J Biomol NMR (2007) 39:107–131
DOI 10.1007/s10858-007-9179-1

ARTICLE

Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins

W. Trent Franks · Kathryn D. Kloepper · Benjamin J. Wylie · Chad M. Rienstra

Received: 26 March 2007 / Accepted: 10 July 2007 / Published online: 9 August 2007

Abstract Chemical shift assignment is the first step in all established protocols for structure determination of uniformly labeled proteins by NMR. The explosive growth in recent years of magic-angle spinning (MAS) solid-state NMR (SSNMR) applications is largely attributable to improved methods for backbone and side-chain chemical shift correlation spectroscopy. However, the techniques developed so far have been applied primarily to proteins in the size range of 5–10 kDa, despite the fact that SSNMR has no inherent molecular weight limits. Rather, the degeneracy inherent to many 2D and 3D SSNMR spectra of larger proteins has prevented complete unambiguous chemical shift assignment. Here we demonstrate the implementation of 4D backbone chemical shift correlation experiments for assignment of solid proteins. The experiments greatly reduce spectral degeneracy at a modest cost in sensitivity, which is accurately described by theory. We consider several possible implementations and investigate the CANCOCX pulse sequence in detail. This experiment involves three cross polarization steps, from H to CA[i], CA[i] to N[i], and N[i] to C′[i−1], followed by a final homonuclear mixing period. With short homonuclear mixing times (<20 ms), backbone correlations are observed with high sensitivity; with longer mixing times (>200 ms), long-range correlations are revealed. For example, a single 4D experiment with 225 ms homonuclear mixing time reveals ∼200 uniquely resolved medium and long-range correlations in the 56-residue protein GB1. In addition to experimental demonstrations in the 56-residue protein GB1, we present a theoretical analysis of anticipated improvements in resolution for much larger proteins and compare these results in detail with the experiments, finding good agreement between experiment and theory under conditions of stable instrumental performance.

Keywords Solid-state NMR · Multidimensional NMR · Protein structure · Magic-angle spinning · Chemical shift assignments

W. T. Franks · K. D. Kloepper · B. J. Wylie · C. M. Rienstra
Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA

C. M. Rienstra
Department of Biochemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA

C. M. Rienstra (✉)
Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA
e-mail: rienstra@scs.uiuc.edu

Introduction

Structural studies of proteins using magic-angle spinning (MAS) solid-state NMR (SSNMR) have rapidly progressed in recent years, based largely on multidimensional chemical shift correlation methods applied to uniformly-^{13}C, ^{15}N-labeled proteins. Early demonstrations of chemical shift assignment methods (McDermott et al. 2000; Pauli et al. 2001) have now been applied in several other contexts to microcrystalline or polyethylene-glycol precipitated samples (Bockmann et al. 2003; Igumenova et al. 2004a, b; Marulanda et al. 2004; Franks et al. 2005; Marulanda et al. 2005; Pintacuda et al. 2006). The most-often used pulse sequences utilize combinations of heteronuclear ^{15}N–^{13}C and ^{13}C–^{13}C polarization transfer

steps. The ^{15}N–^{13}C transfers are typically achieved with transferred echo double resonance (TEDOR) (Hing et al. 1992), double cross polarization (DCP) (Schaefer and Stejskal 1979), adiabatic (Baldus et al. 1996) and/or band-selective SPECIFIC DCP (Baldus et al. 1998) pulse sequence elements. The homonuclear ^{13}C–^{13}C correlations are most often established by radio-frequency driven recoupling (RFDR) (Bennett et al. 1992, 1998), supercycled permutationally offset compensated C5 (SPC-5) (Hohwy et al. 1999), dipolar-assisted rotational resonance (DARR) (Takegoshi et al. 2003) or RF-assisted diffusion (RAD) (Morcombe et al. 2004). Together these heteronuclear and homonuclear mixing schemes are now used routinely to generate 2D and 3D correlation spectra of solid proteins.

These protocols have enabled resolution and assignment of backbone and side-chain signals for several proteins with molecular weights in the range of 5–10 kDa. Structural folds at adequate resolution can then be solved by the addition of empirical dihedral restraints (from TALOS (Cornilescu et al. 1999)) and semi-quantitative distance restraints (from ^{13}C–^{13}C (Castellani et al. 2002) or ^{1}H–^{1}H (Lange et al. 2005), using solid-state experiments closely analogous to NOESY (Kumar et al. 1980)). Examples of complete structures of microcrystalline proteins have now been reported (Castellani et al. 2002; Lange et al. 2005; Zech et al. 2005), and in some membrane protein systems chemical shifts have been assigned (Hiller et al. 2005) and the perturbations in the context of protein interactions observed (Lange et al. 2006). Similarly, the study of fibril structure has been accelerated by application of chemical shift assignment methods to uniformly labeled proteins (Heise et al. 2005; Ritter et al. 2005). These successes bode well for application of multidimensional MAS methods to a large range of important biological problems including atomic-resolution structure determination in many contexts where single crystal and solution NMR methods are not applicable.

To extend these successes to proteins of greater size and complexity with uniform isotopic labeling requires further improvements in assignment methodology. Studies of large membrane proteins (>15 kDa) show promise for success in this context. For example, McDermott and co-workers have recently observed and analyzed site-specific chemical shift perturbations attributed to ligand binding in a 119 kDa P450 enzyme (Jovanovic et al. 2005; Ravindranathan et al. 2007), and we have demonstrated that 2D and 3D spectra of a 144 kDa membrane protein yield linewidths comparable to those observed in small microcrystalline proteins, as well as strong signals for a large number of spin systems in 3D spectra (Frericks et al. 2006). These studies follow earlier work by a number of laboratories in which site-specific measurements have been made in very large solid proteins, such as bacteriorhodopsin (Hu et al. 1995), rhodopsin (Carravetta et al. 2004), EPSP synthase (McDowell et al. 1996a, 2004), and tryptophan synthase (McDowell et al. 1996b). Thus, no fundamental limits inherently prohibit examination of such large systems by NMR. Rather, the primary impediments are the practical and technical details of acquiring high-resolution spectra of proteins in the size range of 20–200 kDa, which contain several hundreds or thousands of residues. Such proteins pose a major challenge for site-specific 2D and even 3D methods, yet are not beyond the capabilities of SSNMR in principle.

In this work, we demonstrate the development and application of 4D chemical shift correlation methods for backbone and side-chain assignment of solid proteins under MAS. We organize the work as follows. First, we illustrate the benefits of 4D spectroscopy by theoretical calculations of resolution in standard chemical shift spectra as a function of linewidth and molecular weight; this is complemented by analysis of unique connectivity information. Second, we discuss the theoretical and experimental requirements of establishing a suitable set of correlations with high sensitivity and reproducibility, including the issue of instrumental (particularly amplifier) stability. Third, we demonstrate the 2D sub-spectra from the optimal polarization transfer pathway; the CA(NCO)CX 2D experiment provides improved resolution and sensitivity (per site) for identifying unique interresidue CA[i]–CA[i−1], CA[i]–C′[i−1] and CA[i]–CB[i−1] correlations. Fourth, we discuss details of acquiring the 4D spectra in an experimentally practical timeframe and processing the data to high digital resolution. Finally, we apply the method to resolve all the backbone CA-N-C′ correlations in the 56-residue protein GB1 in the indirect dimensions of a 4D experiment intended to reveal a large number of long-range correlations in the direct (CX) dimension. Although the 4D technique is not essential for performing the backbone or side-chain assignments of GB1, as previously published (Franks et al. 2005), we show that the uniquely resolved indirect dimension frequencies permit a long homonuclear mixing time to be applied prior to the direct dimension acquisition period. This results in a spectrum with minimal overlap, enabling identification of several hundred unique correlations, including multiple medium- and long-range correlations per residue.

To complement the experimental study in a small protein, we also perform simulations to estimate the improvement in resolution anticipated for larger proteins. In separate studies, we will demonstrate application of the optimized pulse sequences to proteins in the size range of 15–144 kDa being studied in our laboratory, including fibrils of the 15 kDa α-synuclein (Heise et al. 2005; Kloepper et al. 2006), microcrystals of the periplasmic disulfide-bond forming enzyme DsbA (20 kDa), and ultracentrifuged preparations of the membrane proteins

DsbB (20 kDa) (Li et al. 2007) and cytochrome bo_3 oxidase (144 kDa) (Frericks et al. 2006).

Materials and methods

In this work, we use the term CO to refer to the carbonyl region of the ^{13}C spectrum ($\sim$168 to $\sim$186 ppm) and C′ to refer to the backbone carbonyl ^{13}C site of a peptide, as consistent with IUPAC nomenclature (Markley et al. 1998).

Simulations of spectral resolution anticipated from various types of 2D, 3D and 4D spectra were performed using logic similar to that described previously by Tycko (1996). Briefly, in a given *n*-dimensional spectrum, the peak *i* in dimension *m* is considered *resolved* from peak *j* ($j \neq i$) if $|\Omega_{i,m}-\Omega_{j,m}| > lw_m$, where lw_m is the (presumed constant) linewidth of all peaks in dimension *m* and $\Omega_{i,m}$ and $\Omega_{j,m}$ are the frequencies of peaks *i* and *j* in dimension *m*. Further, if the peak *i* is resolved in any dimension from all peaks $j \neq i$, then the peak *i* is considered to be *uniquely resolved*. A spin system is likewise presumed to be resolved if any peak arising from that spin system is resolved in a given experiment. For example, in a CANCOCX experiment ($n = 4$), each spin system *i* gives rise to peaks CA[i]–N[i]–C′[i−1]–C′[i−1], CA[i]–N[i]–C′[i−1]–CA[i−1], CA[i]–N[i]–C′[i−1]–CB[i−1]; if any one of these three peaks is resolved, this particular spin system is counted as resolved. Clearly this is a less stringent requirement than to have all three peaks resolved.

In the same 4D experiment, a resonance from the spin system *i* is defined to provide *unique connectivity* information if at least two inequalities $|\Omega_{i,m} - \Omega_{j,m}| > lw_m$ are satisfied, for $m = 1$ or $m = 2$ and $m = 3$ or $m = 4$, for all possible values of *j* not equal to *i*. That is, a unique connectivity is observed in theory if the source residue correlation (CA[i]–N[i] pair, in F1–F2) and the destination residue correlation (C′[i−1]–CA[i−1], in F3–F4) are *both* resolved. In this case, the pairs of amino acids would be unambiguously linked, although not necessarily assigned to a particular residue type or mapped to a unique position in the protein sequence.

GB1 was prepared in microcrystalline form (Franks et al. 2005) and 18 mg (2.7 μmol) was packed in a 3.2 mm limited speed rotor (Varian, Inc., Fort Collins, Colorado and Palo Alto, California). Spectra were collected at 500 MHz ^1H frequency on a Varian InfinityPlus four-channel spectrometer, with a Balun™ ^1H–^{13}C–^{15}N 3.2 mm MAS probe. Unless otherwise noted, the MAS rate was 11.111 kHz and the temperature of variable temperature gas set to 273 K with 100 scfh flow, resulting in an effective mean sample temperature (due to frictional heating and geometrical offset factors) of 280 ± 4 K, as determined by methanol calibration (Van Geet 1968).

Tangent shaped ramps were applied for all CP periods. In the following paragraphs, for clarity the variation of nutation frequency is indicated following the mean value, and the parameters describing the shape of the ramp, following Eqs. 3 and 4 of Detken et al. (2001), are Δ and β, where we slightly modified the expressions to allow greater flexibility within the InfinityPlus pulse programmer logic, as follows:

$$\omega_1^X(t) = \omega_1^{HH} + |\beta| \cdot \tan\left(\alpha \cdot \left[t - \frac{\tau}{2}\right]\right) \quad (1)$$

defined over the limit $0 \leq t \leq \tau$, where the adiabaticity parameter is consistent with Detken:

$$\alpha = \frac{2}{\tau}\arctan\left(\frac{\Delta}{\beta}\right) \quad (2)$$

We note that the value of Δ used in our experiments is always positive, and the value of β can be positive or negative, corresponding to "up" or "down" ramps as described by Detken. Thus the shape and the direction of the amplitude sweep together are defined by the magnitude and the sign of the parameter β, which we typically optimized after the Hartmann-Hahn condition (ω_1^{HH}) was found and the Δ parameter set. Typical values for $\Delta/2\pi$ were $\sim$10 kHz for ^1H–^{15}N or ^1H–^{13}C transfers and $\sim$1 to 2 kHz for ^{15}N–^{13}C transfers. The values for $\beta/2\pi$ were approximately five times smaller with a sign depending on the direction of polarization transfer as well as the chemical shift of ^{13}C spin involved in the transfer.

Two pulse phase modulation (TPPM) decoupling (Bennett et al. 1995) was used during evolution periods. The total phase difference and pulse width are quoted following the decoupling nutation frequency, with appropriate units. No further improvement from these optimized conditions were observed upon the use of SPINAL-64 decoupling (Fung et al. 2000).

Two-dimensional NCO and NCA correlation spectra were acquired with initial H-N CP periods of 1.5 ms with a constant amplitude $\omega_1^H/2\pi$ of $\sim$61 kHz, and an average $\omega_1^N/2\pi$ of $\sim$50 kHz with a tangent ramp ($\Delta/2\pi$ = 8.3 kHz, $\beta/2\pi$ = 3.3 kHz). CP nutation frequencies quoted here are approximate, based on two-dimensional optimizations of the parameters, typically resulting in conditions matching odd half-integer multiples of the spinning rate. The ^1H decoupling during the N–C CP period was $\sim$102 kHz for both experiments; the ^1H nutation frequency has an uncertainty or range over the sample of $\sim$3 to 5% in this probe. For the NCA experiment, the ^{15}N carrier frequency was $\sim$80 ppm, the ^{13}C carrier was placed at $\sim$55 ppm,

with $\omega_1^N/2\pi = \sim 26$ kHz (or $\sim 5/2\ \omega_R/2\pi$) and $\omega_1^C/2\pi = \sim 15$ kHz ($\sim 3/2\ \omega_R/2\pi$), $\Delta/2\pi = 2.3$ kHz and $\beta/2\pi = -500$ Hz. For the NCO experiment, the ^{13}C carrier was placed at ~ 170 ppm, with $\omega_1^N/2\pi = \sim 16$ kHz ($\sim 3/2\ \omega_R/2\pi$), and $\omega_1^C/2\pi = \sim 27$ kHz, $\Delta/2\pi = 2.3$ kHz and $\beta/2\pi = +500$ Hz. TPPM decoupling (~ 76 kHz, 16°, 6.2 μs) was used during free evolution periods. The indirect dimension was digitized with 512 rows (TPPI) and a dwell time of 60 μs (t_{1max} = 30.72 ms); the direct dimension maximum evolution time was 38.4 ms, with 4 scans per row, 1.5 s pulse delay, 32 steady state scans, yielding a total measurement time of 52 min. per spectrum.

The 2D CC spectra near the weak coupling limit were acquired at an MAS rate of 7.814 kHz (Seidel et al. 2004). The initial H-C CP was 0.75 ms with $\omega_1^H/2\pi = \sim 58$ kHz and $\omega_1^C/2\pi = 66$ kHz, $\Delta/2\pi = 10$ kHz and $\beta/2\pi = +1.5$ kHz. The indirect dimension was digitized to 1280 rows with a dwell time of 12.8 μs (t_{1max} = 16.4 ms), for a total measurement time of 3.3 h. The DARR 2D spectra were acquired at the MAS rate of 11.111 kHz with an initial CP (H–C) of duration 0.75 ms with $\omega_1^H/2\pi = \sim 61$ kHz, and $\omega_1^C/2\pi = \sim 72$ kHz, $\Delta/2\pi = 10$ kHz and $\beta/2\pi = +1.5$ kHz. The indirect dimension was digitized to 1024 rows (TPPI) with a dwell time of 15 μs (t_{1max} = 15.4 ms), with 4 scans per row, for a total measurement time of 2.4 h. TPPM decoupling for both spectra was applied during free evolution and acquisition periods with a $\omega_1^H/2\pi = 74$ kHz (18°, 6.9 μs). The 2D CA(NCO)CX experiment CP conditions are described below in the context of the 3D and 4D experiments. The spectrum was acquired with TPPM decoupling (68 kHz, 15°, 7 μs), and the indirect dimension was digitized to 1024 rows with a dwell of 15 μs (t_{1max} = 15.4 ms), with 4 scans per row, for a total measurement time of 2.4 h.

