Anal. Chem. **2006,** *78,* 3040−3045

Biexponential Fitting of Diffusion-Ordered NMR Data: Practicalities and Limitations

Mathias Nilsson,*,† Mark A. Connell,† Adrian L. Davis,‡ and Gareth A. Morris†

School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK, and Pfizer Global Research and Development, Sandwich, Kent, CT13 9NJ, UK

High resolution diffusion-ordered NMR spectroscopy (HR-DOSY) generally uses monoexponential fitting of the diffusional attenuation of pulsed-field gradient stimulated echo spectra and can, thus, only give proper separation of signals in the diffusion domain where the individual peaks are well-resolved in the spectral domain. In principle, it should be possible to resolve the decays of coincident spectral peaks by multiexponential fitting, but the well-known fundamental difficulties in such a separation are exacerbated by instrumental distortions of the signal decay. The limitations imposed on biexponential fitting by finite signal-to-noise ratio and by the distortion of the theoretical form of the signal attenuation as a result of pulsed-field gradient nonuniformity are explored, and the improvement in DOSY spectra produced using biexponential fitting afforded by compensating for the latter are illustrated.

Diffusion-ordered spectroscopy (DOSY) aims to separate the NMR signals from species of different sizes (hydrodynamic radii) on the basis of their apparent diffusion coeffiecients.[1−3] In a typical DOSY experiment, a series of spectra is recorded with incremented pulsed-field gradient (PFG) amplitudes in a pulsed-field gradient spin−echo (PFGSE) or a pulsed-field gradient stimulated echo (PFGSTE) experiment. The 1D spectra are then converted into a 2D spectrum by fitting the decay of signal as a function of PFG amplitude to the appropriate theoretical expression and constructing a second, diffusion, dimension with Gaussian line shapes centered on the calculated diffusion coefficient (*D*) and with line widths derived from the estimated error of fit.

The decay of the signal in an ideal pulsed-field gradient echo experiment is described by the Stejskal−Tanner equation,[4]

$$S = S_0 e^{-D\gamma^2\delta^2 g^2 \Delta'} \qquad (1)$$

where S is the signal amplitude, S_0 is the echo amplitude that would have resulted had there been no diffusion, D is the diffusion coefficient, δ is the gradient pulse width, γ is the magnetogyric ratio, g is the gradient amplitude, and Δ' is the diffusion time corrected for the effects of finite gradient pulse width. In practice, as noted below, significant deviations from exponential behavior may be observed, notably as a consequence of the spatial variation over the sample of the magnitude of the gradient g.

In the high-resolution DOSY technique (HR-DOSY[5]) the signal decay is fitted to eq 1; if there is no spectral overlap, the best experimental data show standard errors in diffusion coefficient estimated in the fitting process of below 0.2%, corresponding to the resolution of components which differ in diffusion coefficient by ∼0.5%. HR-DOSY gives excellent results where signals are well-resolved, but there is generally some signal overlap in the spectra of all but the simplest mixtures. Fitting superimposed decays to a single exponential will result in a compromise value for the diffusion coefficient, complicating or even preventing proper interpretation. Extending 2D DOSY to a 3D DOSY experiment, in which signal overlap is reduced using, for example, COSY[6,7] or HMQC[8,9] and fitting the decays of the resolved cross-peaks can be a very effective way to circumvent the problem of overlap, but inevitably much longer acquisition times are required. Signal overlap remains a problem, but for more complex mixtures.

A number of approaches have been used to resolve the diffusion dimension in 2D DOSY data of overlapping signals; these may be divided into single-channel[10−12] and multichannel[13−19]

* Corresponding author. E-mail: mathias.nilsson@manchester.ac.uk.
† University of Manchester.
‡ Pfizer Global Research and Development.

