J Biomol NMR (2007) 38:341–351
DOI 10.1007/s10858-007-9173-7

ARTICLE

HIFI-C: a robust and fast method for determining NMR couplings from adaptive 3D to 2D projections

Gabriel Cornilescu · Arash Bahrami · Marco Tonelli · John L. Markley · Hamid R. Eghbalnia

Received: 9 May 2007 / Accepted: 13 June 2007 / Published online: 4 July 2007

Abstract We describe a novel method for the robust, rapid, and reliable determination of J couplings in multidimensional NMR coupling data, including small couplings from larger proteins. The method, "High-resolution Iterative Frequency Identification of Couplings" (HIFI-C) is an extension of the adaptive and intelligent data collection approach introduced earlier in HIFI-NMR. HIFI-C collects one or more optimally tilted two-dimensional (2D) planes of a 3D experiment, identifies peaks, and determines couplings with high resolution and precision. The HIFI-C approach, demonstrated here for the 3D quantitative J method, offers vital features that advance the goal of rapid and robust collection of NMR coupling data. (1) Tilted plane residual dipolar couplings (RDC) data are collected adaptively in order to offer an intelligent trade off between data collection time and accuracy. (2) Data from independent planes can provide a statistical measure of reliability for each measured coupling. (3) Fast data collection enables measurements in cases where sample stability is a limiting factor (for example in the presence of an orienting medium required for residual dipolar coupling measurements). (4) For samples that are stable, or in experiments involving relatively stronger couplings, robust data collection enables more reliable determinations of couplings in shorter time, particularly for larger biomolecules. As a proof of principle, we have applied the HIFI-C approach to the 3D quantitative J experiment to determine N-C′ RDC values for three proteins ranging from 56 to 159 residues (including a homodimer with 111 residues in each subunit). A number of factors influence the robustness and speed of data collection. These factors include the size of the protein, the experimental set up, and the coupling being measured, among others. To exhibit a lower bound on robustness and the potential for time saving, the measurement of dipolar couplings for the N-C′ vector represents a realistic "worst case analysis". These couplings are among the smallest currently measured, and their determination in both isotropic and anisotropic media demands the highest measurement precision. The new approach yielded excellent quantitative agreement with values determined independently by the conventional 3D quantitative J NMR method (in cases where sample stability in oriented media permitted these measurements) but with a factor of 2–5 in time savings. The statistical measure of reliability, measuring the quality of each RDC value, offers valuable adjunct information even in cases where modest time savings may be realized.

Keywords HIFI-C · High-resolution Iterative Frequency Identification of Couplings · Fast NMR data collection ·

Electronic supplementary material The online version of this article (doi:10.1007/s10858-007-9173-7) contains supplementary material, which is available to authorized users.

G. Cornilescu (✉) · A. Bahrami · M. Tonelli · J. L. Markley · H. R. Eghbalnia (✉)
National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, Madison, WI 53706, USA
e-mail: gabrielc@nmrfam.wisc.edu

A. Bahrami · J. L. Markley · H. R. Eghbalnia
Center for Eukaryotic Structural Genomics, University of Wisconsin-Madison, Madison, WI 53706, USA

A. Bahrami · J. L. Markley · H. R. Eghbalnia
Department of Biochemistry, Graduate Program in Biophysics, University of Wisconsin-Madison, Madison, WI 53706, USA

H. R. Eghbalnia
Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA
e-mail: eghbalni@nmrfam.wisc.edu

NMR coupling measurements · RDC · Residual dipolar couplings · Adaptive reduced-dimensionality

Introduction

The accurate determination of small couplings is frequently of importance in biomolecular NMR spectroscopy. Applications include the collection of dihedral angle constraints, the identification and quantification of *trans* hydrogen bond couplings (Cordier and Grzesiek 1999; Cornilescu et al. 1999a, b) and the measurement of residual dipolar couplings (RDCs) (Tjandra and Bax 1997; Tolman et al. 1995). Various approaches developed for acquiring data for coupling measurements include S3E (Meissner et al. 1997), IPAP (Ottiger et al. 1998a), J-modulated (Ottiger et al. 1998b) and quantitative J experiments (Chou et al. 2000). The 3D quantitative J experiment (Chou et al. 2000) was designed specifically for the purpose of measuring the many couplings of interest in proteins and nucleic acids that are significantly smaller than the line widths of the coupled peaks. Such couplings could not be determined with an acceptable degree of precision by earlier 2D methods. Later 2D methods have addressed the resolution demand of ECOSY type of coupling measurements (Permi et al. 1999; Kover and Batta 2004; Hoshino and Otting 2004), but for proteins with extensive overlap in these 2D spectra the 3D quantitative J method is preferable. This experiment utilizes 3D data collection to eliminate peak overlap and, as a consequence, requires lengthy data acquisition times. To circumvent this problem, we have developed a novel reduced dimensionality (RD) experiment that replaces full 3D data collection of a quantitative J experiment by collection of one or more 2D planes at judiciously chosen tilt angles. The desired couplings can be extracted from the resulting planes with great precision.

The extraction of accurate coupling values from collected data, either 2D or 3D, depends on a number of factors. The robustness of computed values depends on the signal-to-noise ratio (S/N) for the data, the degree of signal overlap, and the numerical procedures used in the computational steps. Overlap can be resolved by collecting higher dimensional data. However, with increasing dimensions, problems such as truncation artifacts lead to a progressive trade-off between resolution and accurate line shape. Thus, when peak overlap can be avoided, lower dimensional data with higher resolution may be preferable for reasons beyond time savings. Long acquisition periods may be necessary for measuring small couplings, particularly with larger biomolecules or more dilute samples. Thus there is a need for a method for determining couplings rapidly and with a measure of confidence in the values obtained.

Our approach to solving this problem is to use reduced dimensionality (RD) data collection (Szyperski et al. 2002). RD has the potential of alleviating some of the difficulties of nD NMR by combining information from different evolution periods into a single dimension. One RD approach is to create ^{15}N–^{13}C double- and zero-quantum coherence in a single evolution time (Szyperski et al. 1993). Similarly, 2D versions of HNCA and HNCO triple resonance experiments, called MQ-HNCA and MQ-HNCO have been developed (Simorre et al. 1994). Another RD approach involves the simultaneous evolution of chemical shifts from two or more nuclei as single-quantum coherences in a single indirect domain. For example, 3D experiments can be recorded as 2D planes in which the two indirect chemical shifts are encoded in the second dimension. Although losses resulting from additional polarization transfers during the evolution periods cannot be avoided in RD experiments (Sattler et al. 1999), resolution and sensitivity gains can be realized.

