be determined from one-dimensional distributions. We show first what the resolution limits are in these spectra, and then we provide a method of analyzing the correlation peaks in terms of chemical shift conditional probability distributions. This principle is illustrated with applications to an organophosphorus compound and to a sample of cellulose extracted from wood.

2. Experimental Section

All experimental results in this paper were obtained on a Bruker Avance spectrometer operating at ^{1}H, ^{13}C, and ^{31}P resonance frequencies of 500.1, 125.8, and 202.5 MHz, respectively. A 2.5 mm CP-MAS probe providing MAS frequencies up to 35 kHz was used for the ^{31}P spectra, and ^{13}C spectra were acquired using a 4 mm CP-MAS probe. For two-dimensional experiments, quadrature detection in F_1 was obtained with the States method[17] for ^{31}P spectra and with the TPPI method[18−20] for the ^{13}C spectra. The pulse programs and phase cycles used are available from our website,[21] or upon request. ^{1}H decoupling was achieved using the CW,[22,23] TPPM,[3] and CM[6,7] methods, as discussed in the following. Further specific details are given in the appropriate figure captions.

3. Spectral Resolution of Correlations in Disordered Solids

The basis of the improvement in spectral resolution of disordered solids comes from the fact that chemical shifts between neighboring atoms are usually strongly correlated, in that they will vary in a determined manner in response to any given structural variation. Different techniques, involving either through-space dipolar methods[24−26] or through-bond scalar methods,[14,27−31] can be envisaged to obtain such correlations in two-dimensional spectra, involving either homonuclear or heteronuclear correlations. In this article, we concentrate on homonuclear correlation spectra acquired using the refocused INADEQUATE[14] experiment that provides through-bond correlations. The pulse sequence of this experiment is shown in Figure 1a. This technique yields correlation peaks between nuclei separated by one or two covalent bonds so that connectivity pathways between successive linked nuclei can be directly extracted. For two bonded I_A and I_B spins, the correlation peaks in an INADEQUATE spectrum appear at the double-quantum frequency in F_1, $\Omega_1 = \omega_A + \omega_B$, and at the individual single-quantum frequencies in F_2. This experiment has been shown to be a robust technique to characterize fully labelled[32,33] or natural abundance crystalline[34−36] samples as well as disordered solids,[14,15] with the I nucleus being either ^{13}C, ^{15}N, ^{29}Si, or ^{31}P. One advantage of this method is that it efficiently restricts correlations between bonded neighbors. The analysis we present here is, however, largely applicable to the other correlation methods mentioned above, including simple proton-driven spin diffusion experiments.[10,37]

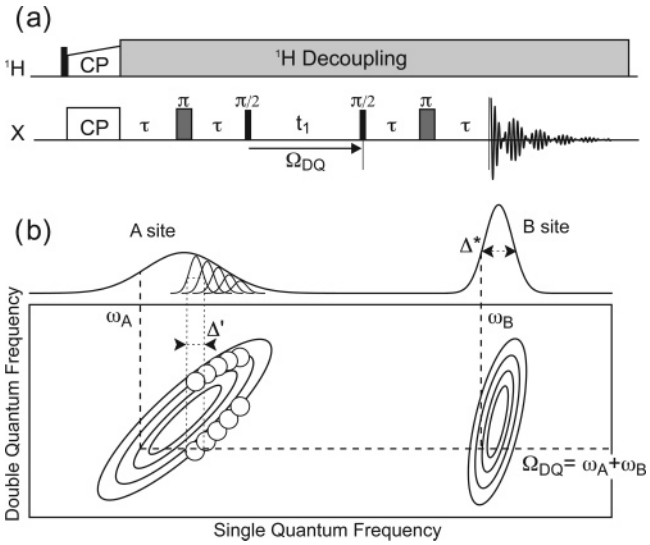


Figure 1. (a) Pulse scheme for the refocused INADEQUATE experiment. After cross polarization from protons, the I spin magnetization evolves during a $\tau-\pi-\tau$ period under the homonuclear J_{II} coupling Hamiltonian. A 90° pulse then creates double-quantum coherences between pairs of bonded nuclei that evolve during the indirect evolution period, t_1. Double-quantum coherences are then converted to anti-phase coherences by another 90° pulse. A second $\tau-\pi-\tau$ period finally converts the anti-phase coherences into in-phase ones before detection in t_2. (b) Schematic refocused INADEQUATE spectrum of a directly bonded two spins model system. The intrinsic line width, Δ', of the NMR resonances (defined in the main text) determines a "pixel" size (circles), whereas the apparent line width, Δ^*, determines the global extension of the correlation peaks. In this drawing, the scales in both dimensions and line shapes are arbitrary.

Figure 1b shows a schematic refocused INADEQUATE spectrum for a compound where the single-quantum frequencies are inhomogeneously distributed over a certain range of chemical shifts. The structure of the correlation peak observed in this spectrum depends on three factors: (i) the shape and the width of the individual chemical shift distributions for each of the two spins, (ii) the probability function that connects the different parts of the two chemical shift distributions, and (iii) the intrinsic width of the NMR resonances (also referred to below as the refocused line width of the NMR resonances).

(15) Fayon, F.; Le Saout, G.; Emsley, L.; Massiot, D. *Chem. Commun.* **2002**, 1702−1703.
(16) Kaji, H.; Schmidt-Rohr, K. *Macromolecules* **2002**, *35*, 7993−8004.
(17) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286−292.
(18) Drobny, G.; Pines, A.; Sinton, S.; Weitekamp, D. P.; Wemmer, D. *Faraday Symp. Chem. Soc.* **1979**, *13*, 49−55.
(19) Bodenhausen, G.; Vold, R. L.; Vold, R. R. *J. Magn. Reson.* **1980**, *37*, 93−106.
(20) Marion, D.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967−974.
(21) http://www.ens-lyon.fr/CHIMIE/Fr/Groupes/NMR/library/.
(22) Bloch, F. *Phys. Rev.* **1956**, *102*, 104−135.
(23) Bloch, F. *Phys. Rev.* **1958**, *111*, 841−853.
(24) Oas, T. G.; Griffin, R. G.; Levitt, M. H. *J. Chem. Phys.* **1988**, *89*, 692−695.
(25) Griffiths, J. M.; Bennett, A. E.; Griffin, R. G. *The Encyclopaedia of NMR*; Wiley: London, 1997; Vol. 4, pp 2390−2393.
(26) Levitt, M. H. *The Encyclopaedia of NMR*; Wiley: London, 2002; Vol. 9, pp 165−196.
(27) Mueller, L. J.; Elliott, D. W.; Kim, K. C.; Reed, C. A.; Boyd, P. D. W. *J. Am. Chem. Soc.* **2002**, *124*, 9360−9361.
(28) Mueller, L. J.; Elliott, D. W.; Leskowitz, G. M.; Struppe, J.; Olsen, R. A.; Kim, K. C.; Reed, C. A. *J. Magn. Reson.* **2004**, *168*, 327−335.
(29) Baldus, M.; Meier, B. H. *J. Magn. Reson., Ser. A* **1996**, *121*, 65−69.
(30) Baldus, M.; Iuliucci, R. J.; Meier, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1121−1124.
(31) Sakellariou, D.; Emsley, L. *The Encyclopaedia of NMR*; Wiley: London, 2002; Vol. 9, pp 196−211.
(32) Grasso, G.; de Swiet, T. M.; Titman, J. J. *J. Phys. Chem. B* **2002**, *106*, 8676−8680.
(33) Brown, S. P.; Perez-Torralba, M.; Sanz, D.; Claramunt, R. M.; Emsley, L. *Chem. Commun.* **2002**, 1852−1853.
(34) Christiansen, S. C.; Zhao, D. Y.; Janicke, M. T.; Landry, C. C.; Stucky, G. D.; Chmelka, B. F. *J. Am. Chem. Soc.* **2001**, *123*, 4519−4529.
(35) Kono, H.; Erata, T.; Takai, M. *Macromolecules* **2003**, *36*, 5131−5138.
(36) De Paepe, G.; Lesage, A.; Steuernagel, S.; Emsley, L. *ChemPhysChem* **2004**, *376*, 869−875.
(37) Suter, D.; Ernst, R. R. *Phys. Rev. B* **1985**, *32*, 5608−5627.