For the CANCOCX experiments throughout this study, the initial H–C CP period was 450 μs with $\omega_1^H/2\pi = 80$ kHz and $\omega_1^C/2\pi \sim 70$ kHz, $\Delta/2\pi = 8$ kHz and $\beta/2\pi = -3.5$ kHz. The CA–N CP period was 6 ms, $\omega_1^C/2\pi$ was ~ 17 kHz, $\Delta/2\pi = 2$ kHz and $\beta/2\pi = 400$ Hz, with constant amplitude on ^{15}N, $\omega_1^N/2\pi$ (~ 28 kHz). The off-resonance N–CO transfer step had a duration of 8 ms with $\omega_1^C/2\pi = 36.5$ kHz, $\Delta/2\pi = 2$ kHz and $\beta/2\pi = -400$ Hz, and $\omega_1^N/2\pi = \sim 28$ kHz. CW decoupling ($\omega_1^H/2\pi = 95$ kHz) was applied during both C–N and N–C CP periods (in the range from ~ 95 to ~ 105 kHz, the polarization transfer efficiency did not change significantly, so the lower end of this range was used in order to minimize risk of probe damage). Homonuclear DARR (Takegoshi et al. 2003; Morcombe et al. 2004) recoupling (18 ms) was applied with the $\omega_1^H/2\pi = 11$ kHz. TPPM heteronuclear decoupling was applied during free evolution periods (72 kHz, 16°, 7.6 μs). The 4D experiments were digitized with maximum evolution times of 4.32 ms in t_1 (CA, 32 × 135 μs), 8.64 ms in t_2 (N, 32 × 270 μs), 5.4 ms in t_3 (CO, 20 × 270 μs) and acquired to 1536 × 15 μs in t_4 with a 2 s recovery delay and 4 scans per row (the minimal phase cycle), for a total of 46.6 h. The experiment with 225 ms DARR mixing was repeated in two identical 51.7 h blocks, for a total measurement time of 103.4 h.

Data were processed with nmrPipe (Delaglio et al. 1995), employing Lorentzian-to-Gaussian apodization, zero filling to all dimensions prior to Fourier transformation, and polynominal baseline (frequency-domain) correction to the direct dimension. Additional processing parameters for each spectrum, including elaboration of the linear prediction procedures, are provided in the respective figure captions. Chemical shifts were referenced according to the DSS scale with adamantane as a secondary standard (Morcombe and Zilm 2003), assuming its downfield signal to resonate at 40.48 ppm.

Spectral peak picking and assignments were performed with Sparky version 3 (T. D. Goddard and D. G. Kneller, University of California, San Francisco), based on our previous assignments (Franks et al. 2005).

Results and discussion

Resolution benefits of 4D MAS NMR

The principal limitations in spectral resolution by MAS NMR have been substantially alleviated in recent years, with the development of improved instrumentation, sample preparation protocols, and development of sequences for homonuclear and heteronuclear correlation and decoupling. Key instrumental developments include probes with balanced circuits (Martin et al. 2003) to provide improved B_1 homogeneity and sensitivity at high field (Paulson et al. 2004) and applications of new resonator designs to enable high-power proton decoupling while minimizing thermal damage (Stringer et al. 2005). Sample preparation protocols include precipitations (Pauli et al. 2000) and microcrystal preparations (Cole and Torchia 1991), which have been systematically analyzed and demonstrated for several proteins (Martin and Zilm 2003). These protocols have now been applied to several soluble, uniformly labeled proteins, in all cases yielding excellent resolution (McDermott et al. 2000; Pauli et al. 2001; Bockmann et al. 2003; Igumenova et al. 2004a, b; Marulanda et al. 2004, 2005; Franks et al. 2005; Pintacuda et al. 2006). The sample preparation methods therefore have general utility to soluble proteins, even in cases where single crystals cannot readily be prepared. Moreover, several types of membrane protein formulations give similar quality spectra (Hiller et al. 2005; Frericks et al. 2006; Etzkorn et al. 2007; Li et al. 2007). Indeed, several systems have been

demonstrated to yield excellent linewidths, on the order of 0.5 ppm or less; often the remaining linewidth is due to choice of digitization conditions, which are indirectly dictated by sensitivity concerns. Clearly if such high-resolution spectra can be acquired, an optimal approach to assignment is to acquire 2D and 3D spectra at the highest possible digital resolution.

Nevertheless, the linewidths depend on the degree of local structural order as well as dynamics, which may vary somewhat among samples or from site to site within a given sample. In cases where inhomogeneous broadening is more significant, line narrowing approaches such as J-decoupling (Igumenova and McDermott 2005; Li et al. 2006), dilute isotopic labeling (Castellani et al. 2002) or higher magnetic fields will not always benefit resolution to the extent desired, since some issues may be beyond the control of the experimentalist. For example, fundamental issues such as the timescale of local molecular dynamics and/or static heterogeneities may be unavoidable; these not only must be tolerated in some instances but in addition are of interest to understand systematically, for example in amyloid systems (Petkova et al. 2004, 2005; Tycko 2004; Havlin and Tycko 2005) where the degree of site-specific order may give insights into the structure and dynamics of fibrils. Thus the ability to resolve and assign signals from proteins in which the linewidths are significantly greater than 0.5 ppm remains a high priority for the SSNMR community, to complement the goal of assigning larger proteins with narrower linewidths.

In order to establish correlations in the multidimensional spectra, the homonuclear (usually ^{13}C–^{13}C) polarization transfer schemes in most common usage are proton-driven spin diffusion (PDSD) (Bloembergen 1949; Suter and Ernst 1985) and its more robust analogs for higher spinning rates, dipolar-assisted rotary resonance (DARR) (Takegoshi et al. 2003) and RF-assisted diffusion (RAD) (Morcombe et al. 2004). These methods offer minimal overall signal loss, with the transfer efficiency of $\sim$25% from the C′ to CA site in a peptide in practical applications. Homonuclear longitudinal adiabatic spin lock transfer schemes such as rotational resonance tickling (R2T) (Costa et al. 1997) and rotary resonance in the tilted rotating frame (R2TR) (Takegoshi et al. 1995) offer comparable signal loss and an improved transfer efficiency of $\sim$50% but are not broadband mixing schemes and are therefore optimal for CO–CA pairs near rotational resonance conditions; this may in some cases not permit coverage of the entire CO–CA region. RFDR (Bennett et al. 1992, 1998) and its finite-pulse variant fpRFDR (Ishii 2001) are preferred in some contexts, such as with fast MAS rates. Double-quantum polarization transfer schemes based on symmetry principles such as C7 (Lee et al. 1995), POST-C7 (Hohwy et al. 1998), SPC-5 (Hohwy et al. 1999) and the R-family of experiments (Carravetta et al. 2000) offer even greater performance in theory, yet present a variety of experimental challenges such as phase transients and proton decoupling that may impair performance at high B_0 field. Most solid protein assignments have been achieved with 2D ^{13}C–^{13}C and ^{15}N–^{13}C experiments involving various combinations of the pulse sequences mentioned above to achieve backbone correlations as depicted in Fig. 1. Ultimately the size range and variety of proteins that can be assigned is a function of both sensitivity and resolution, and tradeoffs are implicitly required when performing multidimensional experiments.

For the correlation of four consecutive ^{13}C and/or ^{15}N nuclei along the protein backbone with ^{13}C detection, there are two possible 4D polarization transfer pathways that extend from the standard NCACX (Fig. 1a) and NCOCX (Fig. 1b) schemes. In the first, CONCACX (Fig. 1c), polarization initiates on the C′[i−1] nucleus (F1), is transferred to the N[i] in the first heteronuclear mixing period, frequency labeled (F2), then transferred to the CA[i] (F3) and to the C′[i] and sidechain carbons of the *i* residue (F4) by the final homonuclear mixing period. This

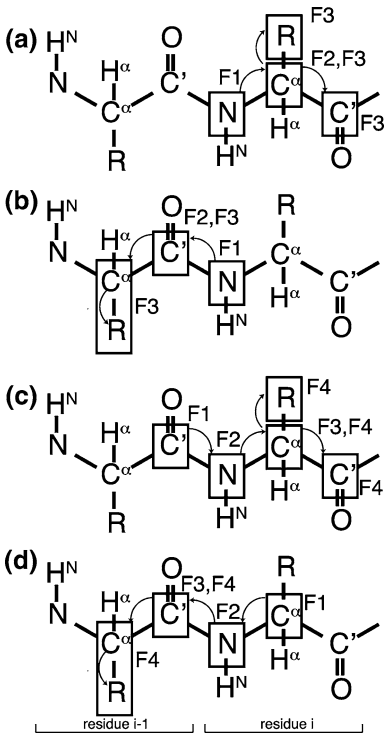

Fig. 1 Polarization transfer pathways utilized for solid-state backbone chemical shift correlation spectroscopy experiments. (**a**) NCACX and (**b**) NCOCX schemes have been utilized for the majority of solid-state NMR assignments performed to date (see references in text) and only involve one polarization transfer between ^{13}C and ^{15}N nuclei. (**c**) CONCACX and (**d**) CANCOCX polarization transfer schemes utilize two selective heteronuclear polarization transfers to correlate the backbone chemical shifts of residues *i* and *i*−1

scheme can be viewed as an NCACX with additional resolution provided by the preceding C′ resonance. The second scheme, CANCOCX (Fig. 1d), is based upon the opposite direction of heteronuclear polarization transfer, starting on the CA[i] site (F1), proceeding through the N[i] (F2) to the C′[i−1] (F3), and then to the CA[i−1] and sidechain carbons (F4); likewise this experiment can be viewed as an NCOCX with an additional indirect (CA) dimension to enhance resolution. The details of the pulse sequence elements will be discussed further below.

First, in order to assess the relative resolution expected from these 3D and 4D schemes, we simulated several types of backbone correlation spectra based on the chemical shifts (N, CA, CB and C′) of proteins previously assigned and deposited in the Biological Magnetic Resonance Data Bank (BMRB). Although this approach does not rigorously disprove the ability to assign proteins in which some peaks are overlapped, nor does it prove that peaks can be assigned when resolved, the approach does provide a relative measure of the assignment difficulty within a given data set or collection of data sets under assumed resolution. Even with badly overlapped spectra, it may sometimes be possible to assign signals based on multiple data sets and thorough analysis. Nevertheless, here we use a simple level of theory to describe the relative percentage of resolved spin systems and assume this value to be correlated in an approximately linear manner to the ease of assignment.

Starting with proteins of rather modest size that have already been studied by MAS NMR (Table 1), we find that even GB1 (56 residues), BPTI (58 residues) and ubiquitin (76 residues) would be only ∼50% resolved in standard types of 3D experiments under the assumption of 2 ppm linewidths. For example, the total number of spin systems resolved in BPTI (based on the solution NMR ^{15}N chemical shifts, which have not yet been fully assigned in the solid state) is only 25 out of 58 for intraresidue NCACX schemes. (Simulations were performed for the NCACO and NCACB experiments and the union of the list of resolved residues in each experiment reported.) The interresidue NCOCX experiment yields slightly better resolution, with 28 unique spin systems identified (using NCOCA and NCOCB simulations). In each case, the 4th dimension improves resolution relative to the 3D at the same digital resolution. Similar results are observed for GB1 and ubiquitin. For these small proteins, under the assumption of 1 ppm linewidths, the 3D spectra are well resolved, and the 4D experiments are not necessary to interpret spectra with short mixing times, where only the CO, CA and CB are observed in the direct dimension. However, one can readily anticipate that with larger proteins and/or those with greater repetition in sequence or secondary structure that the 4D methods will have a substantial benefit. Also the additional resolution will aid in analyzing experiments with longer homonuclear mixing times, which have a larger number of peaks, as discussed further below.

With much larger proteins, the benefits of the 4D experiments are amplified considerably. The assignment of maltose binding protein, at 376 residues, was a major achievement for solution NMR (Gardner et al. 1998). Its simulated 3D SSNMR spectra with 2 ppm linewidths would yield only 57 and 40 resolved spin systems in the NCACX and NCOCX experiments respectively, which increases to slightly more than half the protein assuming 1 ppm linewidths; extension to a fourth dimension would yield 285 (CONCACX) or 332 (CANCOCX) resolved spin systems. Likewise, malate synthase G, the largest monomeric unit so far fully assigned by solution NMR (Tugarinov et al. 2002) would yield less than 10% of the spin systems resolved (59 or 61 out of 723) in the NCACX and NCOCX experiments with 2 ppm resolution, increasing significantly to ∼30% (270 or 291 out of 723) with 1 ppm resolution and approximately doubling again upon extension to 4D with 1 ppm linewidths (505 and 591 spin systems would be resolved respectively in the CONCACX and CANCOCX spectra). In each of these very large (>30 kDa) proteins, the 4D experiments with 2 ppm resolution would give major (79–229%) increases in the numbers of resolved spin systems, as a result of the fact that different amino acid types give significantly different CA–N and/or CA–C′ signatures, and the CANCOCX experiment identifies both of these features. Although these results are not qualitatively surprising, because it is well known that higher dimensionality experiments yield theoretical improvements in resolution, the quantitative analysis of the extent of improvement demonstrates that the benefits are especially valuable for larger proteins. In particular, the benefits with relatively large linewidths are derived from particular types of residue pairs, such as Gly, Ala, Thr, Ser and Pro, where one or more distinct chemical shift value (such as the Gly CA, Ala CB, Thr CB, Ser CB, Pro N) tends to be resolved from all other amino acid types even with linewidths of multiple ppm.

In the limit of greater (∼2 ppm or more) linewidths, relatively small peptides with high sequence repetition, such as amyloid fibrils (Petkova et al. 2004, 2005; Tycko 2004; Havlin and Tycko 2005), might also benefit significantly by extension to 4D. One example of a fibril of significant length that has been mostly assigned is the HET-s prion protein; although the formulation presented in studies by Meier and co-workers demonstrated very high resolution (Siemar et al. 2005, 2006), we consider the limits of greater linewidths, as typically encountered in peptides such as β-amyloid. Although in general overlap increases with a larger linewidth, increasing the dimensionality with the correlation scheme presented here still

Table 1 Theoretical resolution of 3D and 4D backbone correlation spectra

Protein[a]	Size [b]	Simulated LW (ppm)[c]	3D NCACX[d]	4D CONCACX[e]	3D NCOCX[f]	4D CANCOCX[g]
Het-S[h]	45 (43)	3.0	10	17	14	24
		2.0	27	31	34	37
		1.0	39	40	42	42
GB1	56 (56)	3.0	12	16	11	20
		2.0	32	39	31	46
		1.0	54	55	54	55
BPTI	58 (58)	3.0	9	19	18	26
		2.0	25	34	28	40
		1.0	47	55	53	57
Ubiquitin	76 (76)	3.0	22	26	20	40
		2.0	37	46	26	49
		1.0	67	75	65	73
Ubiquitin (SSNMR)	76 (66)	3.0	18	25	16	33
		2.0	39	49	35	50
		1.0	66	65	58	62
Maltose binding protein	376 (355)	3.0	25	32	15	54
		2.0	57	103	40	138
		1.0	215	285	217	332
Malate synthase G	723 (711)	3.0	29	38	22	56
		2.0	59	116	61	189
		1.0	270	505	291	591

[a] The chemical shift values for each protein were derived from the following BMRB entries and references: Het-S fibrils (Siemar et al. 2005, 2006) GB1 in the solid state (BMRB 15156) (Franks et al. 2005); BPTI in solution (BMRB 5358) (McDermott et al. 2000); ubiquitin in solution (BMRB 5387) (Flynn et al. 2002); ubiquitin in the solid state (BMRB 7111) (Igumenova et al. 2004a, b); maltose binding protein in solution (BMRB 4354) (Gardner et al. 1998) malate synthase G in solution (BMRB 5471) (Tugarinov et al. 2002)

[b] The value in parentheses represents the limit of ideal resolution, equivalent to the number of resolved spin systems calculated with the assumption of 0.001 ppm linewidths in all dimensions. Even in this limit, not all residues are resolved in the simulated spectra because the amide ^{1}H sites, as used for the solution NMR assignments, are not used in our calculations

[c] Common linewidths are assumed in each dimension

[d] This column reports the union of the list of resolved residues from the NCACB and NCAC′ experiments. Side-chain assignments beyond the CB are not available for all proteins

[e] This column reports the union of the list of resolved residue pairs from the CONCACB and CONCAC′ experiments. This experiment emphasizes intraresidue correlations within residue *i*, utilizing the C′[i−1] correlation to enhance resolution

[f] This column reports the union of the list of resolved residue pairs from the NCOCA and NCOCB experiments

[g] This column reports the union of the list of resolved residue pairs from the CANCOCA and CANCOCB experiments

[h] The Het-s assignments do not include 10 C′ carbons. For this calculation, the value for these missing chemical shifts was set to the BMRB average shift for the residue

partially alleviates overlap even in cases where the linewidth is large. For example, in Het-S fibrils with 3 ppm linewidths, nearly 55% of residues are resolved in the 4D CANCOCX, while only ~30% are resolved in the 3D NCOCX. The 3D NCACX spectrum is improved from ~25% of all residues resolved, to ~38% resolved in the 4D CONCACX. Here the true benefit of the 4D in fact may be underestimated due to the fact that several C′ resonances were not reported for HET-s.

