JOURNAL OF MAGNETIC RESONANCE **133,** 379–384 (1998)
ARTICLE NO. MN981472

Practical Aspects of the Measurement of the Diffusion of Proteins in Aqueous Solution

Marcus L. Tillett, Lu-Yun Lian, and Timothy J. Norwood*

*Biological NMR Centre, Department of Biochemistry, Leicester University, University Road, Leicester LE1 7RH, United Kingdom; and *Department of Chemistry, Leicester University, University Road, Leicester LE1 7RH, United Kingdom*

Received September 8, 1997; revised April 2, 1998

NMR diffusion measurements have been shown to be a useful tool for investigating the conformation and oligomeric state of proteins. Four main problems associated with making diffusion measurements of proteins in aqueous solution are identified and solutions proposed. The resulting experiment is demonstrated for an aqueous solution of hen egg white lysozyme. © 1998 Academic Press

Key Words: **diffusion; PGSE; proteins; lysozyme.**

NMR diffusion measurements are potentially a valuable tool for probing the conformation and oligomeric state of molecules such as proteins in solution. Unfortunately, attempts to measure diffusion, particularly of dilute species in aqueous solution, are often subject to large errors. Here we examine these problems and propose measures to overcome them.

NMR self-diffusion measurements are sensitive to the shape and size of the diffusing molecule, the viscosity of the medium through which it is diffusing, and the size and geometry of any barriers to its random motion (*1*). While many studies have concentrated on using the latter, which is known as restricted diffusion (*2*), to determine quantities such as cell size and permeability in biological systems and pore size in rock, a number of recent studies have sought to relate the diffusion coefficient to the conformation and oligomeric state of proteins (*3–8*).

Diffusion is usually measured by NMR with pulse sequences based upon either spin or stimulated echoes in which the magnetization is first dephased and subsequently rephased by a linear magnetic field gradient (*9–12*). If a gradient G_z is applied along the *z*-axis, at a coordinate *z*, the magnetization will precess during the application of the gradient at a frequency given by

$$\omega = (1 - \sigma)\gamma(z \cdot G_z + B_0), \qquad [1]$$

where σ is the shielding constant, γ is the magnetogyric ratio, and B_0 is the strength of the static magnetic field. If the gradient is applied during the dephasing period for a time δ, each spin will accumulate a phase $\omega\delta$ which, as can be seen from Eq. [1], will be dependent upon its *z*-coordinate. If no motion occurs along the *z*-axis this phase can be completely reversed by applying a second gradient pulse with the same area as the first during the rephasing period of the experiment, resulting in the formation of an echo. Any attenuation of the signal will be solely due to relaxation. However, any random motion along the *z*-axis, such as that due to diffusion, will lead to incomplete rephasing of the signal and hence to further attenuation of the observed echo. If the time between the start of the two gradient pulses is Δ, the signal attenuation $A(t)$ will be given by

$$A(t) = A(0)\exp(-R(t) - \gamma^2 G^2 D\delta^2(\Delta - \delta/3)), \qquad [2]$$

where $R(t)$ takes into account relaxation attenuation and depends on the structure of the pulse sequence used. Typically a number of experiments are performed using different values of G to change the extent of diffusion attenuation of the observed signal. A value for D can then be obtained by fitting the data with Eq. [2] directly, or by rearranging it so that D can be extracted as the gradient of a semilog plot. All delays are kept constant to maintain a constant value of $R(t)$.

For proteins longitudinal relaxation usually occurs more slowly than transverse relaxation (i.e., $T_1 > T_2$) and it is usually the case, when measuring the diffusion of proteins, that $\delta < \Delta$. This means that $R(t)$, and hence signal loss due to relaxation, can be minimized by storing the magnetization along the *z*-axis whenever it is not evolving under the influence of a magnetic field gradient or being observed. Therefore, the *s*timulated *e*cho *p*ulsed *g*radient spin *e*cho (SE-PGSE) experiment (*10*), Fig. 1A, forms the most suitable basis for an experiment to measure the diffusion of proteins. For this experiment $R(t)$ is given by

$$R(t) = 2\tau/T_2 + \tau_z/T_1. \qquad [3]$$

Before diffusion measurements can be used to provide useful data, it is important that these measurements be made in a reliable and reproducible manner. Many of the papers in the literature which report the use of diffusion measurements do not address the problems that can arise when making them; in

