Journal of Biomolecular NMR, 4 (1994) 483–490
ESCOM

J-Bio NMR 188

Improved resolution in triple-resonance spectra by nonlinear sampling in the constant-time domain

Peter Schmieder[a], Alan S. Stern[b], Gerhard Wagner[a] and Jeffrey C. Hoch[b,*]

[a]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, U.S.A.*
[b]*Rowland Institute for Science, 100 Edwin H. Land Boulevard, Cambridge, MA 02142, U.S.A.*

Received 20 September 1993
Accepted 17 January 1994

Keywords: Nonlinear sampling; Constant-time spectra; Maximum entropy reconstruction; Resolution enhancement; Heteronuclear triple-resonance experiments; HNCO

SUMMARY

Nonlinear sampling along the constant-time dimension is applied to the constant-time HNCO spectrum of the dimerization domain of Gal4. Nonlinear sampling was used for the nitrogen dimension, while the carbon and proton dimensions were sampled linearly. A conventional ct-HNCO spectrum is compared with a nonlinearly sampled spectrum, where the gain in experiment time obtained from nonlinear sampling is used to increase the resolution in the carbonyl dimension. Nonlinearly sampled data are processed by maximum entropy reconstruction. It is shown that the nonlinearly sampled spectrum has a higher resolution, although it was recorded in less time. The constant intensity of the signal in the constant-time dimension allows for a variety of sampling schedules. A schedule of randomly distributed sampling points yields the best results. This general method can be used to significantly increase the quality of heteronuclear constant-time spectra.

INTRODUCTION

The increasing ease with which proteins can be enriched with both ^{13}C and ^{15}N (McIntosh and Dahlquist, 1990) has established 3D heteronuclear double- and triple-resonance experiments (Kay et al., 1990; Montelione and Wagner, 1990; Clore and Gronenborn, 1991) as standard techniques for spectral assignment and the determination of relevant structural parameters. Most of these experiments are very sensitive, and the time to record the spectra is dictated by the desired number of points in the indirectly detected dimensions and the phase cycle, rather than the signal-to-noise requirements. In recent years it has been shown that pulsed field gradients (Bax et al., 1980; Hurd, 1990; Bax and Pochapsky, 1992; Davis et al., 1992; Kay, 1993) can alleviate restrictions imposed by phase cycling.

*To whom correspondence should be addressed.

In many, if not all heteronuclear triple-resonance experiments, one of the evolution times is conveniently executed in a constant-time fashion (Bax et al., 1979; Powers et al., 1991; Clubb et al., 1992; Grzesiek and Bax, 1992, 1993). Relaxation losses are prevented by incorporating the evolution time in a fixed delay necessary for the evolution of coupling in a transfer step. Unlike dimensions in which the signal decays with increasing evolution time, resolution can be increased in the constant-time dimension without loss of sensitivity, by sampling the maximum number of data points allowed by the constant-time period. Since the total experiment time usually is restricted, the number of points in the other indirectly detected dimension of the 3D experiment is necessarily limited. We show here that this problem can be ameliorated by applying nonlinear sampling in the constant- time dimension, an approach that makes it possible to obtain a desired resolution with significantly fewer points than would otherwise be required, just as pulsed field gradients allow one to obtain an adequate signal-to-noise ratio with fewer scans than would be required by phase cycling.

Data collected with a nonlinear sampling schedule cannot be processed by a discrete Fourier transform (DFT). By contrast, the Maximum Entropy Method (MaxEnt) (Sibisi, 1983; Laue et al., 1985, 1986; Stephenson, 1988; Hoch, 1989; Jones and Hore, 1991) is particularly well suited for truncated or incomplete (such as nonlinearly sampled) data. For truncated data, e.g., data from a constant-time experiment, it provides a valuable alternative to extrapolation by linear prediction followed by apodization and DFT. The method has not yet enjoyed widespread use, however, mainly due to its increased computational cost compared to DFT. Recent dramatic developments in computer performance now permit routine applications to 2D and 3D data (Hodgkinson et al., 1993; Schmieder et al., 1993).

