Journal of Biomolecular NMR, **23:** 317–322, 2002.
KLUWER/ESCOM

Clean SEA-HSQC: A method to map solvent exposed amides in large non-deuterated proteins with gradient-enhanced HSQC

Donghai Lin[a,b], Kong Hung Sze[a], Yangfang Cui[a] & Guang Zhu[a,*]
[a]*Department of Biochemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, SAR, P. R. of China;* [b]*Center for Drug Discovery and Design, State Key laboratory of Drug Research, Shanghai Institute of Meteria Medica, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, 294 Taiyuan Rd., Shanghai 200031, P.R. of China*

Received 24 May 2002; Accepted 25 June 2002

Key words: amide hydrogen exchange, CLEAN-PM, HSQC, proteins, resonance overlap, solvent exposed amides

Abstract

The recent introduction of the SEA-TROSY experiment (Pellecchia et al. (2001) *J. Am. Chem. Soc.*, **123**, 4633–4634) can alleviate the problem of resonance overlap in $^{15}N/^{2}H$ labeled proteins. This method selectively observes solvent exposed amide protons with a SEA element. However, SEA-TROSY spectra may be contaminated with exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons and longitudinal relaxation contributions. Furthermore, for non-deuterated proteins or protein-ligand complexes, SEA-TROSY spectra may contain NOE contributions from aliphatic protons. In this communication, a modified version of the SEA element, a Clean SEA element, is introduced to eliminate these artifacts.

Recently, a new method, SEA (Solvent Exposed Amides)-TROSY (Pellecchia et al., 2001), was proposed to resolve the problem of resonance overlap in very large proteins. This new method is based on an idea that, in NMR binding studies, only amides that are exposed to the solvent are of interest, whereas those buried in the interior of the protein are not likely to be involved in intermolecular interactions. Since the solvent exposed amide protons can exchange with water rapidly, surface residues can be readily identified by only recording solvent exchangeable resonances in HSQC-type or TROSY-type experiments.

To achieve this goal, a double ^{15}N filter (Otting and Wüthrich, 1990; Breeze, 2000) is used to eliminate all the magnetization generated from amide protons. In this case, water magnetization is not affected by the ^{15}N filter and allowed to exchange with amide protons during a succeeding mixing period (τ_m). Perdeuteration of the protein sample is a prerequisite for the implementation of the SEA selection in the NMR studies of large proteins. However, for non-deuterated proteins or protein-ligand complexes, SEA-TROSY spectra may contain NOE contributions from aliphatic protons. Furthermore, SEA-TROSY spectra may be contaminated with exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons (Hwang et al., 1997) and artifacts due to the longitudinal relaxation process during τ_m (Gemmecker et al., 1993).

In this communication, we present Clean SEA, a modified version of the SEA element, which not only effectively eliminates these NOE contributions but also suppresses the longitudinal relaxation contributions by using an appropriate phase cycling scheme. NOE contributions are eliminated by applying a spin-echo filter (Mori et al., 1996, 1997) integrated with a double ^{15}N filter for ^{15}N or $^{15}N/^{13}C$ labeled samples, or a double $^{13}C/^{15}N$ filter for $^{15}N/^{13}C$ labeled samples (Ikura et al., 1992; Gemmecker et al., 1992, 1993; Grzesiek et al., 1993) along with the Phase-Modulated CLEAN chemical EXchange (CLEANEX-PM) spin locking sequence during τ_m (Hwang et al., 1997, 1998). Therefore, perdeuteration of the sample is not

*To whom the correspondence should be addressed.
E-mail: gzhu@ust.hk

required. The Clean SEA element, coupled with a water flip-back version of 2D gradient-enhanced HSQC or TROSY scheme (Zhu et al., 1999), can be used to select the signals from solvent exposed residues.

