115 B

JOURNAL OF MAGNETIC RESONANCE 84, 425–430 (1989)

Improved Solvent Suppression in One- and Two-Dimensional NMR Spectra by Convolution of Time-Domain Data

DOMINIQUE MARION,* MITSUHIKO IKURA, AND AD BAX

Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

Received April 25, 1989

Numerous methods are available for suppression of the H_2O resonance in the ^1H NMR spectra of water-soluble compounds (*1*). On average, these methods present a trade-off between the excitation profile and the level of suppression of the water resonance: with few exceptions, methods that yield the best solvent suppression per single scan tend to have a worse excitation profile than the simple "jump-and-return" or "1–1" sequence (*2*) which has a relatively good, sine-bell-shaped excitation profile. For H_2O suppression in 2D experiments, a second and even more important criterion is the reproducibility of H_2O suppression from one scan to the next. Poor reproducibility leads to t_1 noise in the vicinity of the H_2O F_2 frequency and the tails of this unsuppressed H_2O resonance can extend into the spectral region of interest. Here we demonstrate that treatment of the time-domain data is an effective method for removing the residual H_2O signal. This process does not require any operator interaction and therefore it can be easily included in automated data processing.

Our approach utilizes a convolution difference of the time-domain data to remove the undesired solvent signal component from the free induction decay. The process is most easily applied for the case where the solvent signal is on-resonance, although, as discussed later, it can be applied independent of the position of the solvent resonance provided that complex data are acquired. Figure 1A shows the FID obtained with a 1–1 sequence for a 1.5 m*M* solution of the protein calmodulin (16.7 kDa) in 90% H_2O. The residual H_2O signal is responsible for the low-frequency component of the signal. To a good approximation, this low-frequency component of the FID can be calculated by averaging neighboring time-domain data points, which is equivalent to convolution with a rectangular function. The width of the rectangle corresponds to the number of time-domain data points that are averaged. This low-frequency component is then subtracted from the original signal (Fig. 1B). Better removal of the effect of high-frequency components to the calculated low-frequency signal can be obtained by convolution with an appropriately shaped (Gaussian or sine bell) function, i.e., by calculating a weighted average for each time-domain data point. Thus, the low-frequency component, $L(n)$, of the FID, $S(n)$, is calculated according to

* On leave from Centre de Biophysique Moleculaire, Centre National de la Recherche Scientifique, 45071 Orleans Cedex 2, France.

0022-2364/89 $3.00

$$L(n) = [\sum_{k=-K}^{k=K} f(k)S(n+k)]/A \quad [1]$$

with

$$A = \sum_{k=-K}^{k=K} f(k),$$

where n is the number of the sampled data point, and $2K + 1$ is the width of the convolution window. The function, $f(k)$, defines the shape of the convolution window. In practice, we prefer using either a sine-bell-shaped $\{f(k) = \cos[k\pi/(2K + 2)]\}$ or a Gaussian-shaped $[f(k) = \exp(-4k^2/K^2)]$ window function. The low-frequency component, $L(n)$, is subtracted from the real signal, $S(n)$, prior to Fourier transformation. In practice, we typically use window functions with a width of 17–65 (K = 8–32) data points. The real and imaginary channels of the complex time-domain data are treated separately and independently.

The convolution procedure described above should not be used for the first K and the last K data points of the FID. Regular convolution of finite signals assumes periodicity of the signal (*3*) and in our case it would average the first with the last data points of the FID, a clearly undesirable feature. In this case, the Fourier transform of $L(n)$ would be equivalent to multiplication of the regular frequency-domain spectrum by a masking function that corresponds to the Fourier transform of $f(k)$, and apart from distorted resonance intensities, the spectrum would remain unchanged. In contrast, we calculate the first data points of $L(k)$ by linear extrapolation of its data points $L(K)$ and $L(K + M)$: $L(K - k) = L(K) + k[L(K) - L(K + M)]/M$. Similarly, the last K data points are calculated by extrapolation of data points $N - K - M$ and $N - K$, where N is the number of sampled data points.

