Determination of Polypeptide Backbone Dihedral Angles in Solid State NMR by Double Quantum $^{13}$C Chemical Shift Anisotropy Measurements

Francisco J. Blanco and Robert Tycko

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received August 30, 2000; revised November 28, 2000

A solid state NMR technique for the determination of peptide backbone conformations at specific sites in unoriented samples under magic angle spinning (MAS) is described and demonstrated on a doubly labeled polycrystalline sample of the tripeptide AlaGlyGly and a sextuply labeled lyophilized sample of the 17-residue peptide MB($i$ + 4)EK. The technique is applicable to peptides and proteins that are labeled with $^{13}$C at two (or more) consecutive backbone carbonyl sites. Double quantum (DQ) coherences are excited with a radiofrequency-driven recoupling sequence and evolve during a constant-time period at the sum of the two anisotropic chemical shifts. The relative orientation of the two chemical shift anisotropy (CSA) tensors, which depends on the $\phi$ and $\psi$ backbone dihedral angles, determines the $t_1$-dependence of spinning sideband intensities in the DQ-filtered $^{13}$C MAS spectrum. Experiments and simulations show that both dihedral angles can be extracted from a single data set. This technique, called DQCSA spectroscopy, may be especially useful when analyzing the backbone conformation of a polypeptide at a particular doubly labeled site in the presence of additional labeled carbons along the sequence.

Key Words: peptide; backbone structure; solid state NMR; carbon-13 NMR; double quantum filtered spectroscopy; magic angle spinning.

$^{13}$C NMR measurements on carbonyl-labeled polypeptides are particularly attractive for $\phi$ and $\psi$ determinations because the carbonyl chemical shift anisotropy (CSA) is large and its tensor orientation is well characterized (36–42). The relative orientation of the CSA tensors of two carbonyl $^{13}$C nuclei at consecutive residues depends on the dihedral angles of the second residue, as illustrated in Fig. 1. We have previously developed two solid state NMR techniques for the determination of backbone dihedral angles in doubly carbonyl-labeled polypeptides, namely two-dimensional magic angle spinning (2D MAS) exchange spectroscopy (5–7) and constant-time double-quantum-filtered dipolar recoupling (CTDQFD) (8), and have applied these techniques to several structural problems (43–45). Related techniques have been described by other groups (9–26). In this Communication, we describe and demonstrate a new approach, called double quantum CSA (DQCSA) spectroscopy, that provides complementary information and has practical advantages as described below.

DQCSA pulse sequences are shown in Fig. 2. After ramped-amplitude cross-polarization (46) from protons to carbons, a block of radiofrequency-driven recoupling (RFDR) (47–49) generates a nonzero average homonuclear dipole–dipole coupling under MAS. RFDR has the advantageous features of being effective when the CSA is large, insensitive to inhomogeneous broadening of the NMR lines, and extremely sparse (only one $^{13}$C $\pi$ pulse per two MAS rotation periods in our DQCSA experiments), so that effects of pulse imperfections and signal losses due to insufficient proton decoupling are minimized. In a doubly $^{13}$C-labeled sample, $^{13}$C DQ coherences are present after application of a $\pi/2$ pulse at the end of the first RFDR block and are selected by cycling of an overall RF phase shift $\zeta$. DQ coherences are converted to observable magnetization by application of a second $\pi/2$ pulse and a second RFDR block of equal length. Before signal acquisition, a $z$-filter is applied to discard undesired magnetization components (50). DQ coherences evolve under the sum of the two anisotropic, time-dependent $^{13}$C chemical shifts during a constant-time period between the two RFDR blocks. In Fig. 2a, this period is one MAS rotation period $t_R$, and a single $\pi$ pulse of length $t_\pi$ centered at time $t_1 + t_\pi/2$ prevents refocusing of the anisotropic shifts under...
FIG. 1. Segment of a polypeptide chain between α carbons of residues $i - 1$ and $i + 1$. Gray rectangles indicate peptide planes whose relative orientation is defined by the dihedral angles $\phi$ and $\psi$ of residue $i$. The principal axes of the carbonyl $^{13}$C CSA tensor are approximately perpendicular to the peptide plane ($\sigma_{33}$ axis, not shown), in the peptide plane at an angle of approximately $40^\circ$ to the CO–N bond ($\sigma_{11}$ axis), and in the peptide plane at an angle of approximately $10^\circ$ to the C–O double bond ($\sigma_{22}$ axis).