MATERIALS AND METHODS

HNCO spectra (Kay et al., 1990) were recorded with linear sampling as well as nonlinear sampling on a Bruker AMX 600 spectrometer in standard configuration, using a triple-resonance probe equipped with a self-shielded gradient coil. The sample consisted of a 0.8-mM solution of the dimerization domain of Gal4, corresponding to positions 50–106 in the sequence. The protein exists in solution as a homodimer (15 kDa). The measurements were done at 308 K. Both spectra were recorded with a constant-time evolution period for the nitrogen dimension, with the pulse scheme proposed by Grzesiek and Bax (1992). Three pulsed field gradients were added to the pulse sequence, which is shown in Fig. 1. One gradient was applied between the proton and the carbon 90° pulse of the first INEPT step and two gradients were applied during the final refocussing period on both sides of a 3-9-19 180° pulse to achieve WATERGATE water suppression (Sklenar et al., 1993). The TPPI-States method (Marion et al., 1989) was used in both indirect dimensions. Eight scans per FID were collected for both experiments. The spectral widths were 10 000 Hz in the acquisition dimension and 1587 and 1515 Hz in the nitrogen and carbonyl carbon dimensions, respectively. All delays were according to Grzesiek and Bax (1992, see also Fig. 1); 27 ms was used as the constant-time delay, allowing for a maximum of 42 complex points in the nitrogen dimension.

The regular ct-HNCO was recorded with 42 complex points in the nitrogen dimension, 48 complex points in the carbonyl dimension and 1024 complex points in the acquisition dimension, resulting in a total experiment time of 26 h. The spectrum was processed using the FELIX

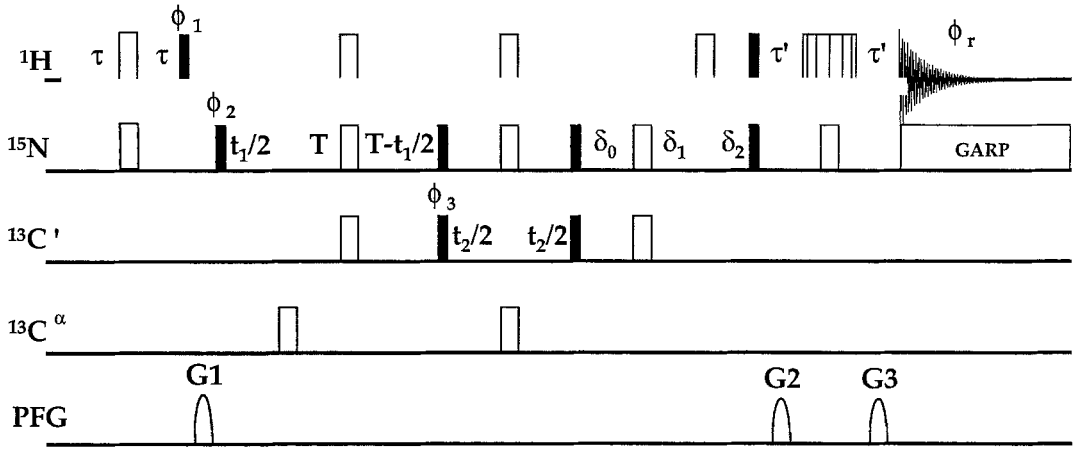

Fig. 1. Pulse sequence for the modified version of a ct-HNCO experiment. The sequence is the same as that given by Grzesiek and Bax (1992), with the exception that three pulsed field gradients were added for water suppression. The gradient G1 had a duration of 5 ms and a strength of 25 G/cm. The other two gradients, together with the 3-9-19 180° pulse in the center of the refocussing period, form a WATERGATE sequence (Sklenar et al., 1993). Both gradients had the same sign and strength (1 ms, 10 G/cm). All gradients were sine-shaped. Solid bars and open bars indicate 90° and 180° pulses, respectively. The delays were set as follows: τ = 2.25 ms, τ' = 1.9 ms, T = 13.5 ms, δ_0 = 13.5 ms, δ_1 = 10.75 ms, δ_2 = 2.75 ms. The phases are ϕ_1 = y,(−y); ϕ_2 = 2(x),2(−x); ϕ_3 = 4(x),4(−x); ϕ_r = (x,−x),2(−x,x),(x,−x); all other phases are x.

software (Biosym Inc.). Following zero-filling, windowing with a squared sine bell shifted by 90°, and Fourier transformation in the acquisition dimension, only the amide region of the spectrum (512 points) was retained. In the carbonyl dimension, a squared sine bell shifted by 90° was applied and the data were zero-filled to 256 points prior to Fourier transformation. The nitrogen dimension was extended to 64 complex points by linear prediction (using 10 poles and 10 peaks with forward linear prediction in FELIX), and additional zero-filling yielded a final data matrix of 512 × 256 × 128 real points.