A number of promising RD techniques have been described recently, including the G-matrix Fourier Transform (GFT) (Kim and Szyperski 2003) and Time-Proportional Phase Incrementation (TPPI) (Ding and Gronenborn 2002) methods. The strength of RD methods lies in the possibility of both reducing the collection time of high-dimensional spectra and increasing their digital resolution. In HIFI-NMR, we combined a tilted plane approach (Kupce and Freeman 2003) with automated peak recognition and intelligent and adaptive selection of mixture ratios to minimize overlap and to avoid the need for reconstructing spectra (Eghbalnia et al. 2005).

By building on the methods proposed in HIFI-NMR, we have developed a novel method (HIFI-C) for the fast collection and analysis of coupling data. The underlying principle behind HIFI-C is to combine the high digital resolution provided by 2D spectra with the ability of tilted-plane data collection to separate overlapped peaks. The quantitative J method, which encodes the coupling in peak amplitudes and thereby avoids the additional signal overlaps present in spectra with chemical shift encoded couplings, is readily adapted to this approach. HIFI-C adaptively chooses the optimal angle for a tilted 2D plane from a 3D quantitative J experiment that minimizes peak overlaps and provides the needed coupling information. If one or more additional planes are required, to resolve remaining peak overlaps, these are chosen adaptively and iteratively. Because the data for each plane are collected and processed independently, results from multiple planes can be used to ascertain the reliability and robustness of individual RDC values. The time savings result from automated termination of data collection at the point where additional acquisitions would not improve the analysis.

Analysis of HIFI-C data for extraction of J-couplings and RDC values has been fully automated.

To illustrate the applicability of the HIFI-C approach, we demonstrate its use in determining RDC values for the N–C′ vector. This is one of the smallest dipolar couplings commonly measured, and its determination requires the highest measurement precision; it thus represents a "worst case" application of the new method. In addition, NC′ is one of the three types of RDCs needed for accurate alignment tensor calculation in the EHM (Extended Histogram Method) developed by Ad Bax (Bryce and Bax 2004). EHM uses the geometry and symmetry of each peptide plane defined by the NH, C′CA and NC′ vectors to compute a multitude of other internuclear vectors (including "synthetic" RDCs) to obtain a much better sampled RDC histogram that yields a considerably more accurate alignment tensor. We compare the $J_{NC'}$ values determined by HIFI-C in isotropic and anisotropic media and the resulting RDCs with those determined by conventional 3D methods, which require 2–5 times longer data acquisition times.

After our studies were completed, another RD approach to fast collection of coupling data, J-GFT, appeared (Atreya et al. 2007). We discuss here the relative advantages and disadvantages of J-GFT and HIFI-C.

Materials and methods

Proteins samples

Three labeled proteins were used in this study. [^{2}H,^{13}C,^{15}N]-GB3 (third IgG-binding domain of protein G) (Gallagher et al. 1994) was a gift from Ad Bax. The 155-residue [^{13}C,^{15}N]-Prp24-12 protein (Reiter et al. 2006) was a gift from Samuel Butcher. Preparation of the 111-residue dimeric protein [^{13}C,^{15}N]-At5g22580.1 has been described previously (Cornilescu et al. 2004).

The GB3 sample in isotropic medium contained ~2 mM [^{2}H,^{13}C,^{15}N]-GB3 in 93% H_2O, 7% D_2O, 50 mM sodium phosphate buffer, pH 6.5, and 0.15 mM NaN_3. The GB3 sample in anisotropic medium contained ~1 mM protein sample under the same solution conditions but with 10 mg/ml filamentous pf1 phage (Hansen et al. 1998).

The Prp24-12 sample in isotropic medium contained ~0.6 mM [^{13}C,^{15}N]-Prp24-12 in 93% H_2O, 7% D_2O, 50 mM potassium phosphate, pH 7.0. The Prp24-12 sample in anisotropic medium contained the same components plus 5% (w/v) phospholipid bicelles, consisting of a 3:1 molar ratio of dimyristoylphosphatidylcholine (DMPC) and dihexanoyl-phosphatidylcholine (DHPC), prepared as described previously (Ottiger and Bax 1998).

The At5g22580.1 sample in isotropic medium contained ~1 mM [^{13}C,^{15}N]-At5g22580.1, 50 mM sodium phosphate buffer, pH 6.5, and 0.1 mM NaN_3 in 93% H_2O and 7% H_2O. The At5g22580.1 sample in aligned medium contained ~0.7 mM protein under the same buffer conditions but with 5% (w/v) phospholipid bicelles, consisting of a 3:1 molar ratio of DMPC/DHPC (Ottiger and Bax 1998).

All samples were contained in 280 µl Shigemi microcells.

Data collection

The 3D experiments were recorded using the original quantitative J method described by Bax (Chou et al. 2000) (http://spin.niddk.nih.gov/bax/pp/trosy-hnco-JNC-3D). The 2D HIFI-C version of this sequence utilized simultaneous evolution of the indirectly detected ^{13}C′ and ^{15}N dimensions as described in detail the HIFI-NMR paper (Eghbalnia et al. 2005) and briefly summarized below.

The 2D tilted planes are collected by evolving both ^{13}C′ and ^{15}N indirect dimensions simultaneously point by point. The ratio between the dwell times of the two simultaneously evolving dimensions determines the angle of the resulting tilted plane. Practically, we set the desired ^{13}C′ and ^{15}N spectral windows and multiply the corresponding dwell times by the sine and the cosine of the chosen tilt angle. To achieve quadrature detection, we use the States method independently for each indirect dimension, resulting in four FIDs recorded interleaved for each time increment. These four FIDs represent all possible combinations of real and imaginary points of the indirect dimensions. The four FIDs are then separated in two spectra, each containing the real and imaginary points for one of the simultaneously evolving dimensions (^{13}C′), but only the real or the imaginary points of the second one (^{15}N). Fourier transformation of these two data sets results in two spectra, 90° out of phase, with duplicated peaks arising from the indistinguishable "+" and "–" ^{15}N frequencies. The sum of these two spectra results in a spectrum that contains a single set of peaks with the frequencies of the two simultaneously evolving nuclei added together ("+ tilt" spectrum). The difference between the two spectra yields a spectrum containing the other set of peaks at the position of the subtracted frequencies ("– tilt" spectrum). The HIFI-C pulse sequence (Fig. 1) and all processing scripts are posted on the web at http://miranda.nmrfam.wisc.edu/HIFI/RDC.