MAS. In Fig. 2b, this period is $2\tau_R$, and three $\pi$ pulses centered at times $t_1 + \pi/2$, $\tau_R$, and $2\tau_R - t_1 - \pi/2$ prevent refocusing of the anisotropic shifts and additionally refocus isotropic shifts ($i.e.$, resonance offsets). In both cases, DQ-filtered MAS spectra are recorded as a function of $t_1$. For doubly carbonyl-labeled polypeptides, the dependence of the MAS sidebands and centerband intensities on $t_1$ can be simulated for any $\phi$ and $\psi$ values, using previously derived expressions for the time-dependent shifts and dipole–dipole couplings (6). Because of the large carbonyl CSA, the signal intensities depend strongly on the $\phi$ and $\psi$ angles, as shown in Fig. 3.

Experimental demonstrations were carried out on a polycrystalline powder sample of the tripeptide L-alanylglycylglycine (AGG), doubly $^{13}$C-labeled at the carbonyls of Ala1 and Gly2, and on a lyophilized powder sample of MB(i + 4)EK, a 17-residue peptide with sequence $N$-acetyl-AEAAAKEAAAKEAAKA-NH$_2$ designed to be highly helical (44, 51), $^{13}$C-labeled at the carbonyls of Ala4, Ala5, Ala8, Ala9, Ala10, and Ala13. Labeled AGG was diluted to 3% in unlabeled AGG by recrystallization (6). Labeled MB(i + 4)EK was not diluted.

The one-dimensional MAS $^{13}$C spectrum of polycrystalline AGG consists of two sets of spinning sidebands with isotropic chemical shifts of 173 (Ala1) and 171 ppm (Gly2) plus natural abundance signals that are efficiently suppressed by double quantum filtering (8). Figures 4a and 4b show experimental and simulated intensities of the centerband and first-order sidebands as functions of $t_1$ over a full rotor period. Experimental intensities are integrated over the two labeled sites and obtained with the DQCSA pulse sequences in Fig. 2a (data in Fig. 4a) and Fig. 2b (data in Fig. 4b). Simulated intensities were obtained by numerical calculations of the quantum dynamics of the two spin system, including finite $\pi$ pulse amplitudes, summed over molecular orientations. For these simulations, the CSA principal values and orientation with respect to the molecular frame were as previously described (6) and the dihedral angles $\phi = -83^\circ$, $\psi = 170^\circ$ for Gly2 were taken from the AGG crystal structure (52, 53). Good agreement between experiments and simulations is obtained, as indicated by $\chi^2$ values of 285 (Fig. 4a) and 416 (Fig. 4b), using the definition...
FIG. 3. Numerical simulations of DQCSA data, based on the pulse sequence in Fig. 2a and assuming CSA parameters and experimental conditions of Fig. 4. Simulated intensities of DQ-filtered MAS centerband (solid lines), upfield first-order sideband (dashed lines), and downfield first-order sideband (dotted lines) are shown for dihedral angles that are characteristic of idealized secondary structures found in proteins.

\[
\chi^2(\phi, \psi) = \frac{1}{\sigma^2} \sum_{i=1}^{N} [E_i - \lambda(\phi, \psi)S_i(\phi, \psi)]^2
\]

where \(\sigma\) is the root-mean-squared noise in the experimental spectra, \(N\) is the number of intensities analyzed (\(N = 192\) in Fig. 4a; \(N = 384\) in Fig. 4b), \(\{E_i\}\) are the experimental intensities, \(\{S_i(\phi, \psi)\}\) are the simulated intensities for the assumed values of \(\phi\) and \(\psi\), and \(\lambda(\phi, \psi)\) is an overall scaling factor adjusted to minimize \(\chi^2\) (6). For a good fit, one expects \(\chi^2\) to be approximately \(N - 1\), with a variance of approximately \(\frac{1}{(N - 1)}\).