Considering the well-known relation between convolution in the time domain and multiplication in the frequency domain, one may wonder whether our procedure could be performed in the frequency domain by multiplication with an appropriate window function that has a null at the solvent frequency. This is not the case. As pointed out above, the difference originates from the fact that our treatment is not a true cyclic convolution, which assumes periodicity of the time-domain signal (*3*), but instead, it uses linear extrapolation to calculate the data points near the ends of the time domain. An operation equivalent to our time-domain data treatment is not easily performed in the frequency domain.

The procedure is illustrated in Fig. 1. The regular free induction decay (Fig. 1A) contains a large low-frequency component, originating from the H_2O signal. This component is removed in Fig. 1B, after application of the modified deconvolution method described above, using the Gaussian function with $K = 8$ and extrapolation with $M = 16$. Figures 1C and 1D compare the spectra corresponding to the untreated and treated FID. As can be seen most clearly on the insets in Figs. 1C and 1D, the broad dispersive tail of the residual H_2O (Fig. 1C) has been completely removed in the spectrum of Fig. 1D, leaving a clean and straight baseline which does not need further frequency-domain manipulation, provided that the early data points of the FID are scaled correctly (*4*) or that linear prediction has been used to calculate them (*5*). The small "distortion" in the vicinity of the H_2O resonance in Fig. 1D is largely

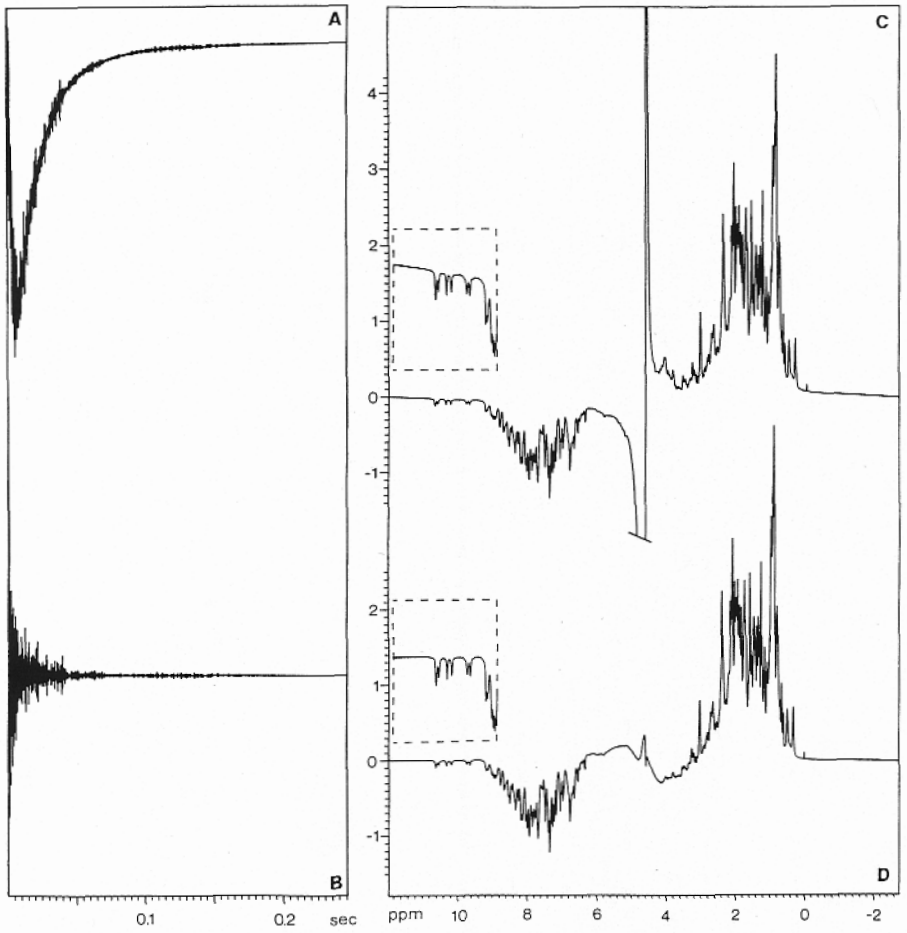

FIG. 1. (A) The real part of the FID of a 1.5 mM sample of calmodulin in 90% H_2O, recorded with a jump-and-return sequence. (B) The corrected FID, using a Gaussian window with $K = 8$ and extrapolation with $M = 16$. (C) and (D) The Fourier transforms of (A) and (B), respectively. The insets in (C) and (D) illustrate that the baseline has been flattened by the time-domain correction procedure.

