Journal of Magnetic Resonance 214 (2012) 91–102

Contents lists available at SciVerse ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

# Suppression of sampling artefacts in high-resolution four-dimensional NMR spectra using signal separation algorithm

Jan Stanek [a], Rafal Augustyniak [b,c,d], Wiktor Koźmiński [a,*]

[a] Faculty of Chemistry, University of Warsaw, Pasteura 1, 02093 Warsaw, Poland
[b] Ecole Normale Supérieure, Département de Chimie, 24 rue Lhomond, 75231 Paris Cedex 05, France
[c] Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France
[d] UMR 7203 Laboratoire des Biomolécules, CNRS-UPMC-ENS, Paris Cedex 05, France

ARTICLE INFO

*Article history:*
Received 25 August 2011
Revised 10 October 2011
Available online 20 October 2011

*Keywords:*
Sparse sampling
Four-dimensional spectroscopy
Signal separation algorithm
CLEAN
HCCH-TOCSY
C, N-edited NOESY
Engrailed 2

ABSTRACT

The development of non-uniform sampling (NUS) strategies permits to obtain high-dimensional spectra with increased resolution in significantly reduced experimental time. We extended a previously proposed signal separation algorithm (SSA) to process sparse four-dimensional NMR data. It is employed for two experiments carried out for a partially unstructured 114-residue construct of chicken Engrailed 2 protein, namely 4D HCCH-TOCSY and 4D C, N-edited NOESY. The SSA allowed us to obtain high-quality spectra using only as little as 0.16% of the available samples, with low sampling artefacts approaching the thermal noise level in most spectral regions. It is demonstrated that NUS 4D HCCH-TOCSY is dominated by sampling noise and requires efficient artefact suppression. On the other hand, 4D C, N-edited NOESY is a particularly attractive experiment for NUS, as the absence of diagonal peaks renders the problem of artefacts less critical. We also present a transverse-relaxation optimized sequence for HMQC that is especially designed for longer evolution periods in the indirectly detected proton dimension in high-dimensional pulse sequences. In conjunction with novel sampling strategies and efficient processing methods, this improvement enabled us to obtain unique structural information about aliphatic-amide contacts.



## 1. Introduction

In recent years, high-dimensional NMR experiments have gained increased importance in research on structure and dynamics of complex biomolecules. Despite continuing hardware improvements, the capabilities of multidimensional NMR seem limited by the inherently slow indirect sampling of chemical shifts [1–3]. The sampling theorem imposes restrictions on the maximum evolution times in indirectly detected dimensions

$$t_{\max} = N\Delta_t, \tag{1}$$

where $N$ is number of data points, $\Delta_t = (sw)^{-1}$, and $sw$ stands for spectral width. In three- and higher-dimensional spectroscopy of biomolecules, this limitation frequently forces the signals to be truncated, thus preventing one from fully utilizing the advantages of relaxation optimization as in TROSY [4]. The conventional way of collecting multidimensional NMR data is widely regarded as inefficient and a variety of alternatives have been proposed [2,5–10].

Random sampling of evolution time space has been widely recognized to be one of the most advantageous approaches [11–13]. Several strategies have been proposed to process such data: maximum entropy [14], multidimensional decomposition [13], non-uniform Fourier Transformation (nuFT) [15], FFT-CLEAN [16], signal separation algorithm (SSA) [17], forward maximum entropy [18], compressed-sensing [19] and Spectroscopy by Integration of Frequency and Time domain information (SIFT) [20]. These methods differ in their level of generality, the assumptions that must be fulfilled, computational complexity, as well as the kind of output. There is still lack of agreement on the best method (if it exists) and the way the results can be compared. Recent reviews provide guidelines on experimental cases where certain methods are most successful, where the capabilities of modern multidimensional NMR can be extended [2,21].

The origin of NUS artefacts in nuFT spectra is the implicit assumption that the intensity of signal at the points that were not measured is zero. The result of such a partial Fourier sum is therefore no longer a Fourier transform of a signal, or – equivalently – *is* a FT of zero augmented data. Thus, the commonly used term 'non-uniform Fourier Transformation' does not appropriately

*Abbreviations:* DFT, Discrete Fourier Transformation; IFT, Inverse Fourier Transformation; FFT, Fast Fourier Transformation; nuFT, non-uniform Fourier Transformation; NUS, non-uniform sampling; APSY, automated projection spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; S/N, signal-to-noise; S/A, signal-to-artefact; MDD, multidimensional decomposition.

* Corresponding author. Fax: +48 822 59 96.
*E-mail address:* kozmin@chem.uw.edu.pl (W. Koźmiński).

1090-7807/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2011.10.009

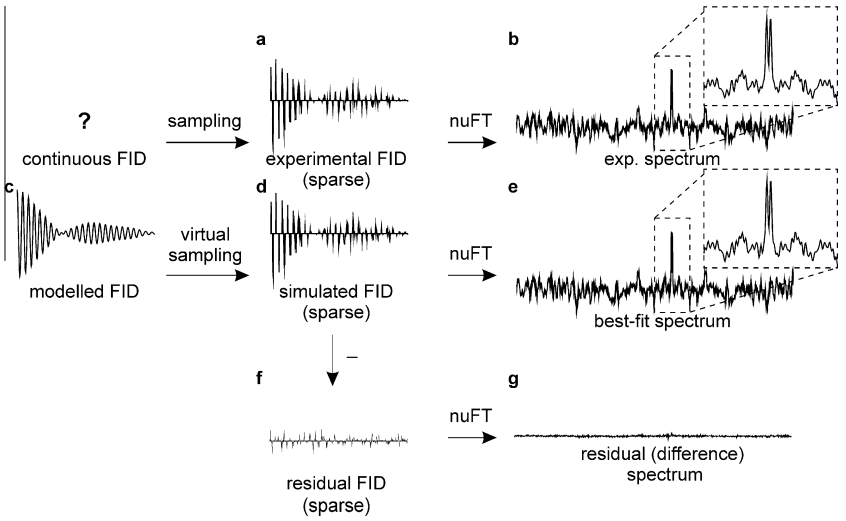


**Fig. 1.** Flowchart illustrating the principle of peak fitting in the frequency domain using a model for the time domain signal. Given a peak localized in the nuFT spectrum (b), the relative amplitudes of the frequency components in the model FID (c) are varied in order to obtain an identical line shape in the simulated nuFT spectrum (e). Finally, a sparse model signal (d) is subtracted from the experimental data (a) yielding a residual FID (f) and a residual spectrum (g). In this simulation the signal represents a doublet with $J = 15$ Hz and $R_2^* = 25\,s^{-1}$. Gaussian noise with $\sigma = 0.02$ (relative to the signal amplitude in the time domain) was added; 100 points were randomly chosen from a 400-point grid, created for $sw = 4$ kHz and $t_{max} = 100$ ms.

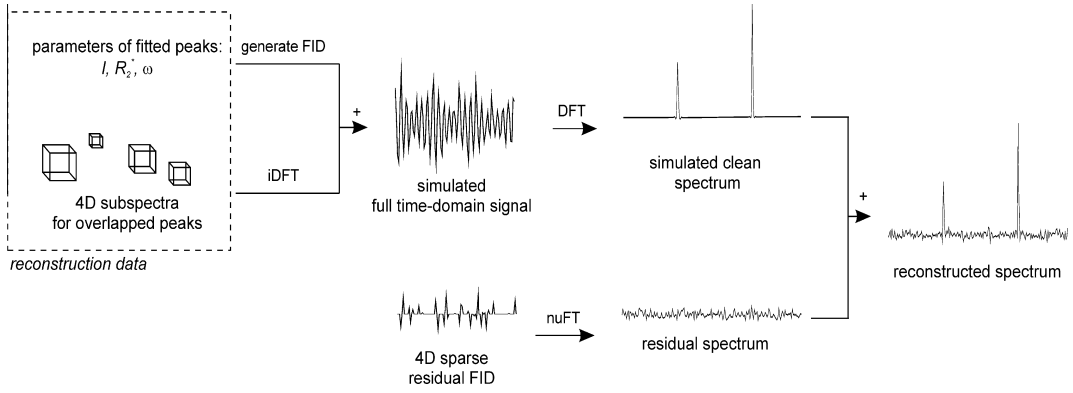


**Fig. 2.** Spectral reconstruction from data generated by the signal separation algorithm (SSA) (see text for details).

reflect the underlying mathematical operation. Nevertheless, we continue to use it across this paper as it is generally accepted in the literature [22].

The aim of this work was to adapt SSA for processing of sparsely sampled four-dimensional spectra. This method can be regarded as a hybrid approach that combines the concepts of CLEAN [23] and manual artefact removal [12]. The goal of SSA is reconstruction of a signal at the points omitted in the sampling schedule, and, as a result, reduction of intensity of NUS artefacts. The successful application of SSA to 3D $^{15}$N- and $^{13}$C-edited NOESY was demonstrated previously for a ubiquitin sample [17].

Side chain resonance assignment represents one of the most laborious tasks during structure determination of proteins. Conventional approaches based on 3D spectra such as H(CCO)NH-TOCSY, (H)C(CO)NH-TOCSY and (H)CCH-TOCSY are also error-prone due to severe spectral overlap of side-chain proton and carbon resonances.

Recently, a couple of higher-dimensional experiments were proposed that provide less ambiguous information for assignment. All of them employ sparse sampling and different processing methods: maximum entropy [24], FFT-CLEAN [16], automated projection spectroscopy (APSY) [25] and nuFT [15]. Further advantages of 4D HCC(CO)NH [24] and 5D HCCCONH [15,25], which offer even better resolution, lie in reduced spectral crowding. However, these experiments are quite insensitive due to the large number of coherence transfer steps. Additionally, nuFT is the only processing method available at present for randomly sampled 5D spectra, and the sensitivity of the experiments is further diminished by the presence of sampling artefacts. The alternative 4D HCCH-TOCSY [16] experiment features higher sensitivity and more redundant information; unfortunately the analysis of such spectra still poses difficulties owing to frequent overlap of resonances. Nevertheless, for residues with long side chains (e.g., Ile, Val, Leu, Lys, Arg) 4D HCCH-TOCSY might be the preferred choice. 4D HCCH-TOCSY is basically less sensitive than its 3D counterparts, H(C)CH- and (H)CCH-TOCSY, however, as argued by Mobli and co-workers [24], 4D experiments provide proton and carbon chemical shifts at once, thus compensating for reduced sensitivity.

