Journal of Biomolecular NMR (2006) 36:157–168
DOI 10.1007/s10858-006-9077-y

Article

Random sampling of evolution time space and Fourier transform processing

Krzysztof Kazimierczuk[a], Anna Zawadzka[a,b], Wiktor Koźmiński[a,*] & Igor Zhukov[c]
[a]*Department of Chemistry, Warsaw University, Pasteura 1, 02-093 Warszawa, Poland;* [b]*Department of Biophysics, Institute of Experimental Physics, Warsaw University, Żwirki i Wigury 93, 02-089 Warszawa, Poland;* [c]*Institute of Biochemistry and Biophysics, Polish Academy of Sciences, ul. Pawińskiego 5a, 02-106 Warszawa, Poland*

Received 28 April 2006; Accepted 10 August 2006

Key words: fast multidimensional NMR, Fourier Transform, random sampling, ubiquitin

Abstract

Application of Fourier Transform for processing 3D NMR spectra with random sampling of evolution time space is presented. The 2D FT is calculated for pairs of frequencies, instead of conventional sequence of one-dimensional transforms. Signal to noise ratios and linewidths for different random distributions were investigated by simulations and experiments. The experimental examples include 3D HNCA, HNCACB and ^{15}N-edited NOESY-HSQC spectra of ^{13}C ^{15}N labeled ubiquitin sample. Obtained results revealed general applicability of proposed method and the significant improvement of resolution in comparison with conventional spectra recorded in the same time.

Introduction

Multidimensional NMR spectroscopy is known to be one of most powerful tools in biomolecular research. It is, however, also one of most time-consuming methods. Despite growing number of high field spectrometers used in biomolecular laboratories, still conventional multidimensional experiment lasts 10s of hours. This is because of sampling requirements (Szyperski et al., 2002), which, contrary to sensitivity limits, increase with B_0 value.

In order to generate N-dimensional data set with $(N-1)$ indirectly sampled time domains one needs to acquire one-dimensional FID for each combination of equally spaced indirectly sampled evolution times. Then, spectrum can be achieved by sequence of N one-dimensional Fourier transforms (Ernst and Anderson, 1966) Equation 1, with respect to each time dimension separately.

$$S(\omega) = \int_{-\infty}^{\infty} dt\, f(t) \exp(-i\omega t) \tag{1}$$

For $N-1$ transforms, each requiring m points, it is necessary to acquire regarding requirements of quadrature $2^{N-1} \times m^{N-1}$ one-dimensional FIDs, which means that time of experiment increases exponentially with number of dimensions.

Miscellaneous methods of non-conventional probing of evolution time space developed in last few years allowed for an acceleration of multidimensional experiments. Amongst them, applications using radial sampling are most often employed and investigated. This kind of sparse sampling is based on the idea of Accordion Spectroscopy (Bodenhausen and Ernst, 1981, 1982) and

*To whom correspondence should be addressed. E-mail: kozmin@chem.uw.edu.pl

backprojection imaging technique (Lauterbur, 1973). The simplest example of recovering of spectral parameters in frequency domain is direct evaluating of frequencies from projection spectra by solving systems of linear equations for each peak in the new variants of Reduced Dimensionality (RD) techniques (Ding and Gronenborn, 2002; Kim and Szyperski, 2003; Koźmiński and Zhukov, 2003). However, owing to direct interpretation and simple analysis, numerous approaches for the recovering spectrum of full dimensionality from the set of projections were proposed (Kupče and Freeman, 2003, 2004a, b; Freeman and Kupče, 2003; Coggins et al., 2004, 2005). One can choose between deterministic and statistical methods, as it is possible to geometrically reconstruct spectrum from projection data (deterministic case) or to fit spectrum to be compatible with set of projections (Yoon et al., 2006). Additive backprojection is the simplest deterministic method, but suffers from line-broadening, ridges and false peaks. More sophisticated deterministic approach to reconstruction is lowest-value algorithm (Freeman and Kupče, 2003), variants of multiway decomposition (Malmodin and Billeter, 2005a, b) and iterative frequency identification (Eghbalnia et al., 2005). Other methods of fast acquisition include: multidimensional decomposition (Orekhov et al., 2003; Luan et al., 2005; Tugarinov et al., 2005), filter diagonalization (Mandelstam et al., 1998; Armstrong et al., 2005), maximum entropy reconstruction (Barna et al., 1987; Rovnyak et al., 2004; Sun et al., 2005; Delsuc and Tramesel, 2006; Frueh et al., 2006), and spatially encoded chemical shift evolution followed by spatially resolved acquisition (Frydman et al., 2002). Also, algorithms of FFT of non-equispaced data based on polynomial interpolation (Dutt and Rokhlin, 1995) were employed to reconstruct spectrum from sparse time domain data sets (Marion, 2005).

In our recent communication (Kazimierczuk et al., 2006) we presented the application of FT with respect to two or more time variables made in a single step as a reliable approach to processing of sparsely sampled multidimensional NMR data sets. For example in 2D case, instead of making sequence of two Fourier transforms:

$$f(t_1,t_2) \overset{\mathrm{FT}_1}{\rightarrow} S_1(t_1,\omega_2) \overset{\mathrm{FT}_2}{\rightarrow} S(\omega_1,\omega_2)$$

these two transforms are calculated in one step, i.e., real discrete cosine transform of signal $f(t_1,t_2)=\cos(\Omega_1 t_1)\cos(\Omega_2 t_2)$ is given by Equation 2 where summation occurs for each pair of frequencies:

$$S(\omega_1,\omega_2) = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} f(t_1,t_2)\cos(\omega_1 t_1)\cos(\omega_2 t_2)w(t_1,t_2) \tag{2}$$

In other words for each point of frequency domain (ω_1,ω_2), the product of time domain signal $f(t_1,t_2)$ and $\cos(\omega_1 t_1)\cos(\omega_2 t_2)$ should be calculated for each point of evolution time space (t_1 and t_2) and summed. In principle, to calculate above integral more accurately for non-equispaced data, weighting terms $w(t_1,t_2)$ corresponding to distribution of (t_1,t_2) points may be used. As shown below the weighting terms are optional. They could be considered as a apodization function or used in integration procedure as two-dimensional time surface elements (triangles).