A more stringent requirement for assessing the ability to assign spectra of a protein is the observation of unique connectivities in the spectra. We define the condition of unique connectivity (see Materials and methods) as observing uniquely resolved peaks both for the source and the destination of polarization in a given experiment. For example, a unique connectivity could be identified between two residues in a CANCOCX experiment if a particular CA-N peak is resolved from all other CA–N peaks *and* the corresponding C′–CX correlation (in this analysis, presumed to be either the C′–CA or C′–CB) is resolved from all other such C′–CX peaks in the spectrum. Similarly for 3D experiments, a unique connectivity is observed if both the ^{15}N dimension and C′–CX (in the NCOCX) or CA–CX (in the NCACX) are resolved from

Table 2 Theoretical unique connectivities of 3D and 4D backbone correlation spectra

Protein	Size[a]	Simulated LW (ppm)[b]	3D NCACX[c]	4D CONCACX[d]	3D NCOCX[e]	4D CANCOCX[f]
Het-S[g]	45 (43)	0.25	24	42	24	40
		0.5	8	33	6	36
		1.0	0	10	0	12
GB1	56 (56)	0.25	14	49	11	52
		0.5	7	41	5	39
		1.0	1	10	0	12
BPTI	58 (54)	0.25	14	48	13	49
		0.5	5	29	5	29
		1.0	2	9	0	6
Ubiquitin	76 (75)	0.25	15	62	14	67
		0.5	6	28	6	36
		1.0	4	7	2	8
Ubiquitin (SSNMR)	76 (64)	0.25	17	55	17	54
		0.5	8	29	9	36
		1.0	2	10	2	15
Maltose binding protein	376 (347)	0.25	8	167	6	196
		0.5	0	23	0	47
		1.0	0	0	0	5
Malate synthase G	723 (683)	0.25	7	165	4	200
		0.5	1	17	3	22
		1.0	0	2	1	3

[a] The value in parentheses represents the limit of ideal resolution, equivalent to the maximum number of connected spin systems calculated with the assumption of 0.001 ppm linewidths in all dimensions. This value in general does not converge to the number of residues in the sequence if ^{1}H dimensions in the spectra are not considered

[b] Common linewidths are assumed in each dimension

[c] Residues in NCACX experiments are counted if the N dimension is resolved and either the CA[i], C′[i], or CB[i] experiment is resolved

[d] Residues in CONCACB experiments are counted if both the C′[i−1] or N[i] is resolved and any of the corresponding CA[i], CO[i], or CB[i] is resolved

[e] Residues in NC′CX experiments are counted if the N dimension is resolved and either the C′[i−1], CA[i−1], or CB[i−1] experiment is resolved

[f] Residues in CANC′CX experiments are counted if both the CA[i] or N[i] is resolved and any of the corresponding C′[i−1], CA[i−1], or CB[i−1] is resolved

[g] The Het-s assignments do not include 10 C′ carbons. For this calculation, the value for these missing chemical shifts was set to the BMRB average shift for the residue

all other peaks in those dimensions of the spectrum. In the hypothetical scenario where unique connectivities for all residues in the protein were available, backbone assignments would be trivial and unambiguous; regardless of starting point in the sequence, a backbone walk could be performed all the way to the end of the sequence. This process could be easily automated without the need for complicated iterative algorithms. Therefore a complete set of unique connectivity information is a wholly sufficient, but not a necessary, criterion for assignment. In other words, one can be certain that assignments would be possible if such data were available.

To derive unique connectivities for all sites throughout even a small protein like GB1 using only 3D experiments would require linewidths of ~0.25 ppm (Table 2). The majority of the residues counted in both the NCACX for GB1 (residues M1, G9, T11, L12, G14, A20, D22, V29, V39, D40, T44, T49, K50, and F52) and NCOCX (excludes M1 and G9) correspond to uniquely resolved ^{15}N chemical shifts, which provide a convenient point to initiate backbone walk assignment protocols. (Indeed this is how we assigned GB1, and significant fractions of the sequence can be assigned with 2D data, despite the fact that unique connectivities are rarely observed in the 2D spectra.) Each CNCC experiment in GB1 provides ~90% of the unique connectivities to perform backbone assignments, assuming 0.25 ppm linewidths. Together the two hypothetical 4D spectra (CANCOCX and CONCACX) provide unique connectivities for the entire protein. With decreasing resolution, the number of unique connectivities in GB1 drops

from ~90% (0.25 ppm) to ~80% (0.5 ppm) to ~20% (1.0 ppm).

With larger proteins such as MBP and MSG, this analysis shows that it is not sufficient to perform 3D experiments to identify unique connectivities, even with ~0.25 ppm linewidths. Most notably, the difference in numbers of unique connectivities changes dramatically. In the case of MBP with 0.25 ppm linewidths, only 8 spin systems are uniquely connected in the NCACX 3D and 6 in the NCOCX 3D, reflecting the fact that very few ^{15}N sites are uniquely resolved in the 1D spectrum. By contrast, utilizing the C′–N plane as a source of polarization (i.e., in F1–F2) and CA–CB and CA–C′ as the destination of polarization (F3–F4), as in the CONCACX experiment, would yield 167 unique connectivities, or 196 in the CANCOCX experiment. Similarly large improvements are observed in the hypothetical MSG spectra. This key result is attributed to the fact that at some resolution, the majority of a given region of a 2D plane (C′–N, CA–CB, CA–C′, etc.) will become resolved at the level of individual peaks; at this threshold, performing assignments becomes trivial. Of course, it is not necessary to achieve this limit in order to make assignments, but the fact that in principle this can be achieved illustrates the potential to assign very large proteins by SSNMR.

This analysis shows that whenever possible, sample preparations should be optimized to give the narrowest possible linewidths. This is the most straightforward way to resolve a large number of resonances and to assign spectra of larger proteins. However, even in the most pessimistic scenarios outlined here, several dozen spin systems (ranging from 30 to 70% in small proteins) will be resolved in 4D experiments despite 3 ppm linewidths. Moreover, to assign very large proteins will require *both* high spectra resolution *and* 4D experiments.

Sensitivity in nD solid-state NMR experiments

Extension of standard SSNMR backbone chemical shift correlation methods (Hong 1999; McDermott et al. 2000; Rienstra et al. 2000; Seidel et al. 2004; Sonnenberg et al. 2004) to additional chemical shift dimensions offers the improvements in resolution as noted above, but in practice this can be achieved only in the limit of high sensitivity and digital resolution. As indirect chemical shift dimensions are added, explicit consideration of the tradeoffs in sensitivity and digital resolution is required in order to assess the value of the higher dimensionality experiments. Therefore in this section we present a formal analysis of the expected sensitivity, presuming parameters typical for SSNMR experiments in proteins, with reasonable estimates of the polarization transfer efficiency and decay parameters in order to derive approximate solutions that aid in parameter selection; later we provide actual experimental values, which agree well with the analysis here.

First we consider the theoretical sensitivity of a 2D ^{13}C–^{13}C experiment (S_{2D}) in comparison to the corresponding 1D ^{13}C experiment (S_{1D}), presuming the direct dimension is acquired and processed identically for both experiments:

$$S_{2D}^{CC}\left(S_{1D},\varepsilon^{CC},t_{1\max},T_2\right) = \frac{\varepsilon^{CC}S_{1D}}{\sqrt{2}t_{1\max}}\int_0^{t_{1\max}}\cos\left(\pi Jt\right)\exp\left[-t/T_2\right]dt \quad (3)$$

where ε^{cc} is the polarization transfer efficiency, the factor of root two loss is due to the hypercomplex sampling of the indirect dimension, and the terms following the integral correspond to the envelope function of the decaying signal of an exemplary ^{13}C site, such as the carbonyl of a peptide. This further presumes that no soft-pulse scalar decoupling methods are applied in the indirect dimension, as is the case throughout this study; straightforward extensions of this analysis are possible by considering that the scalar decoupling method removes the cosine modulation but also causes an overall loss of approximately 20% of the initial signal. Typical values for the parameters are J = 55 Hz and T_2 = 10 ms, and the maximum evolution time would usually be chosen to be approximately equal to (or slightly less than) the point of zero-crossing in the J-modulation, i.e., $t_{1\max} \cong 1/(2J)$. In this regime the envelope function is closely approximated by a linear decay,

$$S_{2D}^{CC}\left(S_{1D},\varepsilon^{CC},t_{1\max},T_2\right) = \frac{\varepsilon^{CC}S_{1D}}{\sqrt{2}t_{1\max}}\int_0^{t_{1\max}}(1-2Jt)\,dt \quad (4)$$

the integral of which evaluates to

$$S_{2D}^{CC}\left(S_{1D},\varepsilon^{CC},t_{1\max},T_2\cong\frac{1}{2J}\right) = \frac{\varepsilon^{CC}S_{1D}}{\sqrt{2}}\left(1-Jt_{1\max}\right) \quad (5)$$

Thus if t_{1max}=1/(2J) and $\varepsilon^{CC} \cong 20\%$, as is typical for one-bond correlations with various homonuclear broadband mixing schemes (DARR, RFDR, SPC-5, etc.), the overall sensitivity of the 2D ^{13}C-^{13}C would be ~7% of the 1D. This is the limit in which most 2D ^{13}C–^{13}C experiments are performed (i.e., t_{1max} = ~9 ms); reducing the t_{1max} value to ~4.5 ms increases the relative sensitivity by 50%, or to 10.6% relative to the 1D experiment. Thus a 2D experiment with equal sensitivity to the 1D would require a factor of 100–200 more acquisition time than the 1D experiment.

In practice, this factor is not prohibitive since with modern instruments 1D ^{13}C spectra can be acquired with individual site sensitivity of at least ~200/(μmol ppm h$^{1/2}$); i.e., a single site with 1-ppm linewidth with a μmol of protein in practice yields a signal-to-noise ratio (SNR) of 100:1 in

approximately 15 min. (~512 scans). Although instrumental performance may vary from this approximation, this scenario is not atypical, and along with the assumptions above would yield a 2D spectrum with one-bond cross peaks having a SNR of ~25:1 in ~3 h (assuming optimal digitization, t_{1max} = 9.0 ms) or less than 1.5 h if inferior digital resolution (t_{1max} = 4.5 ms) is acceptable. The values used here are similar (within a factor of two in most cases) to those reported by several groups, once appropriate corrections are made for B_0 field, sample quantity and linewidths (McDermott et al. 2000; Pauli et al. 2001; Bockmann et al. 2003; Castellani et al. 2003; Martin and Zilm, 2003; Igumenova et al. 2004a, b; Marulanda et al. 2004; Franks et al. 2005; Hiller et al. 2005; Marulanda et al. 2005; Pintacuda et al. 2006; Etzkorn et al. 2007). These studies have demonstrated in many instances that the *acquisition* of 2D spectra of solid proteins ranging in size from 5 to 20 kDa is now relatively routine.

Construction of heteronuclear correlation schemes allows even higher efficiency transfers to be utilized ($\varepsilon \cong 50\%$, as discussed in more detail below) and straightforward elimination of the scalar ^{15}N–^{13}C evolution. Thus the expression describing sensitivity of the ^{15}N–^{13}C 2D is

$$S_{2D}^{CC}\left(S_{1D}, \varepsilon^{CC}, t_{1\,max}, T_2\right) = \frac{\varepsilon^{NC} S_{1D}}{\sqrt{2}} \frac{T_2\left(1 - \exp[-t_{1\,max}/T_2]\right)}{t_{1\,max}} \tag{6}$$

Assuming t_{1max} = 2 T_2, this evaluates to ~15% S_{1D}, or ~22% S_{1D} if shorter maximum evolution times are employed (t_{1max} = T_2). Even with lower performance in the ^{15}N–^{13}C transfer step than assumed here, such heteronuclear spectra therefore can readily be acquired in a few hours or less (Wylie et al. 2006), and straightforwardly enable extension to 3D by combination with the homonuclear schemes. Thus the analogous NCC 3D experiments (Fig. 1a, b) results in an expression for theoretical sensitivity as follows, combining Eqs. 5 and 6 to yield

$$S_{3D}^{NCC} = \frac{\varepsilon^{NC}\varepsilon^{CC} S_{1D}}{2} \frac{T_2^N\left(1 - \exp[-t_{1\,max}/T_2]\right)}{t_{1\,max}^N} \left(1 - Jt_{2\,max}^C\right) \tag{7}$$

Typical digitization parameters for excellent digital resolution would be t_{1max} = 2T_2 for ^{15}N and t_{2max} = 1/(2 $^1J_{CC}$) for the indirect ^{13}C dimension, yielding overall sensitivity of the 3D experiment per unit time of ~1.1%, and to observe the principal cross peaks (e.g., N[i]–C′[i−1]–CA[i−] in the NCOCX experiment) with SNR = 10 with the assumed parameters would require ~21.5 h. The assumed parameters correspond to reasonably high digital resolution; reducing the evolution times to t_{1max} = T_2 for ^{15}N and t_{2max} = 1/(4 $^1J_{CC}$) for ^{13}C would recover a factor of two in sensitivity and therefore reduce the required measurement time, and/or enhance the sensitivity of ^{13}C signals further along the sidechain that may be valuable for confirming assignments.

Extending further to four dimensions, we presume the additional indirect dimension involves a ^{13}C spin, considering Fig. 1c and d, where the first CP period places polarization on the ^{13}C site. This enables two, highly efficient heteronuclear transfers to be utilized, and minimizes the required sampling bandwidth in the indirect ^{13}C dimensions. This situation would yield an overall sensitivity of

$$S_{4D}^{CNCC} = \frac{\varepsilon^{CN}\varepsilon^{NC}\varepsilon^{CC} S_{1D}}{2\sqrt{2}} \left(1 - Jt_{1\,max}^C\right) \frac{T_2^N\left(1 - \exp[-t_{2\,max}/T_2^N]\right)}{t_{2\,max}^N} \left(1 - Jt_{3\,max}^C\right) \tag{8}$$

Assuming again that the heteronuclear transfers have 50% efficiency, the homonuclear transfer has 20% efficiency, that t_{1max} = t_{3max} = 1/(4 J), and that t_{2max} = T_2^N, the calculated overall efficiency is 0.6%. In practice, one might prefer to reduce the maximum evolution times in each dimension further, to recover sensitivity such that it is comparable to the 3D, at the cost of digital resolution (which is slight and may be recovered by linear prediction schemes in the data processing if desired). For example, using the values t_{1max} = t_{3max} = 1/(8 J) and t_{2max} = 0.5 T_2^N approximately doubles the sensitivity. Alternatively, the 4D spectrum could be acquired with greater digital resolution, of course requiring longer signal averaging times.

In each sample the optimal tradeoff of resolution and sensitivity will depend upon the measured decay constants and linewidths, as typically determined in 2D spectra (with partially overlapping peaks an upper bound can be placed on this value in combination with the known fine structure and decay constants). Here we have assumed the case of small inhomogeneous broadening, as observed with most microscopically well-ordered solid proteins, where the value T_2 is approximately equal to T_2*. With optimized heteronuclear decoupling, the T_2 value furthermore tends to be ~10 ms or more. The 4D scheme of course allows simple adjustment in the limit of greater inhomogeneous broadening, where t_{max} values would be reduced accordingly to recover sensitivity yet increase resolution greatly relatively to the corresponding 2D or 3D experiments.