In this communication, we report the application of SSA to crowded 4D HCCH-TOCSY and C, N-edited NOESY spectra. Both experiments can provide or complement side-chain assignments. The experiments were carried out for a 13.4 kDa fragment of Engrailed 2 homeoprotein, that contains a globular homeodomain and 54-residue intrinsically disordered N-terminal extension that is functionally important. The preparation of the sample and resonance assignments have been recently reported [26].

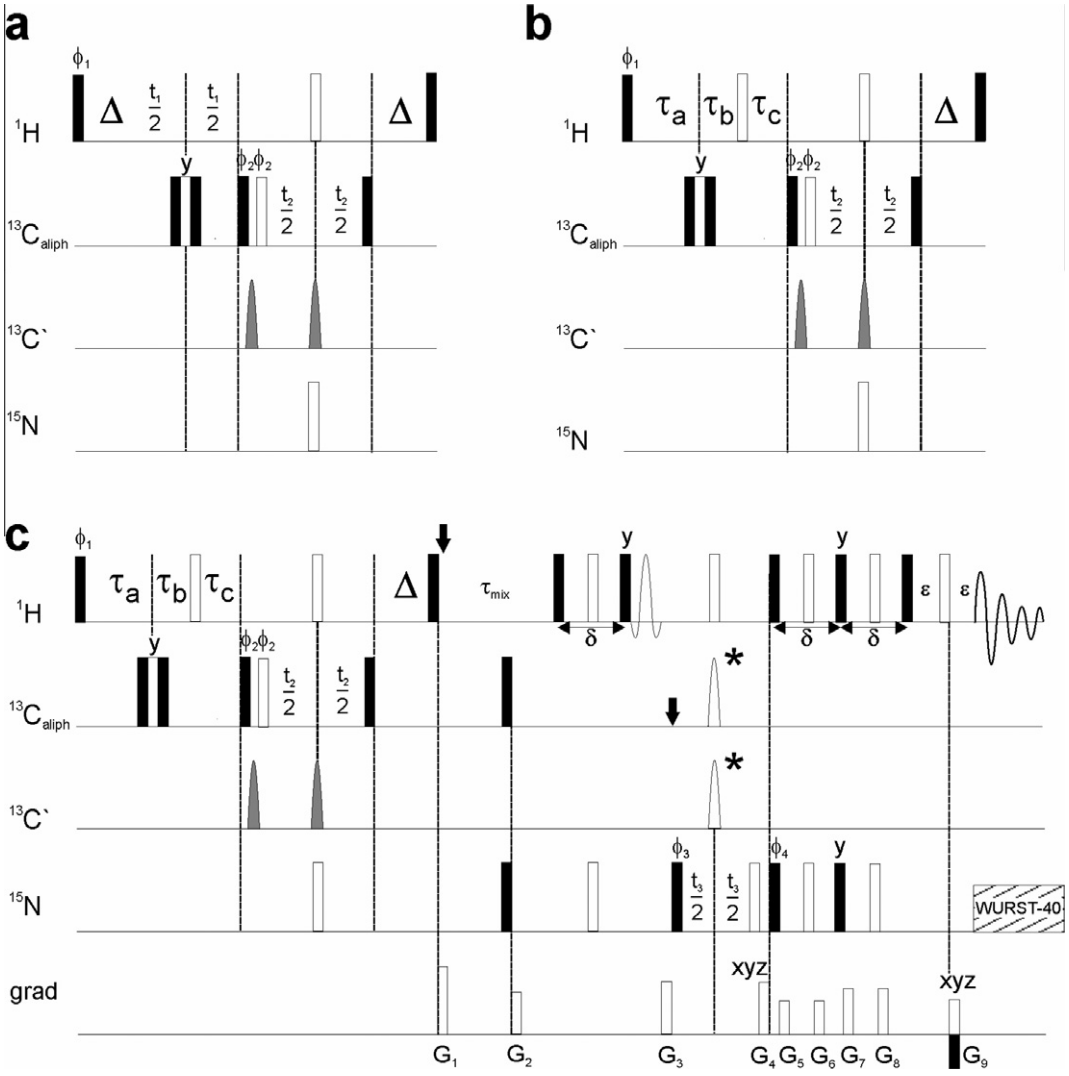


**Fig. 3.** Comparison of the standard (a) and optimized (b) HMQC blocks. (c) Pulse sequence for 4D C, N-edited HMQC-NOESY-HSQC including optimized HMQC. Solid and open bars represent 'hard' 90° and 180° pulses, respectively. All pulses are applied along the *x*-axis of the rotating frame unless indicated otherwise. Selective sinc-shaped 180° pulses, with $\gamma B_1$ adjusted to obtain inversion of C' spin with no effect at $C^{\alpha}$, are represented by gray bell-shaped pulses. Adiabatic hyperbolic secant inversion pulse of 200 ppm bandwidth and duration of 1 ms is indicated with a star. Open bell-shaped pulse represents water flip-back sinc pulse with duration of 1.2 ms. Quadrature detection in $t_1$, $t_2$ and $t_3$ is accomplished by altering $\varphi_1$, $\varphi_2$ and $\varphi_4$, respectively, according to the States-TPPI procedure. Axial peaks in the indirect frequency dimensions are shifted to the edges of spectrum by increasing $\varphi_1$, $\varphi_2$ and $\varphi_3$ and the receiver phase by 180° for even-numbered points in $t_1$, $t_2$ and $t_3$. The phase cycle is: $\varphi_1 = -x$; $\varphi_2 = -x, x$; $\varphi_3 = 2(x), 2(-x)$; $\varphi_4 = -x$, $\varphi_{rec} = x, -x, -x, x$. The delays are $\Delta = 3.57$ ms, $\delta = 5.38$ ms, $\varepsilon = 0.35$ ms, $\zeta = 2.15$ ms, and $\tau_{mix} = 160$ ms. For the semi-constant time evolution in $t_1$ the delays $\tau_a$, $\tau_b$ and $\tau_c$ are $t_1/2$, $t_1(1 - 2\Delta/t_{1,max})/2$ and $\Delta(1 - t_1/t_{1,max})$, respectively. Gradients are applied as follows: $G_1$ (0.5 ms, 18.5 G/cm), $G_2$ (0.5 ms, 15.5 G/cm), $G_3$ (0.5 ms, 20.3 G/cm), $G_4$ (2 ms, 31.9 G/cm), $G_5 = G_6$ (0.5 ms, 3.7 G/cm), $G_7 = G_8$ (0.5 ms, 5.5 G/cm), $G_9$ (0.2 ms, ±32.3 G/cm). The proton carrier is shifted from 3.17 to 4.77 ppm at the beginning of the mixing period. The carbon carrier is centered at 37.4 ppm, and shifted to 117.4 ppm after the gradient $G_3$ as denoted by arrows.

As noted previously [27,28] intrinsically disordered proteins, which have been extensively studied in recent years [29], represent a very good example of applications of sparse sampling for at least two reasons: (i) poor chemical shift dispersion leading to heavy resonance overlap, and (ii) favourable relaxation properties that allow one to use high-dimensional sequences.

## 2. Methods

### 2.1. Processing of 4D spectra using SSA

The signal separation algorithm was recently described in detail for three-dimensional spectra [17]; here only a brief summary of the processing scheme is given. The algorithm comprises three stages:

(1) Peak detection.
(2) Simulation of FID signals that correspond to the peaks found in step 1.
(3) Subtraction of simulated FID signals from the experimental ones.

These steps are repeated until no further intense peaks are detectable; typically, up to ten iterations are performed. Similar algorithms that gradually build a model of systematic portion of data (i.e. different from noise) are known as maximum likelihood (ML) method [30] and CLEAN [23]. As for SSA, both ML and CLEAN use a threshold-based criterion for peak detection. The most apparent distinction between SSA and model fitting based on ML principle is that the latter aims at extrapolation of uniformly sampled data to greater evolution times. In contrast, SSA and CLEAN achieve resolution enhancement by effective interpolation of non-uniformly sampled signal.