To enable benchmarking and comparison of results from standard 3D quantitative J and HIFI-C experiments, data for each protein were collected in 3D and HIFI mode in both isotropic and aligned states as outlined in Table 1. The $^1J_{C'N}$ coupling was derived from the relative intensity in two interleaved 3D spectra or from one or more HIFI 2D gradient and sensitivity enhanced HNCO spectra with $^1J_{C'N}$ dephasing intervals of approximately $1/(2\ ^1J_{C'N})$ (reference intensity) and approximately $1/^1J_{C'N}$ (attenuated intensity).

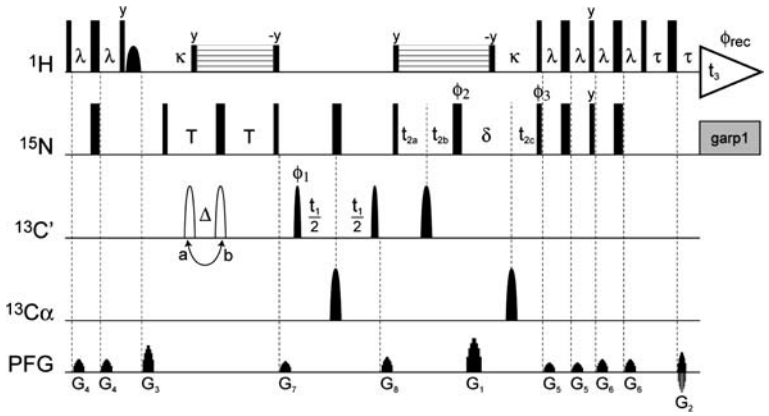

Fig. 1 Pulse sequence scheme for the 2D HIFI HNCO quantitative $J_{C'N}$ experiment. Narrow and wide pulses correspond to flip angles of 90° and 180°, respectively. All pulse phases are *x*, unless specified otherwise. All $^{13}C'$ pulses have the shape of the center lobe of a sin*x*/*x* function. Narrow $^{13}C'$ pulses bracketing the t_1 evolution time are 90° pulses, all other $^{13}C'$ pulses are 180°. $^{13}C^{\alpha}$ pulses are soft squared pulses with a flip angle of 180°. GARP1 decoupling was used on ^{15}N during acquisition. Delay durations: $\lambda = 2.4$ ms; $\kappa = 5.3$ ms; T = 33.7 ms; $\Delta = 16.6$ ms; $\delta = 14$ ms; $\tau = 0.8$ ms. For the reference spectrum, the first 180° pulse on $^{13}C'$ is given at position a; for the attenuated spectrum this pulse is given at position b. Constant time evolution is used for the ^{15}N dimension when $t_2/2 < \delta$: $t_{2a} = \delta - (t_2/2)$, $t_{2b} = 0.0$, $t_{2c} = (t_2/2)$; Semi-constant time ^{15}N evolution is used when $t_2/2 > \delta$: $t_{2a} = 0.0$, $t_{2b} = (t_2/2) - \delta$, $t_{2c} = (t_2/2)$. For collecting HIFI tilted planes: $t_1 = [(i - 1)/swC'] \times \cos\theta$ and $t_2 = [(i - 1)/swN] \times \sin\theta$; i is the increment counter common to both simultaneously evolving indirect dimensions; swC′ and swN are the spectral windows of the $^{13}C'$ and ^{15}N dimensions respectively; θ is the projection angle of the tilted plane. Phase cycling: $\varphi_1 = x,-x$, $\varphi_2 = x,x,-x,-x$, $\varphi_3 = x$, $\varphi_{rec} = x,-x$. Four FIDs are collected for each increment to accomplish quadrature detection for both $^{13}C'$ and ^{15}N indirect dimensions. Quadrature detection for $^{13}C'$ dimension is achieved by changing the φ_1 phase to *y*,–*y*. Quadrature detection for ^{15}N dimension is achieved by changing φ_3 to –*x* and inverting the sign of the rephasing gradient G_2. For processing of the HIFI tilted planes refer to the text and references therein. Gradients are 1 ms WURST shaped *z*-axis gradients, except for the dephasing and rephasing gradients, G1 and G2 that are set to 2.3 ms and 0.5 ms, respectively. Gradient strengths (G/cm): G1: 19.6, G2: +/– 9.14, G3: 17, G4: 2, G5: 3, G6: 4, G7: –3, G8: 5

Table 1 Experimental details

Protein	Medium	Data collection mode	Field (MHz)	Number of transients	Signal-to-noise		Total time (h)
					Mean	SD	
GB3	Isotropic	3D J	900	2	650	243	7.9
GB3	Isotropic	HIFI-C	900	8	239	94	1.7
GB3	pf1	3D J	600	4	315	161	17.5
GB3	pf1	HIFI-C	600	16	88	45	3.1
PRP24-12	Isotropic	3D J	800	8	64	64	44.1
PRP24-12	Isotropic	HIFI-C	600	48	19	16	14.6
At5g22580.1	Isotropic	3D J	600	8	133	289	31.5
At5g22580.1	Isotropic	HIFI-C	600	64	58	131	17.7
At5g22580.1	Bicelles	3D J	600	16	35	50	63.1
At5g22580.1	Bicelles	HIFI-C	600	192	27	37	48.6

The random error in the $^1J_{C'N}$ value extracted in this manner is inversely related to the S/N in the reference spectrum. As described in the 3D quantitative J paper, a S/N of 30:1 in the reference spectrum yields random errors of less than 0.2 Hz in the extracted $^1J_{C'N}$ value (Chou et al. 2000).

NMR spectra of the GB3 samples were collected at 25°C. NMR spectra of Prp24-12 in isotropic medium were acquired at 30°C and in anisotropic medium at 32°C. NMR spectra of At5g22580.1 in isotropic medium were acquired at 25°C and in anisotropic medium at 35°C.