1090-7807/98 $25.00

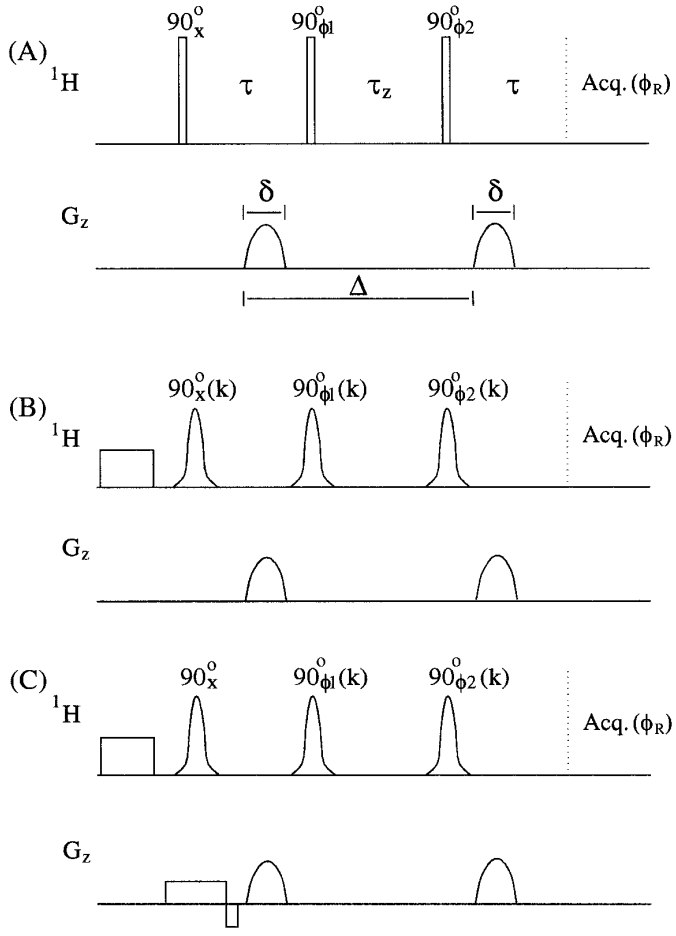

FIG. 1. Stimulated echo-based PGSE pulse sequences for measuring the diffusion of proteins. (A) Basic pulse sequence. (B) Pulse sequence with water suppression. (C) Pulse sequence with water suppression and slice selection. For (B)–(C) the rectangular pulses denote presaturation and the selective pulses labeled 90°(*k*) perturb only the region *k* of the spectrum to be observed. For (C) the first selective pulse is responsible for slice selection in the direction of the gradient. Phase cycle: $\phi 1 = x, y, -x, -y$; $\phi 2 = x, y, -x, -y, -x, -y, x, y$; $\phi R = x, -x, x, -x, -x, x, -x, x$.

some cases the data presented may contain large errors. Additionally, if an aqueous solution of a protein is also to be used for structural or dynamic studies, it is usually most convenient to use H_2O as a solvent since the amide protons may be replaced with deuterium in D_2O.

Four main problems arise when one tries to implement this sequence for measuring the diffusion of dilute species in aqueous solution: spatially dependent variations in the gradient strength, water suppression, restricted diffusion if any interfaces in the direction of the gradient are present within the receiver coil, and thermal currents within the sample. The latter problem does not seem to be widely recognized but can result in large errors. In addition to the above problems, the switching of magnetic field gradients generates eddy currents that may distort the data. While this has in the past been a major problem, it has now been largely overcome by the use of shielded magnetic field gradients. Eddy currents may not be totally suppressed by some older shielded magnetic field gradients; pulse sequences that take this into account by incorporating delays into the pulse sequence to allow eddy currents to die away (*11, 12*), or use bipolar gradients to reduce the eddy currents produced by pairs of short closely spaced gradient pulses (*13*), have been described. We discuss each of the four problems noted above and propose possible solutions for them.

