Journal of Biomolecular NMR (2005) 32: 55–60
DOI 10.1007/s10858-005-3339-y

Article

High-resolution aliphatic side-chain assignments in 3D HCcoNH experiments with joint H–C evolution and non-uniform sampling

Zhen-Yu J. Sun[a], Sven G. Hyberts[a], David Rovnyak[a,b], Sunghyouk Park[a], Alan S. Stern[c], Jeffrey C. Hoch[d] & Gerhard Wagner[a,*]

[a]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115;* [b]*Department of Chemistry, Bucknell University, Lewisburg, PA 17837;* [c]*Rowland Institute at Harvard, Cambridge, MA 02142;* [d]*Department of Molecular, Microbial, and Structural Biology, University of Connecticut Health Center, Farmington, CT 06030*

Received 20 December 2004; Accepted 15 February 2005

Key words: automated NMR assignment, maximum entropy reconstruction, non-uniform sampling, reduced-dimensionality, side-chain assignment

Abstract

We describe an efficient NMR triple resonance approach that correlates, at high resolution, protein side-chain and backbone resonances. It relies on the combination of two strategies: joint evolution of aliphatic side-chain proton/carbon coherences using a backbone N–H based HCcoNH reduced dimensionality (RD) experiment and non-uniform sampling (NUS) in two indirect dimensions. A typical data set containing such correlation information can be acquired in 2 days, at very high resolution unfeasible for conventional 4D HCcoNH-TOCSY experiments. The resonances of the aliphatic side-chain protons are unambiguously assigned to their attached carbons through the analysis of the 'sum' and 'difference' spectra. This approach circumvents the tedious process of manual resonance assignments using HCcH-TOCSY data, while providing additional resolving power of backbone N–H signals. A simple peak-list based algorithm has been implemented in the IBIS software for rapid automated backbone and side-chain assignments.

Introduction

Whereas efficient methods are available for protein backbone assignments, obtaining side-chain C–H correlations is still a serious bottleneck. A set of backbone triple-resonance experiments sufficient for assignment of proteins up to 20 kDa can now be recorded within a single day (Rovnyak et al., 2004a), when using non-uniform sampling (NUS) (Schmieder et al., 1993; Hoch and Stern, 1996; Rovnyak et al., 2003). In addition, suites of reduced-dimensionality (RD) (Szyperski et al., 1993; Simorre et al., 1994; Freeman and Kupce 2003) experiments have been proposed to speed up sequential backbone assignments (Ding and Gronenborn, 2002; Szyperski et al., 2002; Bersch et al., 2003; Atreya and Szyperski, 2004 Kim and Szyperski, 2004), and automated assignment programs, such as GARANT (Bartels et al., 1996), AutoAssign (Zimmerman et al., 1997), and IBIS (Hyberts and Wagner, 2003), can typically derive backbone assignments in a few hours. With backbone assignments known, 3D hCcoNH- and HccoNH-TOCSY experiments (Montelione et al., 1992; Grzesiek et al., 1993a; Gardner et al., 1996; Lin and Wagner, 1999) can then correlate the N–H spin pair with either ^{13}C or ^{1}H signals from the side-chains, but do not yield direct side-chain C–H

*To whom correspondence should be addressed. E-mail: gerhard_wagner@hms.harvard.edu

correlations. The 3D HCcH-TOCSY experiment (Kay et al., 1993) can provide such correlations but does not connect to the backbone NH group, hindering automated assignments.

The 4D HCcoNH-TOCSY experiments (Logan et al., 1992; Clowes et al., 1993) would correlate both side-chain ^{13}C and ^{1}H with each other and with backbone N–H spin pairs; however, these have to be recorded at low resolution to fit within typically available instrument time (Table 1). Recently published 2D RD-HCcoNH type experiments take less time and use DEPT editing to minimize spectral overcrowding in the 2D spectra (Brutscher, 2004). However, separate experiments are required to correlate the Hα/Hβ and methyl C–H with backbone N–H, making the side-chain assignments incomplete and less straightforward.