The nonlinearly sampled ct-HNCO was recorded with 16 points in the nitrogen dimension, selected out of 42 by a random number generator. The sampled points were 1, 4, 7, 11, 13, 16, 21, 22, 25, 27, 28, 33, 34, 35, 41 and 42. In the carbonyl and acquisition dimensions, 90 and 1024 complex points were linearly sampled, respectively, resulting in a total experiment time of 19 h. Note that not all the time gained from the nonlinear sampling scheme was used to increase the resolution in the second indirect dimension, because the evolution time in the carbonyl dimension should be kept shorter than 65 ms to avoid signal loss, caused by heteronuclear long-range coupling between carbonyl carbons and protons (Grzesiek and Bax, 1992). Processing in the first two dimensions was performed as with the linearly sampled data, except that 90 points were used in the carbonyl dimension. The data were then transferred to the Rowland NMR Toolkit for MaxEnt reconstruction and back to FELIX afterwards for analysis.

The MaxEnt algorithm we employed is a modified implementation of the 'Cambridge' algorithm of Skilling and Bryan (1984), and was discussed in Schmieder et al. (1993). The value of χ^2 was estimated from the noise level in a blank region of the spectrum, following the Fourier transform processing steps and prior to MaxEnt reconstruction. The value of def, the scaling parameter, was chosen empirically by adjusting the value while iteratively reconstructing a single representative

row until an acceptable result was obtained. (We know of no algorithmic procedure for choosing the value of def; fortunately the reconstruction is relatively insensitive to this value.) Convergence was determined by monitoring the value

$$\text{Test} = \left| \frac{\nabla S}{|\nabla S|} - \frac{\nabla C}{|\nabla C|} \right|$$

which is a measure of the extent to which the gradients of the entropy (S) and the constraint (C) fail to be parallel. MaxEnt reconstruction was performed on each row separately, iterating until the value of Test was less than 10^{-4} or for a maximum of 50 iterations. Using a value of 5000 for χ^2 and 10 for def, the average value of Test was 1.21×10^{-4}, and the largest value was 5.16×10^{-2}. The reconstruction took 3 min on a Silicon Graphics 4D/480 (eight R3000 CPUs, 256 Mb of memory) or 12 min on a Silicon Graphics Indigo XZ4000 (one R4000 CPU, 128 Mb of memory).

RESULTS AND DISCUSSION

To make best use of nonlinearly sampled data, the design of a well-suited sampling schedule is crucial. A procedure for the design of the sampling schedule has been presented (Schmieder et al., 1993), based on earlier work (Barna et al., 1987; Robin et al., 1991). The method takes into account the decay of intensity caused by relaxation and is thus not applicable to constant-time evolution. Since there is no decay in this setting, the method would generate a schedule consisting of uniformly spaced delays, effectively equivalent to linear sampling with a reduced spectral width, and would lead to poor results. Instead, the schedule should be designed for optimal MaxEnt reconstruction. Most importantly, the first and the last point should be collected; all other points can be chosen more-or-less freely. For the selection of the other points we tested a variety of sampling schedules.

A nonlinearly sampled data set can be described as the result of multiplying a linearly sampled data set by a sampling function, which is equal to one for the sampled points and zero for other points. In the linear world of the discrete Fourier transform, this operation would result in the convolution of the spectrum with the transform of the sampling function. This is not true for MaxEnt reconstruction, but nevertheless, one can gain insight into the artifacts caused by a particular sampling schedule, by examining the reconstruction of its sampling function. (This reconstruction is in fact the spectrum that one would obtain if the signal consisted of a single peak at zero frequency.) Figure 2 illustrates the relation between the discrete Fourier transform of a random sample schedule and the MaxEnt reconstruction. Note that the additional features in Fig. 2b are not noise, but artifacts resulting from the sampling schedule. The intensity of these artifacts is substantially attenuated in the MaxEnt reconstruction (Fig. 2c), but close examination reveals that the structure of the artifacts is preserved. The smallest artifacts can, of course, be expected from a linear schedule that samples all the points, which would result in one single peak in the center of a noise-free spectrum. The best nonlinear sampling schedule will be the one that yields a reconstructed spectrum as close as possible to this ideal one.