Experimental implementation: polarization transfer along the peptide backbone

As described above, in theory both CNCC 4D correlation schemes provide similar sensitivity. However, in

experimental practice (Fig. 2) the CANCOCX scheme is expected to provide greater sensitivity than the CONCACX for two major reasons. First, the initial cross polarization from protons proceeds with higher efficiency from the HA to CA than it does from remote protons that polarize the carbonyl site; in practice, we find that this effect is not prohibitively large (~25% difference), but favors the CANCOCX version of the experiment. Second, in the CANCOCX experiment, the final homonuclear mixing scheme results in a greater percentage of polarization on its primary destination (F4) spins, the CA[i-1], than does the CONCACX on its destination spins (CB[i+1] and C′[i+1]); i.e., the CA site is a bifurcation point in the polarization transfer pathway, allowing transfer either to the C′ or the CB and subsequent side-chain carbons in the CONCACX experiment. Instead, with polarization resident on the C′ in t_3 in the CANCOCX experiment, transfer to the CA site is more efficient. Although the CONCACX provides a greater number of one-bond correlations within the i+1 residue, the cost is approximately a factor of two in sensitivity in the limit of short homonuclear mixing (<20 ms). One could also envision 4D schemes initiating on ^{15}N and resolving multiple side-chain carbons in different dimensions (e.g., NCACBCG); these sequences would require customized homonuclear polarization transfer schemes for each set of amino acid side-chain topologies, and therefore they have less general utility. Although these experiments will have great utility for challenging assignment problems, they will require further specific development and are not discussed further in this study. Clearly all of the possible 4D schemes offer complementary information and might be combined in a complete suite of experiments for assignments of large proteins.

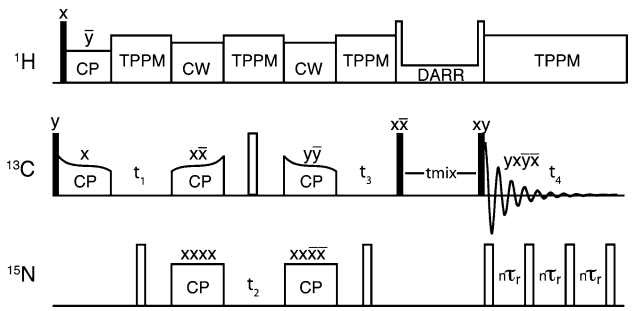

Fig. 2 Pulse scheme for CNCC chemical shift correlation experiments. Solid and open rectangles represent $\pi/2$ and π pulses respectively. Band-selective ^{13}C–^{15}N and ^{15}N–^{13}C polarization transfers are achieved with tangent-ramped SPECIFIC CP (Baldus et al. 1998), and DARR (Takegoshi et al. 2003) or RAD (Morcombe et al. 2004) homonuclear ^{13}C–^{13}C mixing. Two-pulse phase modulated (TPPM) decoupling (Bennett et al. 1995) was applied during the evolution and acquisition periods. Time-proportional phase incrementation (TPPI) (Marion and Wüthrich 1983) was applied to the pulse sequence elements immediately following the frequency encoding periods

For this study, we chose the CANCOCX for detailed experimental analysis, examining the polarization transfer efficiency in a well-studied protein, GB1, at 500 MHz. In our experience, heteronuclear polarization transfers involving strong, one-bond couplings such as the N–CA and N–C′ proceed with roughly 50% efficiency, when comparing the absolute intensity in the ^{13}C-observed double cross polarization (DCP) (Schaefer and Stejskal 1979) experiment using tangent-ramped SPECIFIC CP (Baldus et al. 1998) (with transfer from ^{1}H to ^{15}N to ^{13}C) to the direct polarization transfer from ^{1}H to ^{13}C. Here the 1D reference spectrum (Fig. 3a) has been thoroughly optimized for maximal overall intensity, and the ^{1}H–^{13}C contact time adjusted for greatest signal intensity in the C′ region. Under these conditions, the carbonyl region integrates to a larger total spectral intensity than the major region of CA intensity, due in part to the fact that some Thr (T25), Gly (G9, G14) and Val (V29) CA peaks do not lie in the canonical range of CA shifts used for integration here (i.e., 45–65 ppm where there are convenient notches in the spectrum of GB1), and therefore do not contribute to the spectral intensity of the CA region in the reference spectrum. This accounts for most of the difference in integrated intensity between the CO and CA regions, despite the loss of intensity to the sidebands. In addition, Glu and Asp sidechain carbonyl intensity is included in the direct CP integration. Overall, the ^{1}H–^{13}C CP efficiency is approximately three, when compared with direct ^{13}C Bloch decay spectra with sufficiently long pulse delays to achieve thermal equilibrium. Among proteins studied in our laboratory, the CP efficiency measured in this manner ranges from ~2.0 to 3.0, with GB1 at 500 MHz at the high end of this range. Thus the ^{1}H–^{13}C polarization transfer is judged to be close to optimal, and the H–N–CO polarization transfer pathway actually results in a larger total signal intensity than the Bloch decay CO spectrum at thermal equilibrium.

With this reference spectrum accurately evaluated, we proceeded to measure the efficiency of heteronuclear transfer. The N–C′ transfer is close to 60% of the reference spectrum intensity when performed with on-resonance ^{13}C spin lock fields of approximately 7/2 ω_R. In this example, the spinning rate is 11.1 kHz, so the ^{15}N spin lock field (on resonance with the amide region) is 5/2 ω_R (i.e., expressed as a nutation frequency, 5/2 $\omega_R/2\pi$ = ~27.8 kHz) with the C′ spin lock of ~39 kHz (adiabatic ramp shapes and mixing times are quoted in the figure caption). In the N-C′ transfer (Fig. 3b), some polarization transfer to the CA region is observed. This incidental transfer is accepted as part of the tradeoff in achieving optimal adiabatic transfer efficiency for the C′ sites. Notably, the observed *experimental* transfer efficiency here is several percent higher than the *theoretical* maximum of the TEDOR method

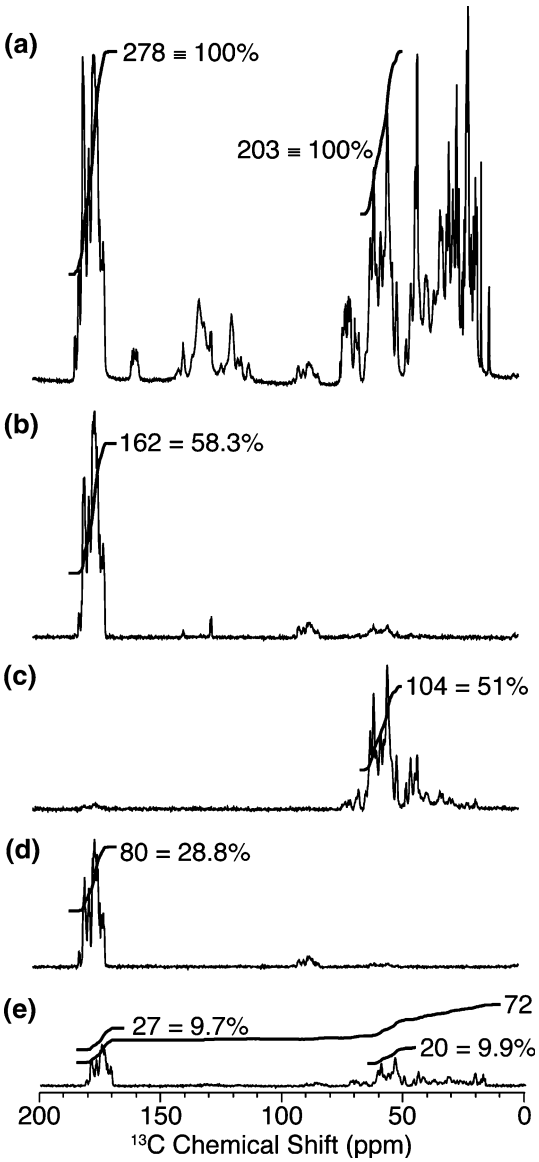

Fig. 3 ^{13}C reference spectra with signal intensity derived from the polarization transfer schemes of Fig. 1. (**a**) ^{13}C CP-MAS spectrum with CP directly from the proton reservoir, using adiabatic CP (Hediger et al. 1995). Total spectral intensity was first optimized, and then the contact time adjusted to maximize total signal for the CO region. This intensity was used to define 100% transfer efficiency. (**b**) ^{13}C-detected spectrum with H–N and N–C′ CP transfers. The ^{13}C carrier frequency was placed at 55 ppm with $\omega_N^{eff} = 5/2\ \omega_R$ and $\omega_C^{eff} = 3/2\ \omega_R$ with a tangent ramp of amplitude ±0.1 ω_r on the ^{13}C channel. (**c**) ^{13}C-detected spectrum with H–N and N–CA CP transfers. The ^{13}C carrier frequency was placed at 55 ppm with $\omega_N^{eff} = 5/2\ \omega_R$ and $\omega_C^{eff} = 7/2\ \omega_R$ with a tangent ramp of amplitude ±0.1 ω_r on the ^{13}C channel. (**d**) ^{13}C-detected spectrum with consecutive H–CA, CA–N and N–C′ transfers. (**e**) ^{13}C-detected spectrum with H–CA, CA–N and N–C′ CP steps followed by 18 ms homonuclear DARR mixing (Takegoshi et al. 2003). Spectral intensities in all cases are reported as integrated signals over the canonical ranges of chemical shifts for CO (168–182 ppm) and CA (48–64 ppm)

(52%); TEDOR is experimentally more robust with respect to instrumental imperfections such as B_1 field inhomogeneity and amplifier instabilities. However, when conditions of stable, adiabatic, band-selective CP can be achieved, we observe this to be optimal (*vide infra*).

As with the N–C′ transfer, SPECIFIC CP configured for transfer from the amide N to CA region proceeds with similar efficiency. Here the CP transfer is performed with lower amplitude spin lock on the ^{13}C channel (3/2 ω_R, or a nutation frequency of ∼16 kHz) with the carrier placed at 55 ppm; on the ^{15}N channel, similar spin-lock parameters are used as were used for the N–C′ transfer (i.e., $\omega_N = 5/2\ \omega_R$). These choices of field amplitude minimize the effect of signal loss due to insufficient ^{1}H decoupling (Ishii et al. 1995; Bennett et al. 1998), while retaining sufficient bandwidth. Notably, even with the lower amplitude spin lock field, the effective field varies by <1% over the ±15 ppm region on interest; for example, with the ^{13}C carrier placed at 55 ppm, the most downfield or upfield CA peaks would have maximum offsets of 15 ppm, corresponding to an effective field of $\sqrt{16.0^2 + (15.0 \times 0.125)^2} = 16.1$ kHz. Because the tangent ramp amplitude (approximately ±1.5 kHz) is significantly larger than the offset of the effective fields (i.e. 100 Hz), the overall efficiency of polarization transfer does not depend significantly on the offset of the carrier from resonance within the CA bandwidth; i.e., the transfer is stable over the entire spectral region of interest. Therefore the edges of the spectrum, where important and generally well-resolved resonances are observed (e.g., Thr, Ser, Pro, Val, Ala, Gly) are not missing or significantly attenuated in intensity when compared to peaks within the center of the CA bandwidth. This result was demonstrated for GB1 in our previous study of chemical shift anisotropy tensors (Wylie et al. 2006) and evaluated in a site-resolved manner in the following section.

The result of combining the two polarization transfer steps into a single experiment, with H–CA and N–C′ transfers, is illustrated in Fig. 3d. The overall transfer efficiency, again normalized to the reference spectrum (Fig. 3a), is almost 29% for the C′ region, agreeing well with the expectation of 58.3% * 51.0% = 29.7%. The small difference in overall efficiency is attributed to the fact that the N–C′ transfer step is performed off resonance for ^{13}C, with the carrier centered at 55 ppm. Thus an off-resonance spin-locking condition is achieved, with typical parameters including $\omega_{C,eff} = \sqrt{35.8^2 + (120.0 \times 0.125)^2} = 38.9$ kHz. The ^{13}C carrier offset from the carbonyl resonance of 120 ppm (15 kHz) enables a simple separation of effective fields

experienced by the C′ and CA sites, matching the N–C′ CP condition ($\omega_N = 5/2\ \omega_R$; $\omega_{CO}^{eff} = 7/2\ \omega_R$) optimally but encountering the N–CA ($\omega_N = 5/2\ \omega_R$; $\omega_{CA}^{eff} = 7/2\ \omega_R$) condition only at the very top (end) of the ^{13}C field tangent ramp. Figure 3e shows the result of performing the three consecutive heteronuclear transfers, with a final longitudinal DARR (Takegoshi et al. 2003) or RAD (Morcombe et al. 2004) mixing period of 18 ms. The resulting polarization transfer to the CA region is 9.9%, in good agreement with, or better than, the estimates above. The majority of polarization (~90%) is confirmed, through analysis of the 4D experiments below, to arise from the CA[i]–N[i]–C′[i−1]–CA[i−1] pathway, with a smaller but significant (~5% each) amount in each of the other two pathways, CA[i]–N[i+1]–C′[i]–CA[i] and CA[i]–N[i]–CA[i]–CA[i].

We have reproduced these polarization transfer schemes with several proteins studied in our laboratory at ^1H frequencies of 500, 600 and 750 MHz instruments by appropriate adaptation of the spin rate and RF field strengths in relative terms. The parameters used here correspond to 11.1 kHz MAS rate at 500 MHz, which can be scaled accordingly to 13.333 kHz at 600 MHz or 16.666 kHz at 750 MHz. The efficiency of the CANC′ scheme typically is reduced at higher B_0 field, ranging from ~30% at 500 MHz to 20% at 750 MHz. Considering the improved detection sensitivity and initial Boltzmann polarization at higher field, the overall sensitivity of the experiment is approximately constant over this range of B_0 field, although the resolution improves greatly with higher field. Further pulse sequence and instrumentation advances will ensure that greater benefits are derived from the high B_0 field configuration, as now observed most dramatically for broadband ^{13}C–^{13}C experiments.

Instrumental stability

Successful and reproducible implementation of three CP steps over many days to weeks of acquisition time requires appropriately optimized instrumentation. The 1D reference spectra in Fig. 3 were acquired in approximately a minute each and therefore demonstrate optimal short-term performance but give no indication of long-term instrumental stability. Variations of ~1% in the RF field amplitudes from the nominal CP spin lock amplitudes severely compromise CP efficiency. Fluctuations in CP efficiency over the course of hours to days of a multidimensional experiment impede resolution and sensitivity. Drifts in the CP amplitude yield an artificially imposed decay that cannot be separated from natural sources of relaxation; therefore indirect dimension linewidths under such circumstances are inflated. Oscillations or sudden changes in the amplitude likewise would produce modulations in the indirect time domains or "t_1 noise". This is a common problem even in solution NMR, where the signal intensity variation from pulse imperfections is linear with respect to the change in RF amplitude (i.e., an 88° pulse excites only ~2% less signal than a 90° pulse). The problem is much more severe for SSNMR spectra than in solution NMR, due to the fact that the CP signal can change by ~50% of its maximal amplitude due to only a ~2% change in RF amplitude.