Data were recorded on Varian INOVA spectrometers with cryogenic probes operating at 600 MHz ^1H frequency (GB3, Prp24-12, At5g22580.1), 800 MHz (Prp24-12), and 900 MHz (GB3). The lowest field (600 MHz) would have provided the highest quality HNCO data because of fast $^{13}C'$ relaxation effects; the other spectrometers were used simply because they were available. Comparisons between 3D quantitative J and HIFI-C were made at the same fields (with the exception of Prp24-12 data). The 3D spectra were recorded as 38–45*(t_1, ^{15}N) × 38–45*(t_2, ^{13}C) × 769–1161* (t_3, ^1H) data matrices. Acquisition times were 12.4–23.8 ms,

10.1–18.2 ms and 80.0 ms in the t_1, t_2 and t_3 dimensions, respectively. The 2D-HIFI spectra were collected as 70–110*(t_1, $^{15}N + ^{13}C$) × 769–1161*(t_3, ^{1}H) data matrices. Acquisition times were 30.0–64.1 ms, 21.2–54.7 ms and 80.0 ms in the t_1, and t_2 dimensions, respectively.

The adaptive decision process for optimal data collection protocol enables the user to balance the competing needs among short data collection times, sample stability and robust analysis of the extracted couplings (Fig. 2). In the HIFI method, two tilted planes are obtained after each data collection run (equivalent to a projection of the 3D signals onto a 2D plane tilted by an angle α). One plane contains a single set of peaks with the frequencies of the two simultaneously evolving nuclei added together ($+\alpha$ spectrum). The second plane contains the other set of peaks at the position of the subtracted frequencies ($-\alpha$ spectrum). The S/N in either case would be equivalent to that of a single spectrum acquired over the sum of the data acquisition times of the two spectra.

Each plane is processed independently and may result in signals with peaks with slightly different line shapes and amplitudes; this effect is more pronounced in the case of weak signals that are typical for the attenuated spectrum when collecting data by quantitative J method. When data collection from m different tilt angles is needed, $2m$ planes are obtained. The $2m$ extracted coupling values, obtained from m independent observations and $2m$ independent computations, show statistical fluctuations around the mean coupling for each residue. The variance in fluctuations of all residues can be statistically screened in order to obtain information about the reliability of extracted coupling values.

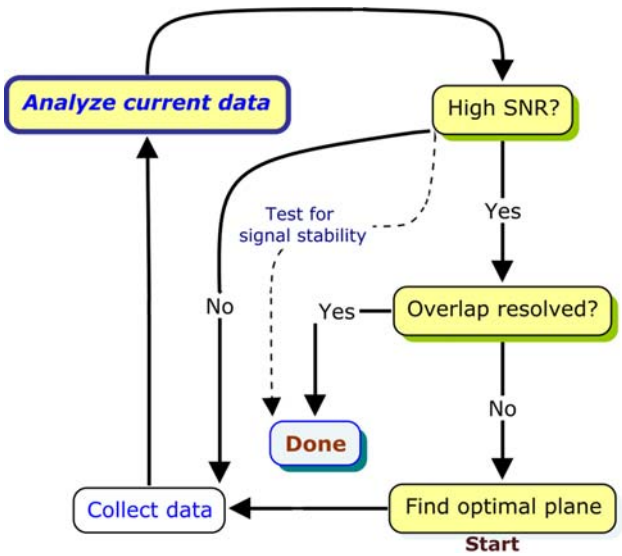

Fig. 2 The "adaptive decision process for optimal data collection" protocol used by HIFI-C starts with HN, N, and C′ peak positions determined from a previous HNCO experiment (standard or HIFI-NMR (Eghbalnia et al. 2005). From this information, the optimal plane with minimal peak overlap is computed, and data for this plane are collected and analyzed for satisfactory stability in the signal-to-noise-ratio (S/N) over time. Significant changes in S/N may be an indication of sample instability and may suggest that the collection of further data may degrade the results. If the S/N is stable, the choice is made to collect an additional plane (if the signals are overlapped) or to collect additional scans (if the signals are not overlapped). The process continues until the desired measure of accuracy and precision in the data has been achieved

The main step in the HIFI-C approach is to collect data for the plane with the least signal overlap from the combined +/– HIFI planes. If isotropic data have been collected by the HIFI approach, these angles are already available. Otherwise, the HIFI dispersion function can be used to determine this angle from a 3D HNCO peak list.

Extraction of RDC values

Scalar and dipolar couplings from both experiments are extracted in the same quantitative J style, i.e., from peak intensity ratios rather than splittings. Because the S/N in the attenuated spectrum is very low, the precise peak positions are picked from the reference spectrum (with high S/N), and the intensity at those exact positions are measured in the attenuated spectrum by using the 3D Fourier interpolation feature in NMRPipe (Delaglio et al. 1995). An in-house developed script (available on the web at http://miranda.nmrfam.wisc.edu/HIFI/RDC) can be used to map previously determined HNCO assignments onto these HIFI planes.

For each HIFI tilt angle, the two planes provide independent measurements of peak amplitudes. We calculated each J or J + D coupling from these measurements and took the average of the two when differences were smaller than 0.5 Hz. When the difference was larger than 0.5 Hz and peak overlap occurs in one of the planes we used the value from the plane with the non-overlapping peaks. For the three proteins used to test the HIFI-C method, a single HIFI plane was sufficient to obtain all or more than 95% of the measurable NC′ couplings. However, additional tilt planes (with optimal dispersion) can be acquired if needed to resolve overlaps. Each additional plane provides one more pair of measured values of peak intensities (if no overlap occurs), which provide cross-validation and increased precision in the extracted couplings.

Results and discussion

Proteins investigated

We used three proteins on hand in the laboratory for testing the HIFI-C method. GB3 is a benchmark protein of 55 amino acids with high resolution crystal and high quality

NMR refined structures available, allowing not only the comparison of the RDCs extracted by HIFI-C and 3D, but also fitting them to a structure. Although the GB3 sample we used happened to be perdeuterated (because it had been prepared for another type of experiment), this labeling pattern was expected to have no effect, positive or negative, on the data collected. PRP24-12 was chosen as an example of a protein of medium size (159 residues); however, because the protein proved to be unstable in the oriented medium, only isotropic couplings could be obtained. We report the results as an example of measurement of small couplings for a medium sized protein. The third protein, At5g22580.1, a 25 kDa homodimer, provided an example of a larger protein for which a comparison of RDCs obtained from HIFI-C and conventional 3D experiments was possible.

Test of the HIFI-C approach under isotropic conditions

Accurate extraction of couplings requires the careful consideration of S/N in the reference spectrum, which determines the error of measurement. The correlation between the isotropic couplings for GB3 extracted from HIFI and conventional 3D spectra (Fig. 3A) yielded a Pearson's correlation factor R of 99.8% and a pairwise root mean square deviations (rmsd) of 0.15 Hz.