The simplistic schedule mentioned above (consisting of equally spaced sample points) would result in a reduced spectral width and aliasing. Another 'naive' schedule would be one where an equal number of points is sampled at the beginning and at the end. This schedule yields poor results as well, showing features similar to the sinc wiggles that result from truncation of the FID

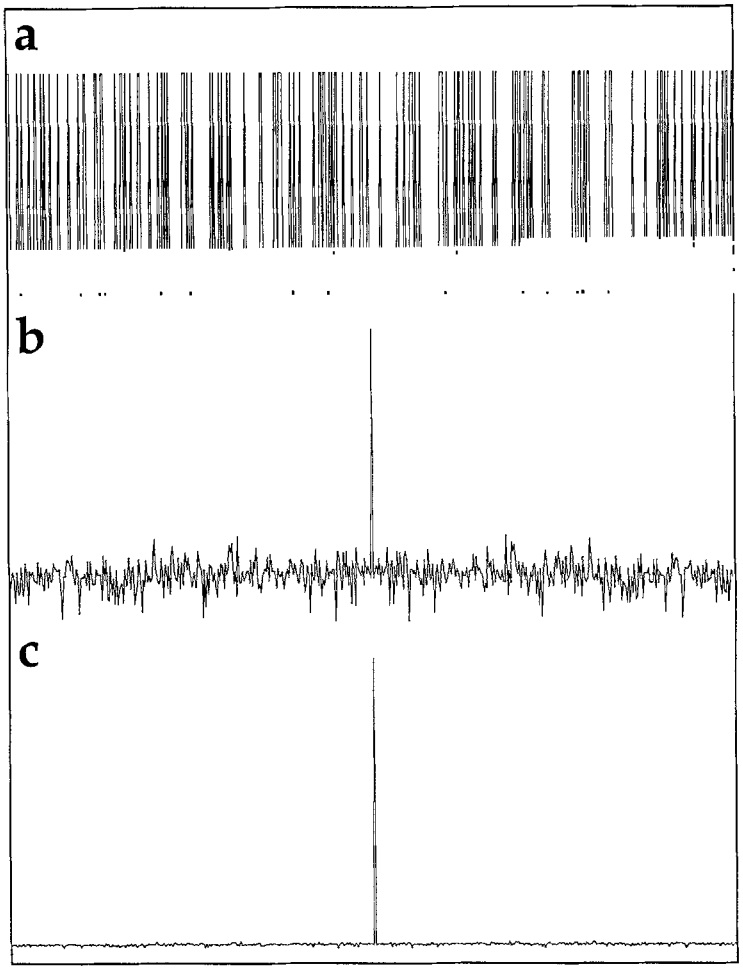

Fig. 2. Comparison of the DFT (b) and the MaxEnt reconstruction (c) of a sampling schedule of 128 points, selected at random out of 512 points shown in (a). The processing of the sampling function alone gives an idea of the artifacts that can be expected when using the corresponding sampling schedule. Although the structure of the artifacts is preserved, their intensity is substantially lower in the MaxEnt reconstruction.

in a DFT. Two more promising possibilities are a schedule designed according to the procedure of Barna et al. (1987) and a schedule with points selected by a random number generator. Barna et al. (1987) described an exponential schedule, using an initial sampling density equal to one. Figure 3 shows two comparisons of the artifacts resulting from these two types of schedules, one for a selection of 128 out of 512 points (a,b) and one for a selection of 16 out of 42 points (c,d). The random schedule is advantageous for a higher number of points, because the structure due to the sampling function is spread more uniformly and is less likely to interfere with the interpretation of the spectrum. The schedules become more similar for a smaller number of points, mainly because there is less distinction between 'random' and other samples for small sample sizes.

