Journal of Magnetic Resonance 212 (2011) 289–298

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

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

Multi-dimensional NMR without coherence transfer: Minimizing losses in large systems

Yizhou Liu, James H. Prestegard *

Complex Carbohydrate Research Center, The University of Georgia, Athens, GA 30602, USA

ARTICLE INFO

Article history:
Received 8 May 2011
Revised 25 June 2011
Available online 10 August 2011

Keywords:
HSQC
Large proteins
Frequency encoding
Sensitivity improvement

ABSTRACT

Most multi-dimensional solution NMR experiments connect one dimension to another using coherence transfer steps that involve evolution under scalar couplings. While experiments of this type have been a boon to biomolecular NMR the need to work on ever larger systems pushes the limits of these procedures. Spin relaxation during transfer periods for even the most efficient ^{15}N–^{1}H HSQC experiments can result in more than an order of magnitude loss in sensitivity for molecules in the 100 kDa range. A relatively unexploited approach to preventing signal loss is to avoid coherence transfer steps entirely. Here we describe a scheme for multi-dimensional NMR spectroscopy that relies on direct frequency encoding of a second dimension by multi-frequency decoupling during acquisition, a technique that we call MD-DIRECT. A substantial improvement in sensitivity of ^{15}N–^{1}H correlation spectra is illustrated with application to the 21 kDa ADP ribosylation factor (ARF) labeled with ^{15}N in all alanine residues. Operation at 4 °C mimics observation of a 50 kDa protein at 35 °C.

1. Introduction

Multidimensional NMR experiments based on coherence transfer through spin–spin coupling interactions have become a mainstay in the structural characterization of biomolecules [1]. For example, assignment of resonances in proteins most frequently begins with experiments that transfer coherence from amide protons to amide nitrogens, to alpha carbons and back, allowing evolution in each of the three ^{13}C, ^{15}N, and ^{1}H dimensions. However, these experiments can be costly due to sensitivity loss during transfer steps. The problem arises in that transfer times are usually significant and fixed by the sizes of scalar or dipolar couplings. Spin relaxation occurs in these periods resulting in substantial losses of sensitivity, particularly for larger systems. Even in the simplest ^{15}N, ^{1}H HSQC (heteronuclear single quantum coherence) experiment, there are two magnetization transfer/refocusing periods of about 5 ms each in which proton transverse spin relaxation (R_2) can cause substantial signal loss. For a fully protonated 50 kDa protein at 35 °C, transverse relaxation of an amide proton is about 100 s^{-1} ($\Delta\upsilon_{1/2}$ = 32 Hz). Loss of sensitivity would amount to about 60%. Here we present a method that avoids losses due to spin relaxation in coherence transfer steps. It stems from one of the earliest means of identifying the resonant frequency of a nucleus spin-coupled to a directly observed nucleus, namely perturbation of observed spectra by selective rf irradiation in a second frequency domain [2,3]. In these early double resonance experiments, the indirect frequencies were seldom incremented in a systematic fashion, as in current Fourier transform methods, but the basis of a two dimensional method certainly existed [3]. Here we systematize the exploration of the indirect frequency domain and make exploration more efficient by multiplexing band selective decoupling and processing data in a Hadamard style experiment [4]. This approach is now facilitated by spectrometer hardware development and modern decoupling methods. Sensitivity gains of a factor of more than 2 over conventional 2D experiments such as HSQC are experimentally illustrated.

Improvement in sensitivity of the basic HSQC experiment has a long history and substantial gains have been made by other means. It is easy to show that optimal coherence transfer efficiency depends on the ratio of the transfer rate to a competing relaxation rate. Transfers are typically through scalar couplings or rotating frame cross-relaxation; the relaxation rates of interest are transverse relaxation rates R_2 or R_1^{ρ}. For transfers through cross-relaxation in the lab frame, a smaller longitudinal relaxation rate R_1 is encountered, but lab frame cross-relaxation is significant only between homonuclei. Over the years, most effort at improving efficiency has been focused on reducing the effective transverse relaxation rate by making use of relaxation interference. The TROSY and CRIPT/CRINETP experiments fall in this category [5–7]. Of course these experiments are only applicable to spin systems with large relaxation interference effects. More recently, optimal control theory has been applied to maximize coherence transfer efficiency in NMR, resulting in the ROPE and CROP schemes for transfer

* Corresponding author. Fax: +1 706 542 4412.
E-mail address: jpresteg@ccrc.uga.edu (J.H. Prestegard).

1090-7807/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2011.07.007

[8–10]. ROPE and CROP are based on an older idea of coherence transfer by selective population inversion (SPI) [11], but utilize optimal control theory to maximize transfer efficiency. CROP is in principle the most transfer efficient method, but broadband transfer is not easily achieved. The method proposed here is broad-banded, requiring only J or dipolar couplings between the spins. It does not depend on relaxation interference, although significant relaxation interference will increase the sensitivity further.

2. Experimental design

The experimental setup and data processing for a ^{15}N–^{1}H 2D-DIRECT experiment is illustrated in Fig. 1. An FID is recorded immediately following a proton excitation pulse. During acquisition, ^{15}N decoupling is selectively turned on or off for a subset of frequencies. If we had just one ^{15}N site and we knew its frequency, we would acquire two experiments, one with decoupling at this frequency and one without. The difference would suppress all non-coupled lines and produce a clear signature for the presence of the site of interest, a triplet with central and outer lines of opposite phase.

In general, our objective is to determine frequencies for a number of sites (n) with distinct and unknown frequencies. We could do this by successively collecting pairs of spectra each using a distinct decoupling frequency with a band width that ultimately dictates the resolution of the experiment, as in the early double irradiation experiments [3]. But this is not very efficient. Instead, following the philosophy of Hadamard experiments [4], FIDs are recorded, each with a unique combination of frequencies. 2^n combinations are required to cover a spectral width, SW, with a resolution band width, BW, where n = SW/BW. Alternatively, if approximate frequencies of interest are known, n is simply the number of frequencies of interest. The 2^n FIDs can be classified into two groups of 2^{n-1} FIDs each, according to whether a particular ^{15}N frequency is decoupled or not. Since the 2^{n-1} FIDs in each group correspond to all possible decoupling patterns for the other $n-1$ ^{15}N frequencies, subtracting signals of the two groups cancels out signals associated with those $n-1$ frequencies and any other signals that are not modulated by decoupling (such as from protons bonded to carbons). In contrast, signals associated with the particular ^{15}N frequency will not cancel, as its decoupling state is either all on or all off for the two groups.

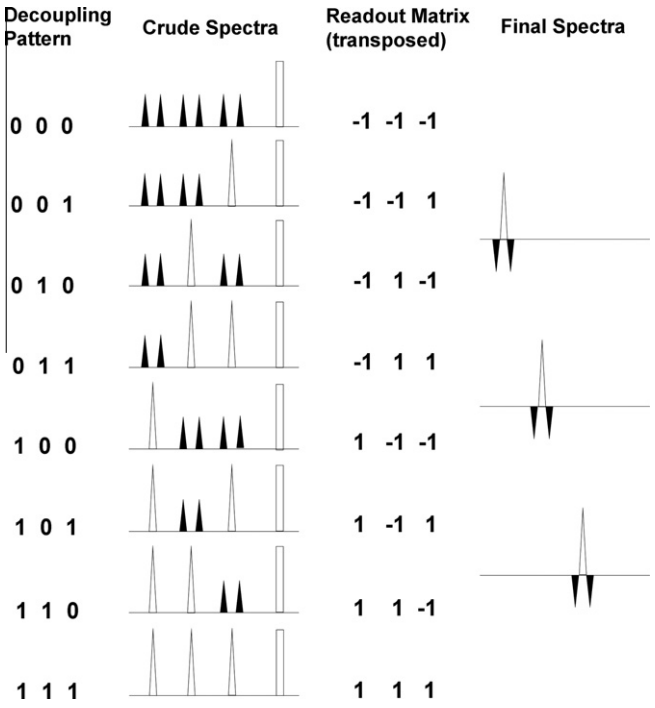

Fig. 1. Experimental design of 2D-DIRECT. In "decoupling pattern", 1 and 0 represent decoupling on and off respectively. The three digits correspond to the decoupling states for the three peaks in "crude spectra" in the same order to facilitate visualization. In reality the binary digits are uniquely mapped to ^{15}N frequencies but have nothing to do with ^{1}H peak positions. A blank triangle represents a decoupled singlet; a pair of solid triangles represents a coupled doublet; a blank square represents a proton signal that is not affected by decoupling. Manipulation of the crude spectra by a readout matrix (transposed from Eq. (2)) produces the final spectra showing triplet peaks vertically separated by ^{15}N frequencies.