The actively shielded gradients provided by many manufacturers do not provide a gradient with a constant value across the full length of the sample within the receiver coil. This gives rise to a number of potential problems. The analysis conventionally applied to NMR diffusion data assumes that the gradient used is linear. The data obtained with a *non*linear gradient would decay as the sum of a series of exponentials with different decay rates and fitting with Eq. [2] would yield an erroneous value for the diffusion coefficient. Furthermore, if the strength of the applied B_1-field also varies across the sample volume from which signal is detected then the apparent value of a diffusion coefficient may also appear to vary with the pulse sequence used. The analysis of diffusion data acquired with some geometries of a nonlinear gradient have been discussed in the literature (*14*). Similarly, variations in sample volume and height within the receiver coil may also result in variations in the apparent values of the diffusion coefficient.

The simplest way to overcome this nonlinearity of the applied gradient might seem to be to use a height of liquid within the sample tube that is positioned so that it corresponds to a region over which the gradient has a uniform value. Unfortunately this brings both the glass–water and the air–water interfaces at the ends of the sample into the receiver coil, causing a number of additional problems. The change in magnetic susceptibilities at the interface gives rise to intense internal magnetic field gradients which because of their symmetry cannot be effectively shimmed out, hence resulting in broad and asymmetric lines in the spectrum. This makes water suppression in aqueous solutions particularly problematic. Effective water suppression is important because even the relatively small tail of a residual water peak can distort the intensity of the protein resonances which it underlies.

Since the diffusion coefficient of water is large compared to that of proteins diffusion weighting can in principle be used to suppress the water signal. However, in practice using parameter values sufficient to suppress water may also result in unacceptable attenuation of the protein resonances. Furthermore, it will restrict the range of gradient values that can be used by imposing as a minimum that value necessary to suppress water. We find that a more satisfactory alternative is to use presaturation of the water resonance and to replace the rectangular nonselective radiofrequency pulses of the sequence given in Fig. 1A with shaped semiselective ones that excite only the region of the spectrum from which the measurement is to be made, Fig. 1B. This latter approach reduces the amount of signal arising from any residual water magnetization. To

avoid interference from water, peak intensities should be measured in regions of the spectrum that are distant from it. When the diffusion of proteins is being studied the methyl region is most suited to this purpose; its peaks are often relatively intense and, unlike the amide protons, methyl protons do not undergo chemical exchange with those of water. This procedure is relatively satisfactory when the sample extends outside the receiver coil. However, when its ends are within the receiver coil, and the shimming problems noted above are present, a combination of presaturation and selective pulses becomes ineffective and in practice diffusion measurements are found to be unreliable. The shimming and water suppression problem can be resolved by using a susceptibility matched sample tube with a plunger so that there is no air–water interface. However, one problem remains: restriction of diffusion by the interfaces at either end of the sample in the direction of the applied gradient. This phenomenon can also distort the apparent value of the diffusion coefficient obtained.

An alternative method of localizing the region of the sample from which signal is observed, but does not suffer from these above problems, is slice selection (*15*). This entails applying a selective pulse in the presence of a magnetic field gradient. From Eq. [1] it can be seen that in the presence of a magnetic field gradient, and ignoring chemical shift, frequency is proportional to spatial location in the direction of the gradient. Thus, the band of frequencies excited by the selective pulse corresponds to a specific region of space in the direction of the gradient. Slice selection is implemented for the first pulse of the pulse sequence given in Fig. 1B, as shown in Fig. 1C. The negative gradient pulse applied immediately after slice selection serves to refocus any dephasing of the magnetization brought about by the latter. Using this sequence reliable measurements can be made even from samples that do not extend outside the receiver coil provided that they *do* extend beyond the region excited by the slice selection procedure. Slice selection has previously been used to limit the effects of B_1 inhomogeneity (*16*).