In the approach presented here, we combine two different but complimentary selective sampling strategies to overcome these problems. First, using the RD technique, we convert the 4D HCcoNH experiment into a 3D experiment that jointly evolves the ^{1}H and ^{13}C spins in synchronized increments. Combining these RD spectra together with either one of the conventional hCcoNH and HccoNH experiments, as the 'basic' and 'central-peak' experiments according to GFT terminology (Kim and Szyperski, 2003), we can retrieve all information that would be obtained from a 4D HCcoNH experiment. In addition, we apply the NUS technique to record these RD and conventional 3D experiments at high resolution in both $^{1}H/^{13}C$ and ^{15}N indirect dimensions within a reasonably short overall spectrometer time. The combination of RD, which encodes information of multiple spins into a lower-dimension spectral window, and NUS, which reduces the amount of data points required to characterize a spectral window, enables this and potentially other methods that would otherwise be impractical.

Table 1. Comparison of NUS-RD-HCcoNH (500 MHz spectrometer) and 4D HCcoNH (600 MHz spectrometers) experimental parameters

Experiments	Complex points	t_{max}(ms)	Resolution (ppm)
	1600 (2D)	32.0 (^{1}H)	0.06
NUS-RD-HCcoNH	[160 ($^{1}H/^{13}C$),	6.4 (^{13}C)	1.2
	40 (^{15}N)][a]	23.7 (^{15}N)	0.83
	40 (^{1}H)	6.4	0.26
4D HCcoNH	8 (^{13}C)[b]	2.43	2.7
(Logan et al., 1992)	8 (^{15}N)	3.76	4.4
	20 (^{1}H)	6.2	0.27
4D HCcoNH	8 (^{13}C)[b]	2.67	2.5
(Clowes et al., 1993)	24 (^{15}N)[b]	23.6	0.70

[a]Maximum points for each dimension in the 2D NUS schedule.
[b] Reduced spectral width with folded/aliased peaks.

Methods

The pulse sequence for the 3D RD-HCcoNH experiment is shown in Figure 1. In addition to the typical coherence transfer steps, it employs simultaneously incremented ^{13}C constant-time (CT) and ^{1}H semi-CT evolution periods. To allow for maximum ^{1}H resolution during the joint evolution, we scaled down the ^{13}C evolution time increments to 20% relative to the ^{1}H evolution. This ensures that a maximum of 160 complex points are sampled within the 6.6 ms ($\sim 1/4J_{cc}$) ^{13}C CT evolution period as well as within the ^{1}H evolution period to reach a ^{1}H maximum evolution time (t_{max}) of 32 ms. The reduced ^{13}C resolution is still sufficient for the unambiguous assignment of the widely dispersed ^{13}C resonances. To maximize signals from atoms near the end of the side-chains, a 20 ms ^{13}C TOCSY mixing was used employing an adiabatic pulse train (Bennett et al., 2003).

In Figure 1, the phase Φ_1 of the first ^{1}H 90° pulse, relative to the phase Φ_2 of the first ^{13}C 90° pulse determines whether it is the $^{1}H/^{13}C$ 'in-phase' ($HCcoNH_I$) or 'anti-phase' ($HCcoNH_A$) experiment type. The 'in-phase' spectrum contains pairs of peaks at the sum and difference of the $^{1}H/^{13}C$ frequencies with the same signs, while the 'anti-phase' spectrum contains the peak pairs with opposite signs. Either ^{1}H or ^{13}C coherence can be used as the central-peak reference and phase incremented in States-TPPI fashion, while the other alternating between 0° and 90° phase producing the 'in-phase' and 'anti-phase' RD-HCcoNH spectra. We chose to record the RD-HCcoNH spectra with ^{1}H phase incremented in a modified States-TPPI fashion, (phase incremented 90°, vs. 180° in the conventional States-TPPI), to shift the center of the effective ^{1}H spectra by 1/4 spectral width to the middle of the aliphatic region while leaving the ^{1}H carrier frequency centered on water. This allows the reduction of the overall combined $^{1}H/^{13}C$ spectral width by half