a real feature, caused by the inhomogeneous broadening of the H_2O resonance and the finite attenuation adjacent to the carrier frequency. When a Gaussian function is used as described above, the attenuation window in the frequency domain also has a Gaussian profile, with 50% attenuation at $\pm 0.416 SW/K$ from the carrier frequency, where SW is the total spectral width.

The main application of our convolution method is found in the treatment of 2D and 3D data that have been recorded in H_2O solution. Frequency-domain baseline corrections are not always straightforward in these types of experiments since it may be difficult to define signal-free regions, necessary for defining the baseline. It also should be noted that the distortion introduced by the dispersive tails of H_2O causes a curved baseline that cannot be removed completely by linear baseline correction.

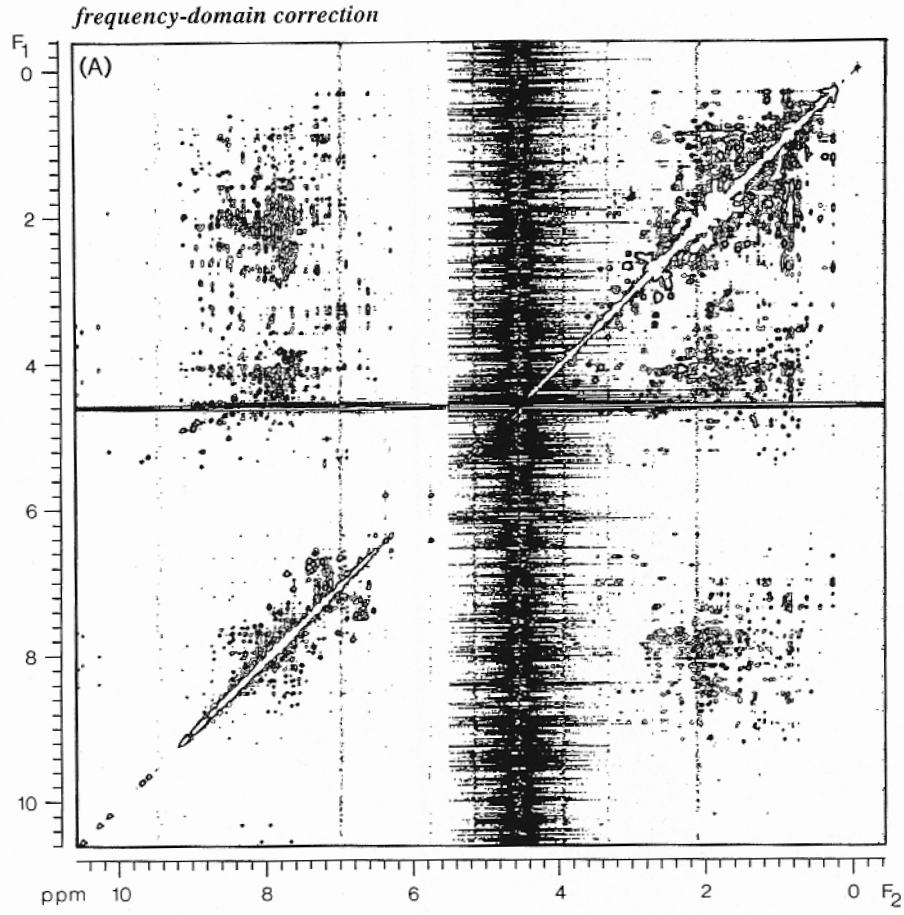

FIG. 2. The 500 MHz 2D NOESY spectra of 1.5 mM calmodulin in 90% H_2O, 47°C, 130 ms mixing, 64 scans per t_1 value; acquisition times were 140 (t_2) and 66 ms (t_1). Digital resolution in the final spectrum is 7 (F_2) and 6 Hz (F_1). The total measuring time was 17 h. Data have been acquired in the hypercomplex format. Data have been recorded with the NOESY pulse scheme, with the last pulse replaced by a jump-and-return pulse (*6*). Both spectra have been obtained from the same set of acquired data, using identical digital filtering (60° shifted sine square bell) and multiplication of the first t_2 data point by an empirically determined factor of 4.4 (*4*). (A) Spectrum obtained with regular processing and application of a linear baseline correction in the F_2 dimension for the region 10.5–5.6 ppm. (B) Spectrum obtained with time-domain deconvolution without frequency-domain baseline correction. Data were first Fourier transformed with respect to t_1. For t_2 traces less than 0.1 ppm removed from the F_1 carrier frequency, a Gaussian window with $K = 8$ was used, and for all other t_2 traces $K = 32$. For extrapolation, $M = 16$ for all traces.

Our approach avoids these baseline distortions altogether and does not require any operator interaction for defining baseline positions. As an example, Fig. 2 compares the NOESY spectra of calmodulin obtained with a NOESY experiment that did not utilize any presaturation of the H_2O but employed a 1–1 read pulse at the end of the mixing period (*6*). Frequency-domain baseline correction has been applied to the amide region of the regular spectrum (Fig. 2A); time-domain deconvolution but no frequency-domain manipulation has been used for the processing of spectrum B.