(1) Antalek, B. *Concepts Magn. Reson.* **2002**, *14*, 225−258.
(2) Johnson, C. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203−256.
(3) Morris, G. A. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; John Wiley & Sons, Ltd: Chichester, 2002; Vol. 9, pp 35−44.
(4) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288−292.
(5) Barjat, H.; Morris, G. A.; Smart, S.; Swanson, A. G.; Williams, S. C. R. *J. Magn. Reson., Ser. B* **1995**, *108*, 170−172.
(6) Nilsson, M.; Gil, A. M.; Delgadillo, I.; Morris, G. A. *Chem. Commun.* **2005**, 1737−1739.
(7) Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson., Ser. A* **1996**, *121*, 88−91.
(8) Barjat, H.; Morris, G. A.; Swanson, A. G. *J. Magn. Reson.* **1998**, *131*, 131−138.
(9) Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson., Ser. A* **1996**, *123*, 215−218.
(10) Ambrosone, L.; Ceglie, A.; Colafemmina, G.; Palazzo, G. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3040−3047.
(11) Delsuc, M. A.; Malliavin, T. E. *Anal. Chem.* **1998**, *70*, 2146−2148.
(12) Morris, K. F.; Johnson, C. S. *J. Am. Chem. Soc.* **1993**, *115*, 4291−4299.
(13) Armstrong, G. S.; Loening, N. M.; Curtis, J. E.; Shaka, A. J.; Mandelshtam, V. A. *J. Magn. Reson.* **2003**, *163*, 139−148.
(14) Huo, R.; Wehrens, R.; Buydens, L. M. C. *J. Magn. Reson.* **2004**, *169*, 257−269.
(15) Huo, R.; Wehrens, R.; van Duynhoven, J.; Buydens, L. M. C. *Anal. Chim. Acta* **2003**, *490*, 231−251.
(16) Stilbs, P.; Paulsen, K.; Griffiths, P. C. *J. Phys. Chem.* **1996**, *100*, 8180−8189.
(17) Van Gorkom, L. C. M.; Hancewicz, T. M. *J. Magn. Reson.* **1998**, *130*, 125−130.

10.1021/ac060034a CCC: $33.50
Published on Web 03/18/2006

methods. In single-channel methods, the notoriously hard task of resolving superimposed exponentials[20] is accomplished by some approximation, more or less crude, of the inverse Laplace transform, and in the multichannel methods by using assumptions as to the form of the decay, the number of components present, or both. Perhaps the simplest approach to resolving superimposed exponential decays is to assume a sum of two discrete exponentials, in a biexponential fitting of the decay of each peak,

$$S = S_{0A}e^{-D_A\gamma^2\delta^2g^2\Delta'} + S_{0B}e^{-D_B\gamma^2\delta^2g^2\Delta'} \quad (2)$$

where D_A, D_B, S_{0A}, and S_{0B} are the diffusion coefficients and amplitudes of the signals originating from the two species A and B, respectively.

In practice, the results of biexponential fitting of DOSY data are frequently very disappointing unless the overlapping signals differ markedly in diffusion coefficient. Two common reasons for this are the limitations imposed by instrumental noise, and systematic deviations from the Stejskal−Tanner equation. Although there are many potential sources of the latter, for example, electronic nonlinearity of the gradient amplifier/gradient coil system, in typical high resolution instrumentation such as that used here the dominant source is spatial nonuniformity of the pulsed-field gradient. Because different regions of the sample experience slightly different gradient amplitude, and hence show slightly different diffusional attenuation as a function of nominal gradient strength,[21] the overall signal decay measured is the integral of a continuous distribution of exponential functions rather than a simple exponential. In standard HR-DOSY experiments, this deviation from a pure exponential decay will normally manifest itself in distorted values for the fitted diffusion coefficients and in systematically higher error estimates. Because the signal decay that results from a distribution of gradient strengths is indistinguishable from that which is produced by a distribution of diffusion coefficients, many of the more sophisticated methods for analyzing DOSY data in terms of multiple components will be adversely affected, and indeed in some cases vitiated, by the effects of field gradient nonuniformity. The standard forms of the CONTIN[22] (*contin*uous distributions of exponentials) and SPLMOD[23] (*spl*ine *mod*el; discrete sum of exponentials) algorithms assume pure exponential decay, although SPLMOD can be adapted to deal with other one-parameter functions. It is a fundamental requirement for DECRA[18] (*d*irect *e*xponential *c*urve *r*esolution *a*lgorithm; based on the generalized rank annihilation method) that the data be a discrete sum of pure exponential decays, and problems with impure exponential data have been noted previously.[1,24] The standard form of CORE[16,19] (*co*mponent *re*solved NMR; based on standard fitting procedures) also assumes exponential decays (or distributions of exponentials, or both), but it could, in principle, be used for any form of decay. In standard MCR[17] (*m*ultivariate *c*urve *r*esolution; based on principal component analysis), no assumption as to the shape of the decay is necessary.

In principle, the impact of nonuniform field gradients can be reduced greatly by using selective pulses in the presence of field gradients to acquire data only from some part of the sample where the field gradient is a good approximation to linear.[25] Unfortunately, such slice selection methods can only reduce, not eliminate, the effects of gradient nonuniformity, and then only at the expense of a large sacrifice in signal-to-noise ratio. It is, therefore, important to seek data analysis methods that can compensate for field gradient nonuniformity, allowing the full sensitivity of DOSY experiments to be retained. At least two such methods have been described, in which the spatial distributions of the signal strength and gradient strength are mapped and the expected nonexponential form of the signal decay due to diffusion is calculated. This synthetic decay can then be parametrized as a sum of error functions[26] or the exponential of a power series[3] to produce a characteristic decay function for a given experimental configuration. Fitting HR-DOSY data to such a modified decay function can then yield highly accurate diffusion results.