In principle a gradient pulse could be applied during the period τ_z to dephase any unwanted magnetization. However, in practice we have not found this to be necessary. Scalar coupling evolution may occur during the experiment. This will distort the phase of peaks in the observed spectrum, though since it is a constant effect it will not affect the quality of the data.

Before any reliable measure of diffusion can be made it is necessary to characterize the magnetic field gradient since in practice it will vary from probe to probe. The easiest way to accomplish this is to use it to make a series of diffusion-weighted images of a sample of water (*17*). One-dimensional diffusion-weighted *z*-images of an H_2O sample extending beyond the region defined by the receiver coil are presented in Fig. 2A. The images were acquired with two different diffusion weightings. These images were obtained with the pulse sequence given in Fig. 1A with a *z*-gradient applied during acquisition to convert the spectrum into an image. If the magnitude of the gradient were uniform along the sample one would expect the two images to be identical except for a scaling factor. However, the weighted subtraction of one image from the other, Fig. 2B, shows that this is not the case; when scaled so that the central regions of the images cancel, the outer regions do not. Since the diffusion coefficient of water is known, these data can be used to calculate the gradient strength at any *z*-coordinate within the receiver coil. For the current example the gradient strength is found to vary across the sample by 10.4%. It can be seen from the difference image that the value of G_z is not even constant at the center of the receiver coil. This means that one of the traditional methods of calibrating gradient strength, from the width of an image obtained across a known distance, may not be entirely reliable.

Diffusion data for a small sample of hen egg white lysozyme in H_2O acquired with the pulse sequence given in Fig. 1B are given in Figs. 3A–3B. This pulse sequence suppresses water by using a combination of presaturation and making all pulses selective for the region of the spectrum from which diffusion is to be measured. A non-susceptibility matched sample tube was used. While the data presented in Fig. 3A appear superficially to be free of water signal, a weighted subtraction of one spectrum from the other reveals an underlying hump that has the diffusion coefficient of water. This is a result of the inevitable reduction of the efficiency of solvent suppression that arises when spectral resolution is degraded by bringing the water–air interface of the sample into the receiver coil. The effects of restriction can be assessed by examining the variations in diffusion attenuation across a pair of one-dimensional images with different diffusion weightings, Figs. 3C–3D. By comparing the weighted difference between these two images with that given in Fig. 2B it can be seen that the data obtained with the small height of liquid (approximately 6 mm) is identical to that obtained with the sample extending outside the receiver coil (Fig. 2B) except for where the sample ends. The sharp spikes in the difference image at these points arise from water molecules whose motion has been restricted by contact with the interfaces at the ends of the sample and therefore appear not to diffuse as quickly as the bulk of the solution. This effect has previously been observed in diffusion-weighted imaging studies of capillary tubes (*18*). Since proteins diffuse more slowly than water this effect will be correspondingly less for proteins, typically by an order of magnitude. However, the largest errors with this method are likely to arise due to variations in sample height and position within the gradient coil; this will affect the measurement of all diffusing species equally.

Slice selection results in loss of signal intensity. This reflects the percentage of the sample within the receiver coil excited by the slice selection procedure, 22% in the current example. The use of selective pulses was found to result in a further 5% signal loss. Missetting the refocusing lobe of the slice selection