The Clean SEA-HSQC pulse scheme shown in Figure 1A starts with a double $^{15}N/^{13}C$ filter (Gemmecker et al., 1993), which serves to eliminate all the magnetization generated from protons attached to nitrogen and carbon in $^{15}N/^{13}C$ labeled proteins. A pair of gradient pulses, g1, is applied at the beginning and the end of the double ^{15}N filter to prevent water magnetization loss caused by radiation damping. After the filter, a gradient pulse, g2, is used to remove non-zero-order coherences. Water magnetization is turned back onto the z axis by the second 90° proton pulse and is not affected by the filter. It is allowed to exchange with amide protons during the mixing period. The CLEANEX-PM mixing scheme is used to eliminate both exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons and intramolecular NOE contributions from aliphatic protons. The underlying principle is that NOE peaks have different polarities in NOE-type experiments and in ROE-type experiments, whereas the exchange peaks are of the same sign. For macromolecules tumbling in the slow-motion limit, the ratio of cross-relaxation rates for ROE versus NOE is 2 to -1. Proper manipulation of magnetization trajectories can remove intramolecular NOEs/ROEs and exchange-relayed NOEs/ROEs simultaneously, leaving only pure exchange contributions (Griesinger et al., 1988). It has been pointed out that NOEs from slowly-moving bound water, which is integrated inside the protein structure, are suppressed in CLEANEX-PM, while NOEs from hydration water, which tumbles fast and has a short residence time in contacting protein surfaces, should appear as negative peaks (Hwang et al., 1997).

The pulse trains in CLEANEX-PM shown in Figure 1A are started from the x direction, while water magnetization prior to τ_m is aligned along the y axis. Under the action of CLEANEX-PM, the final positions of the spins are locked in the yz plane toward the y axis (>90%, Hwang et al., 1997). Radiation damping during τ_m is removed by a weak gradient, g3 (Sklenar et al., 1995). In Clean SEA-HSQC spectra, in which only solvent exposed amide protons are visible, the intensities, I, are related to their exchange rates.

The exchange rates depend on a number of factors including the secondary and tertiary structures in proteins, the amino acid type, solvent accessibility, as well as the pH and temperature of the solution (Bai et al., 1993; Dempsey, 2001). In general, the residues located in the loop region on the protein surface are in very fast exchange with water and appear as stronger peaks in the spectrum, whereas residues located in hydrogen-bonded secondary structures or buried in the interior of proteins are in relatively slow exchange with water and appear as weaker peaks or do not appear in the spectrum.

Figure 1B shows another version of the Clean SEA-HSQC pulse scheme, which integrates a double ^{15}N filter with a spin-echo filter (Mori et al., 1996, 1997). This scheme can be applied to ^{15}N or $^{13}C/^{15}N$ labeled proteins. The length of the double ^{15}N filter is extended to $6\Delta_1 = 6/^1J_{HN} = 33$ ms, during which most of the protein magnetization decays away completely because of J-coupling evolution and much shorter longitudinal relaxation time when compared with that of water protons. A bipolar gradient, g1, is applied to suppress radiation damping in the spin-echo filter. In comparison, peak intensities in the spectrum recorded with the pulse scheme of Figure 1A are about 10% stronger than those in the spectrum recorded with the pulse scheme of Figure 1B. The longitudinal relaxation of protons in water is responsible for the signal losses of the exchange peaks (Mori et al., 1996). However, the spin-echo filter in Figure 1B and the double $^{13}C/^{15}N$ filter in Figure 1A are equally effective in suppressing NOE contributions from aliphatic protons (data not shown).

The longitudinal relaxation process during the mixing period can also contribute to the exchange spectra (Gemmecker et al., 1993). Amide protons tend to restore the Boltzmann equilibrium magnetization during τ_m by longitudinal relaxation, which normally obscures the effect of the exchange with water and distorts the peak intensities, or gives rise to some non-labile amide peaks in the spectrum. The artifact can be effectively suppressed by applying appropriate phase cycling to the first 90° proton pulse and the receiver (Figure 1). The water magnetization prior to τ_m is alternatively aligned along the $+y$ and $-y$ axis in subsequent scans. This leads to the alternating exchange of positive and negative y magnetization for labile amide protons, while longitudinal relaxation always creates signals of the same sign. Subtraction of alternating scans then eliminates the longitudinal relaxation contributions, leaving only signals caused by exchange with the water (Gemmecker et al., 1993).