To address this problem, we have customized our spectrometers and pulse sequences and monitored the intensity of the CP signal (Fig. 4) to evaluate overall instrumental performance. The principal contribution to variations in B_1 field amplitude is drift of the room temperature, which can affect both the amplifier gain as well as probe tuning; maintaining temperature to within ±0.5°C is preferred, with a constant recirculation of air through the room to avoid periodic oscillations. Slightly larger variations in temperature (±1°C) can be tolerated with amplifiers specifically designed to minimize the temperature-dependent gain (e.g., with passive compensation of temperature drifts). Active B_1 field compensation as used for many years by Schaefer and co-workers (Stueber et al. 2006) is expected to enhance performance further in the context of nD experiments. In addition to amplifier stability, probe tuning stability is essential; all pulse sequences used here have a constant total overall RF duty cycle, and the probe is tuned in the steady state under the experimental MAS and variable temperature settings. While optimizing the spin lock amplitudes, a sufficient number of pulses is employed at the beginning of each array to ensure

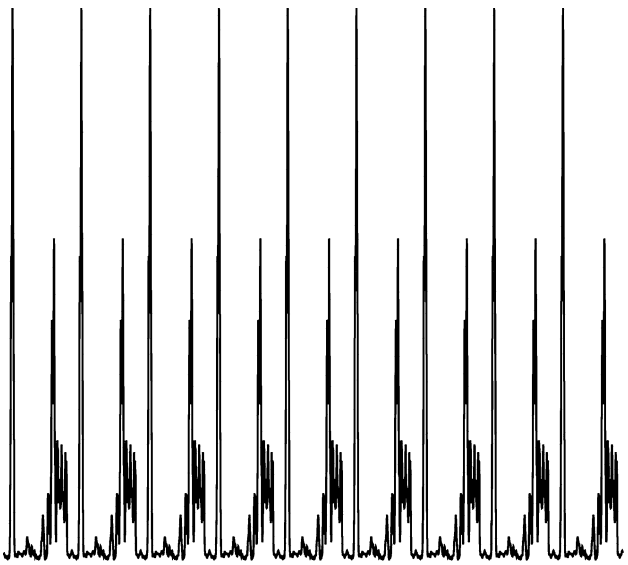

Fig. 4 Representative ^{13}C-detected CANCOCX reference spectra. The nine spectra acquired periodically over 24 h demonstrate instrumental stability and are presented as a stacked plot at the same absolute intensity scale for each spectrum. Integration of the full spectral width varies by <2% within this group of experiments, consistent with <0.2% variations in amplifier power

that the steady-state tuning is achieved; although this number varies depending on the details of probe performance and sample ionic strength, in our experience 16–32 pulses are sufficient. Optimal amplitude values are confirmed with multiple iterations of the array, and fine-tuned with each new sample.

The CP reference spectrum amplitude was monitored periodically (Fig. 4) throughout the course of the nD experiments described further below. The spectra showed variations of at most ±2% in integrated intensity, corresponding to less than 0.2% variations in any of the RF channels over the period of the 48-h experiment. Comparable performance is observed for many days, provided that the room temperature stays within the range specified above. As necessary, data acquired during periods of substandard instrumental performance is discarded and the relevant data blocks acquired again. Data in the 4D acquisition is deemed acceptable only if the reference spectra before and after the acquisition block are within the stated tolerance above. More sophisticated software approaches to analyzing bad data blocks are also possible, although the optimal approach is to avoid the fluctuations with a stable instrument.

Cross-polarization bandwidth

An experimental concern that relates closely to the issue of instrumental stability is the bandwidth of the ^{15}N–^{13}C CP conditions, which must be sufficiently narrow (following SPECIFIC CP principles) (Baldus et al. 1998) to transfer signal primarily from N to CA or C′, but also sufficiently broad to cover the entire CA or C′ bandwidth in the experiment of interest. We show (Fig. 5) that using the CP conditions described above that the signal intensities of outlying resonances, such as Val, Thr and Gly N–CA correlations, have comparable sensitivity to the resonances in the central region of the spectrum. The signals on the periphery of the spectrum have particularly high value in the context of performing assignments, since they are often (even in larger proteins) resolved in 2D planes and have side-chain spin systems that can readily be classified to a particular amino acid type. In addition to the examples shown here for GB1, Pro and Ser residues (which are not present in GB1) have well-resolved signatures and again tend to fall on the edges of the nominal CA bandwidth.

Beyond the qualitative appearance of the spectra in Fig. 5, quantitative analysis further supports this idea. In the NCO 2D (Fig. 5a), the upfield ^{15}N region has several individually resolved peaks, including T49-A48, G14-K13, T11-K10, G9-N8, T44-W43, G41-D40 and G38-N37; although G41-D40 and G38-N37 are significantly overlapped, they are still easily distinguishable in standard integration and SNR analysis routines in Sparky (Goddard and Kneller 2006). Among these seven resonances, the SNR is 59.9 ± 9.8 with outlying values of 43 (T11-K10) and 71 (G14-K13). For another set of several well-resolved resonances in the downfield ^{15}N region (F52-T51, E56-T55, D40-V39, L7-I6, I6-L5, D46-Y45 and T55-V54), the SNR is 52.7 ± 8.5, with individual values ranging from 38 (D40-V39) to 65 (I6-L5). Clearly the variations from site to site, attributable to local dynamics and other variations in the local environment, are as large or larger than the variations due to the dependence of polarization transfer upon ^{15}N chemical shift. The range of C′ chemical shifts is extremely small compared to the typical spin lock fields used on the ^{13}C channel, so likewise it is not surprising that variations as a function of ^{13}C chemical shift within the C′ range are also negligible.

This trend is also observed among the N-CA correlations in Fig. 5b, even when a smaller spin lock field (with

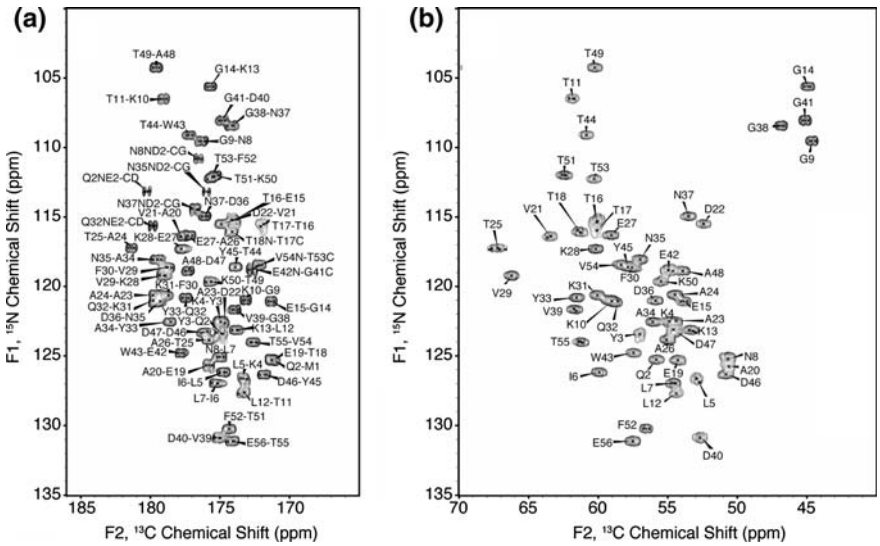

Fig. 5 2D NCO and NCA correlation spectra of GB1. (**a**) NC′ 2D. (**b**) NCA 2D. Peaks are labeled in accordance with prior assignments (Franks et al. 2005). Lorentzian-to-Gaussian line broadening (net +5 Hz) was utilized in each dimension

tangent ramp, as described above) must compensate for a larger (~25 ppm) spectral width. The four Gly residues in the upper right corner of the spectrum (upfield in each dimension) have SNR of 42.8 ± 2.1. Several resonances in the downfield ^{15}N region (F52, D40, E56, I6, L5, E19 and Q2) have SNR of 35.0 ± 9.6. The upper left region of the spectrum, including six individually resolved Thr residues (T49, T11, T44, T51, T53, T25) and two Val (V21 and V29) have 31.2 ± 9.8, with a low value of 18 for T25, which we attribute to the fact that the T25CB resonance has a very similar chemical shift value and therefore competes for polarization with the T25CA site. Although the upper left region of the spectrum has a lower overall SNR than the upper right, it is notable that the two extreme outlying resonances (T49 and V29) have higher sensitivity (37 and 46, respectively), consistent with expectations for the Gly region. Therefore, we again conclude that the bandwidth of the N–CA polarization transfer, or by symmetry the CA–N transfer in the 4D experiment, has equivalent efficiency over the bandwidth of interest, within the experimental error.

Interresidue CA–CA correlations

Previous investigators have demonstrated the utility of proton-driven spin diffusion (PDSD) experiments under conditions approximating rotational resonance for the C′–CA favoring dipolar transfer along the peptide backbone (Seidel et al. 2004). This approach has been applied to challenging assignment problems (Lange et al. 2006; Etzkorn et al. 2007), by virtue of unique sets of interresidue cross peaks that arise in the 2D ^{13}C–^{13}C correlation spectra under such conditions. We therefore compared the PDSD experiment in GB1 at the specified spinning rate of 7.814 kHz (corresponding to ~62.5 ppm for ^{13}C, Fig. 6a) with another commonly used homonuclear correlation scheme, DARR (Takegoshi et al. 2003) (Fig. 6b), at the standard spinning rate (11.1 kHz) that we employ on our 500 MHz instrument. This spinning rate minimizes overlap of C′ sidebands with the aliphatic region, placing the first order sideband downfield of the CA region and the second sideband upfield of the methyl region. Both the PDSD and DARR experiments yield broadband correlations and a high overall efficiency of transfer among the side-chain resonances at the chosen mixing time (200 ms). In both cases, sets of interresidue correlations are well resolved in the near-diagonal region involving the CA resonances (F_1 = 50–65 ppm and F_2 = 50–65 ppm), as well as the interresidue CA–CB (or in general CA–CX) regions uniquely resolving Thr or Val (F_1 = 65–72 ppm), Gly (F_2 = 44–48 ppm), Ile (F_1 = 12–16 ppm) residues, as labeled on Fig. 6. Depending on the exact spinning rate and mixing time, the precise number of cross peaks may vary between these two experiments, but the overall pattern is similar. In general the DARR spectrum appears to give somewhat higher sensitivity under these conditions (despite ~38% shorter total measurement time than the PDSD experiment). This can be attributed in part to the relatively low B_0 field conditions used for this study; previous investigators employed these schemes primarily at 600 to 800 MHz ^{1}H frequencies, where weak coupling conditions are achieved at correspondingly higher MAS rates. Nevertheless, both spectra share the property of substantial overlap in many regions where backbone interresidue correlations are observed. For example the CA–C′ region (F_1 = 50–62 ppm, F_2 = 170–180 ppm) shows very few peaks individually resolved, despite the favorable resolution of the GB1 sample used here; similarly, the major portion of CA-CB region (F_1 = 50–62 ppm, F_2 = 30–40 ppm) is highly congested. We attribute this in both cases to the fact that not only interresidue C[i]–C[i−1], but also C[i]–C[i+1] and intraresidue C[i]–C[i] peaks are observed, and the intraresidue peaks have much greater intensity than the interresidue peaks.

Figure 6c illustrates the analogous CA(NCO)CX experiment, which establishes correlations between the CA[i] and most resonances (C′, CA, CB and additional sidechain sites) within the $i-1$ residue. This polarization transfer pathway mostly alleviates the problem of degeneracy among C[i], C[i+1] and C[i−1] correlations. Although the experiment requires a triple resonance probe and is more demanding in terms of instrumental performance, the benefit is a significantly simplified 2D plane in which only the CA[i]–CX[i−1] peaks are observed at high intensity. Due to imperfections in the CP conditions (as discussed above; namely the transfer from N[i] to CA[i] during the second step, which is intended to transfer from N[i] to C′ [i−1]), intraresidue correlations are also observed, but in general they are much lower in intensity than the interresidue correlations (<10% of the total signal). In addition to the weak autocorrelations, the signal along the diagonal near F_1 = F_2 = 55–60 ppm also includes a contribution from pairs of amino acids that coincidentally have similar CA shifts. Even the more congested CA[i]–C′[i−1] region of the spectrum shows most peaks individually resolved. Notably, the Gly CA[i] region is also well represented in terms of signal intensity, for reasons described above.

Relatively few individual interresidue cross peaks are resolved either in the PDSD and DARR spectra; among those identified uniquely, most have a SNR of 4:1 to 6:1. For example, V29CA-K28CA, V39CA-G38CA and G38CA-N37CA are some of the best-resolved peaks in the PDSD spectrum, with SNR values of 6, 5, and 4 respectively. The same peaks in the DARR spectrum have greater

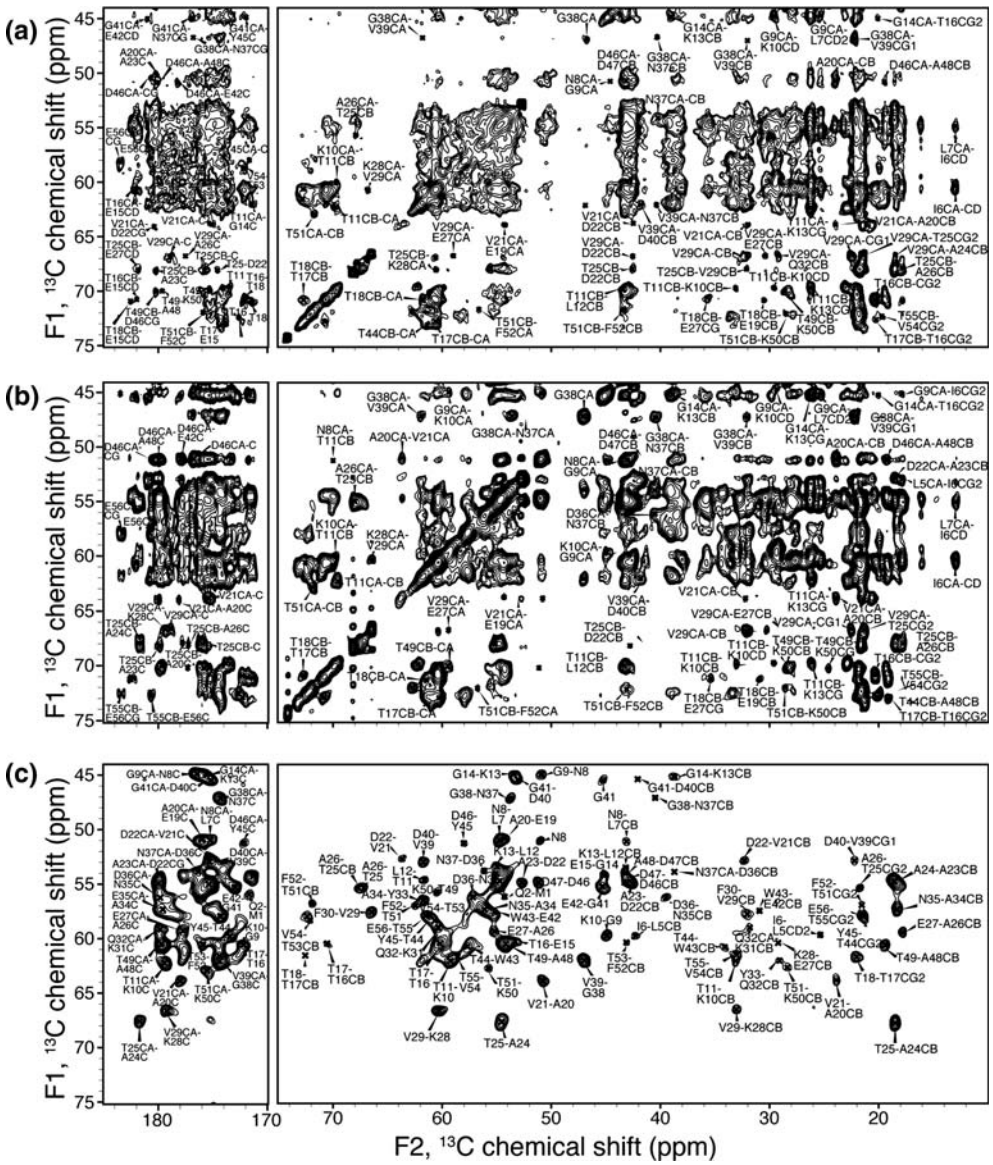

Fig. 6 ^{13}C–^{13}C chemical shift correlation spectra of GB1. The three experiments were performed consecutively on the same sample at 500 MHz ^{1}H frequency. (**a**) 200 ms PDSD (Bloembergen 1949) mixing with slower MAS rate (7.814 kHz) to achieve the weak-coupling limit (Seidel et al. 2004). (**b**) 200 ms DARR mixing at 11.111 kHz spinning. (**c**) CA(NCO)CX mixing with 6 ms CA–N, 6 ms N–C′ and 18 ms DARR (Takegoshi et al. 2003) mixing at 11.111 kHz spinning. All spectra were acquired consecutively under optimized CP conditions (see discussion of Fig. 3)

sensitivity (8, 13 and 14), which is comparable to the CA(NCO)CX (12, 13, and 7). Furthermore, several interresidue cross peaks in the CA(NCO)CX spectrum, such as V21CA-A20CA and T25CA-A24CB, have greater than 10:1 SNR, whereas they are either not resolved or not observed in the other experiments. Overall, in the CA(NCO)CX we observed 53 of 55 possible CA[i] resonances correlated to their preceding CA[i−1] or CB [i−1], and 40 of 55 resolved CA[i]-C′[i-1] peaks. In contrast, due to the abundance of intraresidue intensity in the PDSD and DARR spectra, very few interresidue peaks can be uniquely identified; in most cases, even when peaks are resolved, without prior knowledge it would not be possible to identify them correctly as i+1 or i-1 correlations.