Since, in the most of NMR experiments, the frequency offset is placed at the center of frequency region of interest the quadrature detection is necessary for determination of sign of peak frequencies. Natural extension of well known 1D quadrature detection (described by real part of complex FT) for such 2D transform is acquiring four sine–cosine modulations of indirectly measured time dimension signal, their appropriate transformation and coaddition (Equation 3a–d):

$$f(t_1,t_2)=\cos(\Omega_1 t_1)\cos(\Omega_2 t_2)\rightarrow S(\omega_1,\omega_2) = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} f(t_1,t_2)\cos(\omega_1 t_1)\cos(\omega_2 t_2)w(t_1,t_2) \tag{3a}$$

$$f(t_1,t_2)=\cos(\Omega_1 t_1)\sin(\Omega_2 t_2)\rightarrow S(\omega_1,\omega_2) = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} f(t_1,t_2)\cos(\omega_1 t_1)\sin(\omega_2 t_2)w(t_1,t_2) \tag{3b}$$

$$f(t_1,t_2)=\sin(\Omega_1 t_1)\sin(\Omega_2 t_2)\rightarrow S(\omega_1,\omega_2) = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} f(t_1,t_2)\sin(\omega_1 t_1)\sin(\omega_2 t_2)w(t_1,t_2) \tag{3c}$$

$$f(t_1,t_2)=\sin(\Omega_1 t_1)\cos(\Omega_2 t_2)\rightarrow S(\omega_1,\omega_2)$$
$$=\sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} f(t_1,t_2)\sin(\omega_1 t_1)\cos(\omega_2 t_2)w(t_1,t_2) \quad (3d)$$

Analogically to 1D case, quadrature notation for 2D 'single step' transform can be presented in Equation 4, using Quaternion numbers (Hamilton, 1847), which are non-commutative four-dimensional equivalents of complex numbers.

$$S(\omega_1,\omega_2)=$$
$$\sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}}\exp(-i\Omega_1 t_1)f(t_1,t_2)\exp(-j\Omega_2 t_2)w(t_1,t_2) \quad (4)$$

where: $i^2=j^2=k^2=-1$ and $ij=-ji=k, jk=-kj=i$, $ki=-ik=j$ and $f(t_1,\ t_2)=\exp(i\Omega_1\ t_1)\exp(j\Omega_2 t_2)$.

Required spectrum with purely absorptive lineshapes is the real part of Equation 4. Note that traditional acquisition of 2D data sets requires recording of the same four modulations for each t_1/t_2 point, however, in this case the data is treated by the sequence of 1D complex FT using the 1D quadrature.

The multidimensional FT performed in a single step, in contrast to the sequence 1D FT does not require data points distributed on a straight lines in evolution time space. Thus demands for time domain sampling are less restrictive than in conventional case. In fact it is possible to sample it in arbitrarily chosen way. Previously we have presented spectra achieved by transformation of data points deployed radially and spirally in evolution time space, showing that the later give significantly better results. In the present work we prove the ability of Fourier Transform as formulated above to process 3D NMR data sets with randomly sampled two time domains.

The idea of random digital sampling is not new, many variants of Digital Alias-free Signal Processing (DASP) methods have been studied for several decades (Marvasti, 2001). Because of lost of data or its inaccessibility, 1D FT of sparse, random data was employed in astronomy, geophysical sciences, medicine, and other disciplines. On the other hand, in some applications (including NMR) non-uniform sampling may be used to perform digital analysis over wide band of frequencies using relatively slow sampling rate (even below Nyquist density). Degradation of spectra achieved from uniform sampling with random deviations was also studied (Berkovitz and Rusnak, 1992), as every spectrum is in fact sampled randomly – and sometimes effects of sampling jitter on FFT processing are strong. Randomly sampled data sets may be processed in several different ways.

First, interpolation by polynomials may allow one to achieve equispaced data points which could be processed by standard FFT algorithms. This approach was discussed recently not only in general papers on signal processing (Marvasti, 1996) but also in the particular case of multidimensional NMR (Marion, 2005). Main disadvantage of interpolation is the presence of artifacts caused by crude approximation of periodic functions by polynomials (even of high order). Moreover, it allows non-uniform sampling in one dimension only (which can be omitted by employing RD methods).

Second approach, which was not employed in multidimensional NMR, is 1D direct Fourier Transformation of data, not using FFT algorithms. Properties of such transformation are quite well known and were discussed in details before (Tarczynski and Allay, 2004; Tarczynski and Qu, 2005). Two kinds of weighting may be used here: weighted samples (WS) and weighted probability (WP). First of them is based on uniform probability density function (PDF) used to generate time points. Then, weighting is applied in spectrum estimator formula analogically to Equation (2). It should be noted, that weighting terms $w(t_1,\ t_2)$ in WS are not connected with accurate values of distance between points. They are rather values of non-uniform PDF (such as exponential or Gaussian) with coordinates of $(t_1,\ t_2)$. This operation is similar to well known in NMR FID weighting procedure – time domain signal is multiplied point after point by values of an arbitrary chosen function of time. Second approach, is Weighted Probability method – sampling PDF is not uniform (but Gaussian or exponential, for example), but no weighting is used during Transformation (i.e., $w(t_1,\ t_2)=1$). Both approaches are presented in experimental examples and simulations below. Moreover, we tried another solution – as the distances between points randomly distributed in time space are known, this information

could be used in some cases for improving accuracy of integration. In other words FT integral can be calculated numerically employing one of well known methods (mid-point or trapezoidal rule in 1D case). For one-step two-dimensional FT, value of integrant is multiplied by appropriate part of evolution time surface. In the case of random sampling the surface can be divided into triangles (Yeh, 1997), each built of three vectors connecting experimental time points (Figure 1). Then, integration could be performed by summing volumes of each truncated prism according to Equation 5:

$$S(\omega_1, \omega_2) = \sum_{n=1}^{\text{n.o.t.}} \Delta^n_{pqr} \frac{1}{3} [\exp(-i\omega_1 t_1^p) f(t_1^p, t_2^p) \exp(-j\omega_2 t_2^p) + \exp(-i\omega_1 t_1^q) f(t_1^q, t_2^q) \exp(-j\omega_2 t_2^q) + \exp(-i\omega_1 t_1^r) f(t_1^r, t_2^r) \exp(-j\omega_2 t_2^r)] \quad (5)$$

where, n.o.t. is the number of triangles, Δ^n_{pqr} means surface of triangle number n with apexes p, q, r, and $f(t_1^p, t_2^p)$, $f(t_1^q, t_2^q)$, $f(t_1^r, t_2^r)$ are signal amplitudes. The Δ^n_{pqr} term corresponds to $w(t_1, t_2)$ in Equation 4. Note, that it is not equal to either WS or WP method. Other methods of processing of non-linearly sampled signals employ theory of uniform almost periodic functions (Ferreira, 1999).