Figures 4c and 4d are contour plots of \(\chi^2(\phi, \psi)\) for the data in Figs. 4a and 4b, calculated using simulations over a grid of \(\phi, \psi\) values with \(10^\circ\) increments. In both contour plots, the crystallographic \(\phi, \psi\) values for AGG lie within the global minimum regions of the \(\chi^2(\phi, \psi)\) surfaces. These regions are roughly circular with widths of approximately \(20^\circ\), indicating that DQCSA measurements place strong constraints on both dihedral angles. The contour plots also display “false” minima (i.e., deep local minima in \(\chi^2(\phi, \psi)\) at incorrect \(\phi, \psi\) values) near \(\phi, \psi = -110^\circ, -120^\circ\) and \(\phi, \psi = -170^\circ, 90^\circ\), but the \(\chi^2\) values at these false minima are significantly greater than at the global minima. Only negative values of \(\phi\) are displayed because DQCSA measurements (like all solid state NMR measurements that depend only on carbonyl \(^{13}\)C nuclei at sequential sites, given that the carbonyl CSA tensor has one principal axis perpendicular to the plane of the peptide bond (6)) are invariant to the substitution \(\phi, \psi \rightarrow -\phi, -\psi\). Apart from this unavoidable symmetry, a single DQCSA data set is sufficient to determine both dihedral angles, provided that the experimental signal-to-noise ratio is sufficiently high.

Although the two labeled carbonyl sites in AGG have resolved NMR signals, the sum of the two signals is analyzed in Figs. 4, and 5. This demonstrates the applicability of the DQCSA technique to systems with poorly resolved carbonyl lines, a more common case in experiments on noncrystalline samples.

The DQCSA pulse sequence in Fig. 2a does not refocus isotropic shifts in the \(t_1\) period except at \(t_1 = t_R/2\). The data are then strongly dependent on the RF carrier frequency, which was set to the average of the isotropic shifts of the labeled carbonyls for the experiment in Fig. 4a. Large inhomogeneous broadening of the carbonyl NMR lines also affects the data, reducing the signal intensities by a \(t_1\)-dependent function, which is equal to \(\exp[-W^2\pi^2(t_R - 2t_1)^2/(4 \ln 2)]\) in the particular case of Gaussian lineshapes with full-width-at-half-maximum \(W\) (Hz) and uncorrelated inhomogeneous broadening at the two carbonyl sites. The pulse sequence in Fig. 2b, which refocuses isotropic shifts at all \(t_1\) values, eliminates the dependence on the RF carrier frequency (to the extent that resonance offsets do not affect the RFDR sequences and the \(\pi\) pulses in the evolution period) and obviates the need for characterization of the inhomogeneous broadening. Figure 5 demonstrates the insensitivity of DQCSA data to resonance offsets with the pulse sequence in Fig. 2b.

In both DQCSA sequences in Fig. 2, DQ coherence is created from transverse \(^{13}\)C magnetization by an RFDR sequence
with one $\pi$ pulse per $2\tau_R$, followed by a $\pi/2$ pulse at the end of the RFDR period. We have found this version of RFDR to be an efficient DQ excitation sequence for doubly carbonyl-labeled polypeptides, both experimentally and in simulations. The experimental DQ filtering efficiency in polycrystalline AGG has previously been reported to be 29% with a 16.256-ms DQ excitation period (8). The distance between labeled carbonyl sites in our AGG sample is 3.19 Å (52, 53). Simulations show that an RFDR sequence with one $\pi$ pulse per $\tau_R$ is less efficient for DQ excitation in doubly carbonyl-labeled polypeptides than the sequence used in our experiments. Other recoupling sequences could be employed. The chief virtue of RFDR in this application is simplicity.

