

Removal of Dispersive Baseline Distortions Caused by Strong Water Signals

MARC ADLER* AND GERHARD WAGNER*

Biophysics Research Division, The University of Michigan, Ann Arbor, Michigan 48109

Received August 31, 1990

A baseline correction routine that eliminates dispersive tails extending from the water t_1 noise band in 2D NMR spectra has been developed. Streaks along the ω_2 axis (Fig. 1a) spreading from the ω_2 position of the water resonance are among the most perturbing artifacts in 2D NMR spectra where the water signal is reduced by presaturation. In our experience, these streaks have dispersive lineshapes and cannot be eliminated with conventional baseplane correction routines, such as polynomial fits. They also add significantly to the apparent noise in the entire 2D spectra. To properly correct for these distortions, an ω_2^{-1} term has been added to the polynomial baseline correction routine of our 2D processing software (FTNMR, Hare Research, Inc.). This term estimates the amount of the dispersive component. The actual correction baseplane is constructed as a sum of a constant offset, a linear term, and a dispersive component where the weight of the latter is calculated from the contribution of the ω_2^{-1} term. With this procedure we have achieved major improvements of the quality of the spectra and of the apparent signal-to-noise in 2D protein spectra in water recorded with presaturation.

Presaturation of the water signal was the first method to record 2D NMR spectra in H₂O (1, 2), and it is still the most common technique for recording 2D NMR spectra of proteins in H₂O. During the recycle delay between accumulations, a long selective pulse is applied to the water resonance. Usually, the transmitter and the decoupler are set to the water frequency in order to minimize distortions in 2D spectra. Phase coherence of transmitter and decoupler is important for good water suppression. Nevertheless, this technique leaves residual artifacts. An example is shown in Figs. 1 and 2. In the 1D spectra of a 5 mM solution of the protein kistrin (3) collected with this technique, the water peak was roughly 50 times higher than the largest methyl signals. A NOESY spectrum acquired under these conditions has large ridges parallel to ω_2 (Fig. 1a). As shown in Fig. 2, these ridges are due to large dispersive artifactual signals centered at the water frequency.

A conventional polynomial baseline correction cannot accurately reproduce a dispersive lineshape. Also, this technique is very sensitive to the choice of reference points for the polynomial fit. Since the reference points for the polynomial fit can only be used from regions of the spectrum where there are no resonances, this procedure is not very practical for crowded protein spectra. An alternative procedure has been

* Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115.