In the experiments shown here, we processed the CA(NCO)CX with a larger amount of net line broadening (50 Hz or 0.4 ppm in each dimension). Despite this, the CA(NCO)CX spectrum has greater resolution than either of the alternatives. The processing with approximately matched line broadening is possible because the spectrum is relatively free of congestion. This yields a greater sensitivity on a per-site basis than the PDSD experiment. However, even when processed with the identical apodization functions as the PDSD and DARR spectra, the sensitivity in the CA(NCO)CX is competitive with DARR and the resolution is dramatically better, owing to the fact that only one-third the number of peaks are present. Under these conditions with a 56-residue protein at 500 MHz, the PDSD and DARR spectra provide relatively little useful interresidue correlations among the C′ and CA sites. The CA(NCO)CX experiment performs optimally in identifying unique sequential correlations. Therefore we use the band-selective transfer scheme for extension to 3D and 4D in the following sections.

Digitization and processing of the indirect chemical shift dimensions

For proteins such as GB1 under the experimental conditions described above, the total experiment time is found to be limited not by sensitivity (for example, the 2D experiments in Fig. 6 were acquired in 2.4 h each), but by the digital sampling of indirect dimensions. This is the "sampling limited" regime as described by Szyperski and co-workers (Szyperski et al. 2002). In the spectra presented here, maximum evolution times of 4.32 ms in the indirect CA dimension (32 × 135 μs), 8.64 ms in the indirect N dimension (32 × 270 μs) and 5.4 ms in the indirect C′ dimension (20 × 27 μs) were employed.

These evolution times are significantly less than the T_2 values for many resonances in GB1, leading to substantial truncation of the interferograms (Fig. 7, boxed portion). This shortcoming may be addressed experimentally by sampling indirect dimensions more efficiently, as recently demonstrated in solution with methods such as GFT (Kim and Szyperski 2003), projection reconstruction (Kupce and Freeman 2004), triangular sampling (Aggarwal and Delsuc 1997) and non-linear sampling (Hoch and Stern 2001). Furthermore, data processing methods to enhance digital resolution are well established, including linear prediction (Barkhuijsen et al. 1987) and maximum entropy reconstruction (Hoch and Stern 2001). Linear prediction has been demonstrated to be generally effective and reliable for analyzing protein spectra with Cartesian sampling. So although other digitization protocols are likely to be useful for SSNMR applications in the future, we focus on linear prediction here.

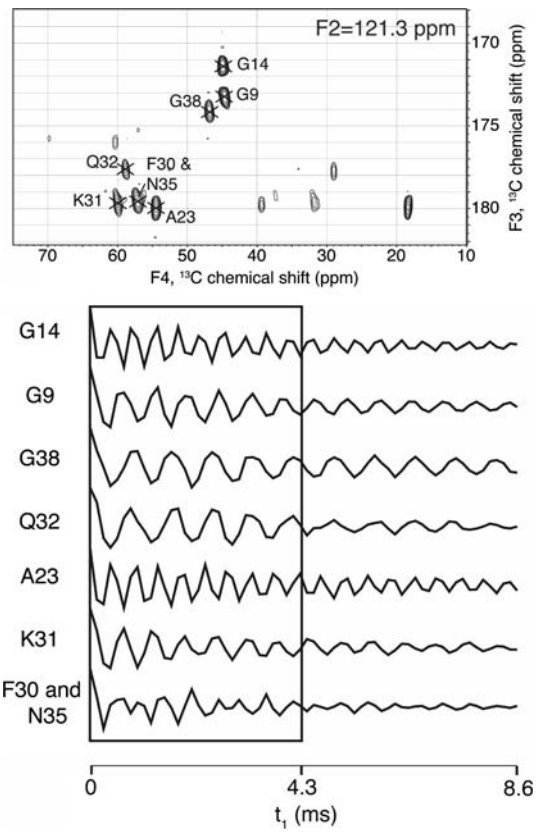

Fig. 7 Indirect CA dimension (t_1) interferograms from a 4D CANCOCX experiment on GB1. The t_4, t_3 and t_2 dimensions were processed with matched Lorentzian-to-Gaussian apodization. The t_1 dimension is illustrated with the acquired data extended by a factor of two with the standard linear prediction parameters in nmrPipe

For example, in Fig. 7 the t_1 dimension interferograms for several resonances are shown. At the maximum acquired evolution time of 4.3 ms, approximately ~50% of the initial intensity remains. Notably, the Gly CA resonances lack the CA–CB scalar coupling and so have greater intensity, which is extended to more than 8 ms with the linear prediction. Other signals show slightly less benefit, but overall the linewidth reduction is almost the full factor of two as anticipated. Processing without linear prediction resulted in an average CA linewidth (among ~20 well-resolved peaks) of 1.5 ± 0.3 ppm, whereas with linear prediction the average linewidth was reduced to 0.8 ± 0.2 ppm. Likewise, the average N linewidth decreased from 1.7 ± 0.2 ppm to 1.1 ± 0.1 ppm upon linear prediction of the t_2 dimension. The benefits of predicting the CA and N indirect dimensions is illustrated graphically in Fig. 8, with 2D skyline projections extracted from the 4D data set. In the CA–CX projections (Fig. 8a, b), beyond the cosmetic improvements at least two congested regions are significantly enhanced upon linear prediction, enabling the resolution of A48-D47 from A24-A23 and K13-L12. The overall resolution is comparable to the 2D planes shown in Fig. 6, which were acquired with much greater digital resolution (i.e., a maximum evolution time of 15 ms). Without scalar decoupling, therefore, the digital resolution of the 4D spectra is comparable to the analogous 2D spectra in the CA–CA dimensions.

Likewise in the N–CX projections (Fig. 8c, d), nearly all peaks are already resolved without linear prediction, yet the improvements with linear prediction are apparent in several resonances. For example, for T25-A24, A48-D47 and A24-A23 the partial overlap is alleviated, and there is a marked improvement to the separation of K31-F30, D36-N35 and K4-Y3. In the case of the t_3 dimension, linear prediction was performed by processing all four dimensions completely (including t_4 with linear prediction as above), and then performing a Hilbert and inverse Fourier transform of F3, followed by linear prediction, apodization and a forward Fourier transform. This process can in general be repeated for F2.

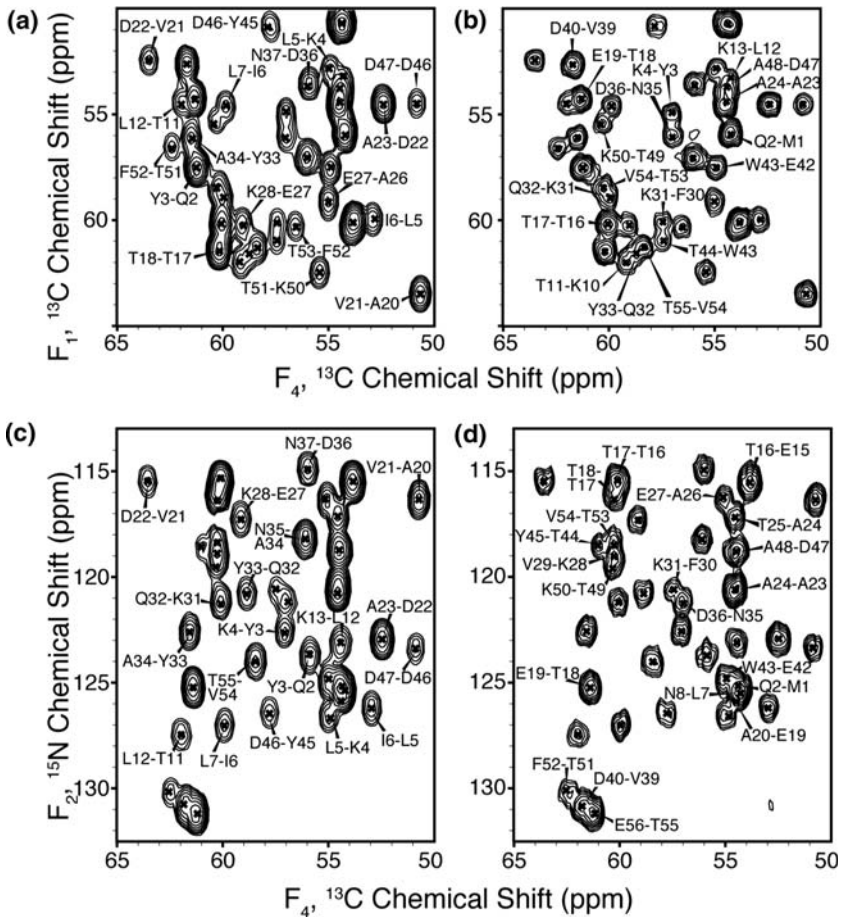

Fig. 8 2D projections illustrating indirect dimension linewidths. (**a**) CA–CA (F1–F4) 2D plane derived from skyline projection of 4D data, processed without linear prediction. (**b**) CA–CA projection derived from 4D processed with linear prediction in the F1 dimension. (**c**) N–CA (F2–F4) 2D plane derived from skyline projection of 4D data, processed without linear prediction. (**d**) N-CA projection derived from 4D processed with linear prediction in the F2 dimension. To avoid truncation artifacts, the spectra in (**a**) and (**c**) were processed with sine bell apodization. Identical sine bell functions were applied to the spectra in (**b**) and (**d**), but after linear prediction

Spectral completeness and sensitivity in GB1

The optimally processed 4D data exhibits excellent resolution, as anticipated from the preceding discussion and simulations. Clearly the 4D experiment is not essential in order to perform GB1 assignments at 500 MHz; yet this sample provides a good test case to evaluate spectral sensitivity. The strip plot in Fig. 9 illustrates some interesting features from several different residues types. All expected CA–CA correlations are observed with 18 ms DARR mixing, with an average SNR of 74 ± 22 throughout the protein. The variation is largely due to differences in the topology of the $i-1$ amino acid, which determines the distribution of polarization during the final homonuclear mixing period. For example, smaller residues have greater CA SNR values; Gly residues have 62% greater (119 ± 15) than the average SNR for CA sites, and Ala residues 8% greater (80 ± 14). As residue size increases, the values steadily decrease, with Thr (68 ± 18) and Val (63 ± 17), with Lys (64 ± 8), Leu and Ile (together 62 ± 2) and the aromatics (among all 3 Tyr, 2 Phe and 1 Trp, 55 ± 8). With the 18 ms DARR mixing time used here, polarization transfer beyond the CG site is minimal, so relatively little deterioration in performance is observed as the number of side-chain ^{13}C sites increases from ~5 to 8 or more. The SNR values all follow sensible trends based on the amino acid size yet are not prohibitively low in any case. Given that the experiment was performed with ~3 μmol of protein in 45.5 h, these final values are in good agreement with the theoretical analysis above, and we have found the values to be reasonable for extrapolation to larger proteins.

In GB1, correlations to all CB [i−1] and most CG[i−1] are also observed at the 18 ms mixing time. Among the CB resonances, the overall average SNR is 41 ± 18, with similar trends with respect to residue size as observed in the CA sites. The Ala residues have CB SNR of 71 ± 14, Thr 45 ± 17 and Val 31 ± 7. The branch point at the CB of Val leads to substantial signal loss, as in 2D ^{13}C–^{13}C spectra of GB1 as we have previously reported (Franks et al. 2005). Furthermore, near rotational resonance (Raleigh et al. 1988) broadening accounts for additional losses in SNR of aromatic residues, particularly Tyr and Phe. The T53-F52 strip in Fig. 9 is drawn with lower contours to illustrate the weak F52CB resonance. Likewise Y33 is among the weakest CB peaks with a SNR value of 9. The overall SNR for aromatic CB resonances is 13 ± 3.

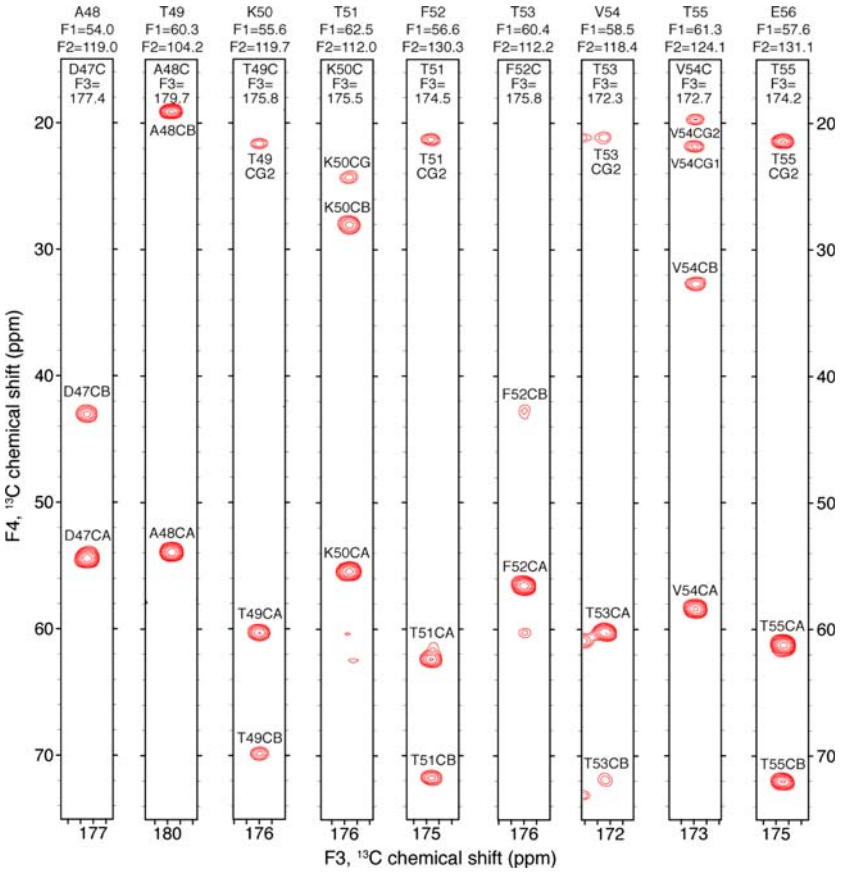

Fig. 9 Strip plots derived from 4D CANCOCX experiment on GB1, illustrating backbone assignments for residues D47–E56. The data set was processed with three points of backwards linear prediction in the direct dimension, apodized with 62.5 Hz (0.5 ppm) net Lorentzian-to-Gaussian broadening and a 72° shifted sine bell, and then zero filled to 2048 points prior to Fourier transformation. The t_3 dimension was apodized with 75 Hz (0.6 ppm) net Lorentzian-to-Gaussian broadening and a 72°-shifted sine bell, zero filled to 128 points, and transformed. The t_2 dimension was apodized with 40 Hz net Lorentzian-to-Gaussian broadening and a 72° phase-shifted sine bell, zero filled to 128 points, and transformed. The t_1 dimension was extended from 32 to 64 points with 12th order linear prediction, apodized with 125 Hz (1.0 ppm) net Lorentzian-to-Gaussian broadening and a 72° phase shifted sine bell, zero filled to 128 points, and transformed. Polynomial baseline correction was applied in the direct dimension only. Contours are drawn for most of the strips starting with the lowest contour at 13 times the RMSD noise level (13 σ), with the exception of T53-F52 (10 σ) and T51-K50 (11 σ), in order to illustrate some of the weaker peaks. Contour spacing is a factor of 1.3

In many instances, weak cross peaks due to interresidue transfer are observed; i.e., CA[i] peaks are observed in the CA[i]–N[i]–C′[i−1] strips drawn at low contour levels, such as T53-F52. In many cases these correlations can be utilized to perform a de novo backbone walk using only the CANCOCX 4D data, although these peaks tend to be observed in this data set only for neighbors of small amino acids (i.e., where the *i*-1 residue is a Gly, Ala, Thr or Val). Overall 32 correlations of the type CA[i]–N[i]–C′[i−1]–CA[i] are observed, with an average SNR of 14 ± 5. Assignment logic based on the 4D data alone is straightforward in such cases, but for larger proteins some type of intraresidue correlation (NCACX or CONCACX) would be required for completeness. Alternatively, one could perform the 4D experiment with longer mixing times, as discussed below, to observe numerous sequential as well as long-range correlations.