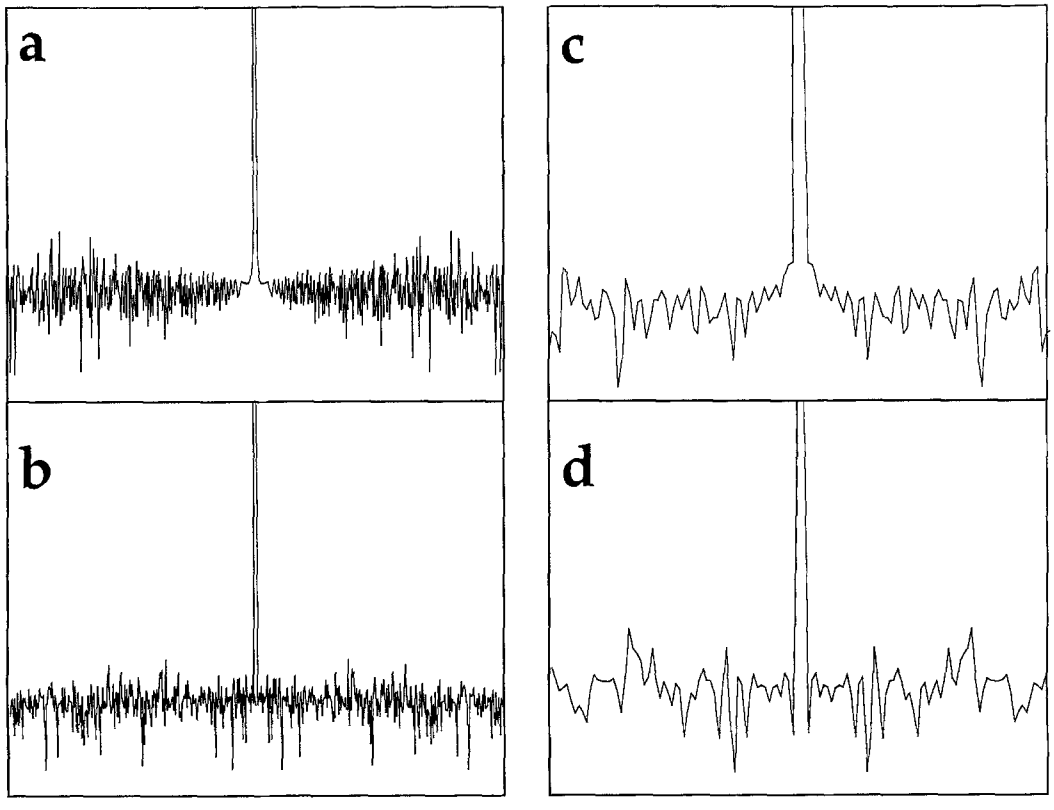

Fig. 3. Comparison of the artifacts resulting from the sampling of data points in a nonlinear fashion. (a,b) Sampling of 128 out of 512 points according to Barna et al. (1987) (a) and by selecting points randomly (b). (c,d) Sampling of 16 out of 42 points according to Barna et al. (c) and by selecting points randomly (d). All spectra have been scaled in the same way; the peak in the center of the spectrum is a factor 1000 larger than the most intense feature and has been cut off (note that the reconstruction in panel b corresponds to the one in Fig. 2c). It can be seen that the random schedule is superior in the case of a high number of points, because it shows less structure. It becomes similar to the schedule according to Barna et al. for a smaller number of points. All reconstructions have been performed with the same MaxEnt parameters.

For our experiment we decided to use random sampling for the selection of 16 out of 42 complex points. Figure 4 shows a comparison between nonlinear sampling and linear sampling for the HNCO spectrum of the dimerization domain of Gal4. Two F1/F2 slices taken at 7.98 ppm in the proton dimension of the spectrum are compared. The peaks are slightly narrower in the ^{15}N dimension (F1) of the MaxEnt spectrum (Fig. 4a), even though the linearly sampled data (Fig. 4b) have been extended with linear prediction from 42 to 64 points. In the ^{13}C dimension (F2), the nonlinearly sampled spectrum (Fig. 4a) shows much greater resolution. More points could be collected in this dimension, although the overall experiment time was shorter, since the number of points sampled in the ^{15}N dimension was significantly reduced. Figure 3 also shows 1D cross sections of the spectra, which demonstrate the difference in signal-to-noise ratio. (However, it should be noted that one of the properties of the MaxEnt reconstruction is that gains in apparent signal-to-noise ratio do not necessarily correspond to increased sensitivity.)