For the PRP24-12 protein, the 3D experiment provided coupling data for 150 assigned peaks, whereas the pair of HIFI-C planes from the first optimal projection angle yielded data for 149 couplings. The 149 isotropic couplings obtained by both methods yielded a pairwise rmsd of 0.29 Hz. When the eight peaks with S/N < 10 were excluded, the rmsd became 0.25 Hz (Fig. 3B); and the correlation factor R was 94.0%.

With the third protein, At5g22580.1, HIFI-C yielded isotropic couplings for 99 residues as compared to 105 from 3D. Over these 99, the pairwise rmsd was 0.3 Hz. The largest difference (1.3 Hz) was for a signal affected by peak overlap in both (+ and –) HIFI planes; when data for this residue were excluded, the rmsd became 0.2 Hz (data not shown).

Test of the HIFI-C approach under conditions of partial alignment

The high stability of GB3 in the pf1 phage alignment medium, along with its high data quality, allowed rigorous comparison of $D_{NC'}$ values determined by the HIFI-C method (single pair of planes selected at 60°) with those from the traditional 3D experiment (Fig. 4A). The [^{13}C,^{15}N]-At5g22580.1 protein partially aligned by bicelles, was stable during the collection of HIFI-C data. A single tilted plane (at 28°) plane was sufficient to extract all 102 NC′ couplings in the aligned state (the same number as obtained from the 3D spectra). HIFI-C data from two additional tilted planes were obtained for comparison purposes. However, during the subsequent recording of 3D quantitative J data, the magnitude of the alignment started to diminish (as determined from decreased quadrupolar splitting of the water signal in 1D deuterium spectra measured between all experiments in the aligned state). Nevertheless, after scaling up the extracted couplings to compensate for the decrease in the alignment tensor, the RDCs extracted by the two methods showed a very good agreement (Fig. 4B) with a pairwise rmsd of 0.6 Hz. Moreover, the additional scatter in Fig. 4B when compared with the correlation plot under isotropic condition in Fig. 3B is due not only to inherent reduced signal-to-noise ratio caused by dipolar peak broadening but also to possible changes in the alignment tensor rhombicity during the 3D experiment. In this case, consistent data from 2 to 3 HIFI planes (i.e. 4–6 independent measurements) constitute a clearly more accurate measurement than a single 3D experiment on less than perfectly stable aligned protein sample. This was further confirmed by the lower pairwise rmsd obtained when comparing (averaged) couplings

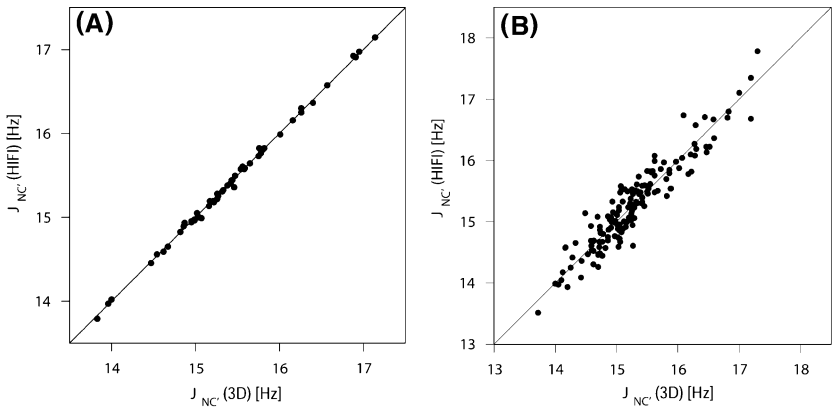

Fig. 3 Comparison between couplings measured by HIFI-C and 3D quantitative J experiments under isotropic conditions for (**A**) the 56 amino acid GB3 protein (R = 99.8%, rmsd = 0.03 Hz) and (**B**) the 159 amino acid PRP24-12 protein (R = 94.0%, rmsd = 0.25 Hz)

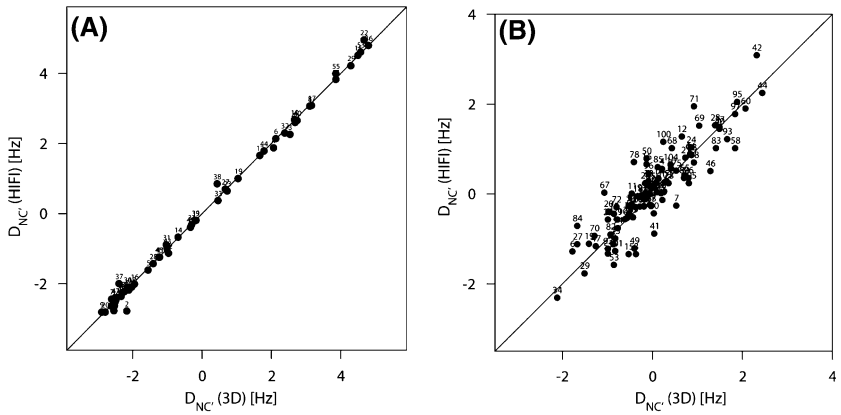

Fig. 4 Comparison between NC′ RDC values measured by HIFI-C (HIFI) and 3D quantitative J (3D) methods. (**A**) Correlation plot (R = 0.998, rmsd = 0.15 Hz) of NC′ RDC values for the GB3 protein. RDC values were obtained for 53 out of 55 residues. Two residues were excluded because of overlap in the 3D spectrum (one was also shifted in the aligned state, possibly due to slightly different pH. (**B**) Correlation plot (R = 0.940, rmsd = 0.6 Hz) of NC′ RDC values for the 25 kDa homodimeric At5g22580.1 protein. The RDC correlation was obtained for all 102 residues, with HIFI-C data from a single tilted plane (±28°)

extracted from measurements with different HIFI tilt angles, i.e. ~0.2 Hz for isotropic state and 0.4 Hz in aligned state (Supplementary material, Fig. 4).

Comparison of RDC values extracted by HIFI-C with those from conventional 3D NMR

Our reliability tests are based on statistical comparison of RDCs derived from HIFI-C and the standard 3D quantitative J experiment. In the case of HIFI-C, data from several independent planes can be used individually, or combined, for comparative analysis.