We first consider the effect of the experimental parameters on the third factor. In this discussion, it is necessary to distinguish between the *apparent line width* (Δ^*) observed in a 1D spectrum, which corresponds to the decay time (T_2^*) of the free induction decay ($\Delta^* = 1/(\pi T_2^*)$), from the *coherence lifetime* T_2', which corresponds to the *refocused line width*, Δ' ($\Delta' = 1/(\pi T_2')$) that would be observed through the time constant of the signal decay in a spin−echo experiment. We have recently published a series of papers discussing the importance and the nature of this difference.[14,36,38,39] Since two-dimensional correlation experiments, such as refocused INADEQUATE, are often based on *coherence transfer echoes*,[40] the resolution in the 2D map is (partly) determined by T_2' rather than by T_2^*. The level of detail that can be observed in a correlation peak for an inhomogeneously broadened system will thus depend on the value of T_2' for this echo, as illustrated schematically in Figure 1b where the circles or "pixels" represent the refocused line widths, Δ'. If T_2' is long (small circles), high resolution of the correlation peaks is expected, whereas short T_2' (large circles) will lead to a blurring of the correlations. We note that since different coherences are involved in t_1 and t_2, Δ' can be different in both dimensions, so that the circles may in fact be ellipses. In most solids, T_2' is not expected to be limited by relaxation due to incoherent fluctuations (T_2), but is expected to be dominated by residual coherent dipolar interactions.[7,38,39] These interactions depend on experimental parameters, such as the MAS frequency, or the quality of heteronuclear decoupling.[7,39] Thus, the optimization of these experimental parameters will lead to improved sensitivity as has been previously shown[36,38] and also to improved resolution as demonstrated experimentally below. Note that under hypothetical ideal decoupling conditions, where residual interactions would be totally removed, T_2' will tend to the limiting value T_2^{solid}, which is only determined by the local field fluctuations. Since motions are much more widespread in liquids, T_2^{solid} could be significantly longer than the corresponding liquid state T_2^{liquid}. Thus, in principle, resolution of the correlation peaks could be expected in a disordered solid (where the 1D spectrum could be very broad) to be higher than that in the spectra of isotropic liquids.

To illustrate the influence of the experimental parameters on the resolution that can be achieved in refocused INADEQUATE spectra, we used as a model system a sample of *N*,*N*-bis-(diphenylphosphino)-*N*-((*S*)-α-methylbenzyl)amine[41] (**1**), a slightly disordered crystalline compound which is shown in Figure 2a along with its ^{31}P cross-polarization magic angle spinning (CP-MAS) spectrum (Figure 2b). This spectrum displays five broad resonances, which actually correspond to eight chemically distinct phosphorus sites arising from four inequivalent molecules per unit cell in the crystal structure. In the following, the eight different sites are labeled x and x', with x ranging from 1 to 4.

Figure 2c,d shows, for two different ^{31}P resonances of (**1**) (respectively the resonances at 57.5 and 51 ppm), the experimentally determined dependence of the transverse dephasing time T_2' on the spinning frequency and on the quality of the heteronuclear decoupling. First, we notice that the general behavior of T_2' as a function of the MAS frequency differs significantly from one resonance to another. In the case of Figure 2d, increasing the MAS frequency leads to a decrease in T_2', whatever the decoupling method and the RF field strength applied. This corresponds to the observation that often decoupling performance decreases with increasing spinning speeds due to interference between ν_R and ν_1.[5,7,24,42,43] In the other case (Figure 2c), the behavior is slightly more complicated. Increasing the MAS frequency causes a decrease in T_2' at low decoupling RF field strength, whereas increasing values of T_2' are observed at higher proton nutation frequencies as a function of the MAS rate. In the latter case, this effect is more or less significant from one decoupling method to another and is more marked with CM or TPPM. In any case, increasing the proton nutation frequency always leads to an increase of the transverse dephasing time and, in particular, at high spinning frequencies. Despite the fact that all the peaks of the 1D CP-MAS spectrum do not reach their best and worst T_2' values under the same experimental conditions, we can still choose a set of experimental parameters that yield relatively long (typically 20 kHz MAS frequency and TPPM decoupling at 140 kHz) or relatively short T_2' (typically 25 kHz MAS frequency and CW decoupling at 60 kHz) for all the ^{31}P NMR resonances of **1**. It is important to note that the one-dimensional spectra for both conditions are virtually identical (data not shown).

Refocused INADEQUATE spectra recorded under both decoupling conditions are shown in Figure 3. The difference between the two spectra is remarkable. First, the sensitivity improvement due to the optimization of transverse dephasing appears clearly (here, a *factor of 20*, taking into account the difference in the number of scans). This is a particularly significant effect in refocused INADEQUATE experiments because of the presence of two relatively long spin−echo periods in the pulse sequence.[36,38] Moreover, we also observe *a noticeable improvement in resolution in both dimensions*, with the different line widths of the resonances extracted along F_1 and F_2 being reduced by up to a factor of 2, *despite no change in the 1D spectra*. The drastic improvement in resolution observed in this example clearly outlines the importance of the optimization of T_2' for disordered systems. Notably, in the short T_2' conditions, the blurring caused by larger pixels (Δ' of 25−40 Hz) leads to line shapes much less representative of the structural information present in the correlation peak. In the optimal case on the other hand, pixel sizes between 2 and 5 Hz were measured, which should not obscure structural information. In the following, we will discuss the other parameters affecting the resolution of the 2D correlation peaks.

4. Chemical Shift Conditional Probability Distributions

In particular, we turn our attention now to two other factors influencing the structure of the correlation peaks: the shape and width of the individual chemical shift distributions for the two spins, and the probability function that connects the different parts of the two distributions. In the following, we will first show how to predict the refocused INADEQUATE spectrum of a directly bonded pair of spins, starting with an arbitrary chemical shift distribution and a conditional probability matrix

(38) De Paepe, G.; Giraud, N.; Lesage, A.; Hodgkinson, P.; Bockmann, A.; Emsley, L. *J. Am. Chem. Soc.* **2003**, *125*, 13938−13939.
(39) De Paepe, G.; Lesage, A.; Emsley, L. *J. Chem. Phys.* **2003**, *119*, 4833−4841.
(40) Maudsley, A. A.; Wokaun, A.; Ernst, R. R. *Chem. Phys. Lett.* **1978**, *55*, 9−14.
(41) Robert, F.; Gimbert, Y.; Averbuch-Pouchot, M. T.; Greene, A. E. *Z. Kristallogr.* **2000**, *215*, 233−236.