Comparison of experimental and theoretical sensitivity

We next compare the experimentally observed sensitivity of the 2D, 3D and 4D experiments through the CANCOCX polarization transfer pathway with the theoretical analysis provided above. Here the SNR values were derived from analysis of the same representative peaks from each spectrum. In the 1D ^{13}C CP-MAS spectrum, we chose an overlapped set of three peaks (N8CA, A20CA and D46CA) at ∼50.7 ppm and determined the average SNR for the CA region. The peak heights and integrated intensities of this region were found to compare well with the major CA region, based on the total number of peaks. Experimental decay constants (T_2) for each nucleus type were determined for the bulk, by 2D Hahn echo experiments. Each spectral region was integrated, and the total integral was then fit to its functional form by multi-variable least

squares fit by assuming appropriate values for homonuclear J-couplings.

Matched linebroadening (or a linebroadening of slightly less than the true linewidth) yielded approximately two-fold increases in the sensitivity; the precise value depended on the exact acquisition time, but this value was kept the same, or data sets truncated for comparison, through this analysis. Overall the agreement between experiment and theory is within the uncertainty of determining the SNR values for the observed set of cross peaks (Table 3). Some variation is observed as a function of processing parameters, most significantly for the direct dimension, which was digitized beyond the average T_2 value for ^{13}C; therefore matched linebroadening yielded a factor of two increase in SNR. In the indirect dimensions, where evolution times were generally shorter than T_2 and therefore relatively little apodization was required to achieve matched linebroadening, the processing details had less impact on the sensitivity values determined. Sine-bell apodization should be used judiciously in the indirect dimensions. This apodization function is generally made to approach zero, which adds an unacceptably large (and usually unrecognized) amount of line broadening especially for data acquired to relatively short t_{max}, Over-broadening a spectrum in this way removes truncation artifacts, but also causes a crucial loss of sensitivity. Therefore one must take care not to add excessive apodization in the indirect dimensions, but rather to perform linear prediction (as described above).

Long-range correlations from 4D CANCOCX

So far most efforts to obtain long-range ^{13}C–^{13}C correlations in uniformly labeled proteins have relied upon high field and/or fractional labeling schemes (Castellani et al. 2002; Zech et al. 2005); one study at relatively low field (400 MHz) utilized 3D experiments in combination with the glycerol pattern labeling scheme (Castellani et al. 2003). Here we demonstrate that sufficient resolution can be obtained from the 4D experiment on GB1 at 500 MHz that unique assignments in many cases of long-range correlations can be achieved. To do so, we performed a second 4D CANCOCX experiment with much longer (225 ms) DARR mixing in order to establish long-range correlations (Fig. 10). The three frequency labels in the indirect dimensions uniquely resolve the origin nucleus of the C′[i−1] residue, using the same polarization transfer strategies described above, prior to the 225 ms CO–CX mixing period. We identified a total of 570 correlations in the aliphatic region of F4 in this data set, including 66 long-range ($|i-j| > 4$), 113 medium-range ($5 > |i-j| > 1$), 215 sequential ($i = j \pm 1$) and 178 intraresidue peaks.

Table 3 Experimental and theoretical sensitivity of nD correlation spectra of GB1

Experiment	Experiment Time (h)	Processing (direct, indirect) (Hz)	SNR[a]	SNR/$\sqrt{}$time	Normalized to 1-D[b](%)	Theoretical[c]
C1D	0.017	0	44	338	100	100%
		50	96	740	100	
NCA 2D	0.9	0, 0	73 ± 18	77	23 ± 5	24.7%
		45, 0	150 ± 39	158	21 ± 5	
		45, 7.5	159 ± 39	167	23 ± 5	
NCO 2D	0.9	0, 0	100 ± 18	105	31 ± 5	29.6%
		40, 0	202 ± 26	213	29 ± 5	
		40, 10	210 ± 22	221	30 ± 5	
CA(NCO)CA 2D	2.4	0, 0	11 ± 5	7.1	2.1 ± 1.0	2.5%
		45 , 0	20 ± 9	13	1.8 ± 1.0	
		45, 45	31 ± 14	20	2.7 ± 1.0	
CAN(CO)CA 3D	13.7	40, 0 , 0	37 ± 12	10	1.4 ± 0.4	1.7%
		40, 7, 40	36 ± 12	10	1.4 ± 0.4	
CANCOCA 4D	45.5	62.5 (t_4), 75 (t_3), 40 (t_2), 125 (t_1)	68 ± 14	10.1	1.4 ± 0.4	1.0%

[a] Average values taken from a representative set of well resonances in each experiment

[b] The peak consisting of N8, A20, and D46 CA (which are within 0.1 ppm) was used as the 1D standard. The root mean squared noise was calculated in Spinsight 4.3.2 (Varian) using 10% of the base-line points, and the reported signal-to-noise ratio divided by three. Separate values with and without apodization of the ^{13}C dimension are used to compare the sensitivity of the nD spectra with and without apodization of the direct ^{13}C dimension. In addition, in some cases the indirect dimensions were apodized to illustrate the (relatively small) effect of apodization, due to the fact that the indirect dimensions demonstrated relatively little decay

[c] Constants were determined on average for GB1 as $\varepsilon^{C-C} = 20\%$; $\varepsilon^{N-CA} = 50\%$; $\varepsilon^{N-C'} = 60\%$; $T_2^N = 40$ ms; $T_2^{CA} = 10$ ms; $T_2^{C'} = 20$ ms; $J^{CA-C'} = 55$ Hz; $J^{CA-CB} = 35$ Hz. Equations 4–7 were used to determine the theoretical sensitivity per unit time

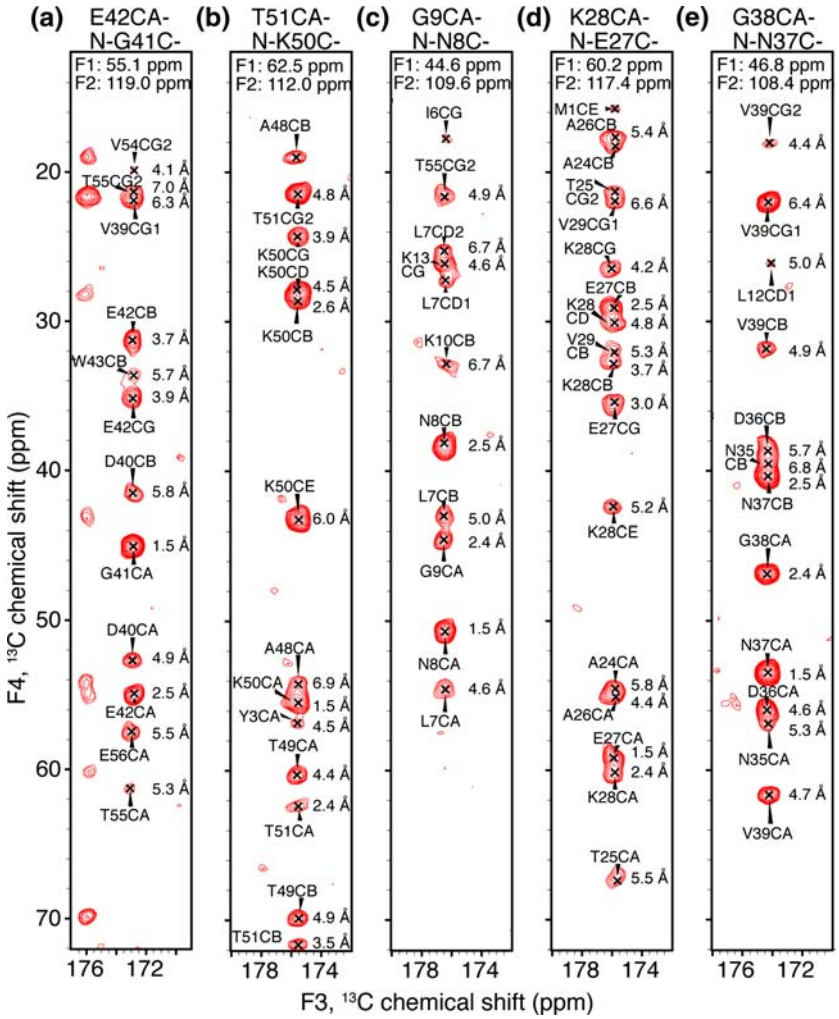

Fig. 10 Strip plot derived from the 4D CANCOCX experiment on GB1 with final 225 ms DARR mixing time. Other parameters are identical to the experiment of Fig. 9. The experiment was repeated twice to observe weaker peaks (total measurement time 103 h) which in many cases correspond to distances of greater than 5 Å. The data set was processed with three points of backwards linear prediction in the direct dimension, apodized with 55 Hz (0.45 ppm) net Lorentzian-to-Gaussian broadening and then zero filled to 2048 points prior to Fourier transformation. The t_3 dimension was apodized with 60 Hz (0.5 ppm) net Lorentzian-to-Gaussian broadening zero filled to 256 points, and transformed. The t_2 dimension was apodized with 30 Hz net Lorentzian-to-Gaussian broadening zero filled to 128 points, and transformed. The t_1 dimension was apodized with 55 Hz (0.45 ppm) net Lorentzian-to-Gaussian broadening zero filled to 128 points, and transformed. Polynomial baseline correction was applied in the direct dimension and in t_1

For example, from G41 (Fig. 10a) correlations are observed to T55CG2 (7.0 Å), T55CA (5.3 Å) and E56CA (5.5 Å). From K50 (Fig. 10b), correlations include Y3CA (4.5 Å) and A48CA (6.9 Å), V54, and E56. From N8 (Fig. 10c), correlations to L12, K13 and T55 are observed. Beyond the value of long-range correlations for potential structure determination, the spectrum confirms weak side-chain correlations in large amino acids such as Lys, which are missing CG, CD and CE correlations in the shorter mixing time data but observed in Fig. 10b and d for K50 and K28 respectively. The latter set of peaks arise from the E27C′, and further illustrate how the data can be utilized to confirm sequential assignments, given a semi-quantitative agreement between peak intensity and proximity to the C′ site. This efficient transfer leads to observation of distances on the order of 7 Å, presumably by multi-step transfers along the side chains, whereas direct, through-space transfers are rarely observed for distances >5 Å (on the basis of comparison with the crystal structure 2GI9 (Franks et al. 2006)).

We observed several instances where the correlations along the backbone had greater intensity than long-range contacts with similar, or even shorter, distances. Polarization transfer was particularly fast in the helical residues of GB1 (e.g., E27 in Fig. 10d), where the strip plots routinely showed i to $i \pm 1$ and $i \pm 2$ correlations, with several showing $i \pm 3$ correlations; e.g., the helical residues E27 and N37 both show strong contacts of this type (Fig. 10d, e). In contrast, in β-sheets the range is limited to two residues. These trends are qualitatively consistent with the relative distances, although the agreement between peak intensity and the inverse cube (or sixth power) of the distance is not at all quantitative. Relaxation and relative orientation of dipole tensors are two effects that would complicate such interpretation. For example, in some cases the relative intensities are qualitatively consistent with the favorable population of crystallites in which the CA–N and N–C′ dipolar couplings would coincidentally have large values. In these crystallites, the projection of the tensors along the peptide backbone would be favored, and thus

enhance transfer to neighboring residues more so than side chains or long-range correlations orthogonal to this director. As previous studies have shown, use of correlation information of this type in a semi-quantitative manner may permit de novo structural folds, although we reserve the detailed analysis of such spectra for future work.

Conclusion

In summary, we have introduced a 4D pulse sequence to establish backbone and side-chain correlations in solid proteins, based on combinations of band-selective heteronuclear polarization transfers with a final broadband homonuclear mixing period. In order to implement this experiment successfully, several fundamental and practical concerns were addressed. Instrumental stability was determined to be especially critical. Whereas the optimal performance of adiabatic transfer schemes exceeds the theoretical efficiency of other experiments, to achieve this level of performance for extended periods of time requires B_1 fields that are stable to within better than 0.5% for tens of hours to several days, depending on the sample. Therefore active feedback control of amplifier power levels, in analogy to MAS rate controllers and B_0 field lock circuits, is likely to be an essential ingredient to the routine application of this type of experiment in the future.

Using modern instrumentation in an environmentally well-controlled room (with less than $\pm 0.5^{o}$C long-term temperature fluctuations), we were able to demonstrate the experiment on the model protein GB1 in this study. The experimental data show cross peaks between CA sites of neighboring residues with high resolution, enabling facile confirmation of backbone and side-chain assignments in GB1 based on the primary polarization transfer pathway CA[i]–N[i]-C′[i−1]–CX[i−1]. In more than half of the residues in the protein, it was possible to perform sequential assignments by analysis of weak correlations in the spectra (e.g., CA[i]–N[i]–C′[i−1]–CA[i]). For greater generality, combination of the CANCOCX 4D experiment with 3D NCACX or 4D CONCACX is expected to provide greater resolution for larger proteins. Beyond the use for backbone assignments, we have also determined that many unique long-range distance restraints can be obtained with a ^{13}C–^{13}C mixing time of 225 ms, despite the uniform ^{13}C labeling at all sites in the protein. The experiment as currently implemented results in cross peaks with a SNR of $\sim$100 in 48 h with 3 μmol of protein, and weaker peaks with SNR of $\sim$10 corresponding to >5 Å distances from experiments conducted over a four day period on the same protein. The sensitivity observed agrees well with anticipations based on theoretical analysis and measured T_2 values.

We can envision several further improvements to the scheme shown here. Constant-time evolution of the indirect chemical shift dimensions is likely to improve resolution further, by virtue of elimination of the scalar couplings; combination with selective pulse schemes is appropriate in this context. The current limitation is the requirement of a four-step phase cycle, which would need to be revised and/or extended to accommodate selective pulses with proper coherence selection. Use of gradients is likely to assist in such efforts. More efficient digitization of indirect chemical shift dimensions is expected to enhance the experimentally acquired resolution.

The 4D correlation experiments are expected to be vital for the assignment of large proteins, proteins with high sequence degeneracy, and those samples with a large linewidths. We anticipate the application of this and similar 4D pulse sequences for efficient assignment and extraction of distance information of solid protein spectra. Studies of several larger proteins (from 15 to $\sim$150 kDa) are underway in our laboratory and will be reported in future publications. Although a primary consideration in attempting to assign protein spectra is to prepare samples with the greatest possible microscopic homogeneity, in order to minimize linewidths and resolve as many peaks as possible in 2D and 3D spectra, we expect that higher dimensionality in general will be beneficial for larger systems, even with linewidths of less than 1 ppm. Indeed, to assign such large proteins will require *both* optimal sample preparation and 4D methods.

Acknowledgments We thank G. Shah and H. Frericks for preparing the GB1 samples used in this study. This work was supported by the University of Illinois, National Science Foundation (CAREER Award MCB0347824 to CMR), Research Corporation (Cottrell Scholars Awards to CMR) and National Institutes of Health (GM073770 to CMR).