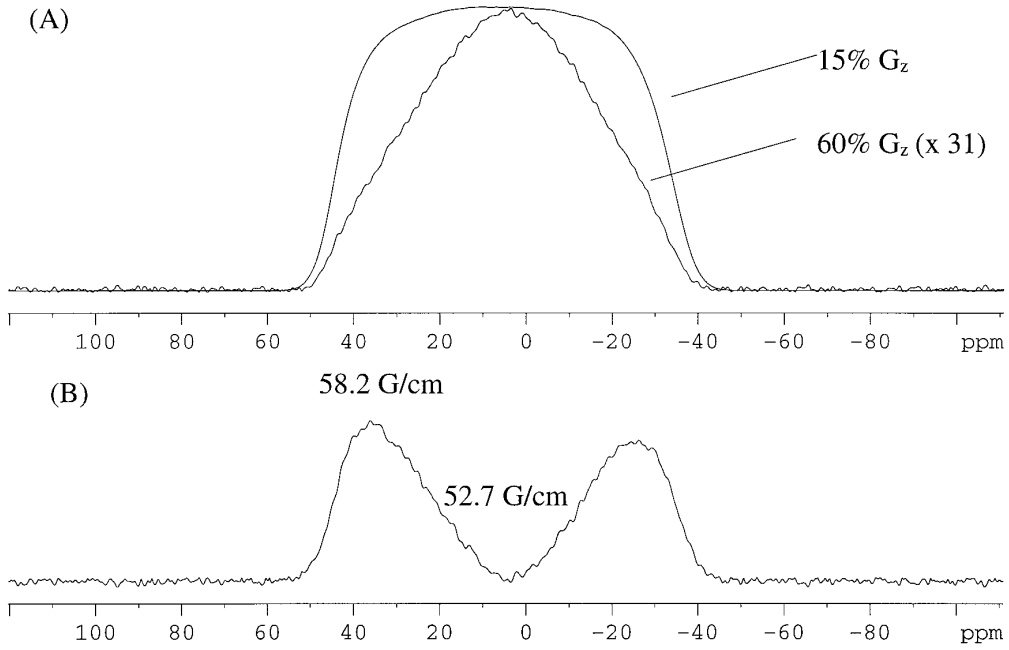

FIG. 2. One-dimensional images of a tube of water extending beyond the receiver coil acquired with the pulse sequence given in Fig. 1A with G_z applied during acquisition. (A) Images acquired with two diffusion weightings; the size of the diffusion-encoding gradient used in each case is indicated as a percentage of the maximum. (B) Weighted difference image of the data given in (A). The lack of complete cancellation across the difference image in (B) indicates a variation of G_z along the sample. Data were acquired using $\Delta = 25.62$ ms, $\delta = 5$ ms, $\tau = 5.6$ ms, $\tau_z = 20$ ms, and $G_z = 8.4$ and 33.6 G cm^{-1} (for diffusion-encoding gradient pulses). For each experiment 64 transients were averaged.

gradient may lead to additional signal loss but was found not to alter the measured value of the diffusion coefficient.

A comparison between diffusion data acquired with and without slice selection is given in Fig. 4A; in all other respects the pulse sequences used were the same in each case. When the diffusion coefficient was measured from the non-slice-selected data it was found to vary depending on the range of gradient values used; the first 10 points yielded a value of $2.307 \pm 0.009 \times 10^{-9}$ m^2 s^{-1}, while the last 10 points yielded a value of $2.224 \pm 0.004 \times 10^{-9}$ m^2 s^{-1}, a difference of 3.6%. This reflects the nonlinear nature of the gradient and the resulting multiexponential decay of the data. Using the slice-selected data the corresponding values calculated for the diffusion coefficients were $2.089 \pm 0.005 \times 10^{-9}$ and $2.068 \pm 0.016 \times 10^{-9}$ m^2 s^{-1}, respectively, a difference of 1%, which is within the experimental error in this case. Clearly, the use of slice selection results in greater accuracy since the apparent value of the diffusion coefficient does not significantly depend on the range of gradient values used and experiments in a linear gradient are much easier to quantify and hence reproduce than those obtained in a nonlinear one.

Data for hen egg white lysozyme in H_2O acquired with the pulse sequence given in Fig. 1C are presented in Fig. 4B. This pulse sequence uses slice selection to restrict the excitation of the sample to a small region, 3 mm in the current example, at the center of the receiver coil. This ensures that only a relatively narrow range of gradient strengths is used in the experiment and eliminates the problems with water suppression associated with using very small sample volumes. The data, obtained at pH values of 8.1 and 1.8, yield diffusion coefficients of $1.116 \pm 0.007 \times 10^{-10}$ and $1.184 \pm 0.008 \times 10^{-10}$ m^2 s^{-1}, respectively. The errors reflect the signal-to-noise of the data which is poor for the larger gradient values used. The difference in the value of the diffusion coefficient at the two pH values probably reflects minor conformational changes, as evident from the spectra (data not shown), and changes in the hydration of the molecule (*19*).