Simulations

Conventional processing of *N*D-NMR data requires equally spaced data points placed on rectangular grid due to sequential usage of 1D FT in the multidimensional spectra and necessities of commonly employed Fast Fourier Transform algorithm. Therefore, in each dimension the maximum distance between two neighboring points is defined by Nyquist Theorem e.g. it is inversely proportional to expected spectral width (SW). On the other hand, the spectral resolution is proportional to maximum value of evolution time period. In consequence it is not possible to increase SW without decreasing spectral resolution (keeping constant number of time points). If Nyquist Theorem is not fulfilled then false (folded) signals appear. This is an intrinsic feature of conventional sampling and independent of processing way ('one-step' 2D transform of such data will give the same effect). However, the new approach allows for any kind of sampling of evolution time space and if it is non-conventional then there is no sharp limit for SW. Instead, even for perfect 2D signal $f(t_1, t_2) = \cos(\Omega_1 t_1)\cos(\Omega_2 t_2)$ without any noise the non-zero value of integral estimated by summation given in Equation 4 would appear for frequency pairs (ω_1, ω_2) other than (Ω_1, Ω_2). This can be shown on the example of 1D FT of conventionally (Figure 2a–e) and randomly (Figure 2f–j) sampled simulated FID of constant length with varying number of points. In the case of equispaced time domain points undersampling, i.e., sampling with the rate lower than Nyquist frequency, results in signal folding. The FT of the same number of input data points but spread at random positions provide spectra with the single peak of correct frequency, however perturbed by

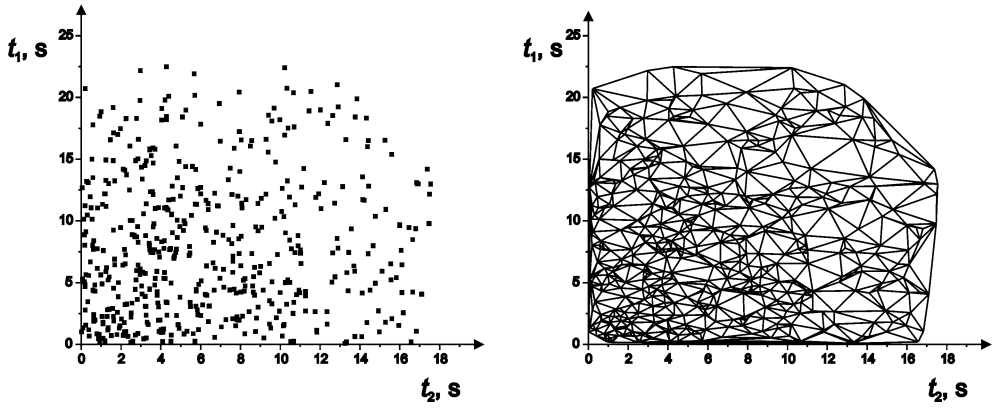

Figure 1. (a) The distribution of 512 time domain data points with probability density function: $\exp(-t^2/\sigma^2), \sigma = 0.5$, (b) triangulation of evolution time surface using points shown in (a).

an additional noise. On the other hand the resolution of conventionally achieved multidimensional spectrum is also limited by Nyquist Theorem. Upper limit for dwell time in each dimension limits the maximum achievable evolution time (t_{max}). Consequently, the resolution is proportional to acquisition time of periodic signal. Random sampling of the evolution time space gives no folded signals and it is possible to increase values of t_{max} together with resolution as much as desired with the expense of increasing noise.

In the next step we investigated the effects of number of points and mode of random distribution on signal to noise ratio in two-dimensional FT. To simulate signal obtained in real experiment, one of modulations was assumed to be achieved from constant-time evolution period and for the second the typical transverse relaxation rate of aliphatic carbon nuclei in proteins was taken. Time domain points were distributed randomly over the entire evolution time space. Both coordinates of time domain points were obtained independently. Various distributions of different numbers of random points with constant t_{1max} and t_{2max} were tested. Calculations of standard deviation σ of spectrum estimator made for WS and WP methods applied to periodic function of frequency and decay rate typical for NMR experiment show that both are unbiased, i.e., frequency independent, estimators with the same σ^2 (Tarczynski and Allay, 2004). Real experimental data, however, contains thermal noise and it is better to use WP method preferring points of smaller time values (where signal to thermal noise ratio is better). The third method based on surface integration improves results for the density of data of points (defined for 2D case in Equation 6) higher than Nyquist density ρ_N (Equation 7):

$$\rho = n/(t_{1\,max} \times t_{2\,max}) \tag{6}$$

$$\rho_N = sw_1 \times sw_2 \tag{7}$$

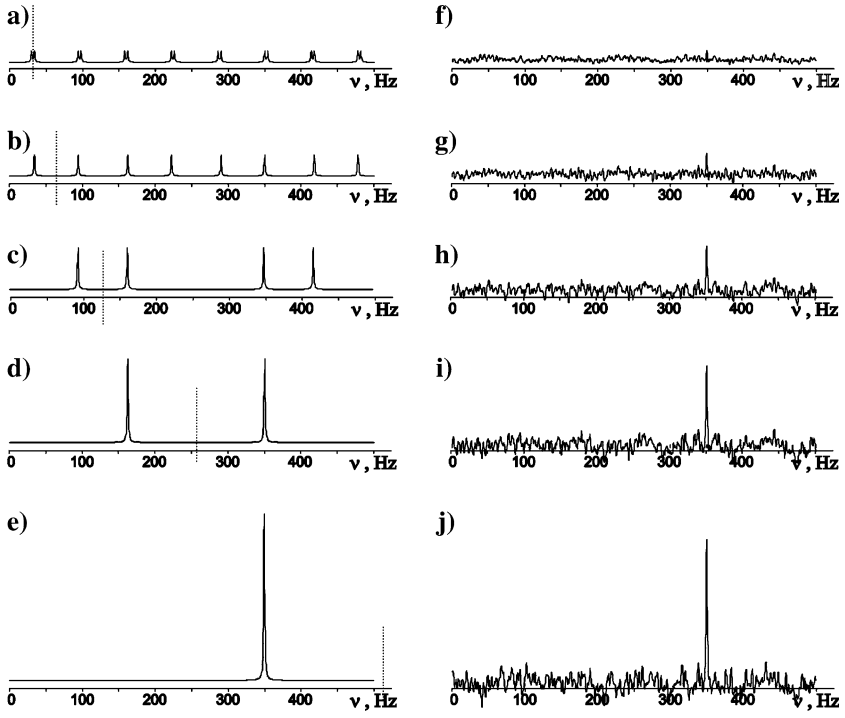

Figure 2. Simulated 1D spectra using different numbers of time domain points distributed from 0 to 1 s. The time domain signal fulfills the formula: $f(t) = \cos(2\pi 350t)\exp(-5t)$. (a–e) Conventional sampling with constant time increments of $1/(n-1)$ seconds, where n is the number of points equal to: 64, 128, 256, 512, and 1024 points, in a, b, c, d, and e, respectively. The spectra shown on traces (f–j) are calculated using the same number of randomly distributed time domain points. Gaussian distribution was used i.e. $\exp(-t^2/\sigma^2), \sigma = 0.5$. The real cosine FT was used, owing to symmetry only positive frequencies are shown. The doted lines in (a–f) denote the maximum frequencies according to Nyquist theorem.

Relative points density, can be defined as:

$$\Theta = \frac{\rho}{\rho_N} \tag{8}$$

where n is number of points, sw_1 and sw_2 are spectral widths in indirect dimensions. However, when the number of sampled data points decreases, surface integration gives more noise than WP or WS method. Moreover this noise is non-uniformly distributed in the frequency domain.