One practical implementation of such a scheme, used in this work, is as follows. An approximate gradient calibration is performed, for example, by measuring the width of the signal profile under a given nominal gradient, and the dependence of the field gradient and signal amplitude on position is mapped using experiments on a sample of known diffusion coefficient D_0. From these data, the form of the dependence of signal attenuation on the square of nominal field gradient amplitude is calculated for a diffusion coefficient D_0 by numerical integration of the Stejskal−Tanner equation (eq 1) over the sample volume. The functional form of this decay is then represented as the exponential of a power series in the normal Stejskal−Tanner exponent σ,

$$S = S_0e^{-\sum_{n=1}^{N}c_n\sigma^n} \quad (3)$$

where

$$\sigma = D\gamma^2\delta^2g^2\Delta' \quad (4)$$

The coefficients c_n may then be found by fitting the calculated signal decay to eq 3; the first coefficient c_1 is the square of the correction factor needed to convert the nominal gradient strength into the true signal-weighted average of the gradient over the sample. In practice, two to four terms generally suffice; the coefficients are characteristic of a particular probe and pulse sequence. Given the coefficients, it is straightforward to replace the exponents in eq 2 to obtain a biexponential fitting function that takes into account the effects of pulsed-field gradient nonuniformity.

$$S = S_{0A}e^{-\sum_{n=1}^{N}c_n\sigma_A^n} + S_{0B}e^{-\sum_{n=1}^{N}c_n\sigma_B^n} \quad (5)$$

It is interesting to speculate on the scope for incorporating correction for nonuniform field gradients into techniques such

(18) Windig, W.; Antalek, B. *Chemom. Intell. Lab.* **1997**, *37*, 241−254.
(19) Stilbs, P.; Paulsen, K. *Rev. Sci. Instrum.* **1996**, *67*, 4380−4386.
(20) Istratov, A. A.; Vyvenko, O. F. *Rev. Sci. Instrum.* **1999**, *70*, 1233−1257.
(21) Hurd, R. E.; Deese, A.; Johnson, M. O.; Sukumar, S.; van Zijl, P. C. M. *J. Magn. Reson., Ser. A* **1996**, *119*, 285−288.
(22) Provencher, S. W. *Comput. Phys. Commun.* **1982**, *27*, 229−242.
(23) Provencher, S. W.; Vogel, R. H. In *Numerical Treatment of Inverse Problems in Differential and Integral Equations*; Deuflhard, P., Hairer, E., Eds.; Birkhäuser: Boston, 1983; pp 304−319.
(24) Alam, T. M.; Alam, M. K. *J. Chemom.* **2003**, *17*, 583−593.
(25) Tillett, M. L.; Lian, L. Y.; Norwood, T. J. *J. Magn. Reson.* **1998**, *133*, 379−384.
(26) Damberg, P.; Jarvet, J.; Gräslund, A. *J. Magn. Reson.* **2001**, *148*, 343−348.

as CONTIN or the RRT[13] (*r*egularized *r*esolvent *t*ransform, a variant of the filter diagonalization method). It should, for example, be relatively straightforward to change the kernel of a maximum entropy algorithm[11] to use a modified decay function of the sort discussed above.

In analyzing experimental data as the sum of two exponential functions, a number of factors affect the ability to characterize successfully the two component decays.[20] These include the signal-to-noise ratio of the data, the extent of any deviations from true exponential behavior, the ratio of the amplitudes of the exponentials, the ratio of their decay constants, the range over which the data are measured, and (as data are acquired discretely) the sampling pattern. In the present investigation, the effects of two of these factors, signal-to-noise ratio limitations and deviations from exponential character caused by nonuniform pulsed-field gradients, on the results obtained from biexponential fitting of DOSY data are studied using simulation and experiment, and some general conclusions are drawn. The two factors are interconnected because biexponential fitting relies on characterizing accurately very small deviations from monoexponential behavior. Separating exponential decays with similar decay constants requires data of very high signal-to-noise ratio, which in turn can expose even very small deviations from exponential character in the component decays. The effect of noise on biexponential fitting is necessarily variable, and where signal-to-noise ratio is marginal, the results of fitting can vary greatly. A small change in the noise can cause fitting to change from converging to a single exponential (or more accurately, an arbitrary mixture of two exponentials with approximately equal decay constants) to converging to two quite different exponentials. The approach taken here is, therefore, to average large numbers of computer simulations to demonstrate the relationship between the most likely fitting results and the signal-to-noise ratio. The consequences of nonexponential signal decay on biexponential fitting can, in contrast, readily be demonstrated experimentally in high signal-to-noise data, and a high-quality experimental dataset is used to show both the florid effects of nonuniform pulsed-field gradients on biexponentially fitted DOSY data and a very effective cure for these.