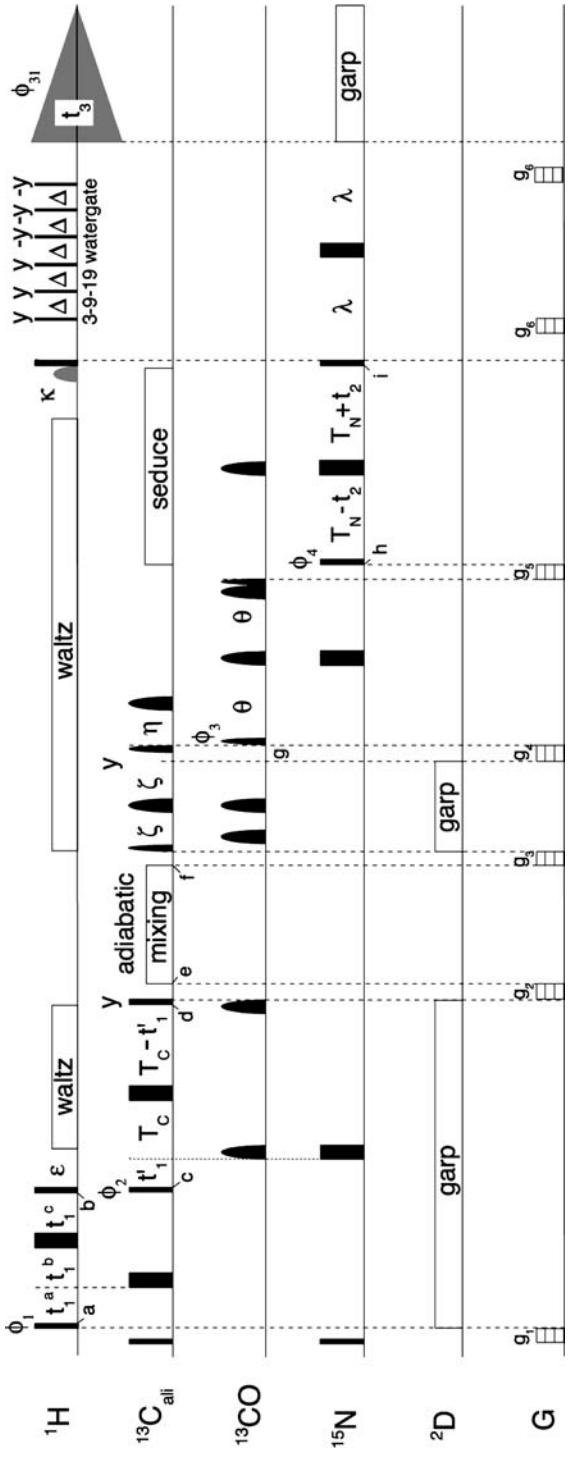

Figure 1. Pulse scheme for the 3D NUS-RD-HCcoNH experiment. Narrow and wide bars and hemi-ellipses indicate 90°, 180° hard pulses and shaped pulses respectively. The delays are: $\varepsilon = 2.1$ ms; $\zeta = 3.5$ ms; $\eta = 4.5$ ms; $\theta = 12.4$ ms; $\kappa = 5.5$ ms; $\lambda = 2.3$ ms; $\Delta = 210$ ms. The ^{1}H coherences evolve in a semi-constant time fashion (points a to b), with the initial delays and increments for the semi-constant time ^{1}H evolution period set according to the literature (Grzesiek and Bax, 1993b). The ^{13}C and ^{15}N coherences evolve in constant-time fashions, with the constant evolution times $T_{C=3.3}$ ms (c to d) and $T_{N=12.4}$ ms (h to i) respectively. The ^{1}H (t_1) and ^{13}C ($t\prime_1$) evolution periods are incremented simultaneously in synchronized steps and the ^{13}C time increments are scaled to 20% of ^{1}H increments. All pulses are applied along x, unless specified otherwise. Phase cycling is as follows: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = 4(x), 4(-x)$; $\phi_4 = 2(x), 2(-x)$; ϕ_{31}(receiver) = $x, -x, -x, x, -x, x, x, -x$. Quadrature detection in $t_1/t\prime_1$uses 1/4 spectral width aliased States-TPPI for ϕ_1 (phase incremented 90° between time increments) while alternating ϕ_2 according to States. Quadrature detection in t_2 (^{15}N) is achieved by incrementing ϕ_4 according to States-TPPI. The ^{1}H WALTZ decoupling uses a decoupling RF field of 5.5 kHz. The ^{13}C TOCSY mixing is achieved by adiabatic mixing - pulse sequence (Bennett et al., 2003). The ^{13}C carrier is placed at 55 ppm on Cα initially, and then switched to 43 ppm to center on aliphatic resonances at point e, until point g where it is switched to 176 ppm on ^{13}CO resonances. The ^{2}D continuous decoupling is implemented for partially deuterated protein samples.