As an application, we performed Clean SEA-HSQC experiments on a 0.8 mM uniformly $^{15}N/^{13}C$ labeled sample of the ribosome-inactivating protein,

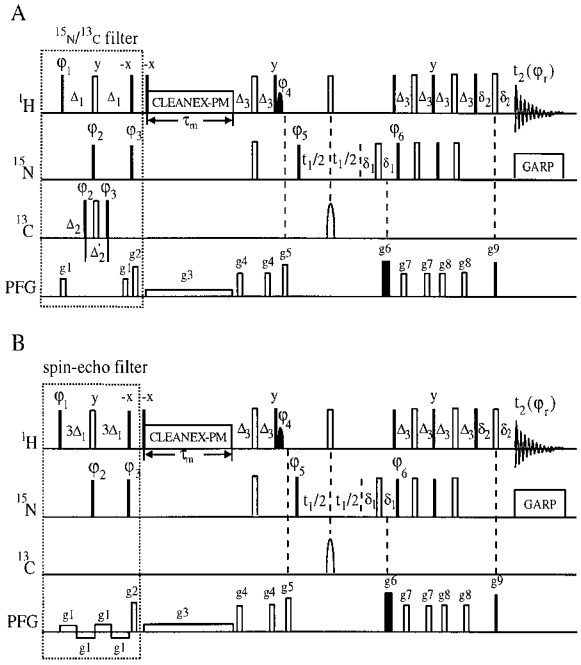

Figure 1. Pulse sequences for selective observation of solvent exposed amide protons with gradient-enhanced HSQC (Clean SEA-HSQC). Narrow and wide squares represent 90° and 180° hard pulses, respectively. Pulse phases are along x, unless otherwise specified. (A) A double $^{15}N/^{13}C$ filter serves to eliminate NOE contributions from aliphatic protons in $^{15}N/^{13}C$ labeled proteins. Radiation damping in the filter is eliminated by a pair of gradient pulses, g1 (1 ms × 1.5 G/cm). Non-zero-order coherences after the filter are removed by a gradient pulse, g2 (1 ms × 15 G/cm). The CLEANEX-PM scheme is applied to eliminate both exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons and intramolecular NOE contributions from aliphatic protons. Radiation damping during τ_m is removed by a weak gradient, g3 (0.1 G/cm). The 7th proton pulse ($\varphi 4$) is a water-selective flip-back 90° pulse (one-lobe sinc function, 1–2 ms). Two gradient pulses, g6 (2.5 ms × 28 G/cm) and g9 (0.25 ms × 27.9 G/cm), serve to select the coherence-transfer pathway. The delays and strengths of the other gradient pulses are g4 = (0.5 ms × 10 G/cm), g5 =(1 ms × 15 G/cm), g7 = (1 ms × 5 G/cm), g8 = (1 ms × 10 G/cm). The phase cycle is as follows: $\varphi 1$ = 8(x), 8(-x); $\varphi 2$ = x, -x; $\varphi 3$ = 2(x), 2(-x); $\varphi 4$ = 8(x), 8(-x); $\varphi 5$ = 4(y), 4(-y); $\varphi 6$ = -x; φ_r = 4(x), 8(-x), 4(x). The delays used in the pulse sequence are $\Delta_1 = 5.5$ ms, $\Delta_2 = 3.7$ ms, $\Delta_2' = 3.6$ ms, $\Delta_3 = 2.3$ ms. The length of the double ^{15}N filter is set to $2\Delta_1 = 1/^1J_{NH} =$ 11.0 ms. The proton carrier is switched during the CLEANEX-PM pulse train from the water resonance to the middle of the NH range. The carbon carrier is placed on the aliphatic carbons (35 ppm) at the beginning of the pulse sequence, then shifted to the center between $^{13}C_\alpha$ and ^{13}CO (116 ppm) after the double $^{15}N/^{13}C$ filter. A selective ^{13}C 180° pulse on $^{13}C_\alpha$ and ^{13}CO is applied to refocus ^{15}N-^{13}C couplings during t_1. The ^{15}N decoupling is accomplished by GARP (Shaka et al., 1985). (B) The spin-echo filter combined with the double ^{15}N filter serves to eliminate NOE contributions from aliphatic protons in ^{15}N or $^{15}N/^{13}C$ labeled proteins. The gradient, g1 (0.1 G/cm), is applied to remove radiation damping in the filter. The length of the double ^{15}N filter is set to $6\Delta_1 = 33.0$ ms.