In comparing $D_{NC'}$ values for GB3 derived from 3D and HIFI-C (Fig. 4A) we observed three outliers (labeled by residue number i of the $N_iC'_{i-1}$ pair) showing differences larger than twice the pairwise rmsd. These small differences stem exclusively from the measurements in the aligned state, because the isotropic values of the couplings measured by the two methods are practically identical (pairwise rmsd is 0.03 Hz). A way of closely inspecting these differences is by comparing the experimental $D_{NC'}$ values with ones back-calculated from the extensively refined 3D structure of GB3 (PDB ID 1P7F). The outlier analysis is presented in the Supplementary material, Table 2. In two cases, those corresponding to larger RDCs (i.e. smaller relative experimental errors), the HIFI-C $D_{NC'}$ values extracted from the single tilt angle (60°) are closer to the values back-calculated from the structure than the $D_{NC'}$ values from the 3D quantitative J experiment. For the third outlier, that corresponding to the smallest $D_{NC'}$, both experimental $D_{NC'}$ values diverge equally (~0.3 Hz) from the back-calculated value.

The excellent correlations (Figs. 3A and 4A) attest to the accuracy of the HIFI-C method when compared to the traditional 3D quantitative J approach. The precision in measuring the NC′ dipolar couplings (which are among the smallest measurable RDC values) by HIFI-C is virtually identical to that provided by the 3D experiment. Thus, we can safely conclude that larger RDCs (such as D_{NH}, $D_{C^\alpha H^\alpha}$ etc.) can be obtained by the HIFI-C approach with even better precision at a fraction of the experimental time needed for the 3D method.

Similar analysis of the RDC values for At5g22580.1 in terms of its structure was not possible owing to the lower precision of the available structure along with the lower precision of the RDC measurements by both 3D and HIFI-C approaches. Despite the larger number of scans collected, the HNCO spectra of the homodimer have considerably broader peaks and lower S/N than those for GB3 (Table 1). Correlation plots for $J_{NC'}$ values for At5g22580.1 in isotropic and aligned media derived from 3D J and three HIFI-C planes (28°, 57°, and 77°) are presented in the Supplementary material, Fig. 1. Outliers (labeled by i of the $N_iC'_{i+1}$ pair) for At5g22580.1 in the RDC comparison plot in Fig. 4B that are larger than twice the pairwise rmsd (0.45 Hz) are listed in the Supplementary material, Table 3. These outliers correspond to residues that are in regular secondary structural elements: on β strands (8, 41, 42), at an end of a β strand (100), on α-helices (71, 78) and at an end of an α-helix (67).

The HIFI-C extracted J + D couplings in the aligned state of At5g22580.1 have been analyzed statistically (Table 2). The average S/N for the three pairs of HIFI-C planes was approximately 26. Tilted planes were collected in pairs in the "most informative" order, i.e. roughly that with the least amount of peak overlap in each plane. Couplings were extracted automatically from each plane and cross-matched in each pair of planes (+/– tilt angle)

in order to obtain a single list of extracted couplings. The method offers independent measurements from each plane and affords increased precision of the extracted couplings by averaging across all valid measurements, exclusion data from overlapped peaks, and identification of outliers. It is interesting to note that even though the number of overlapping peaks increased in the subsequent optimal planes (from 7 to 16 to 19; Table 2), the redundancy (i.e. number of couplings extracted from both planes) also increased and, more importantly, the pairwise rmsd of the couplings extracted from each pair of planes did not deteriorate (it actually decreased). In fact, the correlation between the couplings extracted from the +77°/–77° pair of planes (the 3rd optimal tilt angle) is at least as good as that of the +28°/–28° pair (the 1st optimal tilt angle) (Supplementary material, Fig. 3). A similar statistical analysis of the isotropic couplings obtained by HIFI-C for At5g22580.1 yielded much smaller rmsd values owing to the sharper peaks in the isotropic state; with an average spectral S/N of approximately 60 (Supplementary material, Table 1).

The observed high correlation of recovered RDC values across different planes in the HIFI-C method with S/N >25, coupled with the similar correlation with data collected by the standard 3D quantitative J method, points out the robustness and accuracy of both methods at high S/N. However, it was not possible to determine the accuracy and precision of RDC values recovered from 3D data sets at lower S/N. The outliers (Supplementary material, Table 3) were analyzed further to see if deviations of recovered RDC values are correlated with S/N values. No such correlation was observed in the data relating S/N to extracted RDCs in from either the 3D or the HIFI-C method. The results suggest that the accuracy of the recovered RDC values does not fall off linearly at S/N values below ~25, but instead is essentially random.

Data from independent HIFI-C planes provide a measure of precision, and this information can be used to determine whether or not to include RDCs as restraints in a structure determination step. Additional discussion and figures regarding the variation of extracted RDC values in different planes is presented in the Supplementary material.

Table 2 Statistical analysis of the couplings extracted from HIFI-C data collected for the At5g22580.1 protein in the aligned state, at three different tilt angles (two planes per tilt angle) collected in the "most informative" order (plane 1 first)

Order	HIFI Plane	Overlap[a]	Couplings extracted[b]	Rmsd: all points included[c]	Outliers[d]	Rmsd: outlier points excluded[e]
	+28°	6	92	0.65 Hz	6	0.45 Hz
	-28°	28	91	(*n* = 81)		(*n* = 75)
1	+28°/-28°	7	102/102	0.71 Hz (*n* = 45)		0.61 Hz (*n* = 75)
	+57°	13	96	0.51 Hz	5	0.38 Hz
	-57°	27	95	(*n* = 91)		(*n* = 86)
2	+57°/-57°	16	100/102	0.48 Hz (*n* = 36)		0.40 Hz (*n* = 82)
	+77°	27	97	0.50 Hz	5	0.38 Hz
	-77°	23	99	(*n* = 95)		(*n* = 90)
3	+77°/-77°	19	101/102	0.42 Hz (*n* = 30)		0.37 Hz (*n* = 83)

[a] For each plane, the overlap column records the number of peaks that overlapped only in that plane; for each pair of planes (shaded rows) the overlap column records the number of peaks that overlapped in both planes

[b] Owing to overlap in individual planes, and the collection of data in pairs of planes, the reported number of couplings extracted is for the pair of planes

[c] Pairwise root mean square deviations (rmsd) for the extracted couplings between pair of planes (+,–): non-shaded rows, for all peak pairs (overlapping and non-overlapping); shaded rows, only for non-overlapping peak pairs. The number of couplings involved is given by *n*

[d] Outliers are those where the difference between couplings extracted from the +/– planes is larger than twice the pairwise rmsd

[e] The rmsd in the shaded rows was calculated only between couplings for which averaging was applied (see text). The number of couplings involved is given by *n*