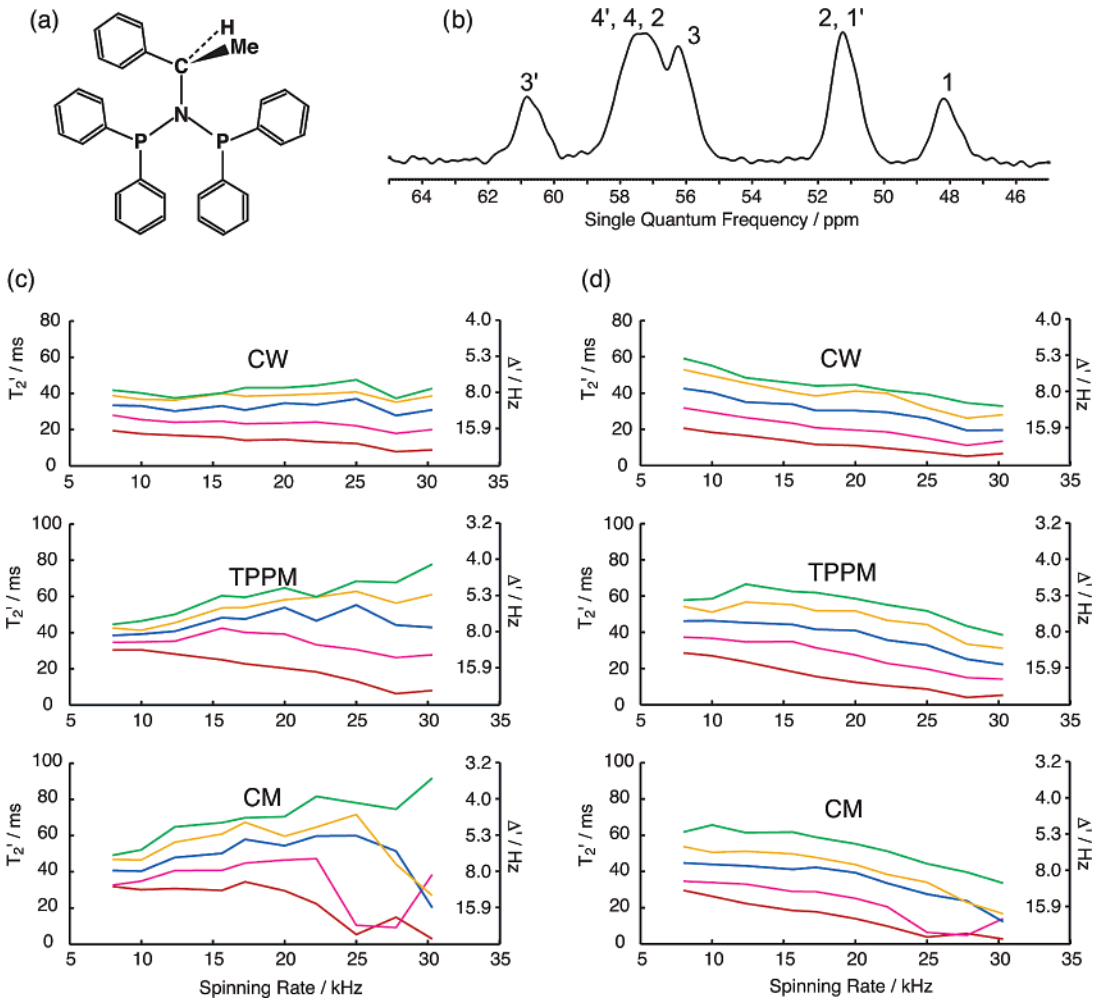


Figure 2. (a) *N*,*N*-bis(diphenylphosphino)-*N*-((*S*)-α-methylbenzyl)amine (**1**). (b) One-dimensional phosphorus-31 CP-MAS spectrum of **1** recorded under TPPM[3] decoupling at 100 kHz and a MAS frequency of 15 kHz. The number of scans was 16, and the contact time for the CP was 1.5 ms. The five broad ^{31}P resonances correspond to eight chemically distinct phosphorus sites arising from four inequivalent molecules per unit cell in the crystal structure. These eight sites were arbitrarily labeled x and x', with x ranging from 1 to 4; the pairs of resonances corresponding to bonded ^{31}P nuclei can be identified from 2D refocused INADEQUATE spectra (see ref 13). Coherence lifetimes, T_2', and refocused line widths, Δ', for the resonance at 57.5 (c) and at 51 ppm (d) as function of the spinning frequency, ν_R, and of the decoupling field strength, $\nu_1 = 60$, 80, 100, 120, and 140 kHz (in red, pink, blue, yellow, and green, respectively). The measurements were done for various decoupling schemes: CW, TPPM-15, and CM. For CM decoupling, the direct spectral optimization[6] was applied for each value of the proton RF field and each spinning frequency. T_2' values were fitted from peak intensity decay curves measured in spin−echo experiments using a monoexponential function. For each T_2' point, 16 values of echo were acquired ranging from 0 to 15 ms.

which correlates each frequency of one distribution to all of the frequencies of the second distribution.

The two chemical shift distributions for each of two connected sites A and B are noted $f(\omega_A^i)$ and $g(\omega_B^k)$. These distributions are considered to be discrete and may have different widths, so that $i \in [1...N]$ and $k \in [1...M]$. Herein, the structural disorder is assumed to be static. We then define $P(\omega_A^i)$ as the probability that one molecule in the sample has an A site which resonates at the frequency ω_A^i, and P_{tot} as the total population of this molecule in the sample. $P(\omega_A^i)$ and $P(\omega_B^k)$ can be written as

$$P(\omega_A^i) = \frac{f(\omega_A^i)}{P_{tot}} \text{ and } P(\omega_B^k) = \frac{g(\omega_B^k)}{P_{tot}} \qquad (1)$$

with the following normalization condition:

$$\sum_{i=1...N} P(\omega_A^i) = 1 \text{ and } \sum_{k=1...M} P(\omega_B^k) = 1 \qquad (2)$$

We now define the conditional probability $P_{\omega_A{}^i}(\omega_B^k)$ as the probability that the B site resonates at ω_B^k given that the A site resonates at ω_A^i (and by permutation for $P_{\omega_B{}^k}(\omega_A^i)$). These conditional probabilities verify the Bayes laws for quasi-full systems:

$$P(\omega_B^k) = \sum_{i=1...N} P_{\omega_A{}^i}(\omega_B^k) P(\omega_A^i) \qquad (3)$$

and

$$P_{\omega_B{}^k}(\omega_A^i) = \frac{P_{\omega_A{}^i}(\omega_B^k) \times P(\omega_A^i)}{\sum_{j=1...N} P_{\omega_A{}^i}(\omega_B^k) \times P(\omega_A^j)} \qquad (4)$$

The successive steps involved in the simulation of a correlation peak are schematically represented in Figure 4. One of the two chemical shift distributions, here $g(\omega_B^k)$ (Figure 4a), and a

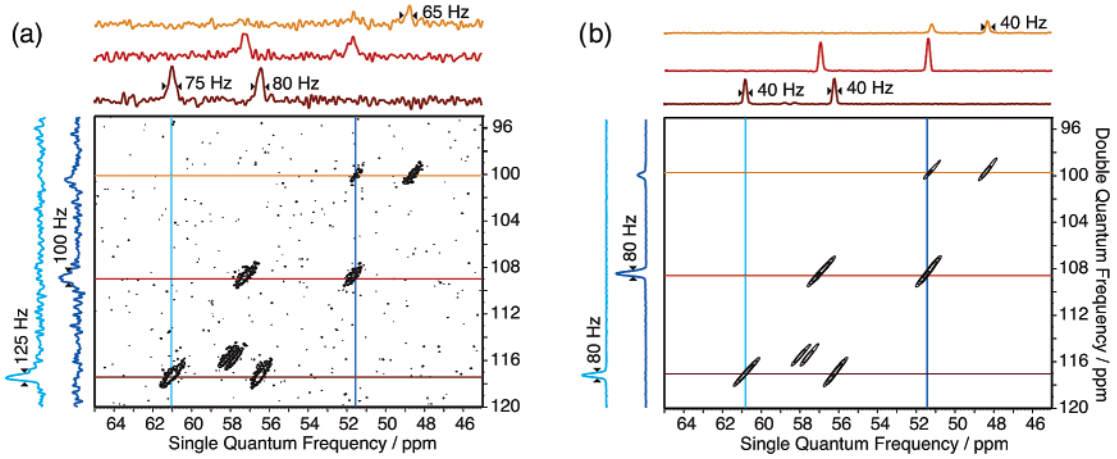


Figure 3. Refocused INADEQUATE spectra of *N*,*N*-bis(diphenylphosphino)-*N*-((*S*)-α-methylbenzyl)amine. All parameters were set identical in both experiments, except that (a) was acquired at $\nu_R = 25$ kHz, under CW ^{1}H decoupling at $\nu_1 = 60$ kHz with 64 scans, and (b) at $\nu_R = 20$ kHz, under TPPM-15 ^{1}H decoupling at $\nu_1 = 140$ kHz and with 32 scans. In both cases, the contact time for the CP step was 1.5 ms; the repetition delay was 3 s, and the length of the echo period, τ, was 5 ms. No apodization function was applied in both dimensions.