For a general protocol, the decoupling flag (1 for on and 0 for off) for each frequency can be represented by a digit of a binary number, as an n digit binary number encodes 2^n possibilities. Let the most significant digit represent the decoupling state of the first ^{15}N frequency, the second digit the second frequency, and so on. The decoupling pattern for the nth FID is then represented by the binary number converted from the decimal number $n-1$. For example, the first FID is acquired with all binary digits set to 0, meaning no decoupling for any frequency, the fourth FID is recorded with the last two digits set to 1 and all others to 0, meaning decoupling only for the last two ^{15}N frequencies, and the last FID is recorded with all digits set to 1, meaning decoupling on for all frequencies. To maintain constant sample heating, decoupling off is actually achieved by decoupling at a large frequency offset. All combinations are listed for n = 3 in Fig. 1.

For processing, the crude FIDs are properly phased and Fourier transformed. The final 2-D spectrum is obtained by multiplying the FT real parts with a readout matrix, as described in the following equation:

$$M_{n\times 2^n}\cdot\begin{pmatrix} R_1^1 & \cdot & \cdot & R_1^{np} \\ R_2^1 & \cdot & \cdot & R_2^{np} \\ \cdot & \cdot & \cdot & \cdot \\ R_{2^n}^1 & \cdot & \cdot & R_{2^n}^{np} \end{pmatrix}_{2^n\times np} = \begin{pmatrix} S_1^1 & \cdot & \cdot & S_1^{np} \\ S_2^1 & \cdot & \cdot & S_2^{np} \\ \cdot & \cdot & \cdot & \cdot \\ S_n^1 & \cdot & \cdot & S_n^2 \end{pmatrix}_{n\times np} \quad (1)$$

R_j^i represents the ith point of the FT real data of the jth FID, S_j^i represents the (i,j) point of the final 2-D spectrum, with i corresponding to the ^{1}H frequency and j to the jth ^{15}N frequency. The $n\times 2^n$ readout matrix M contains only -1 and 1, with:

$$M_{ij} = 2\left\{\left(\text{int}\frac{j-1}{2^{n-i}}\right)\%2\right\} - 1 \quad (i = 1,2,\ldots,n;\quad j = 1,2,\ldots 2^n) \quad (2)$$

Here int represents taking the integer part, and the operator %2 means modulus 2, giving 0 for an even input or 1 for an odd input.

The frequency encoding and data reconstruction process of 2D-DIRECT shares conceptual similarity with Hadamard NMR spectroscopy. However there are some fundamental differences. The backbone of a Hadamard pulse sequence is still the regular nD sequence, including coherence transfer steps, but with multiplexed selective inversion pulses in place of an indirect frequency evolution period. In 2D-DIRECT, the pulse sequence reduces to a single proton excitation pulse followed by multiplexed selective decoupling, thereby avoiding relaxation penalties from coherence transfer. However, in Hadamard, the ability to invert signals, rather than frequency shift signals, can be exploited to optimize its efficiency. All peaks are manipulated in each acquisition, but with a specific combination of plus and minus signs. In the reconstruction process, appropriate linear combinations of acquisitions are made to regenerate spectra representing connections to just a single indirect frequency. In 2D-DIRECT, signals are differentiated by peak shifts between decoupled singlets and coupled doublets of the same phase. These spectra cannot be manipulated to cancel or superimpose through linear operations as in Hadamard.

Examination of the decoupling patterns in Fig. 1 shows that on average only half the frequencies can be probed at a given time, and a larger number of scans are necessary to encode a given number of frequencies.

The choice of a decoupling sequence and decoupling power is important to the success of the experiment as the ^{15}N resolution is limited by the decoupling selectivity. In general, the mean decoupling field should be at least comparable to J for efficient decoupling [12]. However, off-resonance signals within a certain range of the decoupling frequency are subject to modulation and reduction of effective couplings resulting in artifacts. With the "cool" WALTZ16 decoupling method [13], the mean decoupling field can be set roughly equal to J for reasonable on-resonance decoupling with a responsive range of 3–4 times J. An example is shown in Fig. 2. The spectra were acquired for an isolated NH peak in ubiquitin using a ^{15}N selective decoupling profile that simultaneously decouples at four distinct frequencies separated by 600 Hz with a mean decoupling field of 106 Hz. This decoupling pattern can be represented, among other possibilities, by "1010101", which is one instance of a 7-frequency encoding with uniform 300 Hz spacing. A total of 150 spectra were displayed side by side with the difference that the ^{15}N carrier frequencies were shifted in 20 Hz steps. This decoupling profile gives a decoupled span (solid bar) and a coupled span (empty bar) both of ~120 Hz, which represents the resolution under this condition. However, between the decoupled and coupled regions, there is also a frequency range that is neither well decoupled nor coupled (striped bar). Signals with ^{15}N chemical shifts within this range will give weak near-resonance artifacts that do not have the standard triplet pattern. Obviously, MD-DIRECT will benefit from a decoupling sequence that gives a sharp off-resonance drop-off. It is worth noting that the resolution and artifact suppression improves with a higher B_0 field as the resonance frequency separation gets larger while J couplings stay constant. For a similar reason, application to systems of small J couplings but significant chemical shift spreads also enjoys these benefits since a weaker decoupling field is sufficient.

The idea of encoding indirect chemical shifts through decoupling during acquisition also shares some similarity with the SITAR (spin-state selective off-resonance decoupling) method from the Riek laboratory [14,15]. However, DIRECT methods and the SITAR methods are fundamentally different. In SITAR, the J scaling effect, as observed in the reduced splittings of off-resonance decoupled multiplets, is used to retrieve chemical shift information for both observed and decoupled nuclei from a single dimension. It achieves efficiency by reducing the number of indirectly detected dimensions, but sacrifices the resolution gain of an additional dimension in the process. DIRECT, on the other hand, is a true multidimensional method, retaining a level of multidimensional resolution without an indirect evolution period and without coherence transfers.

As our process is described, we will lose a factor of two in signal to noise ratio because intensities from the singlet and the doublet cannot be combined through linear operations. It is tempting to think that this limitation may be removed by a deconvolution technique when coupling patterns are conserved across all sites of interest. However, it is critical to evaluate the effects of deconvolution on both signal and noise in order to consider its consequence. In the following, we will first demonstrate how deconvolution can restore the singlet lineshape, and then analyze the effect of deconvolution on noise and further propose an effective method to solve the noise problem.