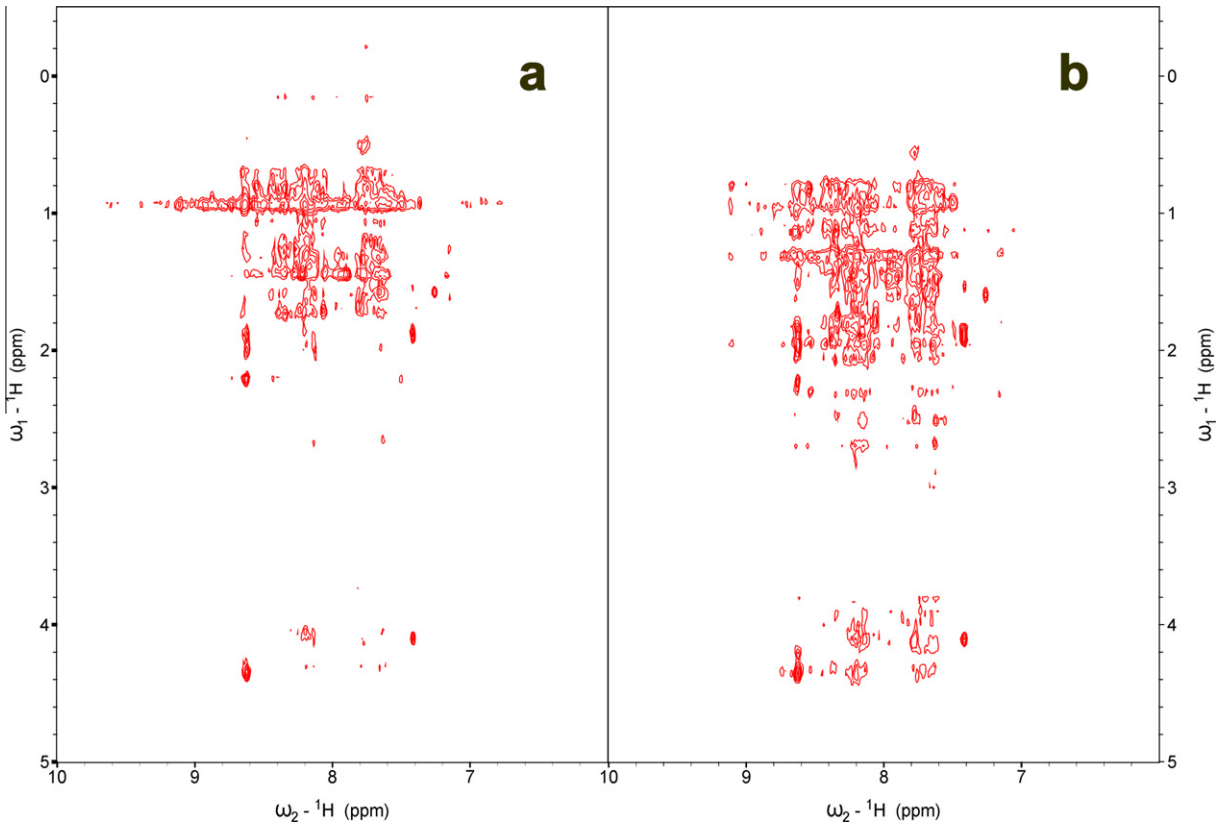


**Fig. 4.** A comparison of the sensitivity of the standard (a) and optimized (b) HMQC sequences, shown for 2D H(CN)H-HMQC-NOESY-HSQC spectra for a 0.5 mM sample of maltose binding protein (42.5 kDa) at 298 K. The pulse sequence for this experiment is shown in Fig. 3. The spectral width for $H_{aliph}$ was 6 kHz in $F_1$, and 128 increments were collected using semi constant-time evolution. In effect, a maximum evolution time of 21.2 ms was achieved. For each increment 32 transients were recorded.

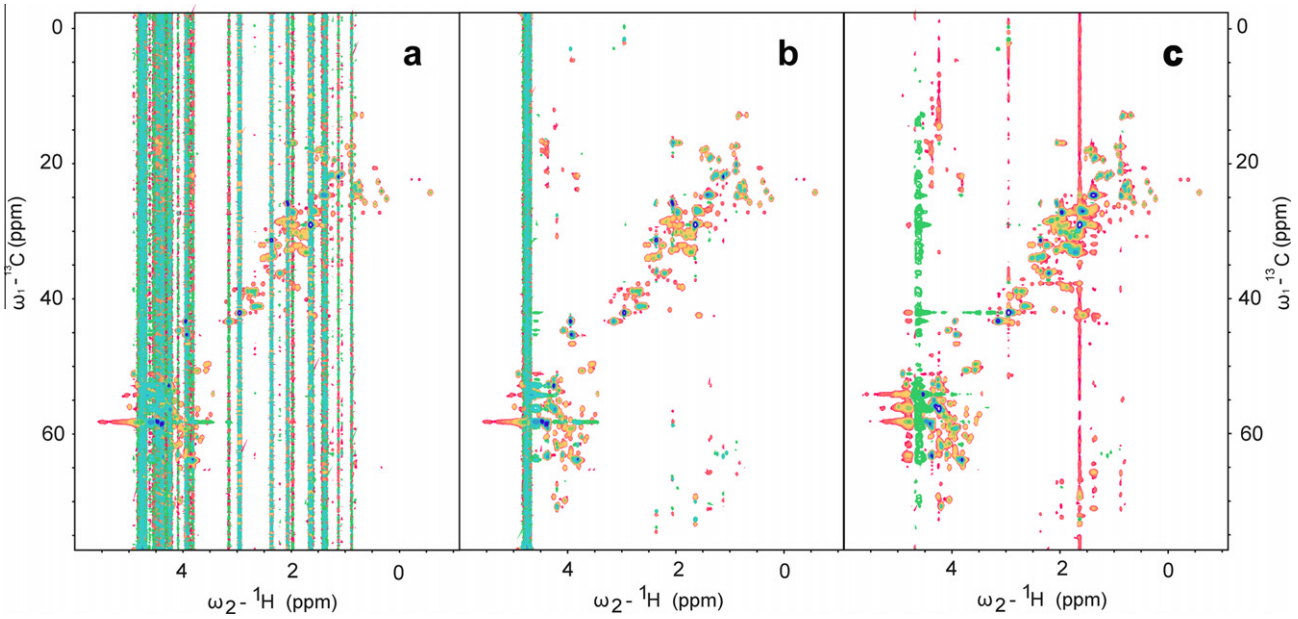


**Fig. 5.** 2D projections onto the $F_3$ (C)/$F_4$ (H) plane of 4D HCCH-TOCSY spectra of 0.5 mM Engrailed (114 amino acids), obtained with nuFT (a) and SSA (b) processing, shown using the same intensity scale. Intense vertical stripes of artefacts present in the nuFT spectrum (a) are suppressed by SSA (b). The projection is compared with a conventionally sampled 2D (HC)CH-TOCSY spectrum using the same pulse sequence (c). The peaks resulting from pulse sequence imperfections and residual water resonance appear in both (c) and (b); thus they are not due to false positives/negatives of SSA processing. The 4D HCCH-TOCSY was recorded using sparse on-grid sampling, with 2900 points randomly chosen on a 90 × 140 × 140 Cartesian grid. A decreasing Gaussian sampling density, $\exp[-(t/t_{max})^2/2\sigma^2]$ for non-negative $t$ and 0 otherwise, with $\sigma = 0.7$ was used. Maximum evolution times of 15, 10 and 10 ms, and spectral widths of 6, 14 and 14 kHz were chosen in the indirectly detected $t_1$ (H), $t_2$ (C) and $t_3$ (C) dimensions, respectively. Effectively, a relative sampling density of 0.16% was achieved. For each $t_1/t_2/t_3$ sampling point eight hypercomplex components were recorded by accumulation of four transients The 4D spectra were transformed with 352 × 512 × 512 points in $F_1$, $F_2$ and $F_3$. The 2D (HC)CH-TOCSY spectrum was recorded immediately after the acquisition of the 4D experiment, using the same pulse sequence and phase cycle. The spectral width was 14 kHz and 140 increments in $F_1$ (C) were recorded to obtain the same spectral resolution in the carbon dimension.

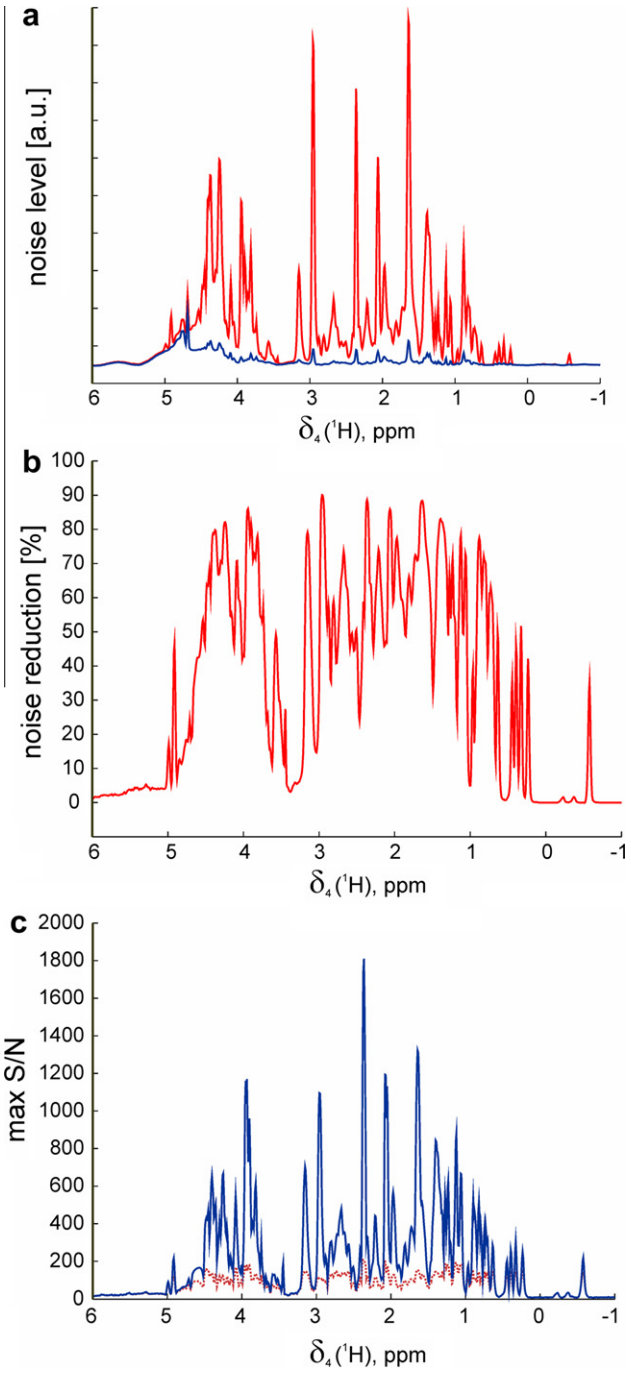


**Fig. 6.** Statistics of SSA processing of a 4D HCCH-TOCSY spectrum (experimental details are given in the caption to Fig. 5). (a) Effective noise level $n_{F4}$ measured in a 3D $F_1 \times F_2 \times F_3$ cube for each point in the acquisition dimension $F_4$. The noise was defined as the median of absolute values of all spectral intensities. Upper (red) and bottom (blue) curves represent noise levels in the nuFT and SSA processed spectra, respectively. (b) Reduction of the noise level (*RNL*), computed from data shown in (a) using the formula $RNL = 100\ (n_{F4}(nuFT) - n_{F4}(SSA))/n_{F4}(nuFT)$. Note that due to the presence of thermal noise the noise reduction can never reach 100%, and achievable values depend on the relative intensities of thermal and artificial sampling noise. (c) Signal-to-noise level of the most intense peak in a 3D $F_1 \times F_2 \times F_3$ cube for each point in $F_4$. The upper (blue solid line) and lower (red dotted line) curves correspond to the SSA and nuFT processed spectra, respectively. The plot provides information on the overall sensitivity of the experiment and illustrates the influence of sampling artefacts on the nuFT spectrum. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The crucial part of SSA is the simulation of FID signals, which are supposed to give the same peak shapes as in a nuFT spectrum. As shown in Fig. 1, SSA is capable of handling arbitrary peak shapes, since the FID signals are not required to be single damped sinusoids. Previously, it has been shown by simulations that SSA works better than the CLEAN algorithm (e.g. [16]) with respect to artefact suppression [17,31]. This is believed to be due to the more accurate reproduction of peak shapes, and the capability of choosing an optimal signal shape, either simple sinusoids or composite signals.