conditional probability matrix, here M_1 (Figure 4b), are taken as the starting point, with M_1 defined as follows:

$$M_1(i,k) = P_{\omega_B{}^k}(\omega_A^i) \tag{5}$$

The M_1 (N,M) dimension matrix can have any form with the only limiting condition being that the sum of all points in one column must be 1. The correlation peak between sites A and B can be obtained by multiplying each point of the M_1 matrix by the corresponding value of the B site chemical shift distribution:

$$S_{\mathrm{pCOSY}}(\omega_A^i,\omega_B^k) = P_{\omega_B{}^k}(\omega_A^i) \times g(\omega_B^k) \tag{6}$$

This correlation peak is referred to as a "pseudo-COSY" cross-peak because it corresponds to a correlation between two single-quantum frequencies (Figure 4c). We note that this contains all of the information about the distributions. Indeed, the second chemical shift distribution can be extracted by taking the horizontal projection of the pseudo-COSY correlation peak (Figure 4d):

$$f(\omega_A^i) = \sum_{k=1\ldots M} S_{\mathrm{pCOSY}}(\omega_A^i,\omega_B^k) \tag{7}$$

The second conditional probability matrix M_2 (Figure 4e), defined as follows, can be obtained by normalizing each row in the pseudo-COSY correlation peak.

$$M_2(i,k) = P_{\omega_A{}^i}(\omega_B^k) = \frac{S_{\mathrm{pCOSY}}(\omega_A^i,\omega_B^k)}{f(\omega_A^i)} \tag{8}$$

It is then possible to construct a complete pseudo-COSY spectrum (Figure 4f) by introducing appropriate "experimental" parameters and by generating the second cross-peak from the transpose of the first one, as shown in Figure 4f. The final step in obtaining an INADEQUATE-type 1Q−2Q correlation spectrum consists of a geometric "shearing transformation" (Figure 4g), which corresponds to a ω_2-dependent frequency shift in ω_1, with the shift being given by

$$\Delta\omega_1{}^{\mathrm{pCOSY}\rightarrow\mathrm{INAD}} = \omega_2 \tag{9}$$

This yields the two INADEQUATE cross-peaks, $S^1_{\mathrm{INAD}}(\Omega_{\mathrm{DQ}},\omega_A)$ and $S^2_{\mathrm{INAD}}(\Omega_{\mathrm{DQ}},\omega_B)$, where $\Omega_{\mathrm{DQ}} = \omega_A + \omega_B$. With the method presented above, we can start from totally arbitrary chemical shift distributions and conditional probability matrices and predict the influence of different parameters on the shape of the refocused INADEQUATE correlation peak.

Figure 5 shows the result obtained with various starting parameters. For each case, we show the M_1 conditional probability matrix and the $g(\omega_B)$ chemical shift distribution of the B site (at about 110 ppm). Those two functions determine the resulting spectrum, as is shown above. Figure 5a−d shows simple examples of the influence of the parameters on the peak shapes. In going from 5a to 5b, we broaden the correlation function in M_1, and this produces a corresponding broadening of the correlation peaks in the spectrum. In 5c, we change the orientation of the correlation function (from $y = kx$ to $y = -kx$, with $k > 0$). It can be seen that if the high frequencies in the A site are correlated to the high frequencies in the B site and low frequencies to low frequencies, then the peaks tend to be more or less parallel (5a), while in the contrary case (5c), they are more or less perpendicular. The correlation peaks of the constructed refocused INADEQUATE spectrum also depend on the chemical shift distribution, as illustrated in Figure 5d, where we have used a narrower distribution. One may think that from these simple cases, the detailed analysis presented here is not particularly necessary since the changes in the parameters produce straightforward changes in the spectra. However, Figure 5e−g illustrates more complex situations. Figure 5e shows a distribution consisting of a sum of two Gaussians, and the effect on the spectrum is clearly observed in comparison to that in Figure 5a, for example. Finally, Figure 5f,g shows the kind of correlation line shapes that can be expected from more complicated conditional probability matrices. The M_1 matrix in Figure 5f,g is made up of three points of maximum probability, which may represent the case of disordered samples with several stable conformations. Such complicated conditional probability matrices will probably also occur in conjunction with more irregular chemical shift distributions (as can be seen, for example, in the 1D spectrum of cellulose in Figure 7a), and this is illustrated in Figure 5g. In these two latter cases, we clearly see that the conditional probability matrices are not