Signal deconvolution is easily appreciated if one realizes that the FIDs are sums of exponentially decaying sinusoids of the form:

$$\sum_i (2\cos(\upsilon_i t) - \cos((\upsilon_i + \pi J/2)t) - \cos((\upsilon_i - \pi J/2)t))$$

and that the $\cos((\upsilon_i \pm \pi J/2)t)$ terms can be rewritten as the sum of $\cos(\upsilon_i t)\cos(\pi J t/2)$ and $\pm\sin(\upsilon_i t)\sin(\pi J t/2)$. Substitution leads to the realization that division by $1-\cos(\pi J t/2)$ will leave a set of frequencies representing only the central line characterized by $2\cos(\upsilon_i t)$. This is a special case of a general time domain deconvolution method based on the convolution theorem, which states: the Fourier transform of a convolved signal equals the product of the Fourier transforms of the signal and convolution functions, or mathematically,

$$F\{f \otimes g\} = F\{f\} \cdot F\{g\} \tag{3}$$

Here the operators $\otimes$ and $\cdot$ represent convolution and point-wise product respectively, and $F\{\}$ represents complex Fourier transformation. In our case, the experimental spectrum (h) can be treated as the convolution of a singlet Lorenzian line (f) and a triplet delta function (g), i.e. $h = f \otimes g$. Since the J coupling is relatively constant at 93 Hz, and the triplet should roughly retain a $-1{:}2{:}-1$ intensity ratio, the function g is known or easily tunable. The objective is to reconstruct the singlet Lorenzian f based on known functions h and g. According to Eq. (3), f can be determined as:

$$f = F^{-1}\left\{\frac{F\{h\}}{F\{g\}}\right\} \tag{4}$$

Here $F^{-1}\{\}$ represents inverse Fourier transformation. As discussed earlier, $F\{g\} = 1-\cos(\pi J t/2)$, which is the Fourier transform of the triplet delta function. An extremely low level of random noise, which is inconsequential, is added to $F\{g\}$ to avoid the division by zero problem inherent with the deconvolution method.

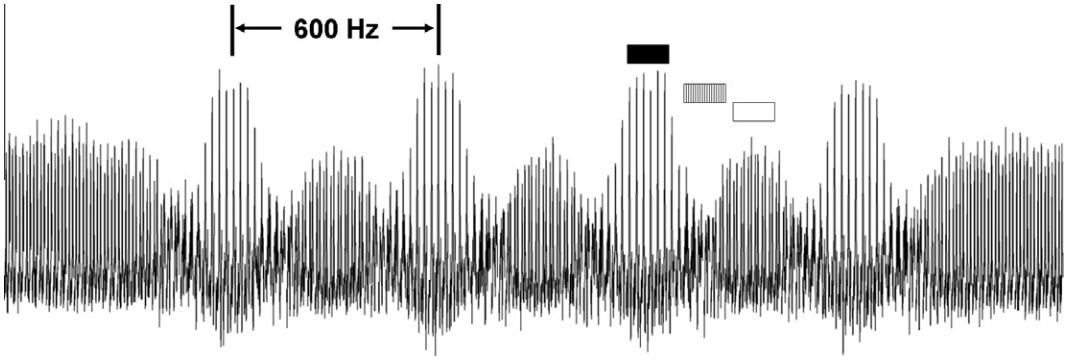

Fig. 2. An experimental demonstration of spectral response to off-resonance effects during selective multi-frequency decoupling. The spectra were acquired as described in the main text. The more intense FIDs are decoupled. The weaker ones are coupled, but the doublet pattern is illegible due to the close proximity in the display.

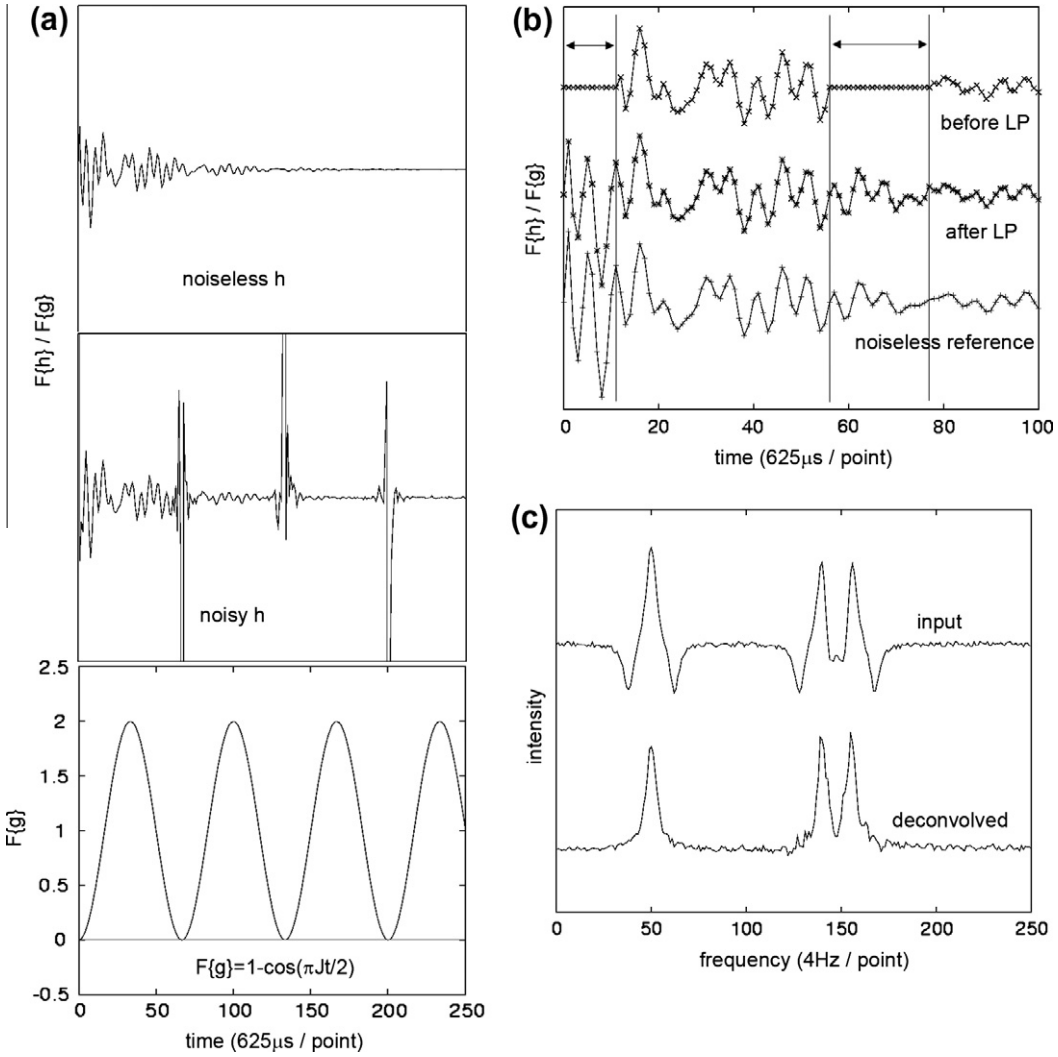

Fig. 3. Linear prediction aided deconvolution of a synthetic spectrum. (a) Top panel: the deconvolved real time-domain data for a noiseless spectrum; middle panel: the deconvolved real time-domain data for the same spectrum as in (a) but with random noise incorporated during spectral synthesis, highlighting the periodic blowing up of deconvolved noise; bottom panel: the deconvolving function $1-\cos(\pi Jt/2)$. (b) Top trace: the deconvolved real time-domain data with noise spikes clipped off for clarity (marked between the double-headed arrow lines); middle trace: data with the clipped region linear predicted; bottom: the deconvolved real time-domain data of a noiseless spectrum, serving as a reference for the evaluation of LP performance. (c) top trace: the noisy input spectrum with triplet lineshapes; bottom trace: the LP deconvolved spectrum.