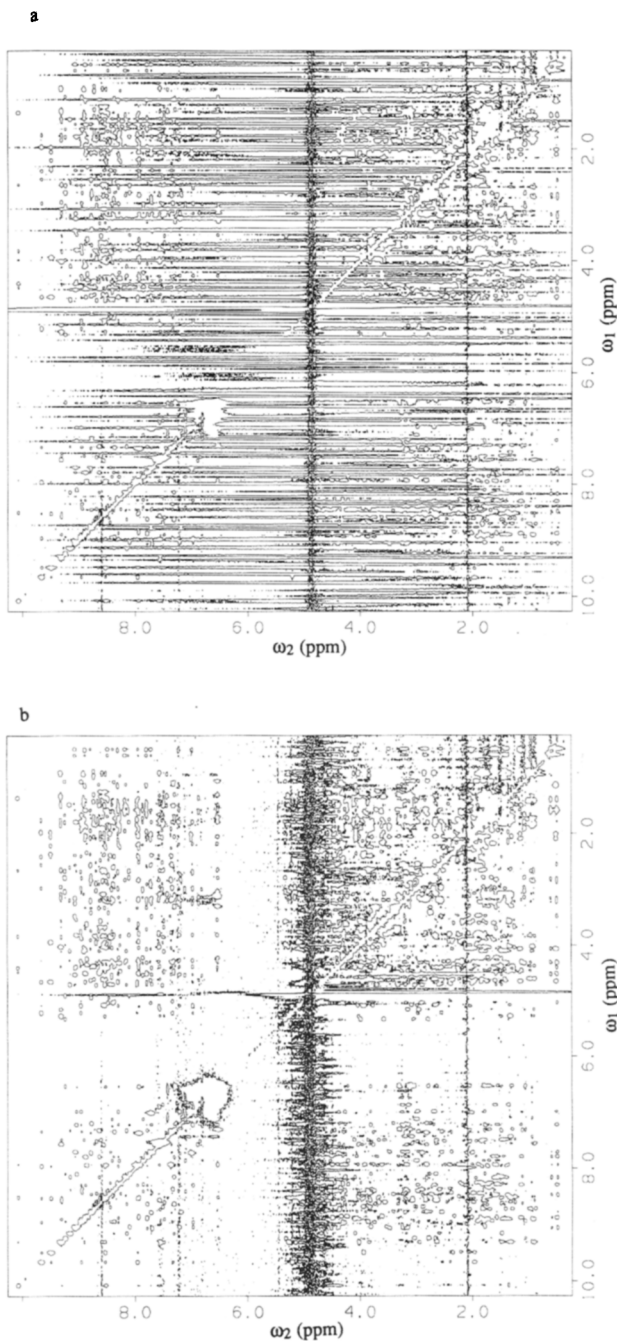


FIG. 1. The 200 ms NOESY spectrum of a 5 mM solution of kistrin in 95% H₂O/5% D₂O at 20°C (a) before and (b) after correction with a dispersive fit. The two spectra share the same scaling and contour level. Water suppression was carried out by selectively irradiating the water resonance during the recycle delay and the mixing period. The spectrum was acquired on a GN-500 NMR spectrometer with 4K complex points in t_2 and 512 real points along t_1 . Time proportional phase incrementation (5) was used to discriminate the sign of the frequencies along t_1 . The data were zero filled to 1K real points along t_1 . A total of 336 scans were collected per FID. Total recycle delay between scans was 0.8 s. A 40° shifted sine-bell window function was used in both dimensions. The method for correcting distortions in the baseline is explained in the text.

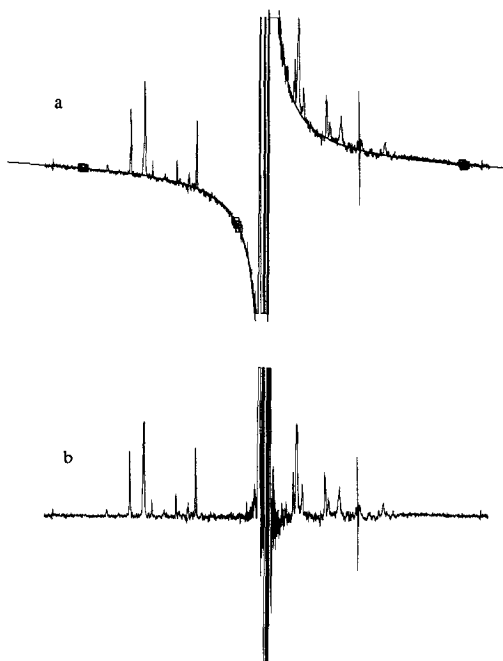


FIG. 2. Effect of the dispersive fit baseline correction shown on a single row (cross section along ω_2) from the spectrum of Fig. 1. At the top is the row after Fourier transform in both dimensions prior to the baseline correction. The boxes are drawn for the reference points selected in the dispersive fit. The solid line drawn through the spectrum is from the function derived from the fitting function. At the bottom is the same row after baseline correction.

proposed to process the original FIDs with some sort of smoothing function to isolate the signal of the water resonance at the carrier frequency and to subtract the smoothed FIDs from the original data (4). This technique, however, also removes any protein peaks which are close to the water. Furthermore, in our experience this technique is very efficient at removing low-frequency components of the water signal, but it tends not to completely remove strong baseline distortions in regions greater than 3 ppm away from the carrier.