EXPERIMENTAL SECTION

All experimental measurements were carried out at 25 °C, nonspinning, on a 400 MHz Varian Inova spectrometer using a 5-mm-diameter indirect detection probe equipped with a gradient coil, allowing gradient pulses of up to 30 G cm^{-1}. A sample containing 1% sucrose, 0.2% propan-1-ol, and 0.2% isopentanol (w/v) in D_2O with sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d_4 (TSP) as a chemical shift reference was used in all experiments. The oneshot sequence[27] was used with an imbalance factor (α) of 0.2 for the diffusion-encoding gradient pulses, with a diffusion delay (Δ) of 0.3 s, a diffusion-encoding pulse width (δ) of 3 ms, and gradient strengths ranging from 3.0 to 27.3 G cm^{-1} in equal steps of gradient squared. For this sequence, Δ' is equal to $\Delta + \delta(\alpha^2 - 2)/6 + \tau(\alpha^2 - 1)/2$, where the two pulses in each bipolar pulse pair have amplitudes in the ratio $(1 + \alpha):(1 - \alpha)$, and τ is the time between the midpoints of the individual gradient pulses in each diffusion−encoding period. For each of 30 gradient amplitudes, 8192 complex data points were acquired. 2D DOSY spectra were constructed as described previously[5] by fitting the decay of each peak in the spectrum to one of eqs 1, 2, or 5. Using an initial gradient calibration carried out according to the manufacturer's instructions, the four parameters c_1−c_4 were 0.928, -9.78×10^{-3}, -3.83×10^{-4}, and 2.51×10^{-5}, respectively. Baseline correction was applied to all spectra prior to DOSY fitting.

To avoid local minima in finding the best fit of the experimental data to the sum of two decay curves (eqs 2 and 5), four-parameter (two amplitudes and two diffusion coefficients) least-squares fitting by the Levenberg−Marquardt algorithm was initiated from multiple starting points, and the fit with lowest χ^2 was chosen. The sampling grid of starting points was set using the results of a two-parameter monoexponential fit to span amplitudes ranging from 0 to 100% of the total signal amplitude and diffusion coefficients ranging from 1/9 to 9 times the diffusion coefficient estimated from the monoexponential fit. For decays in which biexponential fitting yielded a negative or an unrealistically large ($>10^{-8}$ m^2 s^{-1}) diffusion coefficient, a negative peak amplitude, or a standard error exceeding 30% for any parameter, a monoexponential fit was substituted.

The effect of noise on the biexponential fitting of DOSY data was investigated using the program Mathematica[28] by synthesising sampled diffusional decays from the Stejskal−Tanner equation and adding controlled amounts of synthetic white Gaussian noise. The relative amplitudes (1:2.8) and the diffusion coefficients (3.28 and 7.08×10^{-10} m^2 s^{-1}) used for the two components were those measured experimentally for two overlapping signals in the proton spectrum of the sample containing sucrose and propan-1-ol described earlier. The same values of g, Δ, δ, and γ as in the real experiment were used. The decays were then fitted to eqs 1 and 2 using the Levenberg−Marquardt nonlinear least squares algorithm to give the estimated amplitudes and decay constants and their associated standard errors for the two exponential components. No allowance was made for gradient nonuniformity, pure exponential decay as a function of gradient squared being assumed. The result of any one fit depends on the noise component, so to draw general conclusions, 1000 fits, each with different pseudorandom noise, were performed for each value of the ratio of the total initial signal to the root-mean-square noise. For each fit, the corresponding trace through the diffusion dimension of a DOSY spectrum was synthesized as the sum of two Gaussians centered on the respective estimated diffusion coefficients, with areas determined by the fitted amplitudes and with widths determined by the standard errors of the two estimated diffusion coefficients. The average diffusion spectrum obtained by summing the 1000 DOSY traces was then plotted as a function of the signal-to-noise ratio for values ranging from 10 000:1 to 10:1.

RESULTS AND DISCUSSION

DOSY spectra were calculated by fitting the experimental signal decay to an appropriate function, 1, 2, or 5. Figure 1 shows 400 MHz proton DOSY spectra for the spectral region 3.4−4.3 ppm containing signals originating from sucrose, isopentanol, and propan-1-ol obtained using fitting to three different functions.