Next we will examine the effect of deconvolution on noise. Note in the previous analysis that the signal intensity following deconvolution, namely $2\cos(\upsilon_i t)$, represents only half of the total signal; therefore the noise level also has to be reduced by a factor of two in order to regain the full S/N. Obviously, noise in the time-domain data is also divided by $1-\cos(\pi Jt/2)$ during this process. This number varies between 0 and 2. Therefore, the noise level is scaled by a factor between ½ and infinity, depending on time t, i.e. the average noise is infinite. An example is shown in Fig. 3a. The deconvolved real time-domain data for a synthetic noisy spectrum contains infinitely high (if digitized enough) spikes that coincide with the zero-points of $1-\cos(\pi Jt/2)$ (Fig. 3a, middle and bottom panels; the spikes are truncated to show fine details of the signal at a much lower level). These spikes lead to infinite noise power and consequently a disastrous noise level following inverse transformation (data not shown). However, here two facts should be noted that suggest a method to deal with this problem. First, the noise spikes always occur at predictable locations, i.e. where $1-\cos(\pi Jt/2)$ gets close to zero. Second, the signal to be reconstructed consists of a sum of decaying sinusoids, which allows linear prediction (LP), a method widely used in NMR signal processing, to estimate intensities in the future or past based on an available set of data [16–20]. Here we replace the highly noisy points produced during deconvolution with ones predicted based on the more signal-dominated portion of data. An LP method based on singular value decomposition (SVD) [16,21], the so called LPSVD, is employed using a program written in house. LP coefficients are obtained from points 12–55 in Fig. 3b, to recreate data enclosed in the double-headed arrows. Fig. 3b compares the deconvolved time-domain data from a noisy spectrum with spikes clipped off to zero for clarity, a noisy spectrum with the zeroed values refilled by LP, and a noiseless reference spectrum containing identical signals. Comparison of the later two shows that LP is able to recreate the main features of signal oscillation. Fig. 3c displays the deconvolved spectrum after LP in comparison to the simulated input noisy spectrum. As expected, the singlet lineshape is restored, and canceled intensities are also recovered as seen for the 2 overlapping peaks upfield. Finally, it should be mentioned that the noise level in the deconvolved spectrum greatly depends on data processing factors such as the percentage of prediction and accuracy of the LP coefficients. For this reason, we do not consider the application of deconvolution during sensitivity comparison of MD-DIRECT and other conventional methods in what follows.

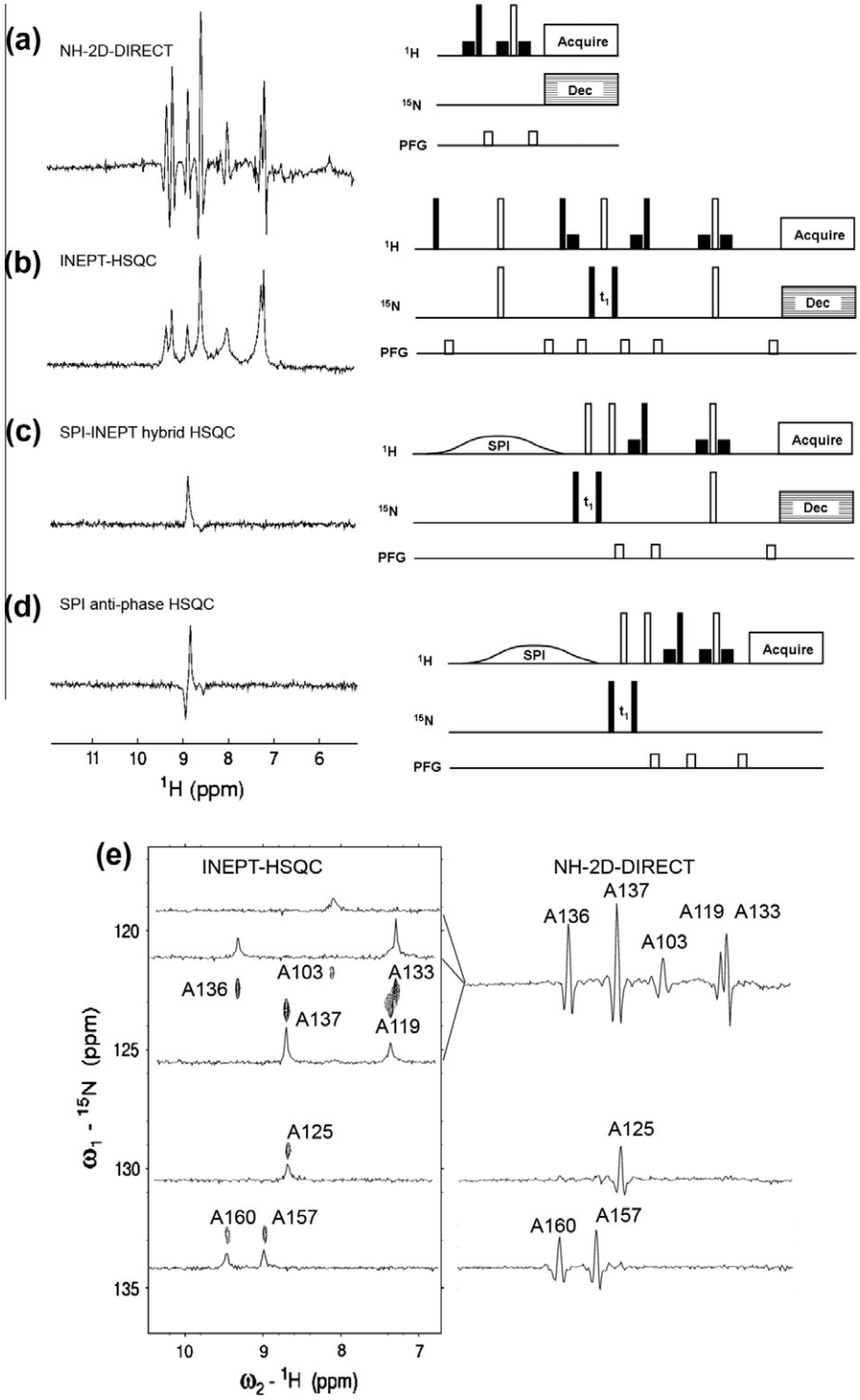

Fig. 4. Experimental sensitivity comparison of 2D-DIRECT and other 2D NH experiments. (a) The "cool" WALTZ16 sequence with a mean decoupling field of 1097 Hz was applied at 126.5 ppm for the decoupled spectrum. For the coupled spectrum, the same decoupling setup was used, but with the decoupling center shifted to 237.6 ppm, to maintain constant sample heating. A water flip-back 2D-DIRECT pulse sequence is shown on the right. A 3-9-19 water-gate is used in the spin echo. (b) The coherence transfer delay was experimentally optimized at 4.38 ms. (c) The experimentally optimized Gaussian inversion pulse has a mean strength of 43.8 Hz and is applied at 49 Hz upfield to the resonance frequency of Ala125. (d) The reverse INEPT is replaced with a much shorter watergate spin-echo for anti-phase SPI-HSQC. Of course ^{15}N decoupling must be turned off. (e) For 2D INEPT-HSQC, a ^{15}N chemical shift evolution time of 20 ms is used. For 2D 2D-DIRECT, three ^{15}N frequencies are resolved at 122.6, 129.3, and 132.8 ppm, with mean decoupling fields of 219, 138, and 138 Hz respectively. The two experiments take the same amount of experimental time (about 30 min).