Comparison of the HIFI-C with the J-GFT reduced dimensionality method

An RD method was recently reported (J-GFT) (Atreya et al. 2007) and applied to the simultaneous measurement of five mutually correlated NMR parameters: ^{15}N backbone chemical shifts and the four one-bond spin-spin couplings, $^{13}C^{\alpha}$–$^{1}H^{\alpha}$, $^{13}C^{\alpha}$–$^{13}C'$, ^{15}N–$^{13}C'$, and ^{15}N–$^{1}H^{N}$. The experiment was demonstrated by measuring RDC values for a small (8 kDa) protein. The J-GFT method has the advantage of measuring several RDC types simultaneously. From the data reported, however, small couplings obtained by the J-GFT approach appear to be less robust than those obtainable by the HIFI-C method. For example, the authors reported a pairwise rmsd of 0.85 Hz for the NC′ couplings (over ~44 of 54 total) obtained by J-GFT at two different magnetic field strengths. In addition, their reported pairwise rmsd values between NC′ RDC values determined by 3D quantitative J and J-GFT at the two magnetic fields were 0.75 Hz and 0.9 Hz, respectively. The relative errors in many of the reported NC′ couplings (as large as 50–100%) suggest that the measurements may not be of practical use as structural constraints. By comparison, the pairwise rmsd between HIFI-C and the traditional 3D quantitative J experiment for a similar sized protein (GB3) was 0.15 Hz (over 53 of 55) (Fig. 4).

Because the J-GFT experiment requires the simultaneous evolution of one chemical shift and four coupling constants in a constant time fashion, with consequent reduction of sensitivity in each added dimension, it is unclear that this approach will be practical for larger proteins. The evolution of the backbone ^{15}N chemical shift alone may not serve to resolve all the peaks for proteins larger than 8 kDa; if so, this would require the co-evolution of additional frequencies and further reduce sensitivity or increase the data collection time. Another uncertainty is how the proposed GFT-J approach for dealing with spin-system-specific phase shifts of cosine-modulated signal components would work for larger systems with weak and overlapped signals. The HIFI-C approach has been demonstrated here with a considerably larger (25 kDa homodimer) protein (Supplementary material, Fig. 2).

Finally, the fully automated data analysis and statistics afforded by HIFI-C provide rapid extraction of coupling data along with a measure of the reliability and robustness of the couplings, not currently available with J-GFT.

Conclusions

The HIFI-C approach described here (Fig. 2), although similar to the original HIFI experiment (Eghbalnia et al. 2005) in its use of adaptively selected optimal planes for reducing 3D data collection to 2D, differs from HIFI-NMR in its fundamental goal. Whereas the HIFI experiment aims simply to identify the positions of peaks in 3D frequency space, the HIFI-C experiment aims to determine the intensity of each non-overlapped peak. We have shown that HIFI-C provides a fast, robust, and automated method for extracting coupling information from 3D spectra. Although we have focused here on examples of residual dipolar coupling measurements, the approach should be applicable to other situations. For example, a TROSY 3D HNCO experiment was used to detect hydrogen bonds in a 30 kDa [^{2}H, ^{13}C, ^{15}N]-protein (Wang et al. 1999). In such an experiment, the optimal repetition rate is given by the $^{1}H^{N}$ T_1 value. To acquire the data in a reasonable period of time (3.8 day), the authors used a repetition rate of 2.6 s instead of the optimal value of 5 s. The HIFI-C approach should permit the use of the optimal (5 s) repetition rate (for highest sensitivity) in the 2D tilt planes and speed up the data collection considerably.

The HIFI-C approach for rapid and robust measurement of RDCs could lead to substantial improvements in the speed and accuracy of NMR structure determinations, refinement, and validation (Andrec et al. 2001; Clore and Schwieters 2003; Cornilescu et al. 1998; Delaglio et al. 2000; Hus et al. 2000; Qu et al. 2004; Tjandra et al. 1997; Wang and Donald 2004; Wedemeyer et al. 2002). Fast data collection can reinforce the potential advantages of RDC-based NMR structure determination, which include the ability to determine structures from a small number of NMR experiments and the promise of overcoming the protein size limit of traditional NMR structure determination approaches.

As shown here, in addition to speeding up data collection, HIFI-C may make it possible acquire information on oriented samples before they degrade. RDCs are measurable only in solution when the protein is in a so-called nematic dilute liquid crystalline form or in anisotropically compressed hydrogels, where large molecules (often much larger than the protein), are added to the solution. These agents align in the magnetic field of the spectrometer due to their own magnetic properties. The interactions between the aligning agent and the protein effectively reduce the number of possible orientations and lead to an “average” orientation of the protein with respect to the magnetic field, which leads to the creation of measurable RDCs. The presence of the agent, however, can introduce conditions that are deleterious to the long-term stability of the protein sample, or the agent itself may become unstable.

The observed correlations in the benchmark protein as well as the other test protein are well within experimental error margins and suggest that HIFI-C results can be used reliably. The inclusion of the RDC terms in the NMR energy function for structure calculation can result in a

highly rippled energy surface with innumerable sharp local minima (Bax 2003), making the search problem exceedingly difficult. However, if the starting model is close to the true structure, and only robust RDC values are included, convergence will be more likely and more reliable. The additional statistics afforded by HIFI-C regarding the reliability and robustness of RDC values should be useful in determining a reliable subset or in weighting the dataset on the basis of reliability in order to achieve better convergence in the structure determination stage.

Acknowledgments We thank Ad Bax and Sam Butcher for allowing us to use their protein samples GB3 and Prp24-12, respectively, in this study. This research was carried out in the National Magnetic Resonance Facility at Madison, which is supported by Biomedical Research Technology Program, National Center for Research Resources, through NIH grant P41 RR02301 and utilized equipment funded by the University of Wisconsin, the NSF Academic Infrastructure Program (BIR-9214394), the NIH Shared Instrumentation Program (RR02781, RR08438), the NIH Research Collaborations to Provide 900 MHz NMR Spectroscopy (GM66326), the NSF Biological Instrumentation Program (DMB-8415048), and the U.S. Department of Agriculture. A.B. received partial support from the National Institute of General Medical Science's Protein Structure Initiative through NIH grant 1 P50 GM64598, which supports the Center for Eukaryotic Structural Genomics. H.E. was supported in part by the National Library of Medicine under grant 1K22LM8992.