(27) Pelta, M. D.; Morris, G. A.; Stchedroff, M. J.; Hammond, S. J. *Magn. Reson. Chem.* **2002**, *40*, S147−S152.

(28) Wolfram, S. *The Mathematica Book*, 4 ed.; Wolfram Media/Cambridge University Press: Champaign, IL, New York, 1999.

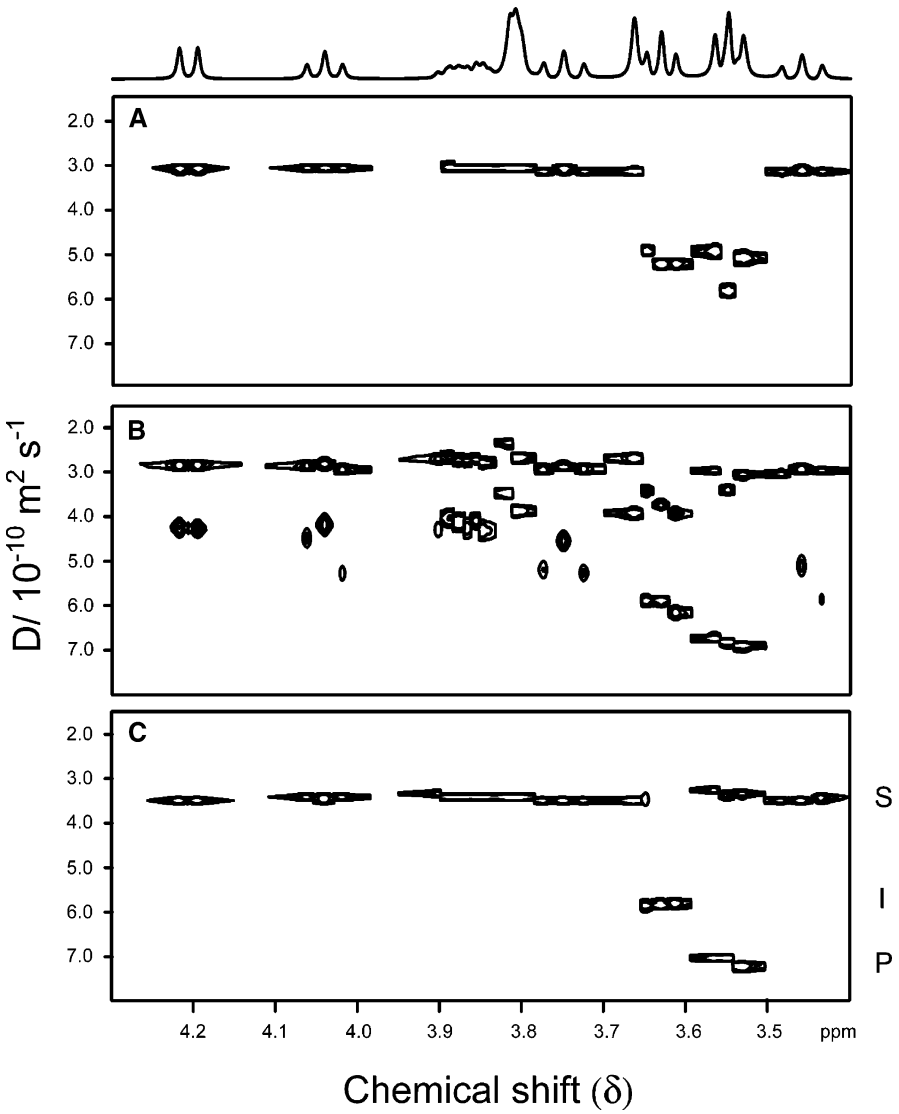

Figure 1. 2D DOSY spectra of an aqueous (D_2O) solution of 0.1% sucrose, 0.2% isopentanol, and 0.2% propan-1-ol, with (top) the least attenuated proton spectrum. (A) DOSY spectrum calculated by fitting to a monoexponential decay (eq 1); (B) calculated by fitting to a biexponential decay (eq 2); and (C) calculated by fitting to the sum of two exponentials of power series (eq 5), to correct for the effects of the spatial nonuniformity of the pulsed-field gradient. The widths of peaks in the vertical (diffusion) dimension reflect the statistical uncertainty in the estimated diffusion coefficient. The origins of the signals are indicated to the right of panel C: S = sucrose, P = propan-1-ol, I = isopentanol.