It is noteworthy that ML-based methods assume that signals are sufficiently separated from each other, otherwise the model of spectrum as linear combination of decaying sinusoids may not be stably fitted. As it is impossible to avoid signal overlap in complex spectra, SSA employs also a more general model for non-Lorentzian signal shapes [17]. Interestingly, multidimensional decomposition (MDD) employs a different general model that is able to reproduce overlapped peaks [13]. However, this flexibility in constructing signal shapes is achieved at the cost of generality, as the decomposition requires that the number of peaks to be known in advance. The interesting difference between SSA and MDD is that the latter can use the shared shape in one or more dimensions for a group of peaks, which improves the quality of decomposition of some kinds of spectra, e.g., 3D $^{15}$N- and $^{13}$C-edited NOESY [13] or 4D HNCACO.

The computational procedure previously applied to 3D spectra can be readily extended to higher dimensionalities, however, a rapid increase in memory requirements and processing times has to be considered, as will be discussed at the end of this section.

In comparison to 3D SSA, a few changes were made for 4D spectra regarding both the automatic identification of peaks and the fitting of signal parameters. Contrary to most existing processing methods (except MDD), which repeat three-dimensional procedures for each point in the directly detected dimension, our algorithm works on *all* dimensions of the 4D spectrum. An advantage of this approach is the increased reliability of peak picking as it is less prone to artefacts or noise peaks. Another benefit comes from the fact that more accurate estimates of peak parameters (amplitudes, frequencies and decay rates) can be obtained if the directly detected dimension is included. This adds only two more parameters to the signal model (a frequency and a decay rate) while the amount of *independent* data increases significantly since the signals typically extend over a certain number of points in the acquisition dimension.

Because of sampling artefacts, the apparent noise level varies along the directly detected dimension. For this reason *chi-square fitting* (also known as *weighted least squares fitting* [32]) of signal parameters was introduced, using a target function of the following form:

$$w(A, \vec{R}_2, \vec{\Omega}) = \sum_i \left(S_i^{sim} - S_i^{\exp}\right)^2 / \sigma_i^2, \tag{2}$$

where $A$, $\vec{R}_2$ and $\vec{\Omega}$ are the parameters of the analytical signal, $S_i^{sim}$ and $S_i^{exp}$ are elements (pixels) of the simulated and experimental nuFT spectra, respectively, and the sum extends over pixels in an isolated "peak frame". The noise level $\sigma_i$ is estimated in a 3D subspectrum spanned by the indirectly detected frequency domains $F_1 \times F_2 \times F_3$, and *may vary* for subsequent points in the acquisition dimension. More specifically, it is defined as the median of absolute values of spectral intensities in a 3D subspectrum. It must be emphasized that all model signals are simulated in *the time domain*, Fourier transformed, and compared with an experimental spectrum in *the frequency domain* employing Eq. (2). This approach ensures that NUS affects simulated and experimental spectra equally, so that corresponding time domain signals can be subtracted directly. In contrast to CLEAN [16], both SSA and ML-based methods (a variant referred to as hybrid time frequency domain maximum likelihood in particular [30]) focus on relatively small regions of spectrum

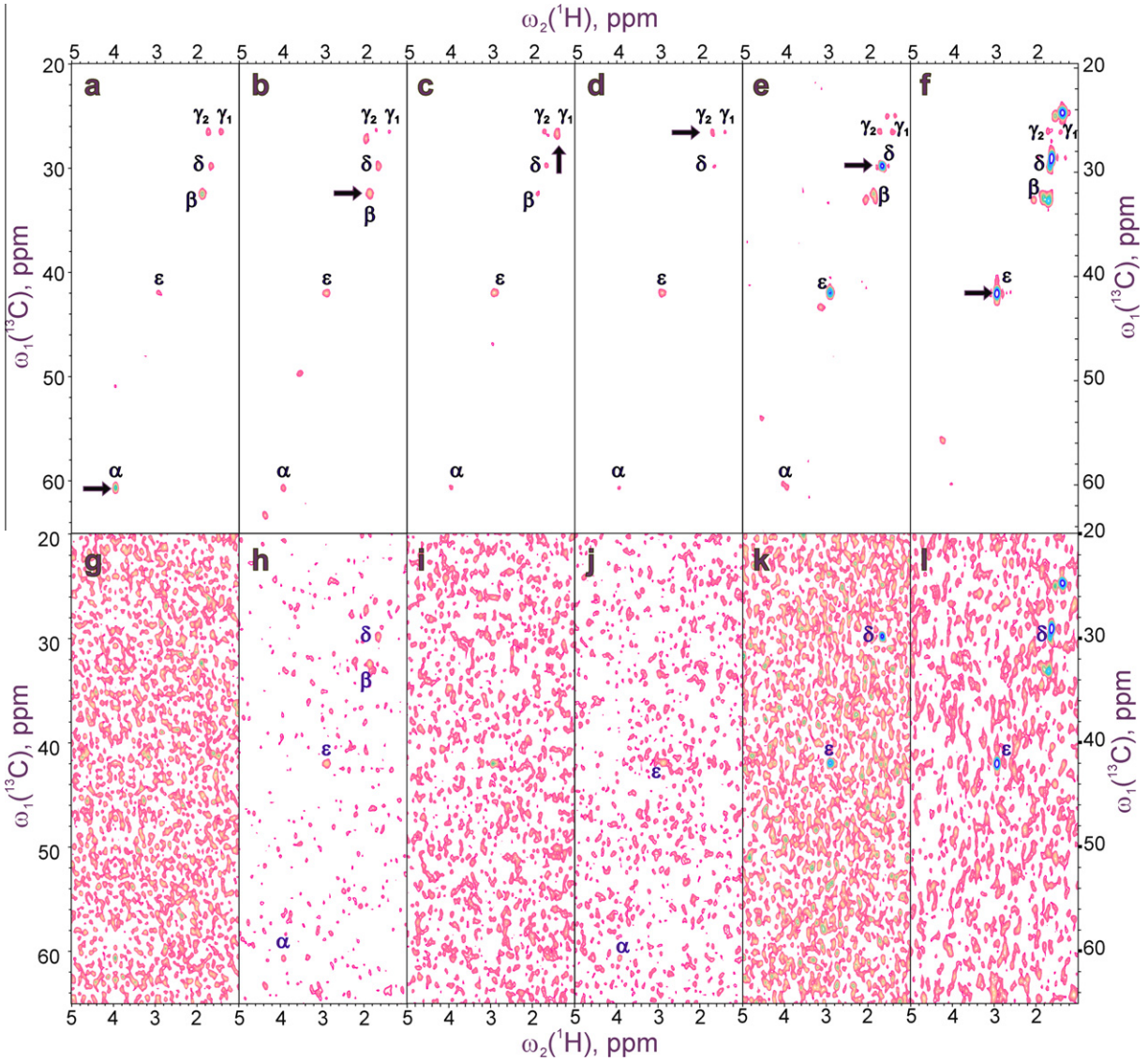


**Fig. 7.** 2D $F_1$ ($H_{aliph}$)/$F_2$ ($C_{aliph}$) cross-sections from a NUS 4D HCCH-TOCSY spectrum, computed at $F_3$ ($C_{aliph}$)/$F_4$ ($H_{aliph}$) coordinates of $C^{\alpha}/H^{\alpha}$ (a, g), $C^{\beta}/H^{\beta}$ (b, h), $C^{\gamma}/H^{\gamma 1}$ (c, i), $C^{\gamma}/H^{\gamma 2}$ (d, e), $C^{\delta}/H^{\delta}$ (e, k) and $C^{\varepsilon}/H^{\varepsilon}$ (f, l) resonances of the side-chain of K216. An arrow denotes the diagonal peak in each cross-section. The corresponding cross-sections from SSA (upper panels) and nuFT spectra (bottom panels) are plotted using the same intensity scales. Only peaks that belong to the K216 spin system and posses S/N > 6 (measured on a particular cross-section) are labelled. Due to intense sampling artefacts in the nuFT spectra (g–l) it is impossible to detect many peaks, in particular, the $H^{\gamma}/C^{\gamma}$ resonances do not appear clearly in any of the cross-sections. On the contrary, almost all cross-peaks are present in the SSA spectrum (a–f), allowing one to unambiguously identify the spin system. Additionally, a high redundancy enables one to exclude peaks belonging to other spin systems, which appear due to overlaps is the C-HSQC close to the position of a particular diagonal peak.

where peaks reside and fit model signals independently to avoid multiple evaluations of full spectrum. However, signals in NUS spectra interfere much stronger than in truncated conventionally sampled ones as point spread functions are less localized. Due to this coupling it is beneficial to repeat fitting procedure in SSA [17].