(equivalent of 10 ppm for ^{1}H), yielding further improvement in resolution. We recorded the 'in-phase' and 'anti-phase' spectra by alternating the ^{13}C Φ_2 phase to generate four interleaving FIDs for each time point of joint ^{1}H/^{13}C evolution. These FIDs were converted by a computer program to generate the 'sum' ($HCcoNH_S$) and 'difference' ($HCcoNH_D$) spectra containing peaks at the sum and difference of the ^{1}H/^{13}C frequencies respectively. The 'sum' and 'difference' spectra are preferred for side-chain assignments as they contain half the number of peaks compared to the 'in-phase' and 'anti-phase' spectra, and the signal-to-noise ratio is improved by 41% for each peak given the same raw data set.

To demonstrate this approach, we have acquired a set of the NUS-RD-HCcoNH spectra for a 0.7 mM uniformly–^{15}N/^{13}C labeled Nck SH3-1 domain (MW 7 kDa) protein in 50 mM $NaPO_4$, 5 mM EDTA, pH 6.5 buffer on a Bruker Avance-500 spectrometer equipped with a cryogenic probe. The total experimental time was 2 days using a high-resolution 1600-point 2D sampling schedule (Table 1), which was generated by using the Rowland NMR Toolkit so that the sampling density corresponds to the decaying signal intensity. Recording the more sensitive NUS-hCcoNH and NUS-HccoNH pair, which uses 14.6 ms and 31 ms semi-constant evolution periods respectively, required one additional day of measurement time. The non-uniformly sampled data were subsequently processed by 2D maximum entropy

(MaxEnt) reconstruction (Hoch and Stern, 1996; Stern et al., 2002) using the Rowland NMR Toolkit. The processing time for the high-resolution spectra with a matrix size of $450(H^N) \times 128(N) \times 1024(H/C)$ complex points is typically around 2.5 hours on a dual processor 3.0 GHz Linux computer. The data were further analyzed using XEASY software.

Results and discussions

Side-chain assignments in proteins are usually pursued with 3D hCcoNH and HccoNH TOCSY experiments that link the side-chain proton and carbon resonances to the backbone N–H spin pairs, followed by the HCcH-TOCSY experiment that further maps out the side-chain spin-systems. Assignment of the ^{13}C resonances to carbon atoms in the side chains is relatively easy since the ^{13}C spectrum is widely dispersed and their distinctive statistical distribution can be found in the BioMagResBank database (http://www.bmrb.wisc.edu). However, aliphatic ^{1}H resonances can only be assigned unambiguously (based on chemical shift information alone) for Gly and eight other residues, including Ala, Asp, Asn, Cys, His, Trp, Phe, and Tyr, with only α and β aliphatic groups. The unambiguous ^{1}H chemical shift assignments of part or all of the aliphatic protons in the remaining 11 residue types rely on the $^{1}H/^{13}C$ correlation identified from experiments such as RD-HCcoNH. Figure 2a shows correlations from one N–H group to three pairs of ^{1}H and ^{13}C signals obtained from the NUS-RD-HCcoNH data set. The correlated $^{1}H/^{13}C$ pairs can be identified when the sum ($\omega_H + k\omega_c$) and difference ($\omega_H - k\omega_c$) of their chemical shift frequencies (where $k = 0.2$ is the ratio between the ^{13}C and ^{1}H evolution time increments) match the peaks in the $HCcoNH_S$ and $HCcoNH_D$ spectra. As illustrated in Figure 2a, the mid-points between the correlated $^{1}H/^{13}C$ peak pairs in the RD-HCcoNH spectra match the peak positions in the HccoNH spectrum, while their chemical shift differences divided by 2 match the peak positions in hCcoNH (scaled by 0.2 as frequency in Hz) relative to the carrier frequency. The identity of the protons can then be unambiguously assigned based on known ^{13}C assignments from IBIS sequential assignment results.