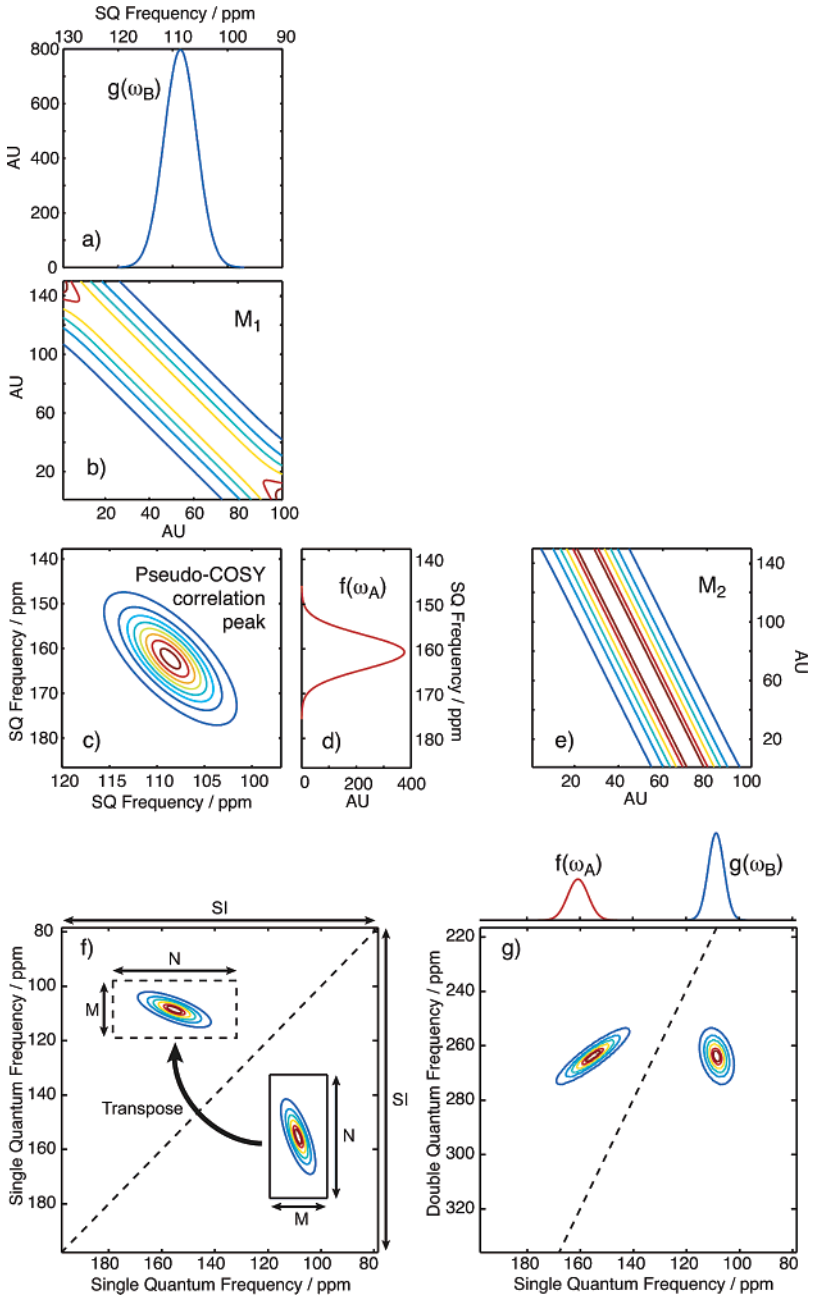


Figure 4. Step-by-step simulation of the refocused INADEQUATE correlation peaks for a directly bonded two-spin system. (a) The starting chemical shift distribution, $g(\omega_B)$, is chosen as a Gaussian having arbitrary width, amplitude, and central chemical shift. (b) The starting conditional probability matrix, M_1, has a linear profile along the $y = -kx$ (k being equal to the N/M ratio) dimension and a Gaussian shape perpendicular to this axis. It is so that the sum of all points in each column is equal to 1. The "pseudo-COSY" correlation peak (c) is obtained by multiplying each column in this matrix by the corresponding intensity of the starting chemical shift distribution. Then, the projection on the F_1 dimension yields the A site chemical shift distribution, $f(\omega_A)$ (e), whereas normalizing the pseudo-COSY cross-peak on the opposite dimension yields the second probability matrix, M_2 (d). (f) To get a pseudo-COSY spectrum, the correlation peak of (c), together with its transpose, is inserted into a 2D map choosing arbitrary values for the spectral widths, the number of points SI and the parts per million scales. The spectral width and the number of points SI are chosen to be equal in both dimensions so that the ratio of the numbers of points N and M determines the height and the width of the correlation peaks in the entire spectrum. The simulated pseudo-COSY spectrum is converted into the corresponding refocused INADEQUATE spectrum (g) by a geometric shearing transformation, with the corresponding 1D spectrum being shown above.

intuitively obvious simply by inspection of the cross-peaks. This highlights the pertinence of this kind of analysis. Note that the second chemical shift distribution $f(\omega_A)$ (not shown) obtained from both f and g of Figure 5 looks like a sum of two Gaussians, which is also not clearly expected from the starting conditional probability matrix. In conclusion, we can see that the chemical shift distribution and the M_1 matrices of eq 5 lead to characteristic variations in the 2D correlation spectrum.

5. Experimental Determination of Conditional Probability Distributions

By reversing the logic of section 4, we should be able to experimentally determine the conditional probability distributions. In this section, we present this in detail using **1** as an example, the corresponding refocused INADEQUATE spectrum being shown in Figure 3b. The method will be then implemented on a disordered cellulose sample, where the conditional prob-

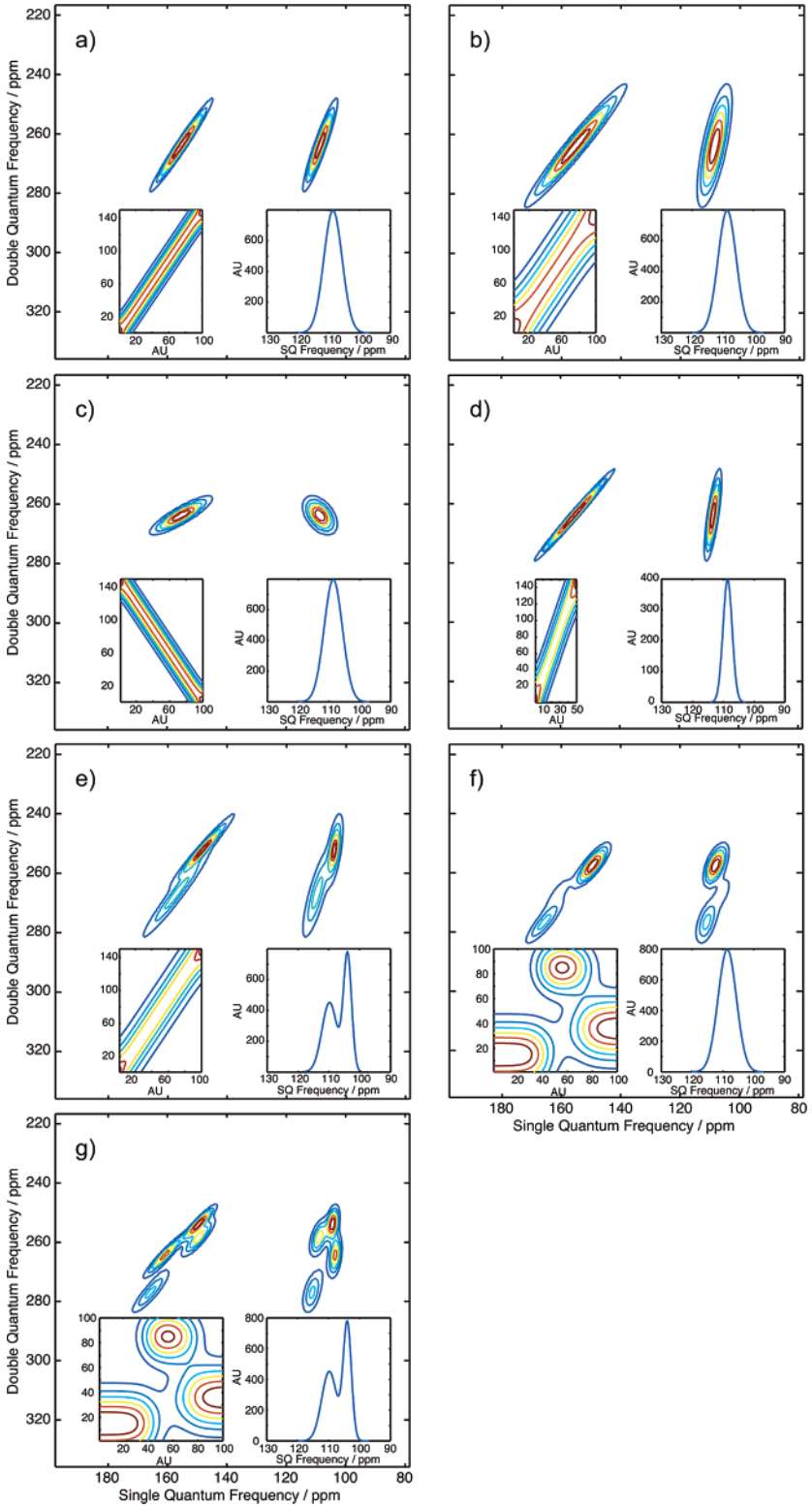


Figure 5. A set of simulated refocused INADEQUATE spectra with their corresponding starting chemical shift distributions and conditional probability matrix are shown. The refocused INADEQUATE spectrum in (a) is constructed from the conditional probability matrix, $M_1(N,M)$, shown in the left corner inset, and a Gaussian chemical shift distribution, $g(\omega_B)$, shown in the right corner inset. For the definition of M_1, the ratio between the number of points in both dimensions N/M was set to 1.5 (which determines the respective width of the two chemical shift distributions). Here, the matrix is constructed as a Gaussian profile around a given linear function. This matrix is then normalized so that the sum of all points in each column is equal to 1. In (b), all parameters are identical to those in (a), except that the Gaussian profile of the correlation function is broader. (c) The slope of the correlation function is reversed with respect to (a) in this example. (d) The starting probability matrix used here is similar to that used in (a), but the ratio of the width, here, is 3 (instead of $^3/_2$). In (e), a more complicated chemical shift distribution is used, consisting of a sum of two Gaussians. In (f), an example is shown with a complicated starting conditional probability matrix. It was obtained from a matrix where the intensity of each point was a sum of three Gaussian functions of its distance to three arbitrary points in the matrix. This matrix was then normalized so that the sum of all points in each column is equal to 1, which gives the conditional probability matrix, M_1. (c) The same M_1 matrix is used here in conjunction with a chemical shift distribution, $g(\omega_B)$, being a sum of two Gaussians.