Figure 1A, a standard high-resolution DOSY spectrum, was constructed using a pure monoexponential fit to eq 1; Figure 1B used a pure biexponential fit to eq 2; and Figure 1C used a fit to two superimposed decays of a shape corrected for the measured spatial nonuniformity of the pulsed-field gradient, eq 5. As discussed below, the residuals for monoexponential fitting of the decay of the signal of a single species are typically <0.1% of the maximum amplitude, whereas those for biexponential fitting are 20 times lower. The region between 3.5 and 3.7 ppm contains signals originating from all three species. In the standard DOSY spectrum (Figure 1A), the signals in the region 3.50−3.60 ppm are distributed over a range of apparent diffusion coefficients, reflecting the different contributions made by overlapping signals of different diffusion coefficients to each peak, rather than appearing at one or other of the diffusion coefficients of the three components of the mixture. Perhaps surprisingly, on applying biexponential fitting (Figure 1B), the spectrum becomes much harder to interpret because most of the spectral peaks are split in the diffusion dimension, even when they originate from a single species, seeming to imply that there are species of many different sizes. Correcting for the nonuniformity of the field gradients by fitting to eq 5, on the other hand, the DOSY spectrum of Figure 1C shows unambiguously the presence of molecules from just three species. The overlapping multiplets of sucrose (glucose H2) and propan-1-ol (methylene 2) at 3.55 and of sucrose (fructose H6) and isopentanol (H2) at 3.63 ppm are correctly resolved into contributions at the two different diffusion coefficients.

The reason for the confused spectrum in Figure 1B is that the biexponential function of eq 2 is an excellent fit to a single decay of the form of eq 3: the variation of PFG strength across the sample causes the same NMR signal in different parts of the sample to decay at different rates, leading to an overall signal decay, which is almost indistinguishable (see Figure 2 below) from a biexponential function. The result is that the nonexponential decay of signal from a single species is represented in the biexponentially fitted DOSY spectrum by two peaks with apparent

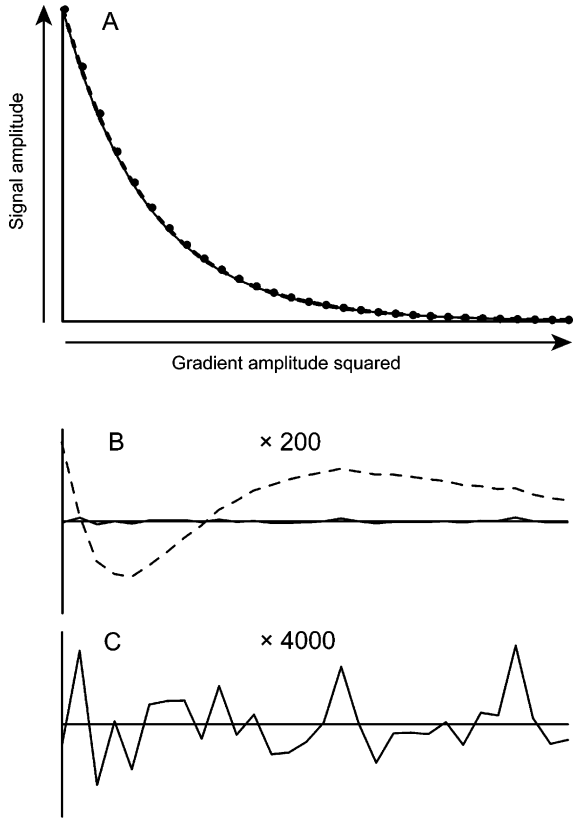

Figure 2. (A) Data points for the decay of the TSP peak (dots), from the experimental data for Figure 1, with the monoexponential fit (dashed line) and biexponential fit (solid line); (B) residuals from the monoexponential fit (dashed line) and the biexponential fit (solid line), expanded 200 times; (C) residuals from the biexponential fit, expanded 4000 times.

diffusion coefficients in the approximate ratio 1:1.4. Although the ratio here is a function of the particular PFG probe used and of the choice of signal attenuation range, any probe suitable for high-resolution NMR will to a greater or lesser extent show the same type of behavior, because of the compromises needed among gradient uniformity, power dissipation, and dynamic characteristics. Biexponential fitting of high signal-to-noise ratio data typically fails to distinguish between the effects of PFG nonuniformity and those of overlap between signals from molecules of different diffusion coefficients. There is, thus, a strong inherent tendency for biexponential fitting to lead to "false positives" when fitting experimental NMR diffusion data; only where overlapping signals have very different diffusion coefficients will biexponential fitting converge to something close to the correct values, because the effects of nonuniform field gradients are experimentally all but indistinguishable from those of two overlapping signals with similar diffusion coefficients.