One further potential source of error should be noted: when an experiment is carried out above room temperature thermal gradients may develop along the sample tube. These arise when the tube is not heated uniformly. Thermal gradients may give rise to convection currents within the sample which may increase the apparent value of the diffusion coefficient. We find that measurements performed using a tube with a plunger gave smaller apparent values for the diffusion coefficient than those without. This suggests that thermal gradients, and hence convection currents, across the region of the tube containing the sample are smaller when a plunger is used. This procedure was used to acquire the data presented in Fig. 4. While a pulse sequence has recently been proposed to suppress convection artifacts (*20*), its length may make it unsuitable for use with proteins.

In conclusion, we have identified the major sources of error encountered when making diffusion measurements on a high-

resolution NMR spectrometer. A pulse sequence and an experimental procedure have been developed that enable these problems to be overcome and accurate measurements to be made, even for dilute species in aqueous solution. Furthermore, we have shown that the diffusion coefficient may be sensitive to the relatively subtle effects resulting from changes in conformation and hydration. Work is currently under way to investigate conformational changes in proteins using diffusion measurements.

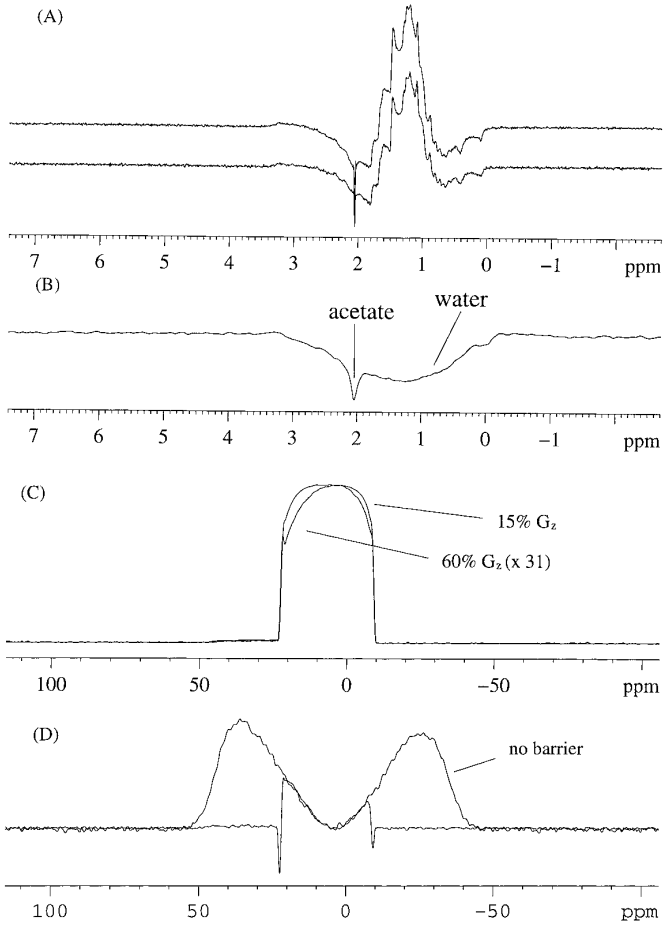

FIG. 3. Diffusion data for samples terminating within the receiver coil. (A) Spectra of 8 mM lysozyme in H_2O acquired using the pulse sequence given in Fig. 1B with two diffusion weightings. (B) Weighted difference of the spectra given in (A). The residual signal arises from acetate impurity (sharp line) and residual water (broad hump). These signals are negative because the spectrum with the least diffusion attenuation has been subtracted from that obtained with the most. (C)–(D) Equivalent of the data given in Fig. 2 for a 6-mm-high H_2O sample terminating within the receiver coil. The size of the diffusion-encoding gradient used in each case is indicated as a percentage of the maximum. The weighted difference of two diffusion-weighted images (C) is given in (D) with the difference image given in Fig. 2B for comparison. The sharp spikes arise from liquid whose diffusion has been restricted by the interfaces at either end of the sample (*18*). For (A)–(B) $\Delta = 209.6$ ms, $\delta = 5$ ms, $\tau = 5.6$ ms, and $\tau_z = 200$ ms with $G_z = 9.35$ and 18.15 G cm^{-1}. Water was presaturated for 1.5 s out of a total relaxation delay of 4 s; eight transients were averaged for each experiment. For (C)–(D) $\Delta = 25.62$ ms, $\delta = 5$ ms, $\tau = 5.6$ ms, and $\tau_z = 20$ ms with $G_z = 8.1$ and 32.4 G cm^{-1}. For each experiment 512 transients were averaged.