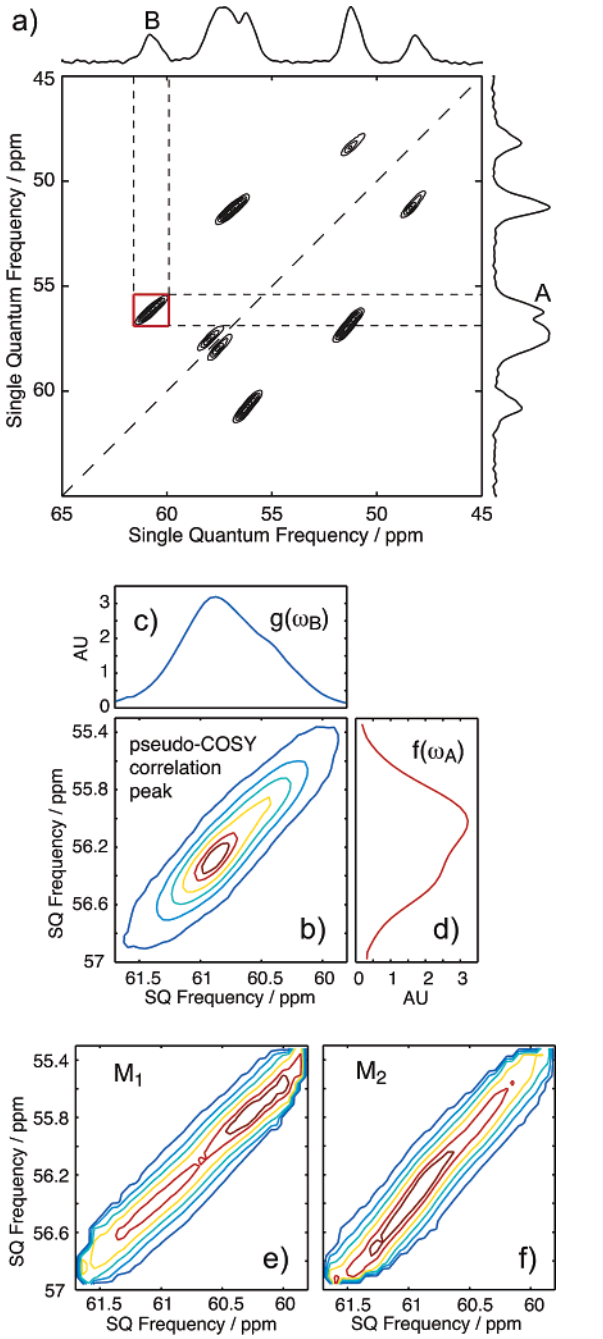


Figure 6. (a) The experimental pseudo-COSY spectrum obtained from the refocused INADEQUATE spectrum of Figure 4b by a shearing transformation. In (b), we show the enlargement of the boxed cross-peak in (a). The extracted chemical shift distributions of sites A and B extracted from (b) are shown in (d) and (c), respectively. The resulting conditional probability matrices, M_1 and M_2, are shown in (e) and (f). Note that the so-called A and B sites actually correspond to the 3 and 3′ sites, respectively, using the notation of Figure 2.

ability distributions are shown to be more complicated, and yield information about structure.

The first step consists of the conversion of the refocused INADEQUATE spectrum of Figure 3b into a pseudo-COSY spectrum using a shearing transformation.[44−46] Starting from

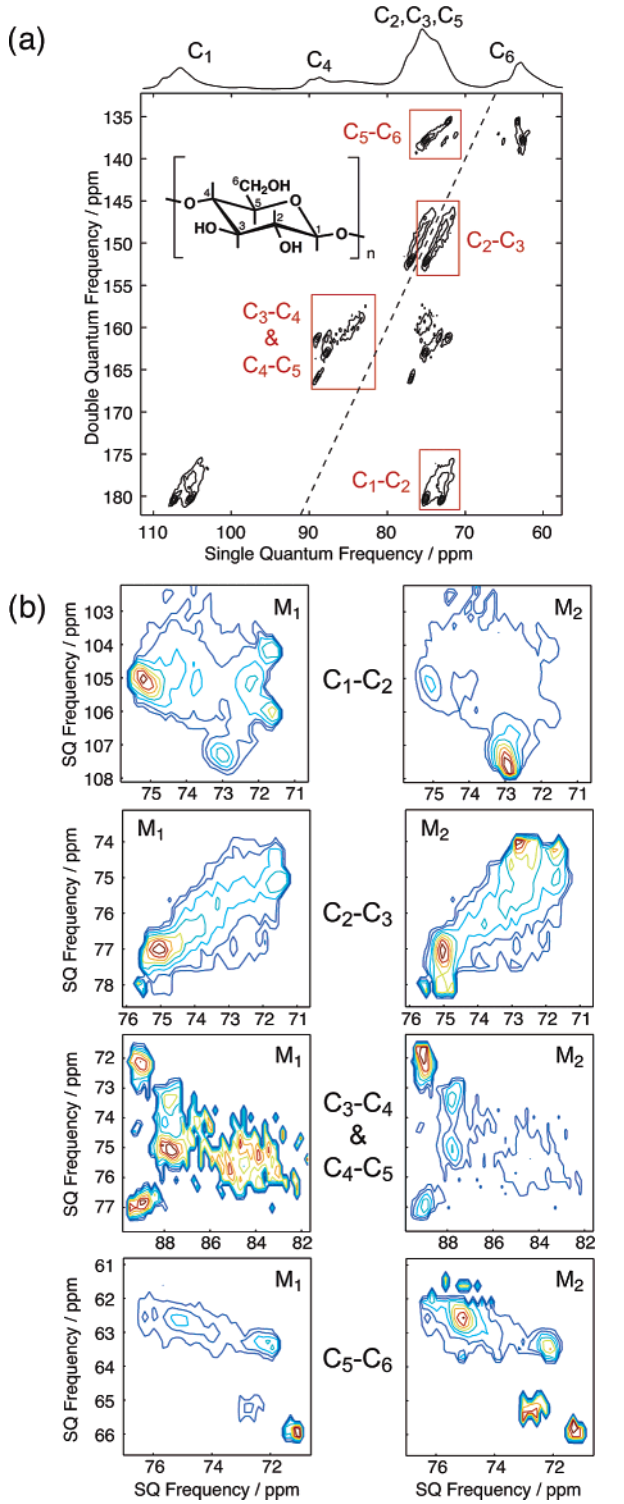


Figure 7. (a) The refocused INADEQUATE spectrum of an 11% ^{13}C-enriched cellulose sample extracted from wood. Acquisitions were made under CM ^{1}H decoupling at $\nu_1 = 100$ kHz and a MAS frequency, $\nu_R = 12.5$ kHz. A total of 288 t_1 increments were used with 384 scans for each one. The CP time was optimized at 1 ms, and the τ period was set to 2.4 ms. In (b), several conditional probability matrices are shown that were extracted from the four boxed correlation peaks in (a). For these pictures, a filter was used so that a point in the pseudo-COSY spectrum having intensity not higher than the noise yields a zero value in both M_1 and M_2 matrices, as discussed in the text. Moreover, for the correlation peak between C_2 and C_3, the part of the symmetric cross-peak which is included in the upper-left corner of the box in (a) was artificially set to zero, so that the corresponding intensity does not appear in the resulting conditional probability matrices.