3. Results

Fig. 4 shows 2D-DIRECT spectra of a 0.5 mM ^{15}N-Ala selectively labeled myristoylated ARF·GDP protein (21 kDa) along with spectra acquired using several other conventional methods. The spectra were collected on a 900 MHz Varian VNMRS spectrometer equipped with a cold probe. The experiments were conducted at 4 °C to create a condition where the ratio of the line width relative to scalar coupling ($R_2/\pi J$) is approximately 0.35 The 2D-DIRECT spectrum (Fig. 4a) clearly gives superior sensitivity to other pulse

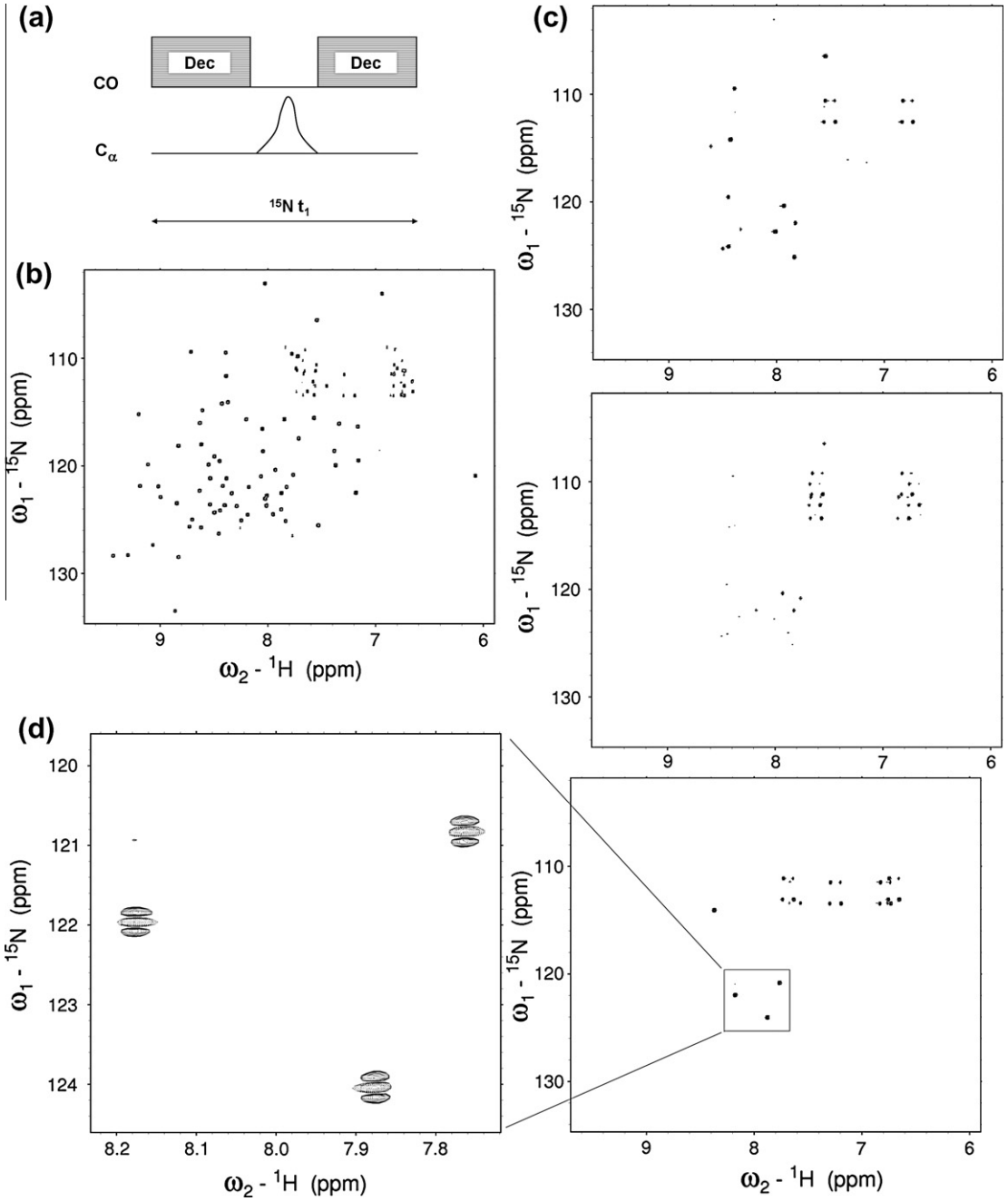

Fig. 5. The 3D MD-DIRECT-HNCO-TROSY experiment. (a) The pulse sequence is applied during ^{15}N chemical shift evolution. A REBURP [31] refocusing pulse centered on 56 ppm covering 50 ppm is used to refocus N–C coupling. Evolution of the small N–C' coupling during this refocusing pulse is negligible. (b) A 2D SE-TROSY spectrum [32] showing all NH peaks. (c) Three NH planes from a 3D MD-DIRECT-HNCO-TROSY. (d) An expanded view of the third NH plane, showing the triplet pattern of the ^{15}N dimension.

sequences tested, including INEPT-HSQC, SPI-INEPT hybrid HSQC, and anti-phase detected SPI-HSQC (Fig. 4b–d). The 2D-DIRECT spectrum in Fig. 4a was actually acquired as a 1D analog. It is the difference of a ^{15}N coupled spectrum and a broadband decoupled spectrum using a "cool" WALTZ16 sequence with a decoupling field of 1097 Hz applied at 126.5 ppm that effectively decouples the entire ^{15}N frequency span. The characteristic triplet lineshape is evident. The other spectra (Fig. 4b–d) were acquired as single FIDs for the first point of the indirect dimension. In principle, the second FID, corresponding to a different ^{15}N phase selection, should be acquired and processed along with the first FID to give a phase-sensitive spectrum. However, in this particular case, in which the ^{15}N chemical shift has not evolved, the second FID actually does not contribute to the real part of the Fourier transformed data. As a result, b–d were collected with just half the time of the 2D-DIRECT spectrum to allow a fair comparison. In general, the second FID must be collected leading to the $\sqrt{2}$ penalty normally encountered in indirect frequency sampling. The transfer delay in INEPT-HSQC was experimentally optimized. In selective population inversion (SPI) experiments (Fig. 4c and d), a 10% truncated Gaussian spin-state selective inversion pulse was applied to Ala157 to generate a coherence transfer, and the bandwidth and frequency offset were experimentally optimized. It was previously shown that an optimized Gaussian inversion pulse achieves comparable transfer efficiency as the CROP sequence [22]. Indeed, comparison of Fig. 4b and c does suggest a considerable advantage

from SPI. In this hybrid experiment (Fig. 4c), the first INEPT was replaced by a SPI pulse. A double-SPI in-phase HSQC can be easily created, but a stand-alone water-gate seems necessary for satisfactory water suppression which makes fair comparison difficult. Interestingly, the anti-phase detected SPI-HSQC (Fig. 4d) gives a substantially stronger signal than in-phase detected experiments despite intensity cancelation in the anti-phased spectrum, mainly due to the absence of the refocusing step. By avoiding both coherence transfer and refocusing steps, 2D-DIRECT (Fig. 4a) achieves the best sensitivity of all. Enhancements relative to in-phase INEPT-HSQC based on Ala157 are 20% for in-phase SPI-INEPT hybrid HSQC, 40% for anti-phase SPI-INEPT hybrid HSQC, and 68% for NH-2D-DIRECT respectively.

Fig. 4e compares 2D INEPT HSQC and a 2D-DIRECT using three band frequency selection as described above. Now, indirect phase discrimination has to be done for the HSQC. So these two experiments take the same amount of spectrometer time; for the 2D-DIRECT, eight FIDs were collected to encode three frequencies, each accumulated over 160 scans; For 2D INEPT HSQC, 40 indirect complex points (80 FIDs) were acquired over a 2000 Hz range giving an acquisition time of 20 ms, and each FID was averaged over 16 transients. Slices from the HSQC through the indirect peak centers are shown on top of the 2D spectrum on the left. On the right are the three slices from the 2D-DIRECT spectrum. The sensitivity advantage from 2D-DIRECT is quite striking. The difference in the 2D versions is more pronounced than 1D scans, due to extra relaxation loss during ^{15}N chemical shift evolution in the 2D HSQC. The effect of relaxation during chemical shift evolution on the FT spectrum can be readily be understood by considering a simple case in which no linear prediction, zero filling or window function is applied. Relaxation causes the intensity to be Lorentzian-distributed between several points on the indirect dimension. This distribution lowers the peak height and accounts for the larger difference seen in 2D experiments than in 1D ones. Although adding up these points, i.e. integrating the indirect lineshape, will restore the full intensity, in doing so, noise is also accumulated, by a factor of $\sqrt{n}$, where n is the number of points. In contrast, in MD-DIRECT, all intensity is focused on a single point. In our case, the gain of DIRECT over HSQC is around a factor of 2.