Trichosanthin Delta C7 (TCS, 247 residues) (Wang et al., 1986) in an NMR buffer (20 mM sodium phosphate, pH 6.4, in 95% H_2O/5% D_2O). The 3D crystal structure of TCS has been solved with a PDB ID of 1J4G. All the 2D NMR experiments were performed on a Varian Unity INOVA 750 spectrometer at 30 °C. The mixing time was 100 ms; the prescan delay was 2.5 s and the spectral widths were 9756 and 2355 Hz for ^1H and ^{15}N dimensions, respectively. 128(t_1) × 1024(t_2) complex points were recorded and processed with NMRPipe (Delaglio et al., 1995). As a reference, the spectrum shown in Figure 2A is a regular gradient-enhanced HSQC spectrum. Severe resonance overlap increases the difficulty in resonance assignment. The spectrum shown in Figure 2B was recorded with a modification of the pulse scheme shown in Figure 4 of the supplementary material which applies the double ^{15}N filter but does not apply the double ^{13}C filter and the CLEANEX-PM scheme so that the water magnetization is aligned along the z axis prior to τ_m. The spectrum shown in Figure 2C is a Clean SEA-HSQC spectrum recorded with the pulse scheme of Figure 1A in which both the double ^{15}N filter/^{13}C filter and the CLEANEX-PM scheme are applied together. The spectrum is simplified significantly because, in principle, only water-accessible amide protons are visible. In comparison with the Clean SEA-HSQC spectrum (Figure 2C), the spectrum shown in Figure 2B is clearly contaminated with both significant intramolecular NOE contributions from aliphatic pro-

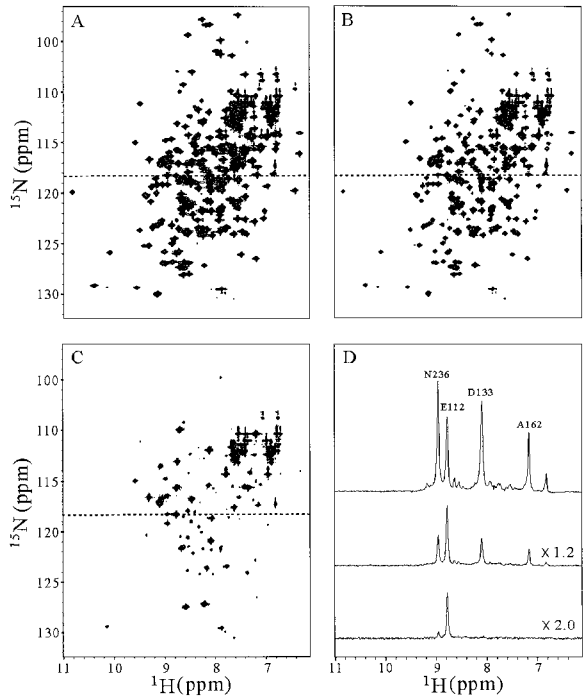

Figure 2. (A) Gradient-enhanced HSQC spectrum. (B) Spectrum recorded with the pulse scheme of Figure 4 in the supplementary material but with the double ^{13}C filter turned off, which clearly identifies the peaks having significant intramolecular NOE contributions from aliphatic protons and exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons as neither the double ^{13}C filter or the CLEANEX-PM scheme is applied. (C) Clean SEA-HSQC spectrum recorded with the pulse scheme of Figure 1A. Spectra B and C were drawn with contour levels being 40% and 20% of that in Spectrum A, respectively. (D) 1D traces along ω_2 (^{1}H) taken at the positions indicated by the dashed lines in Spectra A, B, and C counted from the top. The mixing time was 100 ms.

tons and exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons. These artifacts are effectively suppressed in the Clean SEA-HSQC spectrum. Figure 2D shows 1D traces along ω_2 (^{1}H) taken at the positions indicated by the dashed lines in the three spectra (Figures 2A–C), which go through the peaks of N236, E112, D133 and A162. The 3D structure of TCS shows that E112 is a surface residue, while A162 is buried deeply in the interior of the protein, N236 and D133 are buried but are close to the protein surface. As expected, in the Clean SEA-HSQC spectrum, E112 appears as a strong peak, N236 and D133 appear as very weak peaks, whereas A162 is invisible. The relative water accessibilities corresponding to these peaks are 39.7%, 9.6%, 11.8% and 0.6% calculated with the use of MOLMOL software package (offered by the group of K. Wüthrich at the ETH Zürich).