This paper presents an approach that uses a direct fit to the dispersive lineshape centered at the carrier. The imaginary component of a Lorentzian line provides an excellent approximation for the baseline artifact;

$$I(\omega) = a_{-1} \frac{\omega_0 - \omega}{(1/T_2^*)^2 + (\omega_0 - \omega)^2}, \quad [1]$$

where a_{-1} is a scaling factor and $1/T_2^*$ is the effective linewidth of the water peak. This equation cannot be linearized. Fortunately, for most of the spectrum $\omega_0 - \omega \gg 1/T_2^*$. Applying this approximation we obtain

$$I(\omega) = a_{-1}(1/(\omega_0 - \omega)), \quad [2]$$

where ω_0 is the frequency of the decoupler. Thus, the weight of the dispersive component is estimated from a linear least-squares fit of a hyperbolic function to baseline points that are sufficiently distant from $\omega = \omega_0$. Once the value of a_{-1} in Eq. [2] has been

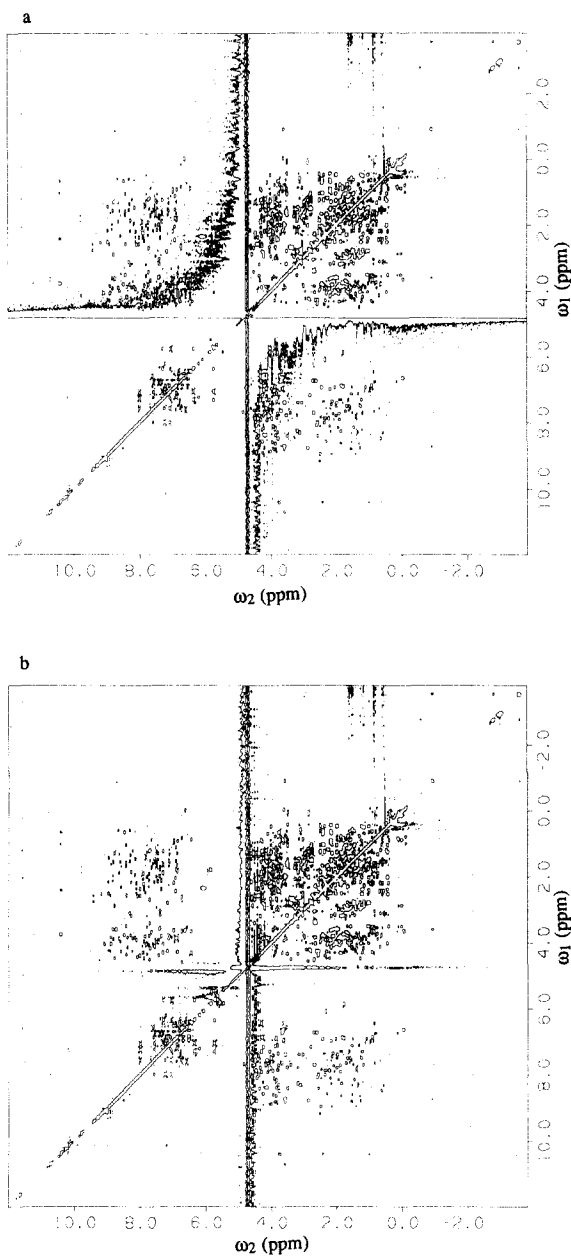


FIG. 3. H_2O TOCSY spectra of Fe(II) cytochrome c_{551} in 50 mM phosphate (pH 6.8) at 25°C (a) before and (b) after correction with a dispersive fit (see text). Four mixing times of 21.7, 30.5, 43.5, and 60.9 ms were coadded to give this spectrum. The spectrum was acquired on a GN-500 NMR spectrometer with 4K complex points in t_2 and 430 real points along t_1 . Time proportional phase incrementation (5) was used to discriminate the sign of the frequencies along t_1 . The data were zero filled to 1K real points along t_1 . A total of 120 scans were collected per FID. Total recycle delay between scans was 1.5 s. Water suppression was carried out by selectively irradiating the water resonance during the recycle delay and the mixing period. A 30° shifted sine-bell window function was used in both dimensions. The original data were collected for work published in Ref. (9).

determined, this constant can be substituted back into Eq. [1]. An approximate value of 10 Hz is used for T_2^* .