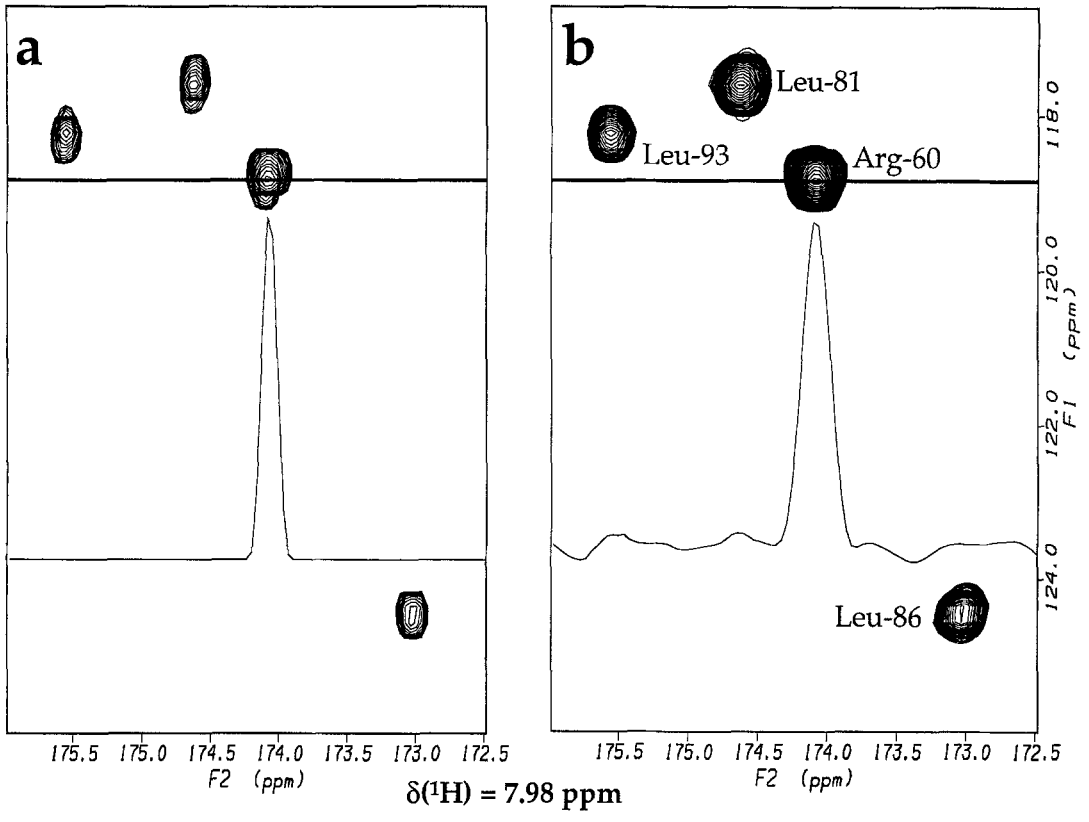

Fig. 4. Comparison between two ^{13}C-^{15}N slices of the ct-HNCO spectrum at a proton chemical shift of 7.98 ppm. (a) Slice from the spectrum obtained by MaxEnt reconstruction of the nonlinearly sampled data set; (b) slice from the spectrum obtained by DFT of the linearly sampled data set, after linear prediction and apodization with a sine bell shifted by 90°. The cross sections along F2 through the peak of Arg^{60} demonstrate the superior signal-to-noise ratio of the MaxEnt spectrum. The resolution in the spectrum resulting from the MaxEnt reconstruction is slightly increased in the ^{15}N dimension and significantly increased in the ^{13}C dimension, although it was recorded in a shorter time.

During the reconstruction, it turned out that the inherent high resolution of the 3D spectrum aids the speed of the MaxEnt reconstruction. Since the program operates independently on each row and there are many rows that are essentially free of signal, the reconstruction proceeds rather quickly. The gain in resolution and the reduction in experiment time certainly outweigh a processing time of less than 15 min on a typical workstation. In addition, the fact that all the signals in the HNCO experiment are of the same or comparable intensity prevents the problems encountered during MaxEnt reconstructions of spectra with large dynamic range (Schmieder et al., 1993).