DQCSA measurements have distinct advantages over related techniques for determining $\phi$, $\psi$ values in doubly carbonyl-labeled polypeptides. As compared to the CTDQFD technique (8), evolution of the DQ coherences under the CSA makes the DQCSA technique much more sensitive to $\psi$, significantly improving the structural resolution. The CTDQFD technique (8) measures the dependences of DQ-filtered signals on the DQ excitation periods, while the DQCSA technique measures the evolution of DQ-filtered signals under the sum of the two carbonyl CSA tensors during $t_1$. Compared with the 2D MAS exchange technique (5–7), the dependence of DQCSA data on RFDR recoupling leads to fewer false minima in the $\chi^2(\phi, \psi)$ surfaces and different locations for these minima. The additional

FIG. 4. (a, b) Experimental and simulated DQCSA data for the tripeptide AGG doubly $^{13}$C-labeled at the carbonyls of Ala1 and Gly2 (see conditions below). The sample contained 3% doubly labeled molecules diluted in unlabeled AGG (approximately 10 $\mu$mol labeled AGG). Integrated experimental intensities (Ala1 plus Gly2) of the centerband (circles), and first-order spinning sidebands (up and down triangles for the upfield and downfield sidebands, respectively), are in units of the root-mean-squared spectral noise, with upfield and downfield first-order sideband intensities shifted vertically for clarity by $+10$ and $-10$ units, respectively. Solid lines are simulated data for $\phi = -83^\circ$, $\psi = 170^\circ$, the crystallographic dihedral angles of Gly2. Simulated intensities are scaled to minimize their $\chi^2$ deviation from the experimental intensities. (c, d) Contour plots of $\chi^2$, illustrating the strong sensitivity of the quality of agreement between experiments and simulations on the $\phi$, $\psi$ values assumed in the simulations. Contour levels increase in increments of 50, with the lowest level enclosing the black regions set to $\chi^2 = 370$ (c) and 670 (d). MAS centerbands, first-order sidebands, and second-order sidebands were analyzed for these plots. White crosses indicate the crystallographic $\phi$ and $\psi$ values. [Conditions: Experimental data were obtained at a $^{13}$C NMR frequency of 100.8 MHz, with the RF carrier frequency set to the average of the two carbonyl isotropic shifts, using a Varian/Chemagnetics Infinity-400 spectrometer and a Varian/Chemagnetics 6-mm MAS probe. Experiments and simulations were performed with the pulse sequences in Figs. 2a (a,c) and 2b (b,d) with $n = 24$ (12-ms recoupling time), $\tau_R = 250$ $\mu$s, $^{13}$C $\pi$ pulse lengths of 13.7 $\mu$s during RFDR blocks and 10.0 $\mu$s in the $t_1$ period, and $^{13}$C $\pi/2$ pulse lengths of 5.0 $\mu$s. $^{1}$H pulses were actively synchronized with a MAS tachometer signal. $^1$H RF fields were 50 kHz during CP, 85 kHz during RFDR blocks and the $t_1$ period, and 60 kHz during FID acquisition. CP time was 1.2 ms. Recycle delay was 1.8 s. The number of $t_1$ points was 64 (a,c) and 128 (b,d). Total acquisition time for each data set was 4.5 h.]
Fitting parameter required in 2D MAS exchange measurements to account for intrasite crosspeaks due to $^{14}$N relaxation when the two carbonyl lines are not resolved (44) is not required in DQCSA measurements. On the other hand, a combination of DQCSA, CTDQFD, and 2D MAS exchange measurements, all of which can be performed on a single sample, provides multiple independent structural constraints that can eliminate ambiguities in dihedral angles due to false $\chi^2$ minima (43, 45) and may permit the characterization of $\phi, \psi$ distributions in partially disordered polypeptides (44).

Because most amino acids in proteins are present in multiple copies, labeling of a consecutive pair of carbonyl sites by biosynthetic methods usually introduces labels at several additional, nonconsecutive carbonyl sites. Possible contributions of these nonconsecutive labels to the NMR data must then be considered. In such cases, DQCSA measurements have the advantage that the DQ excitation period can be kept short (at the cost of reduced signals), so that the observed DQ-filtered signals must arise predominantly from consecutive carbonyl $^{13}$C pairs. In contrast, CTDQFD measurements require longer excitation periods and 2D MAS exchange measurements require longer exchange periods. When there are many nonconsecutive carbonyl labels, strong intrasite crosspeaks (5, 6) from these labels can also compromise the analysis of 2D MAS exchange data.