We have used the following relation to obtain the baseline correction parameters;

$$I(\omega) = a_1x + a_0 + a_{-1}x^{-1} \quad (x = \omega_0 - \omega). \quad [3]$$

Three sets of points are chosen as reference points for a linear least-squares fit (Fig. 2). This includes two sets from the right- and left-hand edge of the spectrum to control for any linear baseline drift. A third set selected from the empty spectral region around 6 ppm determines the extent of the dispersive component. The baseline at these three spectral regions determines the values of the three adjustable parameters in Eq. [3]. Once the a_i values have been determined by a linear least-squares fit, Eq. [1] is then substituted for last term in Eq. [3]. This equation gives an excellent approximation for the dispersive component of the water line. The baseline correction routine was incorporated into our data-processing software package FTNMR (Hare Research, Inc.) by making minor modifications to the polynomial baseline fitting subroutine.

The dispersive fit is a very conservative baseline correction. Since it mostly involves first-order and zero-order terms, it only eliminates dispersive signals at the carrier frequency. In our experience, the technique does not introduce any new distortions if the spectrum already has a flat baseline. However, the accuracy of the baseline correction is strictly dependent on the assumption that the distortion stems from a dispersive signal at the carrier. Other types of errors, such as distortions caused by first-order phasing errors, may cause problems which this technique will not properly address.

The result of the fitting procedure in 1D is shown in Fig. 2. Figures 1 and 3 show NOESY and TOCSY spectra with two different types of baseline distortions from the water. While the NOESY spectrum usually has mainly dispersive streaks along ω_2 at many ω_1 frequencies, the TOCSY spectrum has a characteristic strong dispersive signal centered at the ω_1 frequency of the water. The technique virtually eliminates the distortions in both spectra.

ACKNOWLEDGMENTS

This work was supported by a grant from Dow Chemical Corporation and by NSF (Grant DMB-8616059). We thank Sven Hyberts for his considerable help in implementing the changes to the source code of FTNMR, David Detefson for providing the TOCSY spectrum, and V. Thanabal for his assistance in running the spectrometer. We thank Dr. Robert Lazarus, Genentech, for providing the sample of kistrin. We also thank Dr. Dennis Hare for providing the source code of FTNMR.

REFERENCES

1. A. KUMAR, G. WAGNER, R. R. ERNST, AND K. WÜTHRICH, *Biochem. Biophys. Res. Commun.* **96**, 1156 (1980).
2. G. WIDER, S. MACURA, A. KUMAR, R. R. ERNST, AND K. WÜTHRICH, *J. Magn. Reson.* **56**, 207 (1984).
3. M. S. DENNIS, W. J. HENZEL, R. M. PITTI, M. T. LIPARI, M. A. NAPIER, T. A. DEISHER, S. BUNTING, AND R. A. LAZARUS, *Proc. Natl. Acad. Sci. USA* **84**, 2471 (1989).
4. D. MARION, I. MITSUHIKO, AND A. BAX, *J. Magn. Reson.* **84**, 425 (1989).
5. D. MARION AND K. WÜTHRICH, *Biochem. Biophys. Res. Commun.* **113**, 967 (1983).
6. L. BRAUNSCHEILER AND R. R. ERNST, *J. Magn. Reson.* **53**, 521 (1983).
7. A. BAX AND D. G. DAVIS, *J. Magn. Reson.* **65**, 355 (1986).
8. C. GRIESINGER, G. OTTING, K. WÜTHRICH, AND R. R. ERNST, *J. Am. Chem. Soc.* **110**, 7870 (1988).
9. D. J. DETEFSON, V. THANABAL, V. L. PECORARO, AND G. WAGNER, *Biochemistry* **29**, 9377 (1990).