In order to illustrate the effect of the double ^{13}C filter, a spectrum (Figure 3A) was recorded with the pulse scheme shown in Figure 4 of the supplementary material, which applies both the double ^{15}N filter and the double ^{13}C filter, but does not apply the CLEANEX-PM scheme. Figures 3B and C are portions of the spectra shown in Figures 2B and C, respectively. The 3D structure of TCS shows the stronger peaks in Figure 3C are related to the surface residues (S13, Y32, L41, D78, M100, A117, A144, K165, V175, I201, R222), while the weaker peaks are related to the buried residues that are very close to the protein surface (N20, A150, V213). It should be noted that the peak from a surface residue, I201, is buried in a group of overlapped peaks in the regular HSQC spectrum shown in Figure 2A. However, in the Clean SEA-HSQC spectrum shown in Figure 2C, I201 appears as an isolated peak. Therefore, it is shown that SEA-type experiments may alleviate the problem of resonance overlap and benefit resonance assignment.

The intramolecular NOE peaks can be identified readily in a comparison between Figures 3A and B. These peaks with significant intramolecular NOE contributions from aliphatic protons are labeled with underlines in Figure 3B. Similarly, the exchange-relayed NOE peaks can be also identified easily in a compar-

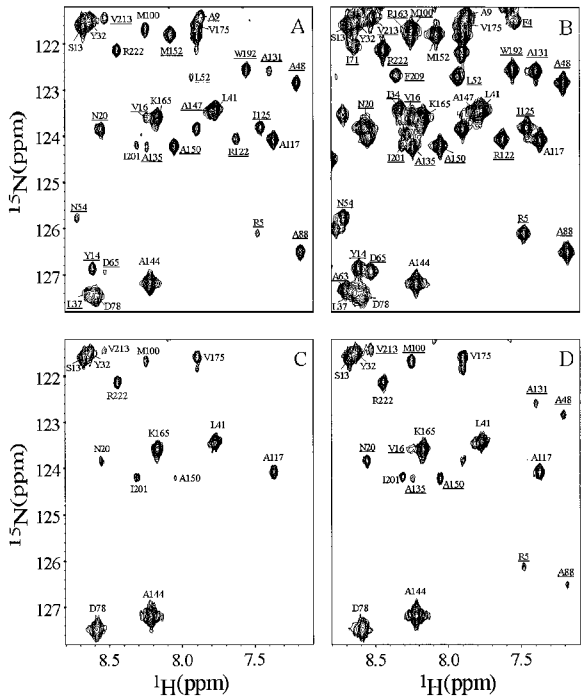

Figure 3. (A) Spectrum recorded with the pulse scheme shown in Figure 4 of the supplementary material, which applies a double $^{15}N/^{13}C$ filter but not the CLEANEX-PM scheme. (B) Spectrum recorded with the same pulse scheme but turning off the double ^{13}C filter. (C) Clean SEA-HSQC spectrum recorded with the pulse scheme of Figure 1A in which both the double $^{15}N/^{13}C$ filter and the CLEANEX-PM scheme are applied together. (D) Spectrum recorded with the pulse scheme of Figure 1A but without phase cycling of the first 90° proton pulse and the receiver. Spectra C and D were drawn with contour levels being 50% of those in Spectra A and B. Spectrum A is contaminated with exchange-relayed NOE contributions from fast exchanging hydroxyl or amine protons (labeled with underlines), whereas spectrum B is contaminated with not only exchange-relayed NOE contributions, but also intramolecular NOE contributions from aliphatic protons (labeled with underlines). Spectrum D is contaminated with longitudinal relaxation contributions during τ_m (labeled with underlines). The 3D structure of TCS shows that the visible residues in the Clean SEA-HSQC spectrum are those on or very near to the protein surface.

ison between Figures 3A and C. These peaks with significant exchange-relayed NOE contributions are labeled with underlines in Figure 3A. These residues with significant NOE contributions are grouped as three categories: (1) The residues buried in the interior of the protein (F4, Y14, I34, L37, L52, N54, D65, I71, A88, I125, A131, A135, A147, M152, R163, W192), which have low water accessibility estimated by the program MOLMOL; (2) the buried residues close to the protein surface (A9, N20, A48, R122, A150, N205, F209, V213); (3) the surface residues that are located in the hydrogen-bonded secondary structures of β sheets (R5) or helixes (V16) and have larger protection factors. It is clear that both intramolecular NOE contributions and exchange-relayed NOE contributions are effectively eliminated in the Clean SEA-HSQC spectrum shown in Figure 3C.