The implementation of 4D SSA has been restricted to sparse *on-grid* distributions. Firstly, this allows one readily to employ the States-TPPI method to shift imperfections to the edges of the spectra. Secondly, such a choice of sampling points enables numerical optimizations of at least two kinds compared to arbitrary off-grid sampling.

The first one regards the possible use of an FFT algorithm, although this is usually only advantageous for the *last* time dimension that must be transformed. In the first stages of the FFT transformation, data arrays basically contain more zeros (omitted points) than recorded data and "slow" FT usually exhibits better performance [17].

A slightly less obvious benefit of on-grid distributions is revealed for larger data sets, with a total number of samples $N$ greater than a few hundred. For simplicity, a case of 2D nuFT will be discussed. For a 2D spectrum with $M_1 \times M_2$ points that is calculated as Fourier sums of $N$ samples, the computational cost scales as (i) an inconveniently large number $M_1M_2N$ for arbitrary off-grid sampling and (ii) a much more reasonable number $M_1N + M_1M_2N_2$ if the points are arranged on an $N_1 \times N_2$ time domain grid. For sparse data sets ($N \ll N_1N_2$), the term $M_1N$ can usually be neglected, since the following relations hold

$$M_1N \ll M_1N_1N_2 \leqslant M_1M_1N_2 \approx M_1M_2N_2.$$

Consequently, processing of on-grid data requires only a small fraction $M_1M_2N_2/M_1M_2N = N_2/N$ of operations compared to arbitrary off-grid sampling. Since the total number of samples $N$ can be considerably greater than number of increments $N_1$ and $N_2$ in either of the two time dimensions

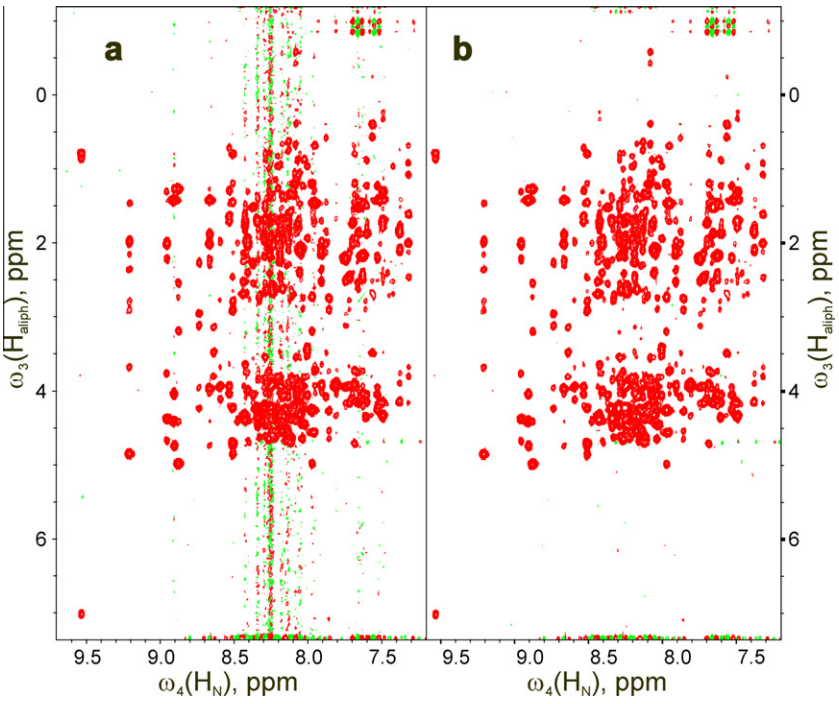


**Fig. 8.** $F_1$ ($H_{aliph}$)/$F_4$ ($H_N$) projections of a sparsely sampled 4D C, N-edited HMQC-NOESY-HSQC spectrum of 0.5 mM Engrailed 2 protein, obtained with nuFT (a) and SSA (b). Sampling artefacts are prominent in the most crowded central amide region only, and are completely removed by SSA processing. Apart from some aliphatic-amide cross-peaks, some axial artefacts appear at the edges of spectrum. The sampling scheme consisted of 6500 points chosen randomly on a 86 × 150 × 100 Cartesian grid in $t_1$, $t_2$, and $t_3$. Maximum evolution times of 14, 10 and 50 ms, and spectral widths of 6, 15 and 2 kHz were used in $t_1$ (H), $t_2$ (C) and $t_3$ (N). Consequently, a relative sampling density of 0.52% was achieved. Four transients were coherently added for each of the eight hypercomplex sampling points. The spectrum was transformed with 336, 592 and 512 points in $F_1$, $F_2$, and $F_3$. Positive and negative contours were plotted in red and green, respectively.

$$N_1, N_2 \ll N,$$

the speed-up factor $N/N_2$ can be quite significant.

### 2.2. Reconstruction of spectra

Most existing processing methods for NUS data aim at reconstructing the full time domain signal, or equivalently, its DFT spectrum. Unfortunately, this approach becomes impractical when dealing with high-resolution four-dimensional spectra. For example, assuming equal maximum evolution times $t_{1,max}$ = $t_{2,max}$ = $t_{3,max}$ = 20 ms and spectral widths $sw_1$ = $sw_2$ = $sw_3$ = 5 kHz, reconstruction of the full time domain grid requires $10^6$ floating point values to be stored *for each point in the acquisition dimension*. Assuming that 500 points in the acquisition dimension are relevant, this yields approximately 2 GB of storage for FID's in single precision, and 16 GB for a 4D spectrum obtained with double zero-filling. Although such files can in principle be stored on modern disks, long retrieval times for different spectral regions make an alternative approach desirable. Another difficulty arises from the fact that, in contrast to our 64-bit processing program, many applications are working in 32-bit environment, and e.g., cannot properly handle large files or operate on large memory arrays.

The solution proposed here for SSA utilizes the following ideas: (i) sparse non-uniform FT [15], and (ii) spectral reconstruction from model shapes as in the MDD method [33]. High-dimensional spectra are usually analyzed using simple experiments, e.g., $^{15}$N or $^{13}$C-HSQC are employed for 4D spectra. In fact, four-dimensional spectra have at least one 2D-projection that resembles $^{15}$N or $^{13}$C-HSQC. This allows one to focus on selected 2D cross-sections (say $F_1 \times F_2$) of a 4D spectrum, for which the remaining two spectral coordinates (e.g., $F_3$ and $F_4$) correspond to a given peak in HSQC. An efficient way of computing 2D cross-sections for SSA output data is described below.

As illustrated in Fig. 2, the final spectrum is obtained from (a) reconstruction data that comprise all extracted peak parameters, plus small spectral sub-regions containing unresolved peaks and (b) residual sparse FID's. The former usually does not require much storage space, and the latter have the same size as the original experimental sparse data. Both files are much smaller than the reconstructed spectra, and can be easily transferred over long distances. For this reason SSA and reconstruction procedures were separated into independent programs.

The most straightforward method to compute a selected cross-section is to simulate known signals on a full time domain grid and to calculate the cross-section using sparse FT (i.e., by direct evaluation of a Fourier sum):

$$A, \vec{R}_2, \vec{\Omega} \overset{\text{simulation}}{\rightarrow} \text{full-grid FID} \overset{\text{sparseFT}}{\rightarrow} 2D \text{ cross-section},$$

$$4D \text{ subspectrum} \overset{\text{IFT}}{\rightarrow} \text{full-grid FID} \overset{\text{sparse FT}}{\rightarrow} 2D \text{ cross-section}.$$

However, this approach is computationally demanding and would require a full time domain grid to be stored in memory as an intermediate result.

The first observation is that peaks represented by the analytical parameters $A$, $\vec{R}_2$, $\vec{\Omega}$ need not be reconstructed in the multidimensional time domain, as they appear in the spectrum as direct products of 1D shapes, which can be rapidly computed in advance.

Another optimization utilizes the fact that sequential FT/IFT (forward/inverse FT) can be performed in arbitrary order. More specifically, let us assume that $F_1 \times F_2$ cross-sections are to be computed, and that the $F_3$ and $F_4$ coordinates (denoted as $\Omega_3$, $\Omega_4$) are known in advance. Then the optimal order of the transformations is:

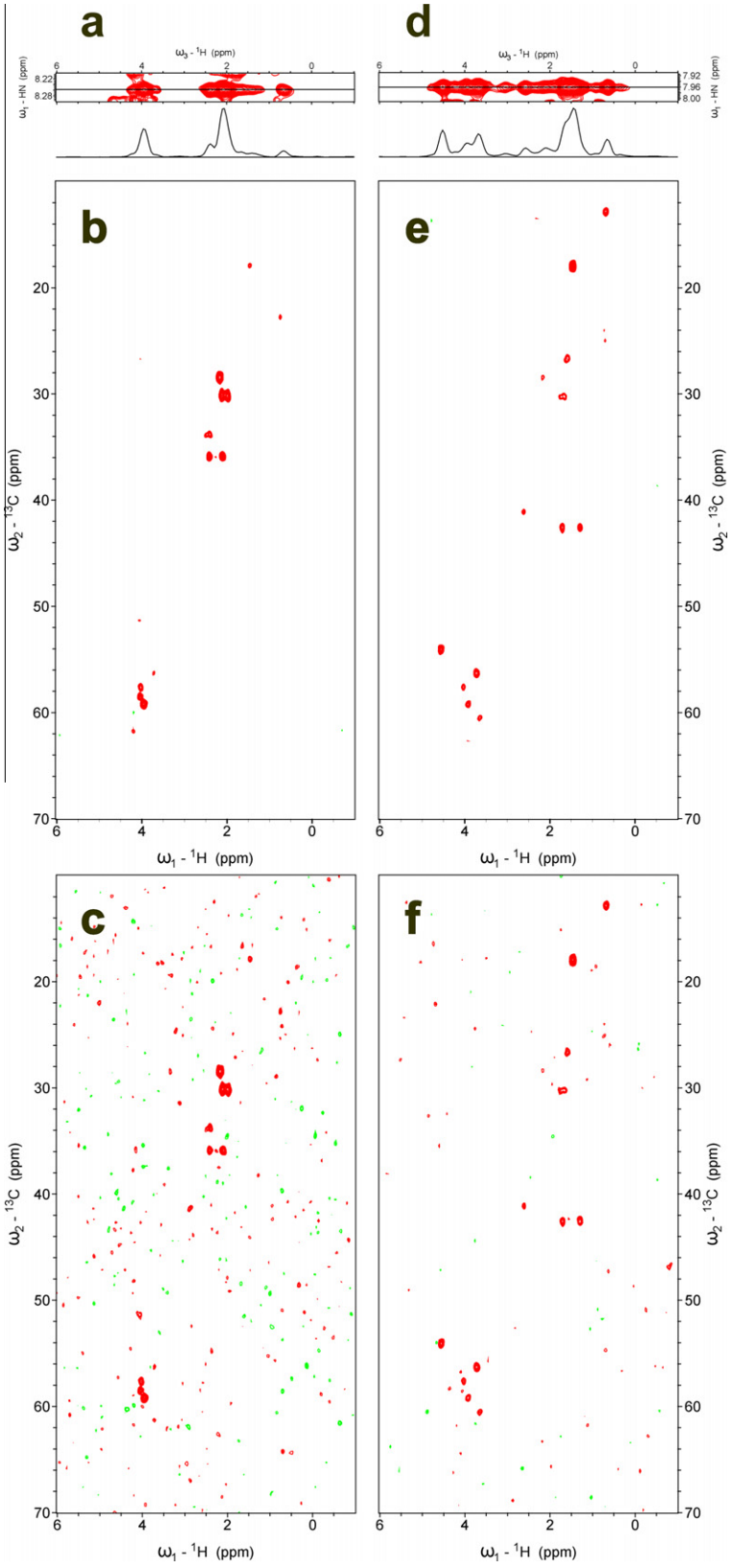


**Fig. 9.** A comparison of spectral resolution between 1D $F_1$ (H) cross-sections of 3D N-edited NOESY-HSQC (a, d) and 2D $F_1$ ($H_{aliph}$)/$F_2$ ($C_{aliph}$) cross-sections of 4D C, N-edited spectra for residues E236 (a–c) and R181 (d–f). Owing to the carbon chemical shift dispersion it is possible to resolve all cross-peaks that overlap in the 3D spectrum. For example, the intra-residue $H^{\gamma}C^{\alpha}$ cross-peak of the E236 residue can be easily distinguished from inter-residue cross-peaks (b). Panels *b* and *e* correspond to the SSA processed spectrum. Cross-sections from a nuFT spectrum are shown in panels *c* and *f* with the same intensity scale as *b* and *e*. The suppression of sampling artefacts enables one to identify weak cross-peaks that are masked in the nuFT spectra (c, f). The cross-sections were taken at $F_3$ (N)/$F_4$ ($H_N$) = 119.64/8.26 ppm (E236) and 121.33/7.955 ppm (R181). Experimental details for 4D C, N-edited HMQC-NOESY-HSQC are given in the caption to Fig. 8. The 3D N-edited NOESY-HSQC was recorded on a 1 GHz spectrometer using conventional sampling on a 64 × 32 grid for $F_1$ (H) and $F_2$ ($^{15}$N), respectively. The spectral widths in the indirect dimensions were 11.9 kHz ($F_1$) and 1.8 kHz ($F_2$). Eight transients were used for each hypercomplex component.

$$4D\,\text{subspectrum}\,S(\omega_1,\omega_2,\omega_3,\omega_4)\overset{\text{IFT-t4}}{\rightarrow}S(\omega_1,\omega_2,\omega_3,t_4)$$
$$\overset{1-\text{point FT-F4}}{\rightarrow}S(\omega_1,\omega_2,\omega_3,\Omega_4)\overset{\text{IFT-t3}}{\rightarrow}S(\omega_1,\omega_2,t_3,\omega_4)$$
$$\overset{1-\text{point FT-F3}}{\rightarrow}S(\omega_1,\omega_2,\Omega_3,\Omega_4)\overset{\text{IFT-t2,IFT-t1}}{\rightarrow}S(t_1,t_2,\Omega_3,\Omega_4)$$
$$\overset{\text{2D DFT}}{\rightarrow}\text{2D cross-section}.$$

As shown above, if sparse FT's (transformations to a single point in the frequency domain) precede two-dimensional DFT, the storage requirements for the intermediate results are greatly reduced. Moreover, the computational burden decreases from

$$\prod_{k=1}^{4}R_k\prod_{k=1}^{4}N_k,$$

where $R_k$ is a number of digital points in $k$th dimension of a 4D subspectrum to:

$$R_1R_2R_3R_4N_4+R_1R_2R_3N_4+R_1R_2R_3N_3+R_1R_2N_3+\text{cost of 2D DFT}.$$

Similarly, the computation of a cross-section from a sparse residual FID ought to be performed as a sequence of sparse FT's (for a single point in the frequency domain) *followed* by 2D-nuFT.

### 2.3. Optimization of HMQC blocks

Until recently, long evolution periods in 4D experiments were inaccessible due to sampling limitations. In contrast, NUS allows one to extend evolution times to the limits of relaxation times [15]. The practical guidelines on the optimal sampling schedules in terms of sensitivity, which rely on prior knowledge of mean $T_2$, were recently given by Rovnyak and co-workers [34]. These results suggest that signals should be sampled at significantly larger evolution times than possible conventionally (up to $\sim 1.26T_2$). The need to minimize relaxation losses during such extended evolutions triggers off the research on pulse sequence improvements. In the following section we demonstrate how HMQC blocks can be modified to extend possible evolution in the constant-time mode (CT) without any further decrease in sensitivity.

The rearrangement of pulses and delays in the initial part of the HMQC sequence (see Fig. 3a and b) is aimed at utilizing *both* $\Delta$ delays for proton evolution (in CT or SCT mode), while in the standard HMQC sequence (see e.g. [35]) only one fixed delay $\Delta = 1/(2J)$ is used. Clearly, proton chemical shifts evolve during $t = \tau_a + \tau_b - \tau_c + \Delta$ in the entire block, and the heteronuclear coupling evolves during $\tau_a - \tau_b + \tau_c = \Delta$ up to the first $\pi/2$ pulse applied to carbon-13. Most importantly, the modification enables a constant-time evolution for protons ($\tau_a = t/2$, $\tau_b = 0$, $\tau_c = \Delta - t/2$) up to $t_{\max} = 2\Delta$, twice as long as in the conventional version. For the case of semi-constant time evolution (SCT), the delays in Fig. 3b should be set to

$$\begin{aligned}\tau_a &= t/2\\ \tau_b &= t/2\cdot(1-2\Delta/t_{\max})\\ \tau_c &= \Delta(1-t/t_{\max}).\end{aligned}\tag{3}$$

One can notice that relaxation losses are also decreased in this case. More formally, the sensitivity for standard evolution mode (Eq. (4a)) and for SCT in the modified (Eq. (4b)) HMQC block, respectively, are given by:

$$\int_0^{t_{\max}}\exp\left(-\frac{2\Delta+\tau}{T_2}\right)d\tau,\ t_{\max}\geq\Delta,\tag{4a}$$

$$\int_0^{t_{\max}}\exp\left(-\frac{2\Delta+\tau(1-2\Delta/t_{\max})}{T_2}\right)d\tau,\ t_{\max}\geq 2\Delta.\tag{4b}$$

It can easily be shown that the largest benefit of the improved HMQC is obtained for $t_{\max} = 2\Delta$, and slightly decreases for longer maximum evolution times. For a molecule with ~50 kDa with a correlation time $\tau_c \approx 20$ ns and $T_2$ ($H_{aliph}$) ≈ 20 ms the maximum sensitivity gain can be estimated to be approximately 18% (assuming $\Delta = 3.6$ ms).

The performance of the modified HMQC sequence was verified for a 2D H(CN)H-HMQC-NOESY-HSQC experiment for a 0.5 mM sample of 42.5 kDa maltose binding protein (see Fig. 3c for the detailed description of the 4D pulse sequence). The acquisition was performed with the standard sequence first, and then with the optimized version (see spectra compared in Fig. 4). An evident, though not overwhelming signal enhancement has been observed, which is in agreement with theoretical predictions.

In this work, the modified HMQC block was employed in a 4D C, N-edited NOESY experiment (see Fig. 3c). It can be applied to 4D C, C-edited NOESY as well, or incorporated into other pulse sequences that include indirect proton and heteronuclear evolution and storage of proton $z$-magnetization both at the beginning and end of the block.

## 3. Results and discussion

### 3.1. 4D HCCH-TOCSY

The 4D sensitivity-enhanced HCCH-TOCSY experiment was recorded in ca. 44 h using a sampling scheme with 2900 points, which corresponds to 0.16% of the samples required for conventional acquisition. As the signal-to-artefact ratio depends solely on the total number of samples [15], and not on the sampling density, the maximum evolution times were adjusted to the relaxation properties. Quite importantly, in the carbon dimensions the signals are modulated by at least one C–C coupling, $J_{CC} \approx 35$ Hz (except for glycine $C^{\alpha}$), so that it is not beneficial to approach $1/(2J_{CC}) \approx 14$ ms. On the other hand, one should attempt to achieve maximum resolution at least in the last carbon dimension to avoid ambiguities in the analysis. Thus, the extension of the maximum evolution time up to 10–12 ms is recommended, approximately twice as long as used previously for this experiment [16]. It should be noted that modulations by C–C couplings act as apodization functions and significantly decaying sampling schemes are not advisable. Here a sampling grid of $90 \times 140 \times 140$ points was used; points were randomly chosen according to a Gaussian probability density with zero mean and $\sigma = 0.7 : p(t) = \sqrt{\frac{2}{\pi\sigma^2}}\exp\left(-\left(\frac{t}{t_{\max}}\right)^2/(2\sigma^2)\right)$ for $t \geqslant 0$ and 0 otherwise (see [9]). The variance of point distribution ($\sigma = 0.7$) was found empirically to match the envelope of $J$-modulated and decaying signal in the carbon dimensions $t_2$ and $t_3$.