The plot of signal to noise ratio in simulated spectrum calculated using WP method and surface integration described by Equation 5 is given in Figure 3. The same behavior could be observed in 1D random case when integration using trapezoidal or mid-point rule is compared to simple summation. It is noteworthy that the sampling noise for direct summation method (WS or WP) is almost independent of relative points density Equation 8, while the S/N increases linearly with $\sqrt{n}$. In the case of Gaussian distribution of time domain points the apparent linewidths obtained in surface integration method are reduced due to overestimation of points for long evolution times. However, the strong truncation artifacts are introduced. Similar effect could be observed by employing inappropriate weighting function in the case of WS or WP method. Therefore, as in all simulated FIDs the average density of time domain points did not extend the Nyquist density, all Fourier Transforms presented here were made using WP and WS method, not integration over triangulated time domain.

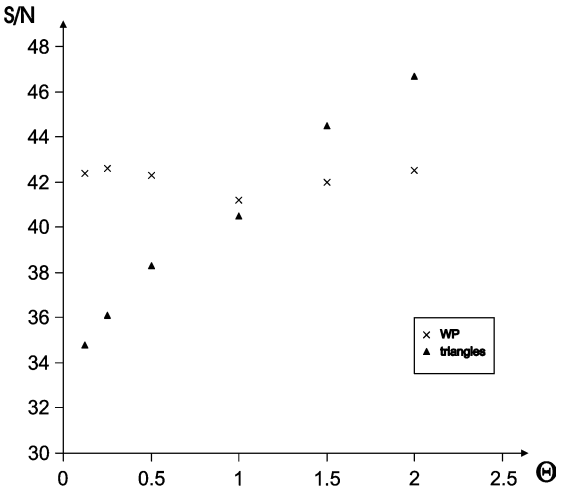

Figure 3. The plot of spectral signal to noise ratio of simulated spectrum $f(t_1, t_2) = \exp(-i \cdot 2\pi \cdot 300\,\text{Hz} \cdot t_1 - j \cdot 2\pi \cdot 300\,\text{Hz} \cdot t_2 - 50\,\text{Hz} \cdot t_2)$ in function of relative density of time domain points ($\Theta = \rho/\rho_N$) comparing: WP method and surface integration procedure (512 evolution time points of Gaussian PDF: $\exp(-t^2/\sigma^2), \sigma = 0.5$). Spectral widths and maximum evolution times were equal $sw_1 = sw_2$, $t_{1max} = t_{2max} = t_{max} = 0.02$ s. Θ was changed by varying both spectral widths (and ρ_N consequently) keeping constant number of points and evolution time surface t^2_{max} (and ρ consequently).

In the case of random 64-point sampling the linewidth is overestimated due to noise. The calculated spectra with corresponding time domain point distribution are shown in Figure 4. Signal to noise ratios and linewidths obtained from the simulations are given in Tables 1 and 2, respectively. The plot of S/N as a function of $\sqrt{n}$ presented in Figure 5a reveal linear behavior in all cases as expected (Tarczynski and Allay, 2004; Tarczynski and Qu, 2005).

The linewidth values obtained in conventional spectra are inversely proportional to maximum evolution times (i.e., to the number of points in each dimension), respectively. In non-conventional spectra obtained by FT from randomly distributed time domain points the linewidth is independent of the number of points. While the resolution achievable from limited number of time domain points with random distribution is significantly improved, the signal to noise ratio is reduced due to 'sampling' noise. Contrary to conventional 2D case, in randomly sampled 2D interferogram each time domain point contributes unique information about both involved frequencies, thus reduces relative amplitude of the 'sampling' noise.

Experimental

All the spectra presented were recorded for 1.5 mM ^{13}C, ^{15}N-double labeled human ubiquitin in 9:1 H_2O/D_2O at pH = 4.5 at 298 K. The 3D HNCA experiments were acquired on a Varian Unity Inova 400 spectrometer equipped with a Performa II z-PFG unit and using the 5 mm ^1H, ^{13}C, ^{15}N – triple resonance probehead with high power ^1H, ^{13}C, and ^{15}N $\pi/2$ pulses of 6.2, 16.0, and 49.0 μs, respectively. For 3D HNCACB and 3D ^{15}N-edited NOESY-HSQC spectra Varian Unity Plus 500 spectrometer equipped with a Performa II z-PFG unit and the 5 mm ^1H, ^{13}C, ^{15}N – triple resonance probehead with high power ^1H, ^{13}C, and ^{15}N $\pi/2$ pulses of 6.3, 13.0, and 36.0 μs, respectively, was

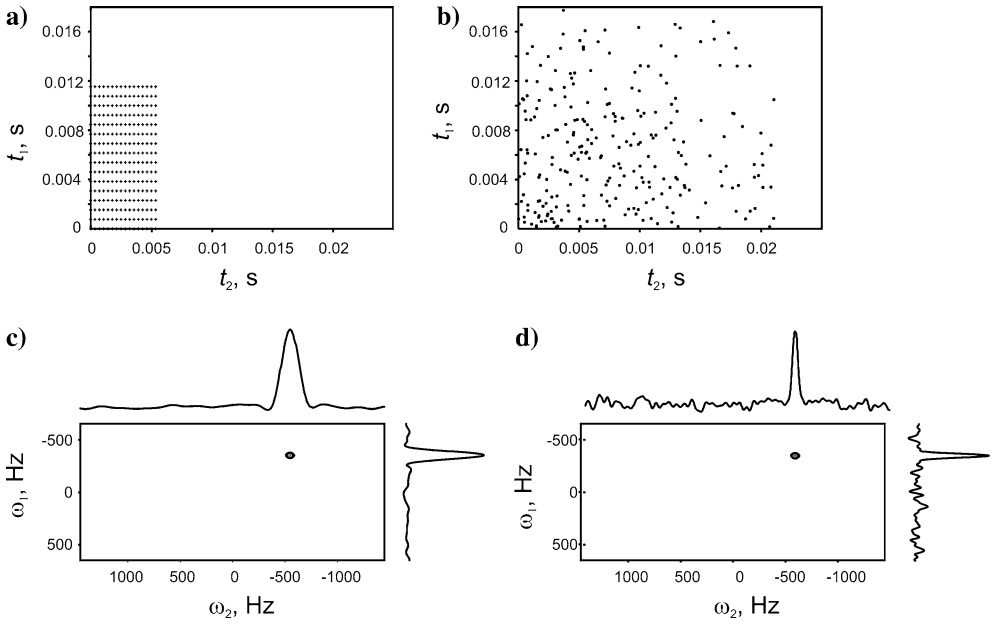

Figure 4. The distribution of time domain data points (a, b) and spectra (c, d) transformed using respective t_1/t_2 interferograms calculated according to the formula: $f(t_1, t_2) = \exp(-i \cdot 2\pi \cdot 350\,\text{Hz} \cdot t_1 - j \cdot 2\pi \cdot 500\,\text{Hz} \cdot t_2 - 50\,\text{Hz} \cdot t_2) + \eta(t_1, t_2)$, where η reflects thermal noise of Gaussian distribution. Gaussian distribution was also used for distribution of time domain data points with probability density: $\exp(-t^2/\sigma^2), \sigma = 0.5$. The number of points equal to 512 in the both cases, however in the case of conventional sampling (a and c) the maximum t_1 and t_2 are limited by number of steps and spectral width (1300×2900 Hz in F_1 and F_2, respectively), while in the case of random sampling (b and d) the maximum t_1 and t_2 are set to 22.5 and 18 ms, respectively. 128×128 frequency domain points were calculated.