For Nck-SH3 containing 60 amino acid residues, we were able to assign 95% of the 223 aliphatic proton resonances from 57 residues. Among these, 105 proton resonances can be assigned unambiguously based solely on their distinctive chemical shifts; 107 (about 48%) proton resonances were assigned unambiguously with the aid of the NUS-RD-HCcoNH data. The remaining 11 ambiguous resonances belong to three Arg/Lys residues that show weak signals. The residue Val51 was not included in the analysis because it precedes Pro52, and the residues Pro52 and Ser53 were also excluded as they precede residue Ser53 and Asn54 which have extremely weak or missing N–H peaks (possibly due to conformational exchange broadening).

Non-uniform sampling and ^{13}C constant-time evolution are two critical factors for achieving high resolution in the indirect ^{1}H dimension in this NUS-RD-HCcoNH experiment. Data acquired with a NUS scheme require about three to five-fold fewer increments for a given resolution. In high-resolution experiments where T_2 relaxation in the indirect dimensions cannot be ignored, significant signal loss will occur at long t_{max} for conventional linearly sampled data. Hence, the overall signal-to-noise ratio may be improved when using a NUS scheme that samples more frequently at the beginning of the evolution time period where the signal is stronger (Rovnyak et al., 2004b). The ^{13}C CT evolution is important because it does not lead to exponentially decaying behavior, so that the linewidths of the peaks in the combined $^{1}H/^{13}C$ dimension is determined by the ^{1}H relaxation only. The drawback of this approach is the limit imposed by the length of the ^{13}C CT period: if using $T_C = 27$ ms ($1/J_{cc}$), the signals become too weak due to the ^{13}C relaxation; with $T_C = 13.5$ ms ($1/2J_{cc}$), then the ^{13}C signals are in anti-phase giving no net signal. As a compromise, the ^{13}C CT was set to 6.6 ms ($1/4J_{cc}$), where the net ^{13}C signals are reduced to 71%. It should be noted that in order to improve both the resolution and sensitivity, the separate 'central-peak' hCcoNH experiment was run instead in a semi-constant time fashion with a ^{13}C t_{max} of 14.6 ms.

In theory, since hCcoNH data have most often already been acquired in assisting backbone sequential assignment, only one additional NUS-RD-HCcoNH experiment (recording both the 'in-phase' and 'anti-phase' data) is needed to complete