The collected data were transformed using 4D SSA yielding a significant improvement in signal-to-artefact ratio. The results can be conveniently presented as 2D projections of the 4D spectrum as a function of $\omega_1$(H) and $\omega_2$(C) (Fig. 5a and b). This projection shows nearly the best resolution achievable for the aliphatic carbon dimension and can be directly compared with $^{13}$C-HSQC. A number of extra peaks were observed that may be attributed to the short 4-step phase cycle. It is apparent from Fig. 5c that the conventionally acquired 2D (HC)CH-TOCSY spectrum exhibits the same distortions.

The quality of processing can be quantified by measuring the noise level in the nuFT processed spectrum and in the final SSA spectrum (Fig. 6a). The noise level was determined as the median of the absolute values of all intensities in a 3D cube for each point in the acquisition dimension $\omega_4$. Initially, the spectrum is heavily distorted by sampling artefacts, which can be 10 times larger than the thermal noise, and 2.7 times on average (in the −1 to 6 ppm region).

A similar experiment with significantly lower resolution was recently presented for the small protein GB1 (56 amino acids,

6.2 kDa) in 1.0 mM concentration [16]. It was reported that CLEAN processing resulted in a 44% decrease of the apparent noise level (from 2.4 to 1.4 times the thermal noise level) on average, and 78% at best. In the case above, SSA processing achieved an average reduction of noise level of 55% (in the region −1 to 6 ppm), and 90% at best. Our results, combined with previous simulations [2,17], suggest that SSA can suppress sampling artefacts more efficiently than FFT-CLEAN method.

One should comment that the mean noise reduction should not be considered as an accurate measure when averaged over the entire spectrum in the presence of thermal noise and residual solvent signals. The latter limit the achievable reduction of the noise level, and the statistics should be corrected in this respect (Fig. 6b). We suggest that a representative region of the acquisition dimension that is reasonably far from the solvent signal be used for statistics. In the case of 4D HCCH-TOCSY the noise intensity is reduced by 71% (from 4.1 to 1.2 times the thermal noise) *in the region between* 1 *and* 3 *ppm*. Obviously, less reduction of the noise level can be obtained in regions with medium or weak peaks, where thermal noise dominates over artefacts (Fig. 6b). This suggests that artefact suppression algorithms are less effective for experiments of low sensitivity, for which nuFT might be employed as well.

It might be questioned whether HCCH-TOCSY requires an advanced method for reduction of sampling artefacts, as the S/N in the nuFT spectrum seems sufficient for peak identification (Fig. 6c). However, even for medium-sized proteins this experiment cannot be successfully analysed without removal of artefacts. Examples of 2D cross-sections (Fig. 7g–l) demonstrate that artefacts obscure most resonances, rendering the analysis inefficient or even impossible. This is due to the fact that sampling artefacts originating from peaks in the same $\omega_4$ region accumulate, decreasing the signal-to-artefact ratio determined by the sampling scheme by a further factor $\approx \sqrt{K}$, where $K$ is the number of resonances of comparable amplitude, which is usually significant in the case of CC-TOCSY experiments. The SSA processed spectrum features a considerably higher S/N ratio and, owing to the high redundancy of information (compare Fig. 7a–f), enables one to resolve side-chain signals even in cases with severe overlap.

The possibility of recording and processing high-resolution 4D HCCH-TOCSY spectra calls for new strategies of analysis. We suggest that spin systems of methyl containing residues be first identified, since the corresponding region of $^{13}$C-HSQC is usually best resolved and $T_2^*$(C) are typically longer for methyl resonances. These resonances can be assigned on the basis of $C^{\alpha}/H^{\alpha}$ (and optionally also $C^{\beta}/H^{\beta}$) chemical shifts observed in the 2D cross-sections at the $^1$H ($F_4$) and $^{13}$C ($F_3$) coordinates of a methyl group. Alternatively, one can start from the $H^{\alpha}/C^{\alpha}$ region, which is also quite well dispersed, however, overlap and the presence of residual solvent signals may hinder the assignment process. Since the manual analysis of such a complex spectrum might be inconvenient, automation that utilizes a high level of data redundancy is desired.

### 3.2. 4D C, N-edited NOESY

The symmetry of four-dimensional NOESY spectra, which correlate two heteronuclear pairs of spins, is currently not fully utilized due to sampling restrictions and prohibitively long measurements. Typically their 3D counterparts (either $^{13}$C or $^{15}$N-edited) are recorded, which suffer from frequent cross-peak overlap, even if optimal spectral resolution is achieved. Very few examples of NUS 4D NOESY spectra were reported so far [36–38]. Indeed, NOESY spectra are quite challenging for processing methods, since they require (i) a faithful linear reproduction of the peak amplitudes, (ii) the preservation of a high dynamic range and (iii) the capability of processing complex spectra with thousands of resonances. 4D NOESY spectra can be processed with the MDD method provided three-dimensional decomposition is employed [36]. Consequently, this approach is limited to moderate sampling densities, typically above ~10% [13]. Another approach was demonstrated by Werner-Allen and co-workers [37] who employed the FFT-CLEAN method for processing 4D amide-amide NOESY spectra with suppression of diagonal peaks. It was shown that neither FFT-CLEAN [37] nor SSA [17] affects the relative intensities of resonances, so that they can provide reliable structural information from NOESY peak amplitudes. Here we present an application of SSA to a 4D C, N-edited NOESY spectrum, which provides valuable amide -aliphatic cross-peaks.

Due to the absence of diagonal peaks and associated artefacts, 4D C, N-edited NOESY spectra are less demanding with respect to signal processing. Indeed, as shown in Fig. 8a. sampling artefacts in the nuFT spectrum tend to dominate over thermal noise only in the region between 8 and 8.5 ppm, where overlapping resonances from the unstructured part of Engrailed accumulate. Fig. 8b illustrates that most sampling artefacts were suppressed efficiently by SSA. Although neither CLEAN nor SSA can remove NUS artefacts completely, a flat baseline showing only thermal noise was achieved. The reduction of the apparent noise level was 5% on average in the region between 7.2 and 9.6 ppm, reaching a maximum of 30% in the most crowded region. It may be argued that artefact cleaning procedures are very useful in the case of NOESY spectra where diagonal peaks are lacking or suppressed [37]. However, even a modest improvement of the signal-to-artefact ratio may reveal new resonances and reduce peak volume distortions.

It is noteworthy that 4D C, N-NOESY should be acquired with high spectral resolution in the nitrogen dimension to avoid ambiguities in signal assignment. In the presented case, a maximum evolution time of 50 ms was used to resolve almost all resonances in the corresponding $^{15}$N-HSQC. For the reasons mentioned above, the maximum evolution time for the aliphatic carbons was limited to 10 ms. For sensitivity reasons, the experiment consisted of 6500 sampling points (86 h). Despite this rather lengthy acquisition, only a fraction of 0.5% of the samples of the corresponding conventional experiment was collected. Due to the fact that the signal-to-artefact ratio does not depend on the sampling density for non-uniform FT [15], both FFT-CLEAN or SSA seem to be methods of choice for *high-resolution* 4D spectra.

The benefits of high dimensionality and high resolution are clearly visible when corresponding cross-sections of 3D and 4D spectra are compared (see Fig. 9). While it is possible to achieve good sensitivity at very high fields (1 GHz in this case), conventional sampling results in severe overlap of cross-peaks in the 3D $^{15}$N-edited NOESY (Fig. 9a and d). The additional carbon-13 dimension in 4D C, N-edited NOESY enables one to resolve overlapping resonances, identify more of them, and improve the accuracy of peak volumes. Moreover, the correlation of proton and carbon frequencies greatly simplifies the assignment of cross-peaks.

4D C, N-edited NOESY is not only useful for structural studies but may also may assist in side-chain resonance assignment, e.g., when long mixing times allow spin-diffusion to be effective [39]. In particular, this experiment can be recorded at the highest field available, while the RF power required for TOCSY-based techniques, which increase proportionally to $B_0^2$, make them technically demanding in very high fields. This approach might also be promising for larger proteins that yield HC(CC-TOCSY)CONH of poor quality and extremely crowded 4D HCCH-TOCSY.

## 4. Experimental

The 4D HCCH-TOCSY and 4D C, N-edited HMQC-NOESY-HSQC spectra have been recorded for 0.5 mM $^{13}$C, $^{15}$N doubly labelled Engrailed protein sample in 9:1 $H_2O/D_2O$ at pH 6.0 at 303 K on a

Varian NMR System 700 spectrometer equipped with a Performa XYZ PFG unit using the 5 mm triple resonance probe. The experimental protocol for expression and purification of Engrailed has been described elsewhere [26]. A mixing time of 160 ms was used for the 4D C, N-edited NOESY experiment. For the 4D HCCH-TOCSY spectrum, z-DIPSI-3 mixing with $\gamma B_1/(2\pi)$ = 7.2 kHz and a duration of 15 ms was used.

The 2D H(CN)H-HMQC-NOESY-HMQC experiment has been performed for a 0.5 mM sample of 0.5 kDa maltose binding protein (MBP, 371 amino acids) in 19:1 $H_2O/D_2O$ at pH 7.2 at 298 K. The sample was purchased from Cambridge Isotope Laboratories Inc.

An acquisition time of 85 ms was used in all cases; a cosine-squared weighting function was applied in the acquisition dimension. Inter-scan delays of 1.2 s and 1.6 s were used for 4D C, N-edited NOESY and HCCH-TOCSY, respectively. In both cases, four scans were coherently added for each acquired FID signal. No weighting functions were employed in the indirect dimensions of the sparsely sampled experiments. The processing times on a 64-bit server running with two Intel Xeon 1.6 GHz CPUs and 8 GB RAM were 27 and 17 h for 4D HCCH-TOCSY and 4D C, N-NOESY, respectively.