The similarity between the signal decay for a single species in a nonuniform field gradient and a biexponential function is demonstrated experimentally in Figure 2A, where the signal decay for the TSP peak in the experiment of Figure 1 is shown together with the fit to a monoexponential (dashed line) and to a biexponential (solid line). The signal-to-noise ratio for the TSP peak in the least attenuated spectrum (corresponding to the first point in Figure 2A) was 7500:1. The two fits are hard to tell apart by eye, but the residuals in Figure 2B clearly show a substantial amount of unfitted signal in the monoexponential fit (dashed line) but not in the biexponential fit (solid line). The residual after the biexponential fit is, as shown in Figure 2C, dominated by noise. It is this almost exact equivalence between the experimental form of the signal decay for a single species with the theoretical form of a composite signal from two species which is to blame for the spurious diffusion peaks in Figure 1B.

The success of a biexponential fit is also critically dependent on the signal-to-noise ratio of the data. This is demonstrated graphically for simulations of perfect experiments (with no nonuniformity of the pulsed-field gradient) in Figure 3, which summarizes the results of fitting simulated biexponential signal decays corresponding to the overlapped sucrose/propan-1-ol signal at 3.55 ppm in Figure 1 to eq 1 and to eq 2, with 200 different amounts of synthetic noise added to give initial signal-to-noise ratios varying from 10 000:1 to 10:1. In the experimental data used to construct Figure 1, the amplitude of this peak in the least attenuated spectrum was 5000 times the root-mean-square noise level, and 25 times in the most attenuated spectrum, so Figure 3 covers the range from slightly higher to very much poorer signal-to-noise ratio than the experiment of Figure 1.The data are shown in the form of a contour plot, with cross sections taken at representative signal-to-noise ratios for the first point of the decay (corresponding to the signal-to-noise ratio of the relevant peak in the first spectrum of a DOSY dataset) of 200:1, 70:1, and 20:1, respectively. The horizontal (diffusion) data domain represents the average of 1000 "diffusion spectra" consisting of two Gaussian peaks, with positions, widths, and areas determined by the fitting algorithm (i.e., cross sections through the diffusion domain of a 2D DOSY spectrum), for each value of the signal-to-noise ratio. At a high signal-to-noise ratio, the fit to a biexponential will successfully yield a DOSY spectrum with two well-defined peaks in the diffusion dimension, but as the signal-to-noise ratio decreases, the fitting procedure is not able to distinguish between the two decays, and a single peak at a compromise value results. Fitting to a single-exponential decay function, as in a conventional HR-DOSY experiment, always yields a single peak in the diffusion dimension at a compromise value, with a width that increases as the signal-to-noise ratio decreases.

There is no single critical value of signal-to-noise ratio above which the two contributions will be separated; rather, as the signal-to-noise ratio increases, so the probability that the component signals will be resolved improves. Thus, for the illustrative case of the data of Figure 3, at a signal-to-noise ratio of 70:1 biexponential fitting will converge sometimes on two components and sometimes on one. To obtain a fit with sufficient accuracy for the diffusion coefficients obtained to have low enough estimated standard errors to be useful in mixture analysis (the purpose of DOSY), a signal-to-noise ratio in excess of 200:1, as in the bottom trace of Figure 3B, is needed. These values of the signal-to-noise ratio are those found for the particular pair of overlapping signals used, which have an amplitude ratio of 1:2.8 and a diffusion coefficient ratio of 1:2.2. The signal-to-noise ratio demands of biexponential fitting increase as the latter ratio decreases and as the former increases; thus, the easiest decays to resolve are those in which the signal amplitudes are similar and the diffusion coefficients are very different.