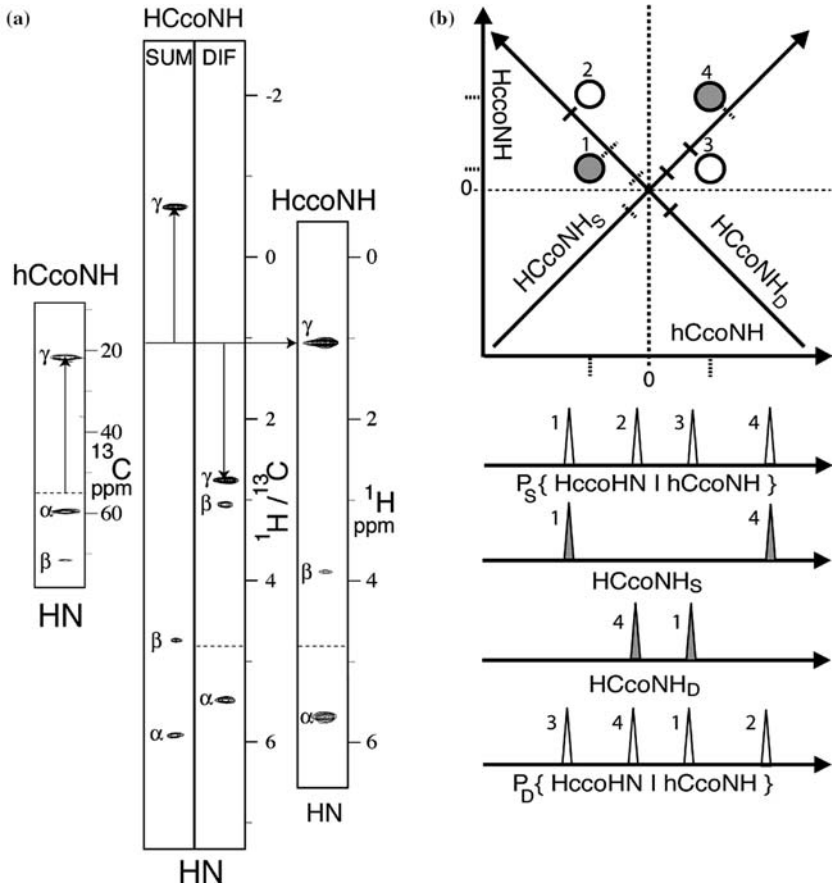

Figure 2. (a) Selected 2D N–H strips (height scaled in Hz) from the non-uniformly sampled RD-HCcoNH, HccoNH, and hCcoNH (spectral width scaled by k = 0.2 in Hz) spectra using a 0.7 mM Nck SH3 domain sample (25 °C). The vectors indicate that the mid-point between the upfield and downfield peaks of the methyl group from a threonine side-chain in RD-HCcoNH match the ^{1}H peak position in HccoNH, while the chemical shift difference (scaled by 0.2 in Hz) matches the ^{13}C peak position in hCcoNH relative to the ^{13}C carrier frequency used in HCcoNH (dotted line). (b) Side-chain ^{1}H/^{13}C resonance assignment scheme. The circles (filled and open) represent a complete combination of possible correlations predicted from identified HccoNH and hCcoNH peaks. The projections onto the combined ^{1}H/^{13}C dimension of the RD-HCcoNH$_S$ and RD-HCcoNH$_D$ spectra (open triangles) are matched to those observed peaks (filled triangles). The final assigned peaks are identified as filled circles. The tilt angle of the projection axes (here exaggerated) depends on the relative scaling factor of the evolution time increments.

the aliphatic side-chain proton assignments. It is preferable, however, to collect the HccoNH data as well to validate the assignments and resolve ambiguities due to occasional weak or missing peaks. As demonstrated with our Nck-SH3 data, for small proteins of less than 10 kDa we can simplify the side-chain proton assignment strategy and eliminate the need for the conventional HCcH-TOCSY experiment with only two extra days of spectrometer time. To obtain equivalent results without NUS-RD, a 4D HCcoNH experiment with the same high-resolution would take months of spectrometer time, which is completely impractical. For large proteins (>10 kDa), longer measuring time and/or partial (65%) deuteration of the protein sample (Lin and Wagner, 1999) may be required to improve the sensitivity of the NUS-RD-HCcoNH experiments.

Another advantage of our approach is that the NUS-RD-HCcoNH data can be used in NH-based automated assignment programs, such as IBIS (Hyberts and Wagner, 2003). We have implemented in IBIS an algorithm related to the 'projection reconstruction' method (Kupce and Freeman, 2004) as illustrated in Figure 2b. The top diagram shows a hypothetical side-chain ^{1}H–^{13}C plane for a particular backbone N–H spin pair in the 4D frequency space. The chemical shift

axes of the RD-HCcoNH and HccoNH, hCcoNH experiments are shown with arrows. We first mark all possible correlations from identified HccoNH and hCcoNH peaks. Then, a 'sum' projection onto the combined $^{1}H/^{13}C$ dimension (P_S{HccoHN — hCcoNH}, with the sums of all possible correlated frequencies) is compared with the $HCcoNH_S$ data and yields the true correlations (filled symbols in Figure 2b) and ^{1}H assignments based on the ^{13}C chemical shifts assigned by IBIS. Finally, a 'difference' projection (P_D{HccoHN — hCcoNH}, with the differences of all possible correlated frequencies) is compared with the $HccoNH_D$ data for consistency and resolving ambiguous assignments. To simplify computation, the entire procedure is implemented using sets of peak lists. Integrating this side-chain assignment strategy with automated backbone assignments and structure calculation can significantly accelerate NMR protein structure determination.

Acknowledgements

This research was supported by the National Institute of Health (grants GM47467 and RR00995 to GW, CA89940 to DR). We are grateful to Drs D. Frueh and M. Sastry for helpful comments.