MD-DIRECT can also be applied to a conventional *nD* method, such as one based on an HSQC readout, to introduce an extra dimension. For example, selective carbonyl decoupling by the MD-DIRECT scheme during ^{15}N chemical shift labeling of an HSQC or TROSY detection scheme can produce a 3D MD-DIRECT-HNCO spectrum (Fig. 5a). Spectrum reconstruction is identical to that in 2D-DIRECT except that the 1D slices (R_j^i) in Eq. (1) are now 2D planes. In MD-DIRECT-HNCO, the long N–C transfer is avoided, and superior sensitivity can be expected when $R_2^N/{}^1J_{NC}$ is large. The ^{15}N TROSY effect can be utilized to sharpen the indirect line and reduce attenuation from subtracting the doublet, simply by running a TROSY-HSQC experiment. Fig. 5c shows three planes from a 3D MD-DIRECT-HNCO-TROSY experiment collected for ubiquitin. The three planes correspond to three ^{13}C′ frequencies separated by 100 Hz decoupled with a mean field of 33 Hz. Note that due to the small size of the NC coupling, a low decoupling field can be used and thereby a relatively high ^{13}C' resolution is possible. The ^{15}N lineshape is the characteristic triplet (Fig. 5d). Because the relaxation interference between ^{15}N chemical shift anisotropy and ^{15}N–^{13}C′ dipolar coupling is negligible, the two coupled peaks have nearly identical linewidths, and therefore the central peak in the triplet can provide an accurate ^{15}N chemical shift measurement, which is critical for frequency-resolved RDC measurements. A low power ^{13}C′ decoupling is preferred during ^{1}H acquisition to decouple a ~5 Hz $^2J_{HC}$ coupling that otherwise leads to an undesired E.COSY pattern.

Finally, we demonstrate the application of LP deconvolution to the experimental MD-DIRECT data to restore the singlet lineshape. The 2D-DIRECT spectrum of ^{15}N-Ala ARF, previously shown in Fig. 4a, is used as the experimental input (Fig. 6a, top trace). During deconvolution, the time-domain points are replaced with predicted values when the denominator $1-\cos(\pi Jt/2)$ is smaller than a cutoff of 0.5, which represents 1/3 of the complete data. A prediction order (number of LP coefficients) of 20 is used. The most reliable region, between 6.5 and 33.3 ms, is used to estimate LP coefficients. In practice, the cutoff value and prediction order can be readily adjusted based on the quality of the deconvolved spectrum. The result is shown as the middle trace of Fig. 6a. Clean singlet lineshapes are restored for a majority of the peaks. However, some baseline wiggling is apparent, particularly near the two upfield peaks. Inspection of the time-domain data from a synthetic source indicates that in the presence of noise, LP tends to systemically overpredict for forward prediction and underpredict for backward prediction. This is illustrated in Fig. 6b, with LP data for a synthetic noisy spectrum and reference data from a noiseless spectrum. The reference-vs-prediction correlations have slopes of 1.29 and 0.91 for forward and backward predictions respectively, with a slope of 1 representing unbiased prediction. These discrepancies are consistent with the relaxation time constant being overpredicted. This overprediction reflects inaccuracies in the LP coefficients that depict signal relaxation. The reason for the inaccuracies may be that LP coefficients in these examples are derived from data within a relatively short time range, during which signal

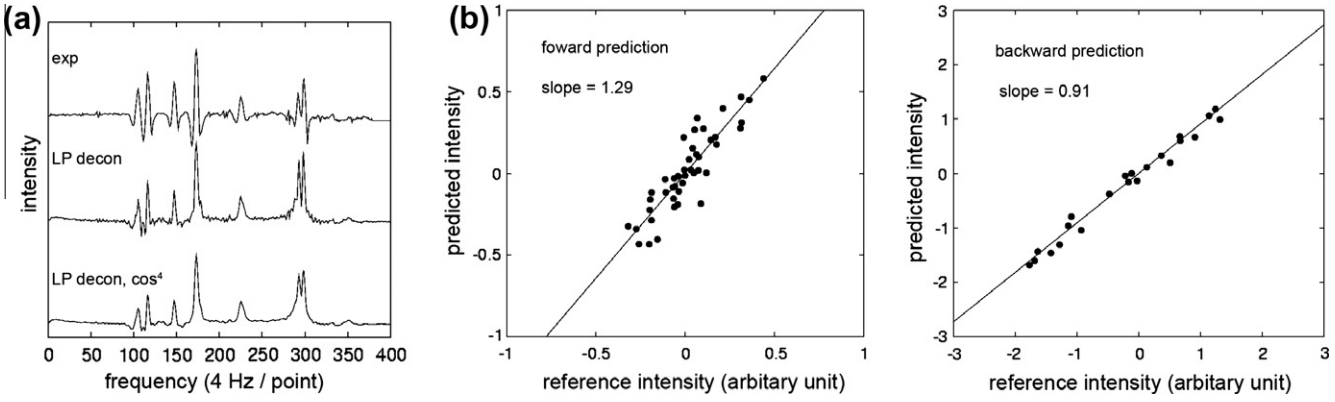

Fig. 6. LP deconvolution of an experimental spectrum, 2D-DIRECT of ^{15}N-Ala ARF1, and an illustration of a systematic error of LP that can lead to truncation artifacts. (a) Top trace: the experimental spectrum, identical to the one shown in Fig. 4a; middle trace: LP deconvolved spectrum displaying truncation artifacts; bottom trace: the spectrum with a cos^4 window function applied after LP. (b) Correlation plots of reference values vs predicted values, for forward prediction (left) and backward prediction (right). The reference values come from the deconvolved time-domain data of a noiseless spectrum. Both real and imaginary parts of the predictions are used to make these plots.

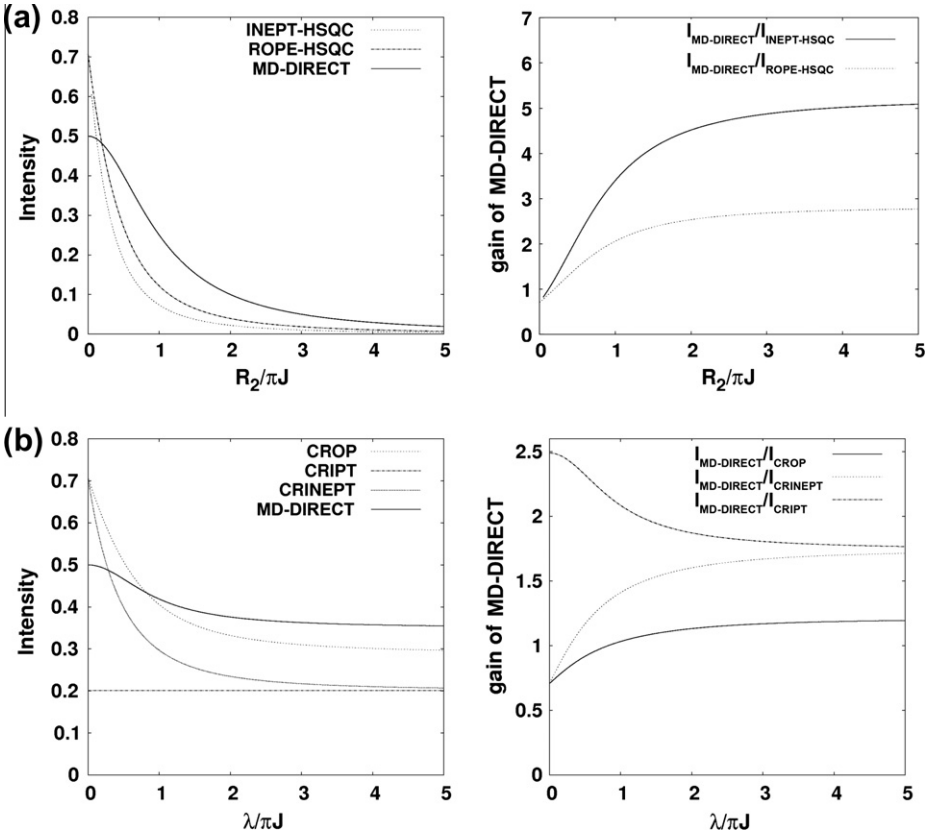

Fig. 7. Theoretical sensitivity comparison of 2D-DIRECT and other 2D NH experiments. (a) In-phase HSQC experiments are considered, which take two coherence transfer steps. Relaxation interference is not taken into account. (b) Anti-phase detected experiments are considered, which require only one coherence transfer step. A high level of relaxation interference ($\eta/\lambda = 0.7$) is taken into account.