(42) Raleigh, D. P.; Kolbert, A. C.; Oas, T. G.; Levitt, M. H.; Griffin, R. G. *J. Chem. Soc., Faraday Trans.* **1988**, *84*, 3691−3711.
(43) Levitt, M. H.; Oas, T. G.; Griffin, R. G. *Isr. J. Chem.* **1988**, *28*, 271−282.
(44) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Oxford: New York, 1987.
(45) Grandinetti, P. J. Private communication.

an INADEQUATE spectrum having two correlation peaks centered at $S^1_{INAD}(\Omega_{DQ},\omega_A)$ and $S^2_{INAD}(\Omega_{DQ},\omega_B)$, we obtain the pseudo-COSY peaks with pseudo-frequencies, $S^1_{pCOSY}(\omega_B,\omega_A)$ and $S^2_{pCOSY}(\omega_A,\omega_B)$. In practice, this corresponds to a first-order phase correction of each column in the F_2 Fourier transformed data, the first-order phase factor being the corresponding ω_2:

$$S_{pCOSY}(t_1(k),\omega_2) = S_{INAD}(t_1(k),\omega_2)\exp(it_1(k)\times\omega_2) \quad (10)$$

We note that the details require taking into account the quadrature detection method, so that F_1 phase corrections can be applied properly to the data. (The program used for data processed using the States quadrature detection method is available on our web site,[21] or by request.) The resulting pseudo-COSY spectrum of **1** is presented in Figure 6a. From the pseudo-COSY spectrum, the chemical shift distributions for each of the two correlated sites can be obtained by isolating either of the two correlation peaks and by projecting them perpendicularly onto the F_1 and F_2 axes. For example, for the correlation peak extracted from the pseudo-COSY spectrum in Figure 6a, which is enlarged in 6b, both chemical shift distributions obtained in this way are shown in Figure 6c,d. Note that this is even possible for the peaks between 55 and 59 ppm in F_2 (data not shown), *which belong to overlapping peaks in the 1D spectrum*. Then, by normalizing each column in the correlation peak for each value of ω_2, such that the sum of all intensities in the column is 1, we obtain the experimental conditional probability matrix M_1 (Figure 6e) as

$$M_1(\omega^i_A,\omega^k_B) = P_{\omega_B{}^k}(\omega^i_A) = \frac{S_{pCOSY}(\omega^i_A,\omega^k_B)}{g(\omega^k_B)} \quad (11)$$

Note that here, M_1 is described as a map, where a point, $M_1(\omega^i_A,\omega^k_B)$, yields the probability that site A resonates at ω^i_A given that site B resonates at ω^k_B. It corresponds to $M_1(i,k)$ described above, but it is more convenient for a chemical analysis to present this map in terms of chemical shifts. (In this description, ω^i_A represents a frequency in the F_1 dimension and appears first in the $M_1(\omega^i_A,\omega^k_B)$ map notation, and i index is a row index in the matrix notation and also appears first.) In the same way, M_2 conditional probabilities are obtained by normalizing each row in the correlation peak, and a point, $M_2(\omega^i_A,\omega^k_B)$, on this map gives the probability that site B resonates at ω^k_B given that site A resonates at ω^i_A (Figure 6f). We note that if the signal-to-noise ratio in the spectrum is particularly poor, then the conditional probability matrices will produce meaningless values (that is, conditional probabilities are also produced for the noise, but these are not interesting). This can be overcome by applying a simple filter. Conditional probabilities are only generated if the signal of a given point in the pseudo-COSY spectrum is above a threshold value determined by the user to be significant. Otherwise the conditional probability is set to zero. This procedure has been applied to determine all the maps shown in Figures 6−8.

For **1**, we note immediately the characteristic linear shape of the experimentally determined probability matrices. In principle, this shape could arise from two potential mechanisms, (i) variations in bulk magnetic susceptibility and (ii) slight structural disorder that would lead to the same chemical shift change for both bonded nuclei. These are not easy to distinguish in this particular case since susceptibility broadening would lead to the same type of stretching of the peaks parallel to the pseudo-COSY diagonal that is observed for this compound. However, susceptibility effects are, in general, not expected to play a significant role in this type of system, where the limiting resolution would typically be around 50 Hz. As this sample is sufficiently crystalline for there to be an X-ray crystal structure,[41] the structural variations that lead to this parts per million broadening must be slight, illustrating that *the chemical shift is clearly a very sensitive probe of disorder*. If this broadening is due to structural variations, then the chemical shift distributions and conditional probabilities determined here provide very strong constraints on the structure since the structural changes lead to identical changes in chemical shifts for both phosphorus atoms in each molecule. We are currently developing simulations using DFT methods to determine the structural changes that could give rise to these variations. Initial results indicate (not surprisingly) that *the structural distortions involved are probably symmetric with respect to the two phosphorus atoms* (data not shown). For example, variation of a single torsion angle in the molecule leads to changes in chemical shifts for the two phosphorus atoms that are different at the two sites (and different from one inequivalent molecule to another). We are now investigating symmetric distortions, and this work will be described elsewhere; however, one can clearly appreciate the structural relevance of the spectra and the spectral analysis presented here. It is most important to note that there are many structural distortions that would lead to spectra compatible with the one-dimensional spectra and one-dimensional chemical shift distributions, but that there are much less that agree with the data from the two-dimensional conditional probability distributions.

In general, more complicated correlation peaks can be expected, so that complex-shaped conditional probability matrices, such as that of Figure 5f,g, are possible and which necessitate the analysis provided above in order to determine the probability distributions. A good example is given in Figure 7, where different experimental conditional probability matrices have been extracted from the pseudo-COSY spectrum (not shown) obtained by shearing the refocused INADEQUATE spectrum of an 11% ^{13}C-labeled cellulose sample extracted from wood,[47] shown in Figure 7a. First of all, we note that these probability distributions present significantly different shapes from one pair to another. For example, the C_1-C_2 conditional probability matrices show a complicated structure with four maximum probability points and a continuum "horse-shoe"-shaped broad distribution. On the contrary, note that the C_2-C_3 correlation has a structure similar to that observed in **1**, with a narrow $y = x$ correlation, ranging over 5 ppm. The C_5-C_6 correlation has what appears more or less to be a $y = -kx$ chemical shift correlation structure, with k being around 0.25, and two isolated high-probability points. Finally, the C_3-C_4 and C_4-C_5 correlations have the most complicated structures, due to the fact that there are two overlapped correlation peaks.

There are many existing studies of cellulose by NMR in the literature, and it corresponds to an extremely challenging

(46) Fayon, F.; King, I. J.; Harris, R. K.; Gover, R. K. B.; Evans, J. S. O.; Massiot, D. *Chem. Mater.* **2003**, *15*, 2234−2239.

(47) Bjorkman, A. *Ind. Eng. Chem.* **1957**, *49*, 1395−1398.

structural system.[35,48−51] For example, there is often coexistence of several stable isomers of the glucose units of cellulose. It has been shown that native cellulose is composed of two forms, I_α and I_β,[52,53] and even in crystalline cellulose, several allomorphs (e.g., I, II, III...) have been characterized by solid-state NMR[49,54,55] or by using X-ray crystallography and neutron diffraction.[56−58] These forms notably differ in the packing of the cellulose chains. Particularly, cellulose I structures are composed of parallel cellulose chains, whereas in cellulose II, the chains are antiparallel. In disordered cellulose, such as the sample we studied here, which has been regenerated after extraction from wood,[47] we expect many different structures to be possibly present depending on the spatial packing of the different chains, on the variations in the conformation of the glycosidic linkage between C_4 and C_1 of two successive units, and on the conformation of the hydroxymethyl moieties. The one-dimensional single-atom distributions that have been studied until now do not allow a determination of the different allomorphs present in the sample since there are several potential sources of disorder that will all lead to similar broad one-dimensional spectra. The two-dimensional probability distributions should be able to distinguish between the different structures.