The 3D $^{15}$N-edited NOESY-HSQC experiment has been recorded for uniformly $^{15}$N labelled Engrailed sample on the Bruker Avance spectrometer operating at proton frequency of 1000 MHz equipped with a cryogenic probe. Recovery delay of 1.5 s and mixing time of 150 ms were used here.

### 4.1. Software availability

The 4D SSA software in the 64-bit binary form can be found on the web site www.nmr700.chem.uw.edu.pl/software.

## 5. Conclusions

We have presented a four-dimensional SSA algorithm applied to sparsely sampled 4D HCCH-TOCSY and 4D C, N-edited NOESY spectra. Both experiments have been performed for a partially unstructured Engrailed protein with 114 amino acids to demonstrate the benefits for the analysis of complex spectra of high dimensionality and very high spectral resolution. It was shown that SSA is capable of processing spectra obtained with a sampling density as low as 1%. In the case of 4D HCCH-TOCSY an efficiency of NUS artefact suppression up to 90% was achieved, yielding a decrease of average noise level by 55%. Additionally, we have proposed the optimization of an HMQC block that minimizes losses due to transverse relaxation during proton evolution. The improved HMQC has been tested for 42.5 kDa MBP and employed to record a 4D C, N-edited NOESY of Engrailed protein. The performance of SSA in applications to demanding four-dimensional spectra with moderate computational times are encouraging for a more common usage of non -uniform sampling in biomolecular research.

## Acknowledgements

This work was supported by the Bio-NMR Project No. 261863 funded by European Commission's Framework Program 7 (FP7). Financial support by the Access to Research Infrastructures activity of FP7 (Contract 228461, EAST-NMR) for conducting research in Warsaw is gratefully acknowledged. Financial support from the TGE RMN THC Fr3050 for conducting the research in Lyon (France) is gratefully acknowledged. The help of Anna Zawadzka-Kazimierczuk (University of Warsaw) in implementing optimized HMQC block is kindly acknowledged. We appreciate Prof. Geoffrey Bodenhausen for his invaluable help with the manuscript improvement.

## References

[1] T. Szyperski, D.C. Yeh, D.K. Sukumaran, H.N. Moseley, G.T. Montelione, Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 8009–8014.
[2] K. Kazimierczuk, J. Stanek, A. Zawadzka-Kazimierczuk, W. Koźminski, Random sampling in multidimensional NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 57 (2010) 420–434.
[3] B.E. Coggins, R.A. Venters, P. Zhou, Radial sampling for fast NMR: concepts and practices over three decades, Prog. Nucl. Magn. Reson. Spectrosc. 57 (2010) 381–419.
[4] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 12366–12371.
[5] L. Frydman, T. Scherf, A. Lupulescu, The acquisition of multidimensional NMR spectra within a single scan, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 15858–15862.
[6] S. Hiller, F. Fiorito, K. Wüthrich, G. Wider, Automated projection spectroscopy (APSY), Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 10876–10881.
[7] V.A. Jaravine, V.Y. Orekhov, Targeted acquisition for real-time NMR spectroscopy, J. Am. Chem. Soc. 128 (2006) 13421–13426.
[8] J.C.J. Barna, E.D. Laue, M.R. Mayger, J. Skilling, S.J.P. Worrall, Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments, J. Magn. Reson. 73 (1987) 69–77.
[9] K. Kazimierczuk, A. Zawadzka, W. Koźmiński, I. Zhukov, Random sampling of evolution time space and Fourier transform processing, J. Biomol. NMR 36 (2006) 157–168.
[10] Ē. Kupče, R. Freeman, Projection-reconstruction of three-dimensional NMR spectra, J. Am. Chem. Soc. 125 (2003) 13958–13959.
[11] M. Mobli, A.S. Stern, J.C. Hoch, Spectral reconstruction methods in fast NMR: reduced dimensionality, random sampling and maximum entropy, J. Magn. Reson. 182 (2006) 96–105.
[12] K. Kazimierczuk, A. Zawadzka, W. Koźmiński, I. Zhukov, Lineshapes and artifacts in Multidimensional Fourier Transform of arbitrary sampled NMR data sets, J. Magn. Reson. 188 (2007) 344–356.
[13] V.Y. Orekhov, I. Ibraghimov, M. Billeter, Optimizing resolution in multidimensional NMR by three-way decomposition, J. Biomol. NMR 27 (2003) 165–173.
[14] J.C. Hoch, A.S. Stern, NMR Data Processing, Wiley-Interscience, New York, 1996.
[15] K. Kazimierczuk, A. Zawadzka, W. Koźmiński, Narrow peaks and high dimensionalities: exploiting the advantages of random sampling, J. Magn. Reson. 197 (2009) 219–228.
[16] B.E. Coggins, P. Zhou, High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN, J. Biomol. NMR 42 (2008) 225–239.
[17] J. Stanek, W. Koźmiński, Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets, J. Biomol. NMR 47 (2010) 65–77.
[18] S.G. Hyberts, D.P. Frueh, H. Arthanari, G. Wagner, FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation, J. Biomol. NMR 45 (2009) 283–294.
[19] K. Kazimierczuk, V.Y. Orekhov, Accelerated NMR spectroscopy by using compressed sensing, Angew. Chem., Int. Ed. 50 (2011) 5556–5559.
[20] Y. Matsuki, M.T. Eddy, R.G. Griffin, J. Herzfeld, Rapid three-dimensional MAS NMR spectroscopy at critical sensitivity, Angew. Chem., Int. Ed. 49 (2010) 9215–9218.
[21] M.W. Maciejewski, A.S. Stern, G.F. King, J.C. Hoch, Nonuniform sampling in biomolecular NMR, in: G.A. Webb (Ed.), Modern Magnetic Resonance, Springer, Dordrecht, The Netherland, 2006, pp. 1287–1293.
[22] M. Mobli, J.C. Hoch, Maximum entropy spectral reconstruction of nonuniformly sampled data, Concepts Magn. Reson. A 32A (2008) 436–448.
[23] J.A. Högbom, Aperture synthesis with a non-regular distribution of interferometer baselines, Astron. Astrophys., Suppl. Ser. 15 (1974) 417–426.
[24] M. Mobli, A.S. Stern, W. Bermel, G.F. King, J.C. Hoch, A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment, J. Magn. Reson. 204 (2010) 160–164.
[25] S. Hiller, R. Joss, G. Wider, Automated NMR assignment of protein side chain resonances using automated projection spectroscopy (APSY), J. Am. Chem. Soc. 130 (2008) 12073–12079.
[26] R. Augustyniak, S. Balayssac, F. Ferrage, G. Bodenhausen, O. Lequin, 1H, 13C and 15N resonance assignment of a 114-residue fragment of Engrailed 2 homeoprotein, a partially disordered protein, Biomol. NMR Assignments 5 (2011) 229–231.
[27] J. Wen, J.H. Wu, P. Zhou, Sparsely sampled high-resolution 4-D experiments for efficient backbone resonance assignment of disordered proteins, J. Magn. Reson. 209 (2011) 94–100.
[28] V. Motáčková, J. Nováček, A. Zawadzka-Kazimierczuk, K. Kazimierczuk, L. Žídek, H. Šanderová, L. Krásný, W. Koźmiński, V. Sklenář, Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments, J. Biomol. NMR 48 (2010) 169–177.
[29] H.J. Dyson, P.E. Wright, Intrinsically unstructured proteins and their functions, Nat. Rev. Mol. Cell Biol. 6 (2005) 197–208.
[30] R.A. Chylla, B.F. Volkman, J.L. Markley, Practical model fitting approaches to the direct extraction of NMR parameters simultaneously from all dimensions of multidimensional NMR spectra, J Biomol. NMR 12 (1998) 277–297.

[31] K. Kazimierczuk, M. Misiak, J. Stanek, A. Zawadzka-Kazimierczuk, W. Koźmiński, Generalized Fourier transform for non-uniform sampled data, Top. Curr. Chem. (2011), doi:10.1007/128_2011_186.
[32] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes: The Art of Scientific Computing, 3rd ed., Cambridge University Press, Cambridge, UK, New York, 2007.
[33] V.A. Jaravine, A.V. Zhuravleva, P. Permi, I. Ibraghimov, V.Y. Orekhov, Hyperdimensional NMR spectroscopy with nonlinear sampling, J. Am. Chem. Soc. 130 (2008) 3927–3936.
[34] D. Rovnyak, M. Sarcone, Z. Jiang, Sensitivity enhancement for maximally resolved two-dimensional NMR by nonuniform sampling, Magn. Res. Chem. 49 (2011) 483–491.
[35] M. Sattler, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, Prog. Nucl. Magn. Reson. Spectrosc. 34 (1999) 93–158.
[36] T. Luan, V. Jaravine, A. Yee, C.H. Arrowsmith, V.Y. Orekhov, Optimization of resolution and sensitivity of 4D NOESY using multi-dimensional decomposition, J. Biomol. NMR 33 (2005) 1–14.
[37] J.W. Werner-Allen, B.E. Coggins, P. Zhou, Fast acquisition of high resolution 4-D amide-amide NOESY with diagonal suppression, sparse sampling and FFT-CLEAN, J. Magn. Reson. 204 (2010) 173–178.
[38] S. Hiller, I. Ibraghimov, G. Wagner, V.Y. Orekhov, Coupled decomposition of four-dimensional NOESY spectra, J. Am. Chem. Soc. 131 (2009) 12970–12978.
[39] Y. Xu, Z. Lin, C. Ho, D. Yang, A general strategy for the assignment of aliphatic side-chain resonances of uniformly 13C,15N-labeled large proteins, J. Am. Chem. Soc. 127 (2005) 11920–11921.