has not undergone sufficient decay, which, coupled with the oscillatory nature of the signal, complicates relaxation time estimation; this can particularly be a problem for slow relaxing signals such as the two upfield peaks in Fig 6a. Indeed, when a smaller cutoff value for $F\{g\}$ is used so that more data are preserved in the prediction template, these discrepancies are migitated (data not shown). But of course, an overly small cutoff runs the risk of keeping too much deconvolution noise in the template. Finally, back to the source of the baseline wiggling problem, when the off-predicted data are interleaved with experimental data with a slight amplitude mismatch, a well-known truncation effect results after Fourier transform. As expected, a damping apodization function largely removes the truncation effect, as shown in the bottom of Fig. 6a, in which a $\cos^4$-bell window function is applied. Of course, the linewidth also gets broader, which is particularly noticeable for the slowly relaxing signals. Finally, it is worth mentioning that other data reconstruction methods, in particular maximum entropy reconstruction [23], can potentially give more reliable results over LP in the presence of low S/N or a small data template size [24].

4. Discussion

It is useful to analyze in more detail the origin of the sensitivity improvements seen in MD-DIRECT. We expect these to come primarily from the avoidance of relaxation losses during INEPT transfer and refocusing steps, but other factors enter as well. For conventional 2D FT experiments, a penalty of $2^{1/2}$ is encountered for indirect dimension frequency determination, because the detected signal comes from the projection of a rotating magnetization vector along one direction. This is avoided in sensitivity-enhanced (SE) experiments in which magnetizations projected on two orthogonal directions are simultaneously detected. However, we do not make comparisons to SE experiments here as the required extra transfer delay causes more relaxation loss for large systems than the potential gain of $2^{1/2}$. Another potential source of signal loss in conventional experiments comes from relaxation during ^{15}N chemical shift evolution. This loss is very dependent on experimental setup and hard to generalize. But importantly, the 2D-DIRECT method described does not suffer losses from either of these effects. This reduces the expected loss from our superimposed doublet and singlet line shapes to less than $2^{1/2}$ in cases where multiplet components are well resolved. In the case where larger line widths degrade resolution, taking the difference of a singlet and a doublet in the data reconstruction process for 2D-DIRECT causes some additional intensity cancelation. Nevertheless, we expect gains due to minimization of relaxation loss to more than compensate for this loss. Moreover, the cancelation due to overlapping signals of opposite sign can be corrected by deconvolution if needed, as shown earlier (Figs. 3c and 6a). But because of the difficulty in quantitating noise in LP deconvolution, this correction is not taken into account in the following comparisons.

In Fig. 7a, the theoretical sensitivity, including expected relaxation losses, for the 2D-DIRECT sequence is compared to those of in-phase INEPT-HSQC and ROPE-HSQC. In the later two experiments, relaxation loss is considered for two coherence transfer/refocusing steps as they occur in these in-phase 2D pulse sequences. Relaxation interference is neglected in these comparisons. The result shows that at a line width to scalar coupling ratio of 1 (a molecular weight of approximately 60 kDa at 25 °C), 2D-DIRECT can be over three times as sensitive as INEPT-HSQC and nearly two times as sensitive as ROPE-HSQC.

In Fig. 7b, relaxation interference is taken into account, and 2D-DIRECT is compared to a number of 2D methods designed to utilize

the interference effect, including CRIPT, CRINEPT, and CROP. A cross-correlation (η) to auto-correlation (λ) ratio of 0.7 is used in the simulation. When interference is strong, it is advantageous to record an anti-phase spectrum in coherence transfer experiments. In doing so, the refocusing step after ^{15}N frequency labeling and transfer to ^{1}H is eliminated to avoid signal loss. In the anti-phase spectrum, maximum intensity occurs near the strong ^{1}H TROSY signal but generally deviates from the exact TROSY frequency due to the superposition of the anti-TROSY signal. Similarly, maximum intensity in MD-DIRECT may no longer occur near the central position, but instead near the ^{1}H TROSY position. The maximum intensity of each experiment is used to make Fig. 7b. The results show that for $\lambda/\pi J$ above 1, 2D-DIRECT is about as sensitive as the most transfer-efficient CROP based sequence without restrictions on frequency selection, and about 30% more sensitive than CRINEPT. Another advantage of 2D-DIRECT is that no change needs to be made to the pulse sequence in order to benefit from relaxation interference, as compared to CRIPT and CRINETP in which the transfer delay must be optimized for a specific cross-correlation rate, and CROP in which the selective inversion pulse should be carefully optimized. Also, one cannot always take full advantage of relaxation interference. If deuteration of protein targets is difficult, as it is in proteins that must be expressed in mammalian cells [25], reduction of linewidths by interference is often minimal. Details of the simulations are given in the supporting information.

The major limitation of MD-DIRECT is not in sensitivity, but in experimental time and resolution. The requirement for 2^n decoupling patterns makes acquisitions long for more than 10 frequencies. If one scan takes 1 s, then experimental time is 2^n seconds, where n is the number of frequencies of interest. So an $n = 10$ experiment minimally requires about 17 min. The examples we have shown require very limited resolution in the indirect dimension. However, these are not unique situations. Cases of sensitivity enhancement in large molecules include a labeled peptide bound to a large protein or a piece of membrane. Also large proteins that must be expressed in mammalian cells are frequently produced in sparsely labeled forms by using a single, or a small number of isotopically labeled amino acids [26–28]. For proteins expressed in mammalian cells use of a limited set of amino acids is often essential because including all amino acids in a uniform labeling medium can be very expensive. Deuteration is also not currently possible in mammalian host cells, removing much of the advantage of TROSY based sequences. New assignment strategies are making it possible to assign sparsely labeled spectra [29,30], and sparse constraints from long range paramagnetic relaxation effects and RDCs are making at least low resolution structure determination possible when supplemented with computational modeling [31–34]. Hence we anticipate an increased number of cases where large systems have relatively small numbers of peaks.

Band selective decoupling requirements also become less limiting if positions of resonances are known. Knowing approximate frequencies for peaks of interest does occur frequently. Multiple point spin relaxation studies, amide proton exchange studies, modulation based measurements of residual dipolar couplings (RDCs) and ligand binding studies are all cases where this occurs.

We have illustrated a case where limited resolution is adequate in 3D studies. High resolution in a uniformly resolved $^{13}C'$ dimension of an HNCO spectrum is impractical for systems of many signals, because to resolve n $^{13}C'$ frequencies, one must acquire 2^n 2-D planes. But in many cases, such as residual dipolar coupling (RDC) and paramagnetic relaxation enhancement measurements, one only needs to resolve through a $^{13}C'$ dimension a small set of signals that overlap in 2D NH experiment. In these cases, MD-DIRECT-HNCO can be readily applied.

We also demonstrated the utility of one deconvolution method aided by LPSVD to restore singlet lineshape for MD-DIRECT experiments. This processing method increases spectral resolution and sensitivity in the presence of overlapping signals with opposite sign. Further improvements using maximum entropy methods may be possible [23,24].