It is possible to use nonlinear sampling in more than one indirect dimension at a time (Robin et al., 1991). However, it is then necessary to perform MaxEnt reconstruction in all those dimensions at once, rather than applying it to each dimension sequentially. While there are advantages to this approach, disadvantages include greater memory requirements and the fact that the sparse nature of most multidimensional spectra can no longer be easily exploited to reduce the computa-

tional cost. The simple procedure described here is more amenable to execution on typical laboratory computers.

CONCLUSIONS

In conclusion, we have shown that the application of nonlinear sampling and reconstruction using MaxEnt forms an easy and straightforward procedure for collecting and processing heteronuclear triple-resonance spectra with constant-time evolution periods. Since nearly all of the commonly used techniques incorporate a constant-time evolution period, the method is of general applicability and can furthermore be extended to 4D spectra. It should help to further increase the resolution attainable in heteronuclear triple-resonance spectra and thus facilitate the automation of the assignment process.

ACKNOWLEDGEMENTS

We thank Ronen Mamorstein and Stephen C. Harrison for providing the sample of the Gal4 dimerization domain. P.S. thanks the Deutsche Forschungsgemeinschaft for a fellowship. This work was supported by NIH Grant GM 47467, The Rowland Institute for Science and the W.M. Keck Foundation.

REFERENCES

Barna, J.C.J., Laue, E.D., Mayger, M.R., Skilling, J. and Worrall, S.J.P. (1987) *J. Magn. Reson.*, **73**, 69–77.
Bax, A., Mehlkopf, A.F. and Smidt, J. (1979) *J. Magn. Reson.*, **35**, 167–169.
Bax, A., De Jong, P.G., Mehlkopf, A.F. and Smidt, J. (1980) *J. Magn. Reson.*, **69**, 567–570.
Bax, A. and Pochapsky, S.S. (1992) *J. Magn. Reson.*, **99**, 638–643.
Clore, G.M. and Gronenborn, A.M. (1991) *Prog. NMR Spectrosc.*, **23**, 43–92.
Clubb, R.T., Thanabal, V. and Wagner, G. (1992) *J. Magn. Reson.*, **97**, 213–217.
Davis, A.L., Boelens, R. and Kaptein, R. (1992) *J. Biomol. NMR*, **2**, 395–400.
Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **96**, 432–440.
Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 185–204.
Hoch, J.C. (1989) *Methods Enzymol.*, **176**, 216–241.
Hodgkinson, P., Mott, H.R., Driscoll, P.C., Jones, J.A. and Hore, P.J. (1993) *J. Magn. Reson. Ser. B*, **101**, 218–222.
Hurd, R.E. (1990) *J. Magn. Reson.*, **87**, 422–428.
Jones, J.A. and Hore, P.J. (1991) *J. Magn. Reson.*, **92**, 276–292.
Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.
Kay, L.E. (1993) *J. Am. Chem. Soc.*, **115**, 2055–2057.
Laue, E.D., Skilling, J., Staunton, J., Sibisi, S. and Bereton, R.G. (1985) *J. Magn. Reson.*, **62**, 437–452.
Laue, E.D., Mayger, M.R., Skilling, J. and Staunton, J. (1986) *J. Magn. Reson.*, **68**, 14–29.
Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.
McIntosh, L.P. and Dahlquist, F.W. (1990) *Q. Rev. Biophys.*, **23**, 1–38.
Montelione, G.T. and Wagner, G. (1990) *J. Magn. Reson.*, **87**, 183–188.
Powers, R., Gronenborn, A.M., Clore, G.M. and Bax, A. (1991) *J. Magn. Reson.*, **94**, 209–213.
Robin, M., Delsuc, M.-A., Guittet, E. and Lallemand, J.-Y. (1991) *J. Magn. Reson.*, **92**, 645–650.
Schmieder, P., Stern, A.S., Wagner, G. and Hoch, J.C. (1993) *J. Biomol. NMR*, **3**, 569–576.
Sibisi, S. (1983) *Nature*, **301**, 134–136.
Skilling, J. and Bryan, R.K. (1984) *Mon. Not. R. Astron. Soc.*, **211**, 111–124.
Sklenar, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J Magn. Reson. Ser. A*, **102**, 241–245.
Stephenson, D.S. (1988) *Prog. NMR Spectrosc.*, **14**, 515–626.