employed. For the random sampling the conventional pulse sequences were adapted from the Varian Userlib BioPack package (Varian Inc., Palo Alto) by setting evolution times to randomly generated values normalized to assumed maximum evolution times. Both coordinates of t_1/t_2 points were calculated independently using generator of pseudo-random numbers implemented in GSL package (Galassi et al., 2003). For all presented spectra four scans were coherently added for four data sets for each t_1/t_2 data point, relaxation delay of 1.4 s, and the t_3 acquisition time of 102 and 85 ms, were used on 400 and 500 MHz spectrometer, respectively. For conventional spectra cosine weighting function was applied prior to Fourier transformation in all dimensions, while in the case of random sampling with exponential or Gaussian data point distribution only in t_3. The data are transformed for $128 \times 256 \times 430$ real points in $F_1(^{15}\text{N})$, $F_2(^{13}\text{C})$ and $F_3(^{1}\text{H})$, respectively, only amide protons region was chosen. Because in all cases average time domain points density as defined in Equation 6 is below Nyquist density, all 2D Fourier Transforms were made using WS (uniform sampling) or WP (Gaussian sampling) formula. The randomly sampled t_1/t_2 interferograms were transformed as described in text employing PC with 3.0 GHz Pentium 4 processor under Linux operating system using C code. The computing time of $n \times m \times 25$ ns was achieved (where n and m are numbers of frequency and time domain points, respectively), i.e., for processing spectrum with $n = 128 \times 256 \times 430$ resolution, with $m = 1024$ recorded t_1/t_2 points requires ca. 6 min. The resulting real parts of 3D spectra were saved in the format of SPARKY program (Goddard and Kneller, 2002). The signal-to-noise ratios were computed using program SPARKY with noise estimated using 10,000 of randomly sampled spectral points.

Results and discussion

To investigate the influence of various sampling modes of evolution time space on obtained spectra, the signal-to-noise ratio of the signal of I36 CA–N–HN correlation was evaluated and listed

Table 1. The comparison of signal to noise ratio obtained by FT of simulated time domain signal with and without thermal noise η described in caption to Figure 4. The probability density of time domain points was defined as $\exp(-t/\sigma)$ and $\exp(-t^2/\sigma^2)$ for exponential and Gaussian distribution, respectively. σ of 0.5 s was used. Spectral width of 1300 Hz and 2900 Hz in F_1 and F_2 was assumed. In the case of random sampling evolution times were normalized to the range of 0–25 ms and 0–20 ms, for t_1 and t_2, respectively

Number of points	Random without thermal noise			Random with thermal noise		
	Uniform	Gaussian	Exponential	Uniform	Gaussian	Exponential
64	11.8	15	15.3	8.6	13.6	12.1
128	17.3	21.5	21.2	12.8	19.8	15.9
256	23.9	30.8	29.7	17.3	27.9	21.6
512	32.7	42.4	41.6	24.5	38.9	29.9
1024	45.8	59.5	59.4	35.0	54.4	42.6

Table 2. The comparison of half-height signal linewidth in Hz, obtained in FT of simulated time domain signal with thermal noise η. The simulation parameters are the same as in Table 1

Number of points[a]	Cartesian grid		Random					
			Uniform		Gaussian		Exponential	
	LW_1	LW_2	LW_1	LW_2	LW_1	LW_2	LW_1	LW_2
64 (8 × 8)	171.5	405.0	48.5	61.7	44.8	63.3	74.5	73.9
128 (8 × 16)	176.6	198.8	45.2	61.0	37.3	60.6	59.3	77.9
256 (16 × 16)	82.2	192.2	47.0	56.8	37.1	52.5	62.3	87.4
512 (16 × 32)	82.9	98.2	43.5	60.3	37.4	54.9	71.8	87.5
1024 (32 × 32)	41.1	97.4	45.3	61.5	35.9	53.9	61.2	82.6

[a]The number of points in t_1 and t_2 for conventionally sampled signal is given in parenthesis.

for comparison in Table 3. The plots of S/N in function of square root of the number of data points are presented in Figure 5b. As expected the experimentally obtained signal to noise ratio reveal similar behavior to the results of simulations. The comparison of contour plots extracted from 3D HNCA spectra obtained in the same experimental time by conventional and random sampling of evolution time space is presented in Figure 6. The F_1F_2 planes for $\omega_3(^1H) = 8.10$ and 8.74 ppm, (a, b) and (c, d), respectively are shown. It is clearly visible that the plots (b and d) obtained by single step Fourier Transform of randomly sampled t_1/t_2 space give rise to proper

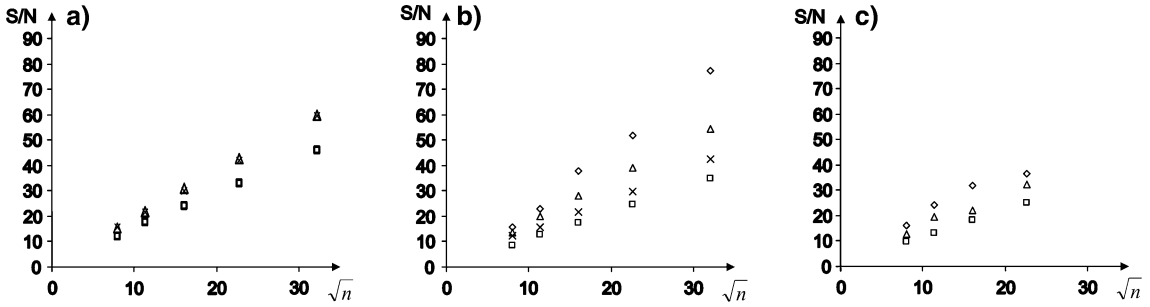

Figure 5. Plots of the signal to noise ratio versus square root from the number of data points *n*, obtained in simulations, according to parameters given in caption to Figure 4: without (a) and with (b) thermal noise, and experimentally (c), measured for ^{13}CA–^{15}N–^{1}HN signal of I36 in 3D HNCA spectrum of ^{13}C, ^{15}N-labeled ubiquitin on 400 MHz spectrometer. Experimental parameters are given in caption to Figure 6. Following sampling schemes were used: conventional (◇), uniform (□), Gaussian (△) and exponential (×).