Figure 8 shows a superposition of the experimental conditional probability matrices from Figure 7b and recent measurements of cellulose chemical shifts for various crystalline polymorphs from the literature.[35,49,50,59] We note that only very recent data are presented since to place unambiguous points on the spectra, complete residue assignments are necessary, and there are only few examples of this in the literature, despite the large number of studies on cellulose NMR. The literature data refer to measurements of chemical shifts for the two magnetically inequivalent residues in crystalline cellulose I_α, I_β, and II. These measurements were made using refocused INADEQUATE spectra. We also show the resonances of the single residue of crystalline cellulose III_I determined unambiguously from a one-dimensional spectrum. (The measurements were made on celluloses of different origin, *Cladophora*, tunicate, *Valonia*, and bacterial celluloses, as described in detail in references.)[35,49,50,59]

It is remarkable that the spread in the literature data is well superposed to the maxima of the conditional probability matrices, including the C2−C3 and C2−C1 matrices. Moreover, the high-probability regions almost all correspond (remarkably well) with resonances of crystalline polymorphs, suggesting that the sample is made up primarily of a superposition of known

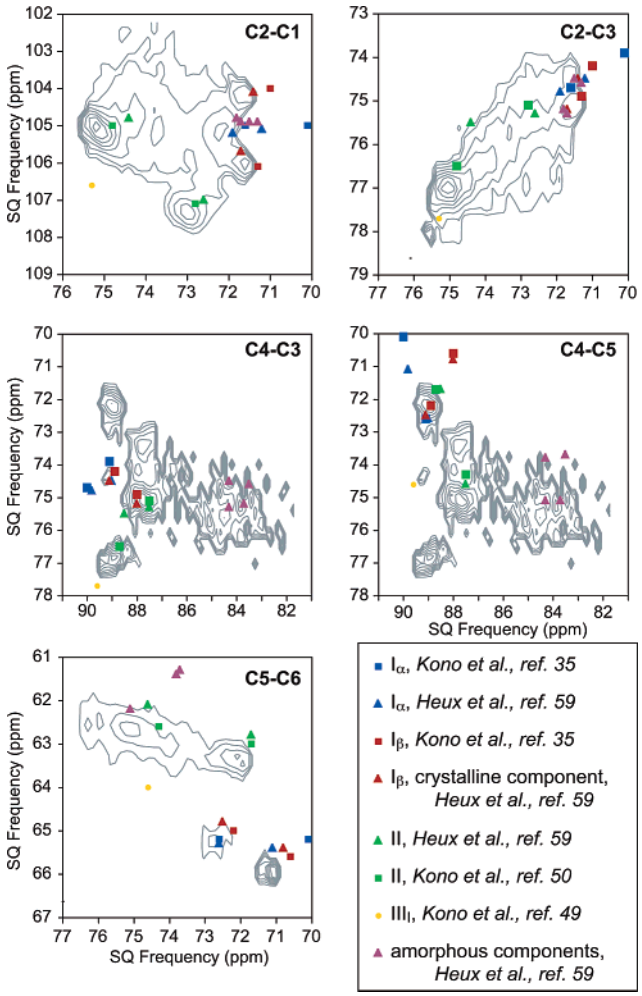


Figure 8. Superposition of the conditional probability matrices of Figure 7b and chemical shifts measured for several crystalline allomorphs of cellulose obtained from the literature.[35,49,50,59] Note that celluloses I_α, I_β, and II all have two inequivalent residues. Cellulose III_I only has one residue.

stable allomorphs. We can immediately determine that the disordered sample is made up of cellulose I_α and I_β, as expected for native cellulose, but that there is clearly a significant amount of cellulose II in the sample. The cellulose II structures were undoubtedly created during the regeneration process.[47] On the contrary, there is apparently no cellulose III_I structure present (as expected since cellulose III_I is generated from cellulose I by treatment with liquid ammonia or ethylenediamine[49]). The C4−C3 and C4−C5 distributions show agreement with previous measurements of "amorphous components",[59] without any structural assignment. In our data, the C1−C2, C2−C3, and C5−C6 conditional probabilities also clearly have intensity in regions outside the crystalline components. The structures of the C2−C3 and C5−C6 correlations (especially that of C5−C6) suggest that the disordered components may be mostly variations of the cellulose II structures (though this hypothesis needs to be confirmed by calculation[51,60,61]).

The above detailed structural analysis is clearly not possible on the basis of the one-dimensional distributions, yet it is straightforward from the two-dimensional analysis.

(48) Atalla, R. H.; VanderHart, D. L. *Solid State Nucl. Magn. Reson.* **1999**, *15*, 1−19.
(49) Kono, H.; Erata, T.; Takai, M. *Macromolecules* **2003**, *36*, 3589−3592.
(50) Kono, H.; Numata, Y.; Erata, T.; Takai, M. *Macromolecules* **2004**, *37*, 5310−5316.
(51) Sternberg, U.; Koch, F. T.; Priess, W.; Witter, R. *Cellulose* **2003**, *10*, 189−199.
(52) Vanderhart, D. L.; Atalla, R. H. *Macromolecules* **1984**, *17*, 1465−1472.
(53) Atalla, R. H.; Vanderhart, D. L. *Science* **1984**, *223*, 283−285.
(54) Fyfe, C. A.; Dudley, R. L.; Stephenson, P. J.; Deslandes, Y.; Hamer, G. K.; Marchessault, R. H. *J. Macromol. Sci.* **1983**, *C23*, 187−216.
(55) Isogai, A.; Usuda, M.; Kato, T.; Uryu, T.; Atalla, R. H. *Macromolecules* **1989**, *22*, 3168−3172.
(56) Langan, P.; Nishiyama, Y.; Chanzy, H. *J. Am. Chem. Soc.* **1999**, *121*, 9940−9946.
(57) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074−9082.
(58) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300−14306.
(59) Heux, L.; Marais, M.-F.; Trierweiler, M. Private communication.
(60) Koch, F. T.; Priess, W.; Witter, R.; Sternberg, U. *Macromol. Chem. Phys.* **2000**, *201*, 1930−1939.
(61) Witter, R.; Priess, W.; Sternberg, U. *J. Comput. Chem.* **2002**, *23*, 298−305.

6. Conclusions

We have shown how chemical shift conditional probability distributions can be extracted from experimental two-dimensional correlation spectra of disordered solids. Notably, we show that the transverse dephasing times (and their optimization)[36,38] are of central importance in determining the resolution (and sensitivity) of these probability distributions. These conditional probability distributions provide a new source of structural information characteristic of disordered solids, which is more sensitive to structure than the individual-atom chemical shift distributions. (We note that a useful side effect of this method is that it provides a way of measuring the individual-atom distributions in the presence of overlap in the 1D spectra.) The information contained in these two-dimensional distributions is clearly potentially extremely rich to understand structural sources of disorder, as this is illustrated with the examples given above where the distributions for cellulose could be directly interpreted in terms of structure. We note that, here, we have considered each pairwise conditional probability in isolation. In systems containing several connected spins, such as the example of cellulose, one can envisage determining the conditional probabilities for the entire spin system.

We are currently working on a way to correlate these conditional probability maps with theoretical studies of chemical shifts as a function of different structural parameters, such as dihedral angles. In this context, we note that the analysis presented here contains no ambiguity since it remains in spectral coordinates. It is well known that the transformation to structural coordinates of this kind of data will lead to ambiguities that require the introduction of prior information.[62−65] Finally, we note that this method of measuring experimental two-dimensional chemical shift correlations to characterize disorder can be applied to studies of a range of different disordered systems, for example, peptides (and proteins), polymers,[16] silicates,[34,66] or glasses,[67] where one-dimensional distributions have already been used in many cases.

Acknowledgment. We are grateful to Y. Gimbert and M. Bardet (Grenoble) for providing us with the samples used in this work, to S. Hediger (Lyon) and J. Baltisberger (Berea) for stimulating discussions, and to L. Heux (Grenoble) for providing us with reprint of reference.[59] Finally, we would like to thank S. Perez (Grenoble) for sharing his wide knowledge of the recent advances in cellulose structures.

JA043698F

(62) Spiess, H. W. *Chem. Rev.* **1991**, *91*, 1321−1338.
(63) Wefing, S.; Spiess, H. W. *J. Chem. Phys.* **1988**, *89*, 1219−1233.
(64) Wefing, S.; Kaufmann, S.; Spiess, H. W. *J. Chem. Phys.* **1988**, *89*, 1234−1244.
(65) van Beek, J. D.; Meier, B. H.; Schafer, H. *J. Magn. Reson.* **2003**, *162*, 141−157.
(66) Hedin, N.; Graf, R.; Christiansen, S. C.; Gervais, C.; Hayward, R. C.; Eckert, J.; Chmelka, B. F. *J. Am. Chem. Soc.* **2004**, *126*, 9425−9432.
(67) Clark, T. M.; Grandinetti, P. J. *J. Phys.: Condes. Matter* **2003**, *15*, S2387−S2395.