References

Atreya, H.S. and Szyperski, T. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 9642–9647.
Bartels, C., Billeter, M., Güntert, P. and Wüthrich, K. (1996) *J. Biomol. NMR*, **7**, 207–213.
Bennett, A.E., Gross, J.D. and Wagner, G. (2003) *J. Magn. Reson.*, **165**, 59–79.
Bersch, B., Rossy, E., Coves, J. and Brutscher, B. (2003) *J. Biomol. NMR*, **27**, 57–67.
Brutscher, B. (2004) *J. Magn. Reson.*, **167**, 178–184.
Clowes, R.T., Boucher, W., Hardman, C.H., Domaille, P.J. and Laue, E.D. (1993) *J. Biomol. NMR*, **3**, 349–354.
Ding, K. and Gronenborn, A.M. (2002) *J. Magn. Reson.*, **156**, 262–268.
Freeman, R. and Kupce, E. (2003) *J. Biomol. NMR*, **27**, 101–113.
Gardner, K.H., Konrat, R., Rosen, M.K. and Kay, L.E. (1996) *J. Biomol. NMR*, **8**, 351–356.
Grzesiek, S., Anglister, J. and Bax, A. (1993a) *J. Magn. Reson.*, **B101**, 114–119.
Grzesiek, S. and Bax, A. (1993b) *J. Biomol. NMR*, **3**, 185–204.
Hoch, J.C. and Stern, A.S. (1996) *NMR data processing*, Wiley-Liss, New York, NY.
Hyberts, S.G. and Wagner, G. (2003) *J. Biomol. NMR*, **26**, 335–344.
Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R. and Forman-Kay, J.D. (1993) *J. Magn. Reson.*, **B101**, 333–337.
Kim, S. and Szyperski, T. (2003) *J. Am. Chem. Soc.*, **125**, 1385–1393.
Kim, S. and Szyperski, T. (2004) *J. Biomol. NMR*, **28**, 117–130.
Kupce, E. and Freeman, R. (2004) *J. Am. Chem. Soc.*, **126**, 6429–6440.
Lin, Y. and Wagner, G. (1999) *J. Biomol. NMR*, **15**, 227–239.
Logan, T.M., Olejniczak, E.T., Xu, R.X. and Fesik, S.W. (1992) *FEBS Lett.*, **314**, 413–418.
Montelione, G.T., Lyons, B.A., Emerson, S.D. and Tashiro, M. (1992) *J. Am. Chem. Soc.*, **114**, 10974–10975.
Rovnyak, D., Filip, C., Itin, B., Stern, A.S., Wagner, G., Griffin, R.G. and Hoch, J.C. (2003) *J. Magn. Reson.*, **161**, 43–55.
Rovnyak, D., Frueh, D.P., Sastry, M., Sun, Z.-Y.J., Stern, A.S., Hoch, J.C. and Wagner, G. (2004a) *J. Magn. Reson.*, **170**, 15–21.
Rovnyak, D., Hoch, J.C., Stern, A.S. and Wagner, G. (2004b) *J. Biomol. NMR*, **30**, 1–10.
Schmieder, P., Stern, A.S., Wagner, G. and Hoch, J.C. (1993) *J. Biomol. NMR*, **3**, 569–576.
Simorre, J.P., Brutscher, B., Caffrey, M.S. and Marion, D. (1994) *J. Biomol. NMR*, **4**, 325–333.
Stern, A.S., Li, K.B. and Hoch, J.C. (2002) *J. Am. Chem. Soc.*, **124**, 1982–1993.
Szyperski, T., Wider, G., Bushweller, J.H. and Wuthrich, K. (1993) *J. Am. Chem. Soc.*, **115**, 9307–9308.
Szyperski, T., Yeh, D.C., Sukumaran, D.K., Moseley, H.N.B. and Montelione, G.T. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 8009–8014.
Zimmerman, D.E., Kulikowski, C.A., Huang, Y., Feng, W., Tashiro, M., Shimotakahara, S., Chien, C., Powers, R. and Montelione, G.T. (1997) *J. Mol. Biol.*, **269**, 592–610.