References

Andrec M, Du P, Levy RM (2001) Protein structural motif recognition via NMR residual dipolar couplings. J Am Chem Soc 123:1222–1229

Atreya HS, Garcia E, Shen Y, Szyperski T (2007) J-GFT NMR for precise measurement of mutually correlated nuclear spin-spin couplings. J Am Chem Soc 129:680–692

Bax A (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci 12:1–16

Bryce DL, Bax A (2004) Application of correlated residual dipolar couplings to the determination of the molecular alignment tensor magnitude of oriented proteins and nucleic acids. J Biomol NMR 28:273–287

Chou JJ, Delaglio F, Bax A (2000) Measurement of one-bond N-15-C-13′ dipolar couplings in medium sized proteins. J Biomol NMR 18:101–105

Clore GM, Schwieters CD (2003) Docking of protein–protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from H-1(N)/N-15 chemical shift mapping and backbone N-15-H-1 residual dipolar couplings using conjoined rigid body/torsion angle dynamics. J Am Chem Soc 125:2902–2912

Cordier F, Grzesiek S (1999) Direct observation of hydrogen bonds in proteins by interresidue $^{3H}J_{NC'}$ scalar couplings. J Am Chem Soc 121:1601–1602

Cornilescu G, Marquardt JL, Ottinger M, Bax A (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J Am Chem Soc 120:6836–6837

Cornilescu G, Hu JS, Bax A (1999a) Identification of the hydrogen bonding network in a protein by scalar couplings. J Am Chem Soc 121:2949–2950

Cornilescu G, Ramirez BE, Frank MK, Clore GM, Gronenborn AM, Bax A (1999b) Correlation between $^{3H}J_{NC'}$ and hydrogen bond length in proteins. J Am Chem Soc 121:6275–6279

Cornilescu G, Cornilescu CC, Zhao Q, Frederick RO, Peterson FC, Thao S, Markley JL (2004) Letter to the Editor: Solution structure of a homodimeric hypothetical protein, At5g22580, a structural genomics target from *Arabidopsis thaliana*. J Biomol NMR 29:387–390

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

Delaglio F, Kontaxis G, Bax A (2000) Protein structure determination using molecular fragment replacement and NMR dipolar couplings. J Am Chem Soc 122:2142–2143

Ding K, Gronenborn AM (2002) Novel 2D triple-resonance NMR experiments for sequential resonance assignments of proteins. J Magn Reson 156:262–268

Eghbalnia HR, Bahrami A, Tonelli M, Hallenga K, Markley JL (2005) High-resolution iterative frequency identification for NMR as a general strategy for multidimensional data collection. J Am Chem Soc 127:12528–12536

Gallagher T, Alexander P, Bryan P, Gilliland GL (1994) 2 crystal-structures of the B1 immunoglobulin-binding domain of streptococcal protein-G and comparison with NMR. Biochemistry 33:4721–4729

Hansen MR, Mueller L, Pardi A (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat Struct Biol 5:1065–1074

Hoshino M, Otting G (2004) Sensitivity-enhanced double-TROSY experiment for simultaneous measurement of one-bond ^{15}N–^{1}H, ^{15}N–$^{13}C'$ and two-bond ^{1}H–$^{13}C'$ couplings. J Magn Reson 171:270–276

Hus JC, Marion D, Blackledge M (2000) De novo determination of protein structure by NMR using orientational and long-range order restraints. J Mol Biol 298:927–936

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

Kover KE, Batta G (2004) More line narrowing in TROSY by decoupling of long-range couplings: shift correlation and $^{1}J_{NC'}$ coupling constant measurements. J Magn Reson 170:184–190

Kupce E, Freeman R (2003) Reconstruction of the three-dimensional NMR spectrum of a protein from a set of plane projections. J Biomol NMR 27:383–387

Meissner A, Duus JO, Sorensen OW (1997) Spin-state-selective excitation. Application for E.COSY-type measurement of J_{HH} coupling constants. J Magn Reson 128:92–97

Ottiger M, Bax A (1998) Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules. J Biomol NMR 12:361–372

Ottiger M, Delaglio F, Bax A (1998a) Measurement of *J* and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson 131:373–378

Ottiger M, Delaglio F, Marquardt JL, Tjandra N, Bax A (1998b) Measurement of dipolar couplings for methylene and methyl sites in weakly oriented macromolecules and their use in structure determination. J Magn Reson 134:365–369

Permi P, Heikkinen S, Kilpelainen I, Annila A (1999) Measurement of $^{1}J_{NC'}$ and $^{2}J_{HNC'}$ couplings from spin-state-selective two-dimensional correlation spectrum. J Magn Reson 140:32–40

Qu YX, Guo JT, Olman V, Xu Y (2004) Protein structure prediction using sparse dipolar coupling data. Nucleic Acids Res 32:551–561

Reiter NJ, Lee D, Wang YX, Tonelli M, Bahrami A, Cornilescu CC, Butcher SE (2006) Resonance assignments for the two

N-terminal RNA recognition motifs (RRM) of the *S. cerevisiae* pre-mRNA processing protein Prp24. J Biomol NMR 36:58

Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc 34:93–158

Simorre JP, Brutscher B, Caffrey MS, Marion D (1994) Assignment of NMR-spectra of proteins using triple-resonance 2-dimensional experiments. J Biomol NMR 4:325–333

Szyperski T, Wider G, Bushweller JH, Wuthrich K (1993) 3D C-13-N-15-heteronuclear 2-spin coherence spectroscopy for polypeptide backbone assignments in C-13-N-15-double-labeled proteins. J Biomol NMR 3:127–132

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

Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278:1111–1114

Tjandra N, Omichinski JG, Gronenborn AM, Clore GM, Bax A (1997) Use of dipolar $^{1}H–^{15}N$ and $^{1}H–^{13}C$ couplings in the structure determination of magnetically oriented macromolecules in solution. Nat Struct Biol 4:732–738

Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. Proc Natl Acad Sci USA 92:9279–9283

Wang LC, Donald BR (2004) Exact solutions for internuclear vectors and backbone dihedral angles from NH residual dipolar couplings in two media, and their application in a systematic search algorithm for determining protein backbone structure. J Biomol NMR 29:223–242

Wang YX, Jacob J, Cordier F, Wingfield PT, Stahl SJ, Lee-Huang S, Torchia DA, Grzesiek S, Bax A (1999) Measurement of $^{3H}J_{NC'}$ connectivities across hydrogen bonds in a 30 kDa protein. J Biomol NMR 14:181–184

Wedemeyer WJ, Rohl CA, Scheraga HA (2002) Exact solutions for chemical bond orientations from residual dipolar couplings. J Biomol NMR 22:137–151