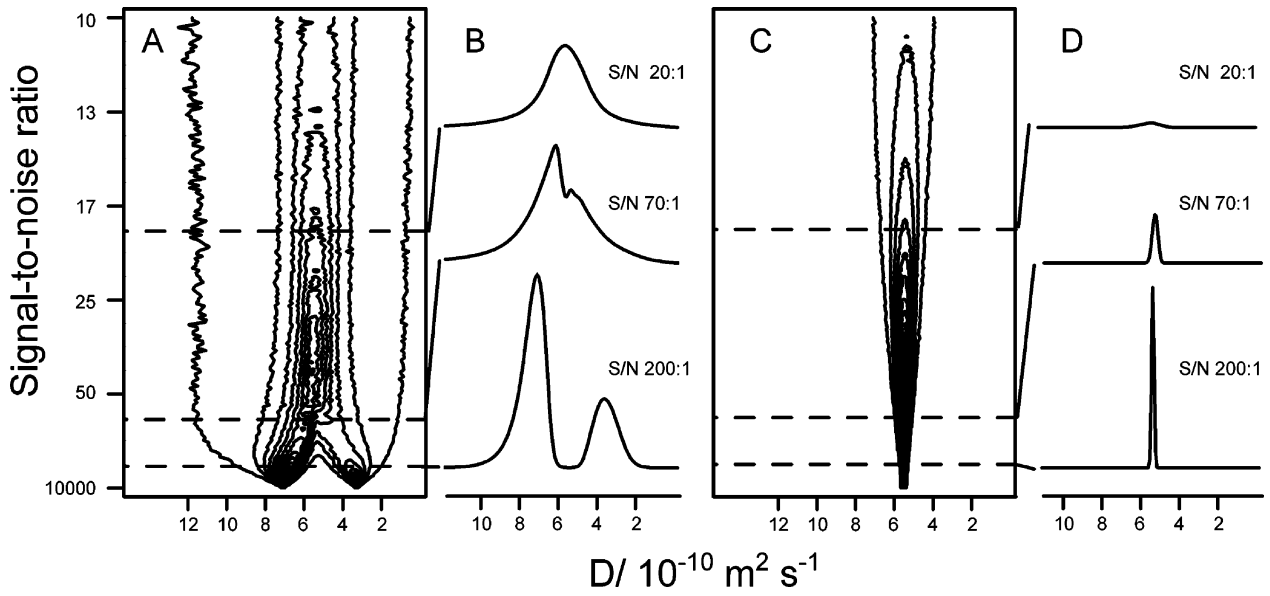

Figure 3. Results of fitting a synthetic decay containing signal of the same relative amplitudes (1:2.8) and diffusion coefficients (3.28 and 7.08 $\times$ 10^{-10} m^2 s^{-1}) as determined experimentally for sucrose and propan-1-ol in the peak at 3.6 ppm in Figure 1A. The decay was calculated using the same values for g, δ, Δ, and γ as in the experiment of Figure 1, and 200 different levels of pseudorandom noise were added to give signal-to-noise ratios ranging from 10000:1 to 10:1. For each signal-to-noise ratio, the diffusion spectra calculated for 1000 fits with different random noise were averaged. (A) Contour plot for fitting to the biexponential function of eq 2; (B) cross sections through (A) showing the average diffusion spectrum at signal-to-noise ratios of 200, 70 and 20:1; (C) contour plot for fitting to the monoexponential function of eq 1; (D) corresponding cross sections through (C). The vertical scale is the same in all traces.

It should be emphasized that the inherent uncertainty involved in fitting noisy data means that for any individual fit, the results obtained in the central region of Figure 3A will vary quite widely; the relatively smooth character of Figure 3A reflects the use of extensive averaging to reveal the underlying trend in the results of fitting. The clear tendency of the fitted diffusion coefficients to converge as the noise level increases, a little reminiscent of the convergence of peaks in chemical exchange, is instructive: the effect of noise here is to reduce the power of biexponential fitting to distinguish between components. As a result, biexponentially fitted DOSY data with experimentally realistic signal-to-noise ratios will *on average* tend to a greater or lesser extent to underestimate the separation of peaks in the diffusion domain.

Real experiments do not yield pure exponential signal decay for the signals from single components. Irrespective of the number of components for which signals overlap, data analyses which take into account field gradient nonuniformity will be more successful than those which do not.

CONCLUSIONS

The model system investigated here is relatively unchallenging to resolve in the diffusion domain, the sucrose and propan-1-ol diffusion coefficients differing by a factor of ~2, so as Figure 3A shows, it is possible to resolve signals of comparable amplitude at signal-to-noise ratios around 100:1. For lower ratios of diffusion coefficient (i.e., for species of more similar size) the signal-to-noise demands rise very sharply, as does the difficulty of distinguishing overlapping signal decays if the effects of nonexponential signal decay are not corrected for. Paradoxically, simple biexponential fitting to eq 2 will work better for low signal-to-noise ratio datasets than for high. Signals with very different diffusion coefficients will be successfully resolved, and the low signal-to-noise ratio will prevent the nonexponential signal decays being interpreted as biexponential. To process reliably high signal-to-noise ratio datasets, such as those needed if the signals of species differing in diffusion coefficient by less than a factor of 2 are to be resolved, it is essential to correct for the effects of pulsed-field gradient nonuniformity.

ACKNOWLEDGMENT

Support from the Engineering and Physical Sciences Research Council (Grant ref GR/S90751/01) is gratefully acknowledged. Pfizer Global Research and Development, Sandwich, Kent, and the EPSRC are thanked for a CASE award to M.A.C.

Received for review January 6, 2006. Accepted February 20, 2006.

AC060034A