References

Aggarwal K, Delsuc M (1997) Triangular sampling of multidimensional NMR data sets. Magn Reson Chem 35:593–596

Baldus M, Geurts DG, Hediger S, Meier BH (1996) Efficient 15N-13C polarization transfer by adiabatic-passage Hartmann-Hahn cross polarization. J Magn Reson A 118:140–144

Baldus M, Petkova AT, Herzfeld JH, Griffin RG (1998) Cross polarization in the tilted frame: Assignment and spectral simplification in heteronuclear spin systems. Mol Phys 95:1197–1207

Barkhuijsen H, Debeer R, Vanormondt D (1987) Improved algorithm for noniterative time-domain model-fitting to exponentially damped magnetic-resonance signals. J Magn Reson 73:553–557

Bennett AE, Ok JH, Griffin RG, Vega S (1992) Chemical shift correlation spectroscopy in rotating solids: Radio-frequency dipolar recoupling and longitudinal exchange. J Chem Phys 96:8624–8627

Bennett AE, Rienstra CM, Auger M, Lakshmi KV, Griffin RG (1995) Heteronuclear decoupling in rotating solids. J Chem Phys 103:6951–6958

Bennett AE, Rienstra CM, Griffiths JM, Zhen W, Lansbury PT Jr, Griffin RG (1998) Homonuclear radio frequency-driven recoupling in rotating solids. J Chem Phys 108:9463–9479

Bloembergen N (1949) On the interaction of nuclear spins in a crystalline lattice. Physica 15:386

Bockmann A, Lange A, Galinier A, Luca S, Giraud N, Juy M, Heise H, Montserret R, Penin F, Baldus M (2003) Solid state NMR sequential resonance assignments and conformational analysis of the 2×10.4 kDa dimeric form of the bacillus subtilis protein crh. J Biomol NMR 27:323–339

Carravetta M, Eden M, Zhao X, Brinkmann A, Levitt MH (2000) Symmetry principles for the design of radiofrequency pulse sequences in the nuclear magnetic resonance of rotating solids. Chem Phys Lett 321:205–215

Carravetta M, Zhao X, Johannessen OG, Lai WC, Verhoeven MA, Bovee-Geurts PHM, Verdegem PJE, Kiihne S, Luthman H, De Groot HJM, DeGrip WJ, Lugtenburg J, Levitt MH (2004) Protein-induced bonding perturbation of the rhodopsin chromophore detected by double-quantum solid-state NMR. J Am Chem Soc 126:3948–3953

Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature 420:98–102

Castellani F, van Rossum BJ, Diehl A, Rehbein K, Oschkinat H (2003) Determination of solid-state NMR structures of proteins by means of three-dimensional 15N-13C-13C dipolar correlation spectroscopy and chemical shift analysis. Biochemistry 42:11476–11483

Cole HBR, Torchia DA (1991) An NMR-study of the backbone dynamics of staphylococcal nuclease in the crystalline state. Chem Phys 158:271–281

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13:289–302

Costa PR, Sun BQ, Griffin RG (1997) Rotational resonance tickling: Accurate internuclear distance measurements in solids. J Am Chem Soc 119:10821–10830

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRpipe - a multidimensional spectral processing system based on unix pipes. J Biomol NMR 6:277–293

Detken A, Hardy EH, Ernst M, Kainosho M, Kawakami T, Aimoto S, Meier BH (2001) Methods for sequential resonance assignment in solid, uniformly C-13, N-15 labelled peptides: Quantification and application to antamanide. J Biomol NMR 20:203–221

Etzkorn M, Martell S, Andronesi OC, Seidel K, Engelhard M, Baldus M (2007) Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. Angew Chem Int Ed 46:459–462

Flynn PF, Milton MJ, Babu CR, Wand AJ (2002) A simple and effective NMR cell for studies of encapsulated proteins dissolved in low viscosity solvents. J Biomol NMR 23:311–316

Franks WT, Wylie BJ, Stellfox SA, Rienstra CM (2006) Backbone conformational constraints in a microcrystalline U-15N-labeled protein by 3D dipolar-shift solid-state NMR spectroscopy. J Am Chem Soc 128:9992–9993

Franks WT, Zhou DH, Wylie BJ, Money BG, Graesser DT, Frericks HL, Sahota G, Rienstra CM (2005) Magic-angle spinning solid-state NMR spectroscopy of the beta 1 immunoglobulin binding domain of protein G (GB1): N-15 and C-13 chemical shift assignments and conformational analysis. J Am Chem Soc 127:12291–12305

Frericks HL, Zhou DH, Yap L, Gennis RB, Rienstra CM (2006) Magic-angle spinning solid-state NMR of a 144 kda membrane protein complex: E. coli cytochrome bo3 oxidase. J Biomol NMR 36:55–71

Fung BM, Khitrin AK, Ermolaev K (2000) An improved broadband decoupling sequence for liquid crystals and solids. J Magn Reson 142:97–101

Gardner KH, Zhang XC, Gehring K, Kay LE (1998) Solution NMR studies of a 42 kda escherichia coli maltose binding protein beta-cyclodextrin complex: Chemical shift assignments and analysis. J Am Chem Soc 120:11738–11748

Goddard TD, Kneller DG (2006) Sparky 2.116. University of California, San Francisco

Havlin RH, Tycko R (2005) Probing site-specific conformational distributions in protein folding with solid-state NMR. Proc Natl Acad Sci USA 102:3284–3289

Hediger S, Meier BH, Ernst RR (1995) Adiabatic passage Hartmann-Hahn cross-polarization in NMR under magic-angle sample-spinning. Chem Phys Lett 240:449–456

Heise H, Hoyer W, Becker S, Andronesi OC, Riedel D, Baldus M (2005) Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. Proc Natl Acad Sci USA 102:15871–15876

Hiller M, Krabben L, Vinothkumar KR, Castellani F, Van Rossum B-J, Kühlbrandt W, Oschkinat H (2005) Solid-state magic-angle spinning NMR of outer-membrane protein g from *Escherichia coli*. ChemBioChem 6:1679–1684

Hing A, Vega S, Schaefer J (1992) Transferred-echo double-resonance NMR. J Magn Reson 96:205–209

Hoch JC, Stern AS (2001) Maximum entropy reconstruction, spectrum analysis and deconvolution in multidimensional nuclear magnetic resonance, Meth Enzym 338:159–178

Hohwy M, Jakobsen HJ, Edén M, Levitt MH, Nielsen NC (1998) Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. J Chem Phys 108:2686–94

Hohwy M, Rienstra CM, Jaroniec CP, Griffin RG (1999) Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy, J Chem Phys 110:7983–7992

Hong M (1999) Resonance assignment of C-13/N-15 labeled solid proteins by two- and three-dimensional magic-angle-spinning NMR. J Biomol NMR 15:1–14

Hu JG, Sun BQ, Griffin RG, Herzfeld J (1995) Solid state NMR detection of multiple m-intermediates in bacteriorhodopsin. Biophys J 68:A332

Igumenova TI, McDermott AE (2005) Homo-nuclear C-13 J-decoupling in uniformly C-13-enriched solid proteins. J Magn Reson 175:11–20

Igumenova TI, McDermott AE, Zilm KW, Martin RW, Paulson EK, Wand AJ (2004a) Assignments of carbon NMR resonances for microcrystalline ubiquitin. J Am Chem Soc 126:6720–6727

Igumenova TI, Wand AJ, McDermott AE (2004b) Assignment of the backbone resonances for microcrystalline ubiquitin. J Am Chem Soc 126:5323–5331

Ishii Y (2001) C-13-C-13 dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. J Chem Phys 114:8473–8483

Ishii Y, Ashida J, Terao T (1995) 13C-1H dipolar recoupling dynamics in 13C multiple-pulse solid-state NMR. Chem Phys Lett 246:439–445

Jovanovic T, Farid R, Friesner RA, McDermott AE (2005) Thermal equilibrium of high- and low-spin forms of cytochrome p450 bm-3: Repositioning of the substrate? J Am Chem Soc 127:13548–13552

Kim S, Szyperski T (2003) GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. J Am Chem Soc 125:1385–1393

Kloepper KD, Woods WS, Winter KA, George JM, Rienstra CM (2006) Preparation of alpha-synuclein fbrils for solid-state NMR: Expression, purification and incubation of wild type and mutant forms. Protein Expr Purif 48:112–117

Kumar A, Ernst RR, Wüthrich K (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem Biophys Res Commun 95:1–6

Kupce E, Freeman R (2004) Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. J Am Chem Soc 126:6429–6440

Lange A, Becker S, Seidel K, Giller K, Pongs O, Baldus M (2005) A concept for rapid protein-structure determination by solid-state NMR spectroscopy. Angew Chem Int Ed 44:2089–2092

Lange A, Giller K, Hornig S, Martin-Eauclaire M-F, Pongs O, Becker S, Baldus M (2006) Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. Nature 440:959–962

Lee YK, Kurur ND, Helmle M, Johannessen OG, Nielsen NC, Levitt MH (1995) Efficient dipolar recoupling in the NMR of rotating solids - a sevenfold symmetrical radiofrequency pulse sequence. Chem Phys Lett 242:304–309

Li Y, Berthold DA, Frericks HL, Gennis RB, Rienstra CM (2007) Partial 13C and 15N chemical shift assignments of the disulfide bond forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy. ChemBioChem 8:434–442

Li Y, Wylie BJ, Rienstra CM (2006) Selective refocusing pulses in magic-angle spinning NMR: Characterization and applications to multidimensional protein spectroscopy. J Magn Reson 179:206–216

Marion D, Wüthrich K (1983) Application of phase sensitive two-dimensional correlated spectroscopy (cosy) for measurements of H-1-H-1 spin-spin coupling-constants in proteins. Biochem Biophys Res Commun 113:967–974

Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich K (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids – (IUPAC Recommendations 1998). Pure Appl Chem 70:117–142

Martin RW, Paulson EK, Zilm KW (2003) Design of a triple resonance magic angle sample spinning probe for high field solid state nuclear magnetic resonance. Rev Sci Instrum 74:3045–3061

Martin RW, Zilm KW (2003) Preparation of protein nanocrystals and their characterization by solid state NMR. J Magn Reson 165:162–174

Marulanda D, Tasayco ML, Cataldi M, Arriaran V, Polenova T (2005) Resonance assignments and secondary structure analysis of e. Coli thioredoxin by magic angle spinning solid-state NMR spectroscopy. J Phys Chem B 109:18135–18145

Marulanda D, Tasayco ML, McDermott A, Cataldi M, Arriaran V, Polenova T (2004) Magic angle spinning solid-state NMR spectroscopy for structural studies of protein interfaces. Resonance assignments of differentially enriched escherichia coli thioredoxin reassembled by fragment complementation. J Am Chem Soc 126:16608–16620

McDermott A, Polenova T, Bockmann A, Zilm KW, Paulsen EK, Martin RW, Montelione GT (2000) Partial NMR assignments for uniformly (C-13, N-15)-enriched BPTI in the solid state. J Biomol NMR 16:209–219

McDowell LM, Klug CA, Beusen DD, Schaefer J (1996a) Ligand geometry of the ternary complex of 5-enolpyruvylshikimate-3-phosphate synthase from rotational-echo double-resonance NMR. Biochemistry 35:5395–5403

McDowell LM, Lee MS, McKay RA, Anderson KS, Schaefer J (1996b) Intersubunit communication in tryptophan synthase by carbon-13 and flourine-19 REDOR NMR. Biochemistry 35:3328–3334

McDowell LM, Poliks B, Studelska DR, O'Connor RD, Beusen DD, Schaefer J (2004) Rotational-echo double-resonance NMR-restrained model of the ternary complex of 5-enolpyruvylshikimate-3-phosphate synthase. J Biomol NMR 28:11–29

Morcombe CR, Gaponenko V, Byrd RA, Zilm KW (2004) Diluting abundant spins by isotope edited radio frequency field assisted diffusion. J Am Chem Soc 126:7196–7197

Morcombe CR, Zilm KW (2003) Chemical shift referencing in mas solid state NMR. J Magn Reson 162:479–486

Pauli J, Baldus M, van Rossum B, de Groot H, Oschkinat H (2001) Backbone and side-chain 13C and 15N resonance assignments of the alpha-spectrin SH3 domain by magic angle spinning solid state NMR at 17.6 tesla. ChemBioChem 2:272–281

Pauli J, van Rossum B, Forster H, de Groot HJM, Oschkinat H (2000) Sample optimization and identification of signal patterns of amino acid side chains in 2D RFDR spectra of the alpha-spectrin SH3 domain. J Magn Reson 143:411–416

Paulson EK, Martin RW, Zilm KW (2004) Cross polarization, radio frequency field homogeneity, and circuit balancing in high field solid state NMR probes. J Magn Reson 171:314–323

Petkova AT, Leapman RD, Guo ZH, Yau WM, Mattson MP, Tycko R (2005) Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. Science 307:262–265

Petkova AT, Leapman RD, Yau WM, Tycko R (2004) Structural investigations of alzheimer's beta-amyloid fibrils by solid state NMR. Biophys J 86:506A-506A

Pintacuda G, Giraud N, Pierattelli R, Böckmann A, Bertini I, Emsley L (2006) Solid-state NMR spectroscopy of a paramagnetic protein: assignment and study of human dimeric oxidized Cu(II)-Zn(II) superoxide dismutase (SOD). Angew Chem Int Ed 46:1079–1082

Raleigh DP, Levitt MH, Griffin RG (1988) Rotational resonance in solid state NMR. Chem Phys Lett 146:71–76

Ravindranathan KP, Gallicchio E, McDermott AE, Levy RM (2007) Conformational dynamics of substrate in the active site of cytochrome p450 bm-3/npg complex: Insights from NMR order parameters. J Am Chem Soc 129:474–475

Rienstra CM, Hohwy M, Hong M, Griffin RG (2000) 2D and 3D N-15-C-13-C-13 NMR chemical shift correlation spectroscopy of solids: Assignment of mas spectra of peptides. J Am Chem Soc 122:10979–10990

Ritter C, Maddelein M-L, Siemer AB, Luhrs T, Ernst M, Meier BH, Saupe SJ, Riek R (2005) Correlation of structural elements and infectivity of the HET-s prion. Nature 435:844–848

Schaefer J, Stejskal EO (1979) Double cross polarization NMR of solids. J Magn Reson 34:443–447

Seidel K, Lange A, Becker S, Hughes CE, Heise H, Baldus M (2004) Protein solid-state NMR resonance assignments from (C-13, C-13) correlation spectroscopy. Phys Chem Chem Phys 6:5090–5093

Siemar AB, Ritter C, Ernst M, Riek R, Meier BH (2005) High-resolution solid-state NMR spectroscopy of the prion protein HET-s in its amyloid conformation. Angew Chem Int Ed 44:2441–2444

Siemer AB, Ritter C, Steinmetz MO, Ernst M, Riek R, Meier BH (2006) 13C, 15N resonance assignment of the het-s prion protein in its amyloid form. J Biomol NMR 34:75–87

Sonnenberg L, Luca S, Baldus M (2004) Multiple-spin analysis of chemical-shift-selective (C-13, C-13) transfer in uniformly labeled biomolecules. J Magn Reson 166:100–110

Stringer JA, Bronnimann CE, Mullen CG, Zhou DHH, Stellfox SA, Li Y, Williams EH, Rienstra CM (2005) Reduction of rf-induced sample heating with a scroll coil resonator structure for solid-state NMR probes. J Magn Reson 173:40–48

Stueber D, Mehta AK, Chen ZY, Wooley KL, Schaefer J (2006) Local order in polycarbonate glasses by C-13{F-19} rotational-echo double-resonance NMR. J Polym Sci Part B: Polym Phys 44:2760–2775

Suter D, Ernst RR (1985) Spin diffusion in resolved solid state NMR spectra. Phys Rev B 32:5608–5627

Szyperski T, Yeh D, Sukumaran D, Moseley H, Montelione G (2002) Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. Proc Natl Acad Sci USA 99:8009–8014

Takegoshi K, Nakamura S, Terao T (2003) C-13-H-1 dipolar-driven C-13-C-13 recoupling without C-13 rf irradiation in nuclear magnetic resonance of rotating solids. J Chem Phys 118:2325–2341

Takegoshi K, Nomura K, Terao T (1995) Rotational resonance in the tilted rotating frame. Chem Phys Lett 232:424–428

Tugarinov V, Muhandiram R, Ayed A, Kay L (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: Chemical shift assignments and secondary structure of malate synthase g. J Am Chem Soc 124:10025–10035

Tycko R (1996) Prospects for resonance assignments in multidimensional solid-state NMR spectra of uniformly labeled proteins. J Biomol NMR 8:239–251

Tycko R (2004) Progress towards a molecular-level structural understanding of amyloid fibrils. Curr Opin Struct Biol 14:96–103

Van Geet AL (1968) Calibration of the methanol and glycol nuclear magnetic resonance thermometers with a static thermistor probe. Anal Chem 40:2227–2229

Wylie BJ, Franks WT, Rienstra CM (2006) Determinations of 15N chemical shift anisotropy magnitudes in a uniformly 15N,13C-labeled microcrystalline proteins by three-dimensional magic-angle spinning NMR spectroscopy. J Phys Chem B 110:10926–10936

Zech SG, Wand AJ, McDermott AE (2005) Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin. J Am Chem Soc 127:8618–8626