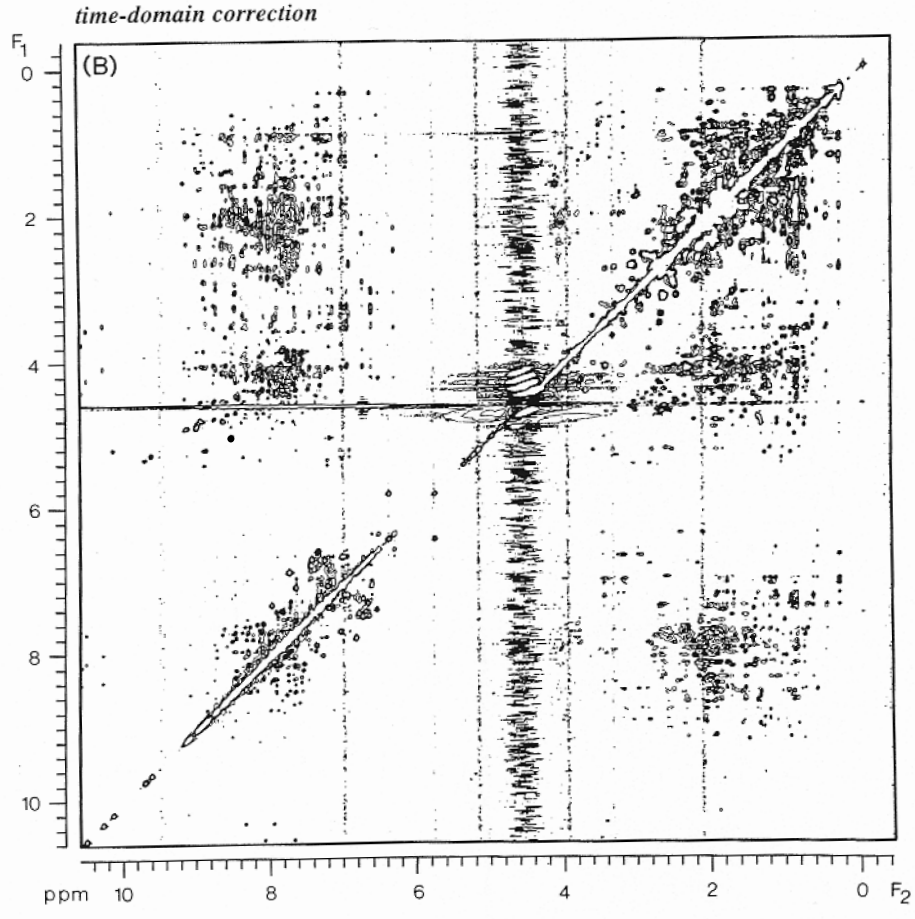

FIG. 2—*Continued*

Of course, the method described here does not make H_2O suppression pulse schemes obsolete. To avoid overload in the RF receiver circuitry or in the ADC converter, suppression of the H_2O signal is needed before it reaches the receiver system. However, the dynamic range of modern NMR spectrometers is such that a 30-fold suppression of the H_2O resonance is often sufficient for obtaining ^1H NMR spectra with near optimal sensitivity. Such a modest level of suppression is easily obtained with very simple suppression schemes such as the 1–1 sequence. Further suppression is then achieved by the convolution method described above.

In the discussion above, we have assumed that the solvent signal is on-resonance. If the solvent signal is at the edge of the spectrum (at the Nyquist frequency), the same convolution procedure can be applied provided that the sign of even-numbered data points is changed prior to the convolution difference operation and changed back once completed. For complex data (with simultaneous acquisition of the real and imaginary channel) the solvent resonance frequency can be changed to any position, including the desired on-resonance position, by applying a linearly time-depen-

dent phase correction to the time-domain data (*7, 8*). After convolution difference, the frequencies can be shifted back to their original values by applying the opposite phase correction.

The present method has some similarity with the data-shift-accumulation (DSA) technique (*9*) which also treats time-domain data, with the main purpose of avoiding problems associated with limited computer word length. This DSA method also results in excellent suppression of the H_2O resonance, including its dispersive tails, and it gives a cosinusoidal modulation of the resonance intensities and of the noise across the frequency-domain spectrum (*9*). Although correct intensities can be restored easily in the frequency domain, the DSA technique introduces a linearly frequency-dependent phase error which results in baseline problems in the 2D spectrum unless a special linear prediction algorithm is used (*5*).