Table 3. The comparison of signal to noise ratio for the signal at $\omega_3 = 6.128$ ppm, obtained by FT of 3D HNCA spectrum of ^{13}C, ^{15}N-ubiquitin. The density of time domain points was defined as $\exp(-t^2/\sigma^2)$ for Gaussian distribution. σ of 0.5 s was used. Spectral width of 1300 and 2900 Hz in F_1 and F_2 was set. In the case of random sampling evolution times were normalized to the range of 0–22.5 ms and 0–18 ms, for t_1 and t_2, respectively. For conventional sampling the maximum evolution times depended on spectral width and the number of points

Number of points[a]	Conventional	Random	
		Uniform	Gaussian
64 (8 × 8)	16.0	9.6	12.7
128 (8 × 16)	24.2	13.3	19.4
256 (16 × 16)	32.0	18.3	22.2
512 (16 × 32)	36.7	25.0	32.1

[a]The number of points in t_1 and t_2 for conventionally sampled signal is given in parenthesis.

signals and retain the full information of conventional 3D spectrum while the resolution achieved in the same measurement time is significantly improved. It is noteworthy that the enhanced F_1F_2 frequency domain resolution increases the actual resolution in F_3. In contrast to previously reported experiments utilizing radial and spiral sampling of evolution time space (Kazimierczuk et al., 2006), the presented spectra reveal no coherent artifacts which could lead to misleading interpretations. In order to evaluate practical usability in spectral analysis of experiments acquired with randomly sampled t_1/t_2 interferograms, we show the comparison of strip plots obtained in 3D experiments: HNCA in Figure 7 and approximately 10 times less sensitive (Sattler et al., 1999) HNCACB in Figure 8.

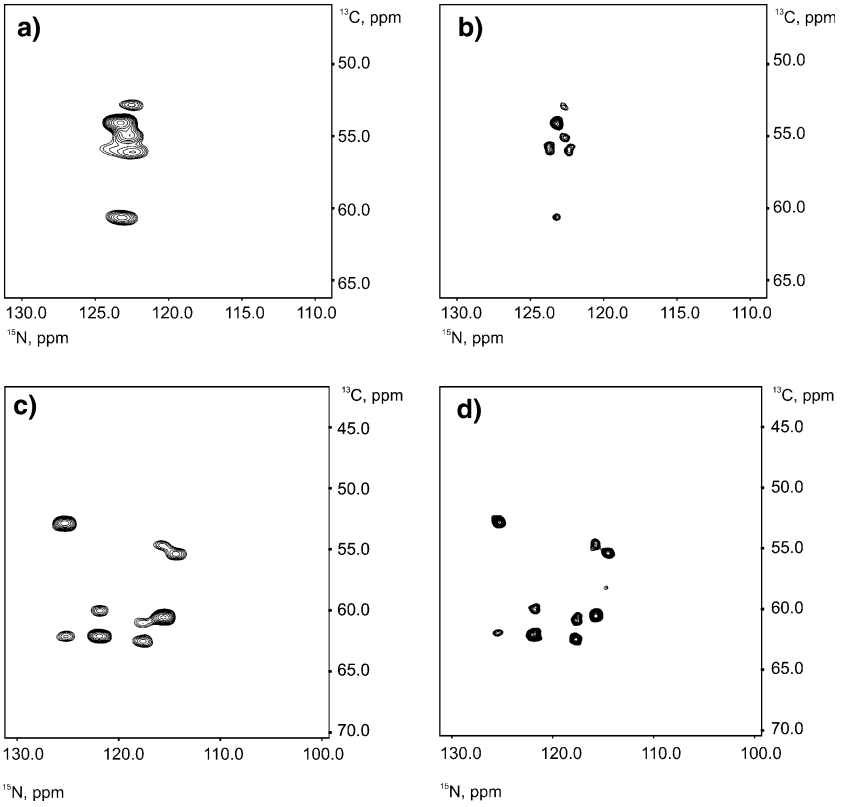

Figure 6. Comparison of contour plots of F_1F_2 planes at $\omega_3(^1H) = 8.10$ (a, b) and 8.74 ppm (c, d) obtained in 3D HNCA experiment for ^{13}C, ^{15}N-labeled ubiquitin on 400 MHz spectrometer, using conventional (a and c) and random sampling (b and d), of evolution time space. The spectral width of 1300 × 2900 × 5000 Hz was set in F_1, F_2, and F_3, respectively. 16 × 32 t_1/t_2 data points was recorded in conventional experiment, i.e., the maximum evolution times t_1 and t_2 of 12.3 and 11.0 ms, respectively were achieved. In the case of experiment with random sampling the maximum evolution times t_1 and t_2 of 22.5 and 18 ms, respectively, were used. Four scans were coherently added in all four data sets for 512 t_1/t_2 data points, thus the acquisition time of both, conventional and randomly sampled experiments were equal. The spectra were transformed with the resolution of 128 × 256 × 1024 points in F_1, F_2, and F_3, respectively.

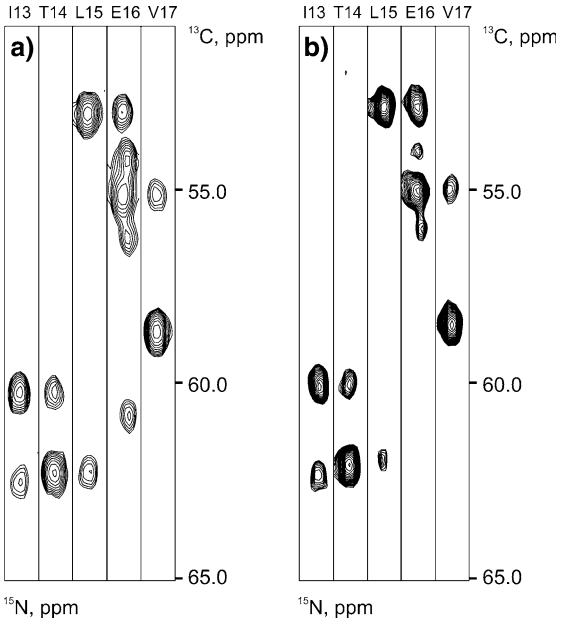

Figure 7. An example of sequence-specific assignments of I13-V17 fragment using 3D HNCA spectrum obtained for ^{13}C, ^{15}N-labeled ubiquitin sample on 9,4 T NMR spectrometer using conventional (a) and random sampling (b) of t_1/t_2 evolution time space. Strips extending over the entire spectral width of ^{13}CA dimension are defined for the amide groups of the residues indicated in each strip. Experimental details are given in caption to Figure 6.