In summary, MD-DIRECT is a general scheme for performing nD NMR spectroscopy that offers significantly higher sensitivity than commonly used coherence transfer based methods. The resolution of the indirect dimension is degraded but it improves with higher field strength or smaller couplings. MD-DIRECT basically requires no optimization to achieve the theoretical maximal sensitivity, unlike the coherence transfer experiments in which careful optimization is needed. It can also be easily grafted to conventional coherence transfer experiments to introduce an extra dimension while keeping the high-resolution feature of the other dimensions, and this can be done with trivial changes to pulse sequences. Systems that benefit most from MD-DIRECT are those with fast transverse relaxation or small couplings.

Acknowledgments

This work was supported by grant R01 GM061268 from the National Institute of Genral Medical Sciences. The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmr.2011.07.007.

References

[1] K. Wuthrich, NMR studies of structure and function of biological macromolecules (Nobel Lecture), Angew. Chem., Int. Ed. 42 (2003) 3340–3363.
[2] R.R. Ernst, Nuclear magnetic double resonance with an incoherent radio-frequency field, J. Chem. Phys. 45 (1966) 3845.
[3] W.A. Anderson, R. Freeman, Influence of a second radiofrequency field on high-resolution nuclear magnetic resonance spectra, J. Chem. Phys. 37 (1962) 85.
[4] E. Kupce, T. Nishida, R. Freeman, Hadamard NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 42 (2003) 95–122.
[5] K. Pervushin, R. Riek, G. Wider, K. Wuthrich, Attenuated T-2 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. USA 94 (1997) 12366–12371.
[6] R. Riek, G. Wider, K. Pervushin, K. Wuthrich, Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules, Proc. Natl. Acad. Sci. USA 96 (1999) 4918–4923.
[7] D.W. Yang, L.E. Kay, Improved (HN)-H-1-detected triple resonance TROSY-based experiments, J. Biomol. NMR 13 (1999) 3–10.
[8] N. Khaneja, T. Reiss, B. Luy, S.J. Glaser, Optimal control of spin dynamics in the presence of relaxation, J. Magn. Reson. 162 (2003) 311–319.
[9] N. Khaneja, J.S. Li, C. Kehlet, B. Luy, S.J. Glaser, Broadband relaxation-optimized polarization transfer in magnetic resonance, Proc. Natl. Acad. Sci. USA 101 (2004) 14742–14747.
[10] N. Khaneja, B. Luy, S.J. Glaser, Boundary of quantum evolution under decoherence, Proc. Natl. Acad. Sci. USA 100 (2003) 13162–13166.
[11] K.G.R. Pachler, P.L. Wessels, Selective population inversion (Spi) – pulsed double-resonance method in FT NMR-spectroscopy equivalent to indor, J. Magn. Reson. 12 (1973) 337–339.
[12] A.J. Shaka, J. Keeler, Broadband spin decoupling in isotropic liquids, Prog. Nucl. Magn. Reson. Spectrosc. 19 (1987) 47–129.
[13] E. Kupce, J. Boyd, I.D. Campbell, Cool decoupling and mixing wave-forms, J. Magn. Reson., Ser. A 110 (1994) 109–112.
[14] C.R.R. Grace, R. Riek, Pseudomultidimensional NMR by spin-state selective off-resonance decoupling, J. Am. Chem. Soc. 125 (2003) 16104–16113.
[15] R. Keller, C.R.R. Grace, R. Riek, Fast multidimensional NMR spectroscopy by spin-state selective off-resonance decoupling (SITAR), Magn. Reson. Chem. 44 (2006) S196–S205.
[16] H. Barkhuijsen, R. Debeer, W.M.M.J. Bovee, D. Vanormondt, Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure, J. Magn. Reson. 61 (1985) 465–481.
[17] H. Barkhuijsen, R. Debeer, D. Vanormondt, Improved algorithm for noniterative time-domain model-fitting to exponentially damped magnetic-resonance signals, J. Magn. Reson. 73 (1987) 553–557.

[18] Y. Zeng, J. Tang, C.A. Bush, J.R. Norris, Enhanced spectral resolution in 2D NMR signal analysis using linear prediction extrapolation and apodization, J. Magn. Reson. 83 (1989) 473–483.
[19] E.T. Olejniczak, H.L. Eaton, Extrapolation of time-domain data with linear prediction increases resolution and sensitivity, J. Magn. Reson. 87 (1990) 628–632.
[20] G. Zhu, A. Bax, Improved linear prediction for truncated signals of known phase, J. Magn. Reson. 90 (1990) 405–410.
[21] D.W. Tufts, R. Kumaresan, Singular value decomposition and improved frequency estimation using linear prediction, IEEE T. Acoust. Speech 30 (1982) 671–675.
[22] K. Bromek, D. Lee, R. Hauhart, M. Krych-Goldberg, J.P. Atkinson, P.N. Barlow, K. Pervushin, Polychromatic selective population inversion for TROSY experiments with large proteins, J. Am. Chem. Soc. 127 (2005) 405–411.
[23] S. Sibisi, J. Skilling, R.G. Brereton, E.D. Laue, J. Staunton, Maximum-entropy signal-processing in practical NMR-spectroscopy, Nature 311 (1984) 446–447.
[24] A.S. Stern, K.B. Li, J.C. Hoch, Modern spectrum analysis in multidimensional NMR spectroscopy: comparison of linear-prediction extrapolation and maximum-entropy reconstruction, J. Am. Chem. Soc. 124 (2002) 1982–1993.
[25] H. Takahashi, I. Shimada, Production of isotopically labeled heterologous proteins in non-*E. coli* prokaryotic and eukaryotic cells, J. Biomol. NMR 46 (2010) 3–10.
[26] K. Werner, C. Richter, J. Klein-Seetharaman, H. Schwalbe, Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy, J. Biomol. NMR 40 (2008) 49–53.
[27] A. Strauss, F. Bitsch, B. Cutting, G. Fendrich, P. Graff, J. Liebetanz, M. Zurini, W. Jahnke, Amino-acid-type selective isotope labeling of proteins expressed in Baculovirus-infected insect cells useful for NMR studies, J. Biomol. NMR 26 (2003) 367–372.
[28] S. Liu, L. Meng, K.W. Moremen, J.H. Prestegard, Nuclear magnetic resonance structural characterization of substrates bound to the alpha-2,6-sialyltransferase, ST6Gal-I, Biochemistry 48 (2009) 11211–11219.
[29] L. Feng, R. Orlando, J.H. Prestegard, Mass spectrometry assisted assignment of NMR resonances in ^{15}N labeled proteins, Abstr. Pap. Am. Chem. S 231 (2006).
[30] W.K. Nkari, J.H. Prestegard, NMR resonance assignments of sparsely labeled proteins: amide proton exchange correlations in native and denatured states, J. Am. Chem. Soc. 131 (2009) 5344–5349.
[31] S. Raman, O.F. Lange, P. Rossi, M. Tyka, X. Wang, J. Aramini, G.H. Liu, T.A. Ramelot, A. Eletsky, T. Szyperski, M.A. Kennedy, J. Prestegard, G.T. Montelione, D. Baker, NMR structure determination for larger proteins using backbone-only data, Science 327 (2010) 1014–1018.
[32] Y. Shen, R. Vernon, D. Baker, A. Bax, De novo protein structure generation from incomplete chemical shift assignments, J. Biomol. NMR 43 (2009) 63–78.
[33] N.G. Sgourakis, O.F. Lange, F. Dimaio, I. Andre, N.C. Fitzkee, P. Rossi, G.T. Montelione, A. Bax, D. Baker, Determination of the structures of symmetric protein oligomers from NMR chemical shifts and residual dipolar couplings, J. Am. Chem. Soc. (2011).
[34] H. Chen, F. Ji, V. Olman, C.K. Mobley, Y. Liu, Y. Zhou, J.H. Bushweller, J.H. Prestegard, Y. Xu, Optimal mutation sites for PRE data collection and membrane protein structure prediction, Structure 19 (2011) 484–495.