In our laboratory, the convolution method described above has proven to be extremely valuable for the processing of 2D and 3D spectra, recorded in H_2O solution. Especially if only few scans are recorded per t_1 value (or per t_1/t_2 value for 3D) or when the water suppression level is not very reproducible for successive increments, the method results in a substantial improvement in spectral quality. The choice of the window function used is somewhat arbitrary: the Gaussian function has the advantage that the attenuation window in the frequency domain also has a Gaussian shape; the sine-bell-shaped window permits the use of smaller K values for the same half-width of the attenuation window, resulting in less error during the linear extrapolation of the first K data points. We have not found significant differences in spectral quality for data sets treated with a Gaussian or with a sine-bell window.

ACKNOWLEDGMENTS

We thank Dr. Claude Klee and Marie Krinks for assistance in preparing the protein sample used in this study and Dr. Ted Becker for useful suggestions during the preparation of this manuscript. This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director, NIH.

REFERENCES

1. P. J. HORE, *in* "Methods in Enzymology" (N. Oppenheimer and T. L. James, Eds.), Vol. 176, Chap. 3, Academic Press, San Diego, California, 1989.
2. P. PLATEAU AND M. GUERON, *J. Am. Chem. Soc.* **104,** 7310 (1982).
3. R. N. BRACEWELL, "The Fourier Transform and Its Application," 2nd ed., p. 362, McGraw–Hill, New York, 1978.
4. G. OTTING, H. WIDMER, G. WAGNER, and K. WÜTHRICH, *J. Magn. Reson.* **66,** 187 (1986).
5. D. MARION AND A. BAX, *J. Magn. Reson.* **83,** 205 (1989).
6. G. M. CLORE, B. J. KIMBER, AND A. M. GRONENBORN, *J. Magn. Reson.* **54,** 170 (1983).
7. A. BAX, R. H. GRIFFEY, AND B. L. HAWKINS, *J. Magn. Reson.* **55,** 301 (1983).
8. A. A. BOTHNER-BY AND J. DADOK, *J. Magn. Reson.* **72,** 540 (1987).
9. K. ROTH, B. J. KIMBER, AND J. FEENEY, *J. Magn. Reson.* **41,** 302 (1980).