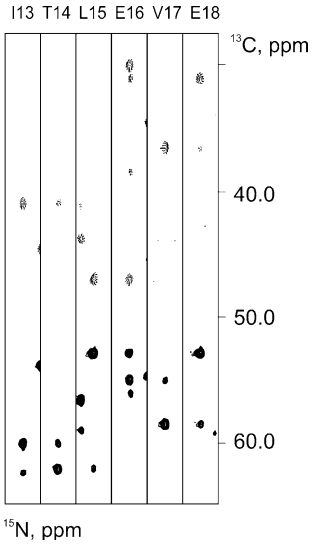

Figure 8. An example of sequence-specific assignments of I13-E18 fragment using 3D HNCACB spectrum obtained for ^{13}C, ^{15}N-labeled ubiquitin sample on 11.7 T NMR spectrometer using random sampling of t_1/t_2 evolution time space. Strips extending over the entire special width of ^{13}CACB dimension are defined for the amide groups of the residues indicated in each strip. The spectral width of $1600 \times 9500 \times 6000$ Hz was set in F_1, F_2, and F_3, respectively. The maximum evolution times t_1 and t_2 of 27.5 and 20 ms, respectively were used in experiment with random sampling. Four scans were coherently added in all four data sets for 4096 t_1/t_2 data points. The spectra were transformed with the resolution of $128 \times 256 \times 1024$ points in F_1, F_2, and F_3, respectively. The negative signals are plotted using dotted line.

Again, the clearly visible advantage of spectra acquired with the random sampling is increased resolution for Glu16 residue cross peaks, overlapping with signals of Val26.

Important problem associated with techniques used for reconstruction of high dimensional spectra from sparse data sets is ability to work with large number of signals of different amplitudes. Theoretically, the sampling noise increases with the number of signals in the same subspectrum, thus reducing the apparent dynamic range of signal intensities. To check this, in some cases significant, limitation we acquired ^{15}N-edited NOESY-HSQC spectrum of ubiquitin sample. The contour plots are presented and compared with conventional spectrum recorded in the same experimental time in Figure 9. The results evidenced improved resolution in the case of spectrum obtained by transformation of randomly sampled t_1/t_2 interferogram. Moreover, as in the case of 1D FT, the relative peak intensities are preserved. Measured noise level in the case of randomly sampled experiment is only approximately two times larger than for conventional spectrum. However, due to reduced linewidths in randomly sampled experiment only slight reduction of resulting S/N ratio is observed. Therefore experiments of high dynamic range of signal amplitudes are possible. It is noteworthy, that so far fast methods of multidimensional NMR were not successfully employed, except some impressive Multi-Dimensional Decomposition (MDD) examples (Orekhov et al., 2003) to process NOESY data.

Above spectra do not reveal any coherent artifacts present in methods of sparse data processing based on polynomial interpolation (Marion, 2005) and, contrary to them, can be easily extended to higher dimensionality by using appropriate extension of QFT. Single step multidimensional transform is nearly as fast as polynomial interpolation-FFT sequence allowing for achieving spectra after few minutes of computing time (depending on resolution and number of time

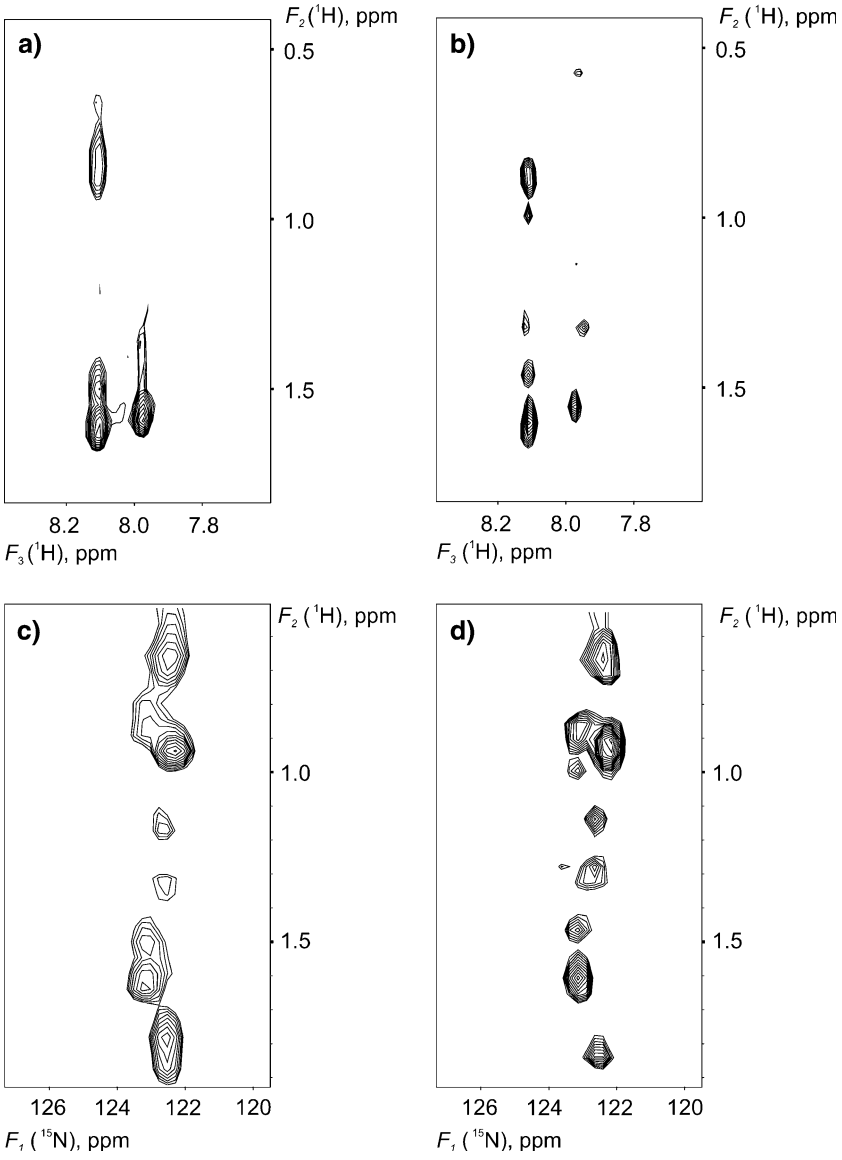

Figure 9. Comparison of contour plots of F_2F_3 planes at $\omega_1(^{15}\mathrm{N}) = 123.2$ ppm (a, b) and F_1F_2 planes at $\omega_3(^1\mathrm{H}) = 8.12$ ppm (c, d), obtained in ^{15}N-edited NOESY-HSQC experiment for ^{13}C, ^{15}N-labeled ubiquitin on 500 MHz spectrometer, using conventional (a and c) and random sampling (b and d), of evolution time space. The spectral width of $1600 \times 6000 \times 6000$ Hz was set in F_1, F_2, and F_3, respectively. 32×128 t_1/t_2 data points was recorded in conventional experiment, i.e., the maximum evolution times t_1 and t_2 of 20.0 and 21.3 ms, respectively were achieved. In the case of experiment with random sampling the maximum evolution times t_1 and t_2 of 25.0 ms, were used. Four scans were coherently added in all four data sets for 4096 t_1/t_2 data points, thus the acquisition time of both, conventional and randomly sampled experiments were equal. The spectra were transformed with the resolution of $128 \times 256 \times 1024$ points in F_1, F_2, and F_3, respectively.

points) on standard PC. It is a great advantage in comparison with MDD and Maximum Entropy Methods (MEM) which require much longer computing times.