The applicability of DQCSA measurements to multiply carbonyl-labeled polypeptides is demonstrated in Fig. 6, which shows experimental data for sextuply labeled MB($i + 4$)EK obtained with $n = 8$ (4-ms RFDR periods). MB($i + 4$)EK has been shown to be highly helical by circular dichroism (44, 51) and solid state NMR measurements (helix content approximately 85% in lyophilized form (44)). The data are analyzed by comparison with simulations for a pair of consecutive carbonyl labels, although the labels are actually at two consecutive residues (Ala4 and Ala5), three consecutive residues (Ala8, Ala9, and Ala10), and one isolated residue (Ala13). A single set of CSA principal values ($\delta = \delta_{11} - \delta_{33} = 155$ ppm; $\eta = (\delta_{11} - \delta_{22})/(\delta_{11} - \delta_{33}) = 0.584$) determined from spinning sideband intensities was assumed for all labeled carbonyl sites. The global minimum in $\chi^2(\phi, \psi)$ for MB($i + 4$)EK occurs in the helical region of the $\phi, \psi$ plane, as expected. Although the value of $\chi^2(\phi, \psi)$ at the minimum ($\chi^2 = 1210$ at $\phi, \psi = -60^\circ, -50^\circ$ in Fig. 6b) is significantly greater than in the analyses of AGG data (see above), indicating a poorer fit relative to the experimental noise level, this observation may be explained by conformational disorder in the noncrystalline, lyophilized MB($i + 4$)EK sample and by the comparatively high signal-to-noise ratio of the MB($i + 4$)EK data, which increases the impact of systematic errors on the $\chi^2(\phi, \psi)$ values.

A technique similar to DQCSA, called DQDRAWS, has been developed and applied to dihedral angle determinations in doubly carbonyl-labeled peptides by Drobny and co-workers (25, 26). The DQCSA technique differs from DQDRAWS in...
signals on propagators (rotation of the relative phases of DQ preparation and DQ mixing rather than from a combination of CSA evolution and modulation). Finally, we analyze and simulate the unnecessary (particularly for the DQCSA sequence in Fig. 2b).

characterization or estimation of DQ linewidths and lineshapes weak, is that it allows us to optimize the efficiency of data collection by restricting the experimental $t_1$ values to ranges where the DQCSA signals are most sensitive to peptide conformation. The DQCSA technique described above may also be viewed as a MAS version of the static DOQSY technique of Schmidt-Rohr (22–24, 56). When sample quantities are limited, as is usually the case in biological systems, the higher signal-to-noise ratio of MAS measurements is advantageous. The higher resolution of MAS measurements may also be important, even with double quantum filtering, in systems with multiple labeled sites or in high-molecular-weight systems where natural-abundance $^{13}$C nuclei outnumber the $^{13}$C labels.

several respects. First, we use the RFDR, rather than the DRAWS (54), recoupling sequence to excite DQ coherences. Demands on spectrometer and probe performance, in particular proton decoupling field strengths and RF inhomogeneity, are thereby reduced. Second, we use a constant-time DQ evolution period, so that both RFDR blocks maintain the same relative rotation of the MAS rotor phase at all $t_1$ values. The dependence of the DQ-filtered signals on $t_1$ then arises solely from evolution under the CSA, rather than from a combination of CSA evolution and modulation of the relative phases of DQ preparation and DQ mixing propagators (55). The constant-time evolution period also makes characterization or estimation of DQ linewidths and lineshapes unnecessary (particularly for the DQCSA sequence in Fig. 2b).

Finally, we analyze and simulate the $t_1$ dependence of DQCSA data in the time domain, rather than the frequency domain. One advantage of time-domain analysis, especially when signals are weak, is that it allows us to optimize the efficiency of data collection by restricting the experimental $t_1$ values to ranges where the DQCSA signals are most sensitive to peptide conformation. The DQCSA technique described above may also be viewed as a MAS version of the static DOQSY technique of Schmidt-Rohr (22–24, 56). When sample quantities are limited, as is usually the case in biological systems, the higher signal-to-noise ratio of MAS measurements is advantageous. The higher resolution of MAS measurements may also be important, even with double quantum filtering, in systems with multiple labeled sites or in high-molecular-weight systems where natural-abundance $^{13}$C nuclei outnumber the $^{13}$C labels.

REFERENCES