To demonstrate the effect of the double ^{13}C filter in Clean SEA-HSQC experiments, we recorded a spectrum with the pulse scheme of Figure 1A but with the double ^{13}C filter turned off. Although the CLEANEX-PM scheme has been used to suppress both intramolecular NOE contributions and exchange-relayed NOE contributions, this spectrum may be still contaminated with some minor artifacts due to intramolecular NOE contributions (data not shown). Therefore, it is suggested that the CLEANEX-PM scheme should be applied with the double ^{13}C filter or the spin-echo filter to eliminate intramolecular NOE contributions more effectively.

Another potential artifact is generated from longitudinal relaxation during τ_m, which is demonstrated in a spectrum (Figure 3D) recorded with the pulse scheme of Figure 1A but without phase cycling of the first 90° proton pulse and the receiver. This spectrum clearly indicates the artifacts due to longitudinal relaxation contributions, which were effectively suppressed in the Clean SEA-HSQC spectrum (Figure 3C). The peaks with significant longitudinal relaxation contributions are labeled with underlines in Figure 3D. Most

of these artifacts are related to the residues buried in the interior of the protein (A88, A135) or the buried residues that are close to the protein surface (N20, A48, A131, A150, V213).

This new Clean SEA element can be readily combined with triple resonance NMR experiments such as 3D HNCA, HNCACB, HNCOCA, HNCOCACB and HNCO to obtain backbone resonance assignment for solvent exposed residues. The pulse scheme for the 3D Clean SEA-HNCO experiment is shown in the supplementary material (Figure 5). For binding studies, the resonance assignment process can be facilitated by the three-dimensional structure of the protein obtained with X-ray crystallography, if available.

Acknowledgements

This work was supported in part by grants from the Research Grants Council of Hong Kong (6199/99M and 6208/00M) to GZ and from the National Natural Science Foundation of China (No 19975038) to DL.

References

Bai, Y., Milne, J.S., Mayne, L. and Englander, S.W. (1993) *Protein Struct. Func. Gen.*, **17**, 75–86.
Breeze, A.L. (2000) *Prog. NMR Spectrosc.*, **36**, 323–372.
Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.
Dempsey, C.E. (2001) *Prog. NMR Spectrosc.*, **39**, 135–170.
Gemmecker, G., Jahnke, W. and Kessler, H. (1993) *J. Am. Chem. Soc.*, **115**, 11620–11621.
Gemmecker, G., Olejniczak, E.T. and Fesik, S.W. (1992) *J. Magn. Reson.*, **96**, 199–204.
Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 627–638.
Hwang, T.-L., Mori, S., S. Shaka, A.J. and van Zijl, P. C.M. (1997) *J. Am. Chem. Soc.*, **119**, 6203–6204.
Hwang, T.-L., van Zijl, P. C.M. and Mori, S. (1998) *J. Biomol. NMR*, **11**, 221–226.
Ikura, M. and Bax, A. (1992) *J. Am. Chem. Soc.*, **114**, 2433–2400.
Mori, S., Abeygunawardana, C., Berg, J.M. and van Zijl, P.C.M. (1997) *J. Am. Chem. Soc.*, **119**, 6844–6852.
Mori, S., Berg, J.M. and van Zijl, P.C.M. (1996) *J. Biomol. NMR*, **7**, 77–82.
Otting, G.. and Wüthrich, K. (1990) *Quart. Rev. Biophys.*, **23**, 39–96.
Pellecchia, M., Meininger, D., Shen, A.L., Jack, R., Kasper, C.B. and Sem, D.S. (2001) *J. Am. Chem. Soc.*, **123**, 4633–4634.
Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.
Sklenar, V. (1995), *J. Magn. Reson.*, **A114**, 132–135.
Wang, Y., Qian, R.Q. and Gu, Z-W. et al. (1986) *Pure Appl. Chem.*, **58**, 789–798.
Zhu, G., Kong, X. and Sze, K. (1999). *J. Biomol. NMR*, **13**, 77–81.