Moreover, our solution does not require adjustment of any parameters, except optimization of sampling PDF, which will be discussed elsewhere.

Conclusions

We have shown that multidimensional single step Fourier Transform allows one to process randomly sampled multidimensional NMR data. This approach enables to optimize the utilization of NMR spectrometers by the recording of

experiments with improved achievable resolution per unit of acquisition time while retaining readability and resolution of high dimensionality. The superior spectral resolution feasible by the single step FT on randomly sampled data sets is of high importance regarding protein NMR spectroscopy. The overall sensitivity in randomly sampled multidimensional experiments is slightly reduced, however, the most of high-dimensionality experiments are limited rather by resolution requirements, especially when modern cryogenically cooled probes at high B_0 field are used. Similarly to conventional data acquisition and sequential FT processing the proposed method does not depend on *a priori* assumptions with regard to the number and shape of the signals, and could be applied to any multidimensional NMR experiment. The described technique could be also applied for processing of unconventionally sampled Magnetic Resonance Imaging signals without necessity of time domain signal interpolation.

Acknowledgements

Authors are grateful to Prof. Andrew R. Byrd (Structural Biophysics Laboratory, National Cancer Institute-Frederick, Frederick, Maryland, USA) for the sample of ^{13}C, ^{15}N-double labeled human ubiquitin.

References

Armstrong, G.S., Mandelshtam, V.A., Shaka, A.J. and Bendiak, B. (2005) *J. Magn. Reson.*, **173**, 160–168.
Barna, J.C.J., Laue, E.D., Mayger, M.R., Skilling, J., Worrall, S.J.P. (1987) *J. Magn. Reson.*, **73**, 69–77.
Berkovitz, A. and Rusnak, I. (1992) *IEEE Trans. Signal Process.*, **40**, 2816–2819.
Bodenhausen, G. and Ernst, R.R. (1981) *J. Magn. Reson.*, **45**, 367–373.
Bodenhausen, G. and Ernst, R.R. (1982) *J. Am. Chem. Soc.*, **104**, 1304–1309.
Coggins, B.E., Venters, R.A. and Zhou, P. (2004) *J. Am. Chem. Soc.*, **126**, 1000–1001.
Coggins, B.E., Venters, R.A. and Zhou, P. (2005) *J. Am. Chem. Soc.*, **127**, 11562–11563.
Delsuc, M.A. and Tramesel, D. (2006) *CR Chim.*, **9**, 364–373.
Ding, K. and Gronenborn, A.M. (2002) *J. Magn. Reson.*, **156**, 262–268.
Dutt, A. and Rokhlin, V. (1995) *Appl. Comp. Harm. Anal.*, **2**, 85–100.
Eghbalnia, H.R., Bahrami, A., Tonelli, M., Hallenga, K. and Markley, J.L. (2005) *J. Am. Chem. Soc.*, **127**, 12528–12536.
Ernst, R.R. and Anderson, W.A. (1966) *Rev. Sci. Instrum.*, **37**, 93–102.
Ferreira, P. (1999) *IEEE Trans. Circuits Syst. II*, **46**, 475–478.
Freeman, R. and Kupče, Ē. (2003) *J. Biomol. NMR*, **27**, 101–113.
Frueh, D.P., Sun, Z.-Y.J., Vosburg, D.A., Walsh, C.T., Hoch, J.C. and Wagner, G. (2006) *J. Am. Chem. Soc.*, **128**, 5757–5763.
Frydman, L., Scherf, T. and Lupulescu, A. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 15858–15662.
Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Booth, M. and Rossi, F. (2003) *GNU Scientific Library Reference Manual*, 2nd edn. Network Theory Ltd. ISBN 0954161734.
Goddard, T.D. and Kneller, D.G. (2002) *SPARKY 3*. University of California, San Francisco, http://www.cgl.ucsf.edu/home/sparky.
Hamilton, W.R. (1847) *Proc. Roy. Irish Acad.*, **3**, 1–16.
Kazimierczuk, K., Koźmiński, W. and Zhukov, I. (2006) *J. Magn. Reson.*, **179**, 323–328.
Kim, S. and Szyperski, T. (2003) *J. Am. Chem. Soc.*, **125**, 1385–1393.
Koźmiński, W. and I. Zhukov, I. (2003) *J. Biomol. NMR*, **26**, 157–166.
Kupče, Ē. and Freeman, R. (2003) *J. Am. Chem. Soc.*, **125**, 13958–13959.
Kupče, Ē. and Freeman, R. (2004a) *J. Biomol. NMR*, **28**, 391–395.
Kupče, Ē. and Freeman, R. (2004b) *Concept Magnetic Res.*, **22A**, 4–11.
Lauterbur, P.C. (1973) *Nature*, **242**, 190–191.
Luan, T., Jaravine, V., Yee, A., Arrowsmith, C.H. and Orekhov, V.Y. (2005) *J. Biomol. NMR*, **33**, 1–14.
Mandelshtam, V.A., Taylor, H.S. and Shaka, A.J. (1998) *J. Magn. Reson.*, **133**, 304–312.
Malmodin, D. and Billeter, M. (2005a) *J. Magn. Reson.*, **176**, 47–53.
Malmodin, D. and Billeter, M. (2005b) *J. Am. Chem. Soc.*, **127**, 13486–13487.
Marion, D. (2005) *J. Biomol. NMR*, **32**, 141–150.
Marvasti, F. (1996) *IEEE Trans. Signal Process.*, **44**, 572–576.
Marvasti, F. (2001) *Nonuniform Sampling Theory and Practice*, Kluwer/Plenum, New York.
Orekhov, V., Ibraghimov, I. and Billeter, M. (2003) *J. Biomol. NMR*, **27**, 165–173.
Rovnyak, D., Frueh, D.P., Sastry, M., Sun, Z.Y.J., Stern, A.S., Hoch, J.C. and Wagner, G. (2004) *J. Magn. Reson.*, **170**, 15–21.
Sattler, M., Schleuchter, J. and Griesinger, C. (1999) *Prog. NMR Spectrosc.*, **34**, 93–158.
Sun, Z.Y.J., Frueh, D.P., Selenko, P., Hoch, J.C. and Wagner, G. (2005) *J. Biomol. NMR*, **33**, 43–50.
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.
Tarczynski, A. and Allay, N. (2004) *IEEE Trans. Signal Process.*, **52**, 3324–3334.
Tarczyński, A. and Qu, D. (2005) *Int. J. Appl. Math. Comput. Sci.*, **15**, 463–469.
Tugarinov, V., Kay, L.E., Ibraghimov, I. and Orekhov, V.Y. (2005) *J. Am. Chem. Soc.*, **127**, 2767–2775.
Yeh, T. (1997) *Appl. Math. Comp.*, **87**, 227–246.
Yoon, J.W., Godsill, S., Kupče, Ē. and Freeman, R. (2006) *Magn. Reson. Chem.*, **44**, 197–209.