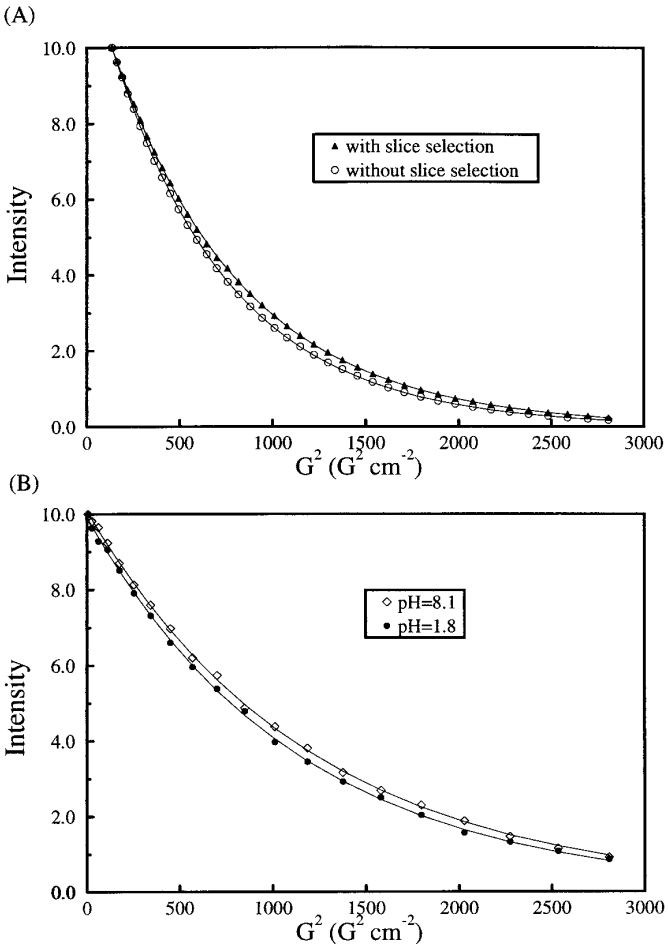

FIG. 4. (A) Diffusion data of 6:4 H_2O:D_2O acquired with the pulse sequences given in Figs. 1B and 1C (without presaturation). (B) Diffusion data for 2 mM lysozyme in H_2O obtained using the pulse sequence given in Fig. 1C with the sample contained in a sample tube with a plunger. Data were acquired at pH values of 8.1 and 1.8. For (A) data points were acquired with 40 gradient values in the range 22–100% of maximum using $\delta = 3$ ms, $\Delta = 26.6$ ms, $\tau = 5.6$ ms, and $\tau_z = 15$ ms; for each experiment 16 transients were averaged. For (B) the data were acquired using $\Delta = 162.6$ ms, $\delta = 4$ ms, $\tau = 6.6$ ms, $\tau_z = 150$ ms, and 20 values of G_z between 2.65 and 53 G cm^{-1}. For each experiment 32 transients were averaged. All data were acquired at 300 K, 5° above room temperature on a Bruker DRX-400 NMR spectrometer operating at 400 MHz for ^{1}H using a triple-resonance inverse probe equipped with an actively shielded *z*-gradient. All diffusion-encoding magnetic field gradient pulses were modulated with a sine function.

ACKNOWLEDGMENTS

We thank Jonathon Boyd (Oxford University) for his invaluable advice on thermal gradients, Gerry Griffith for assistance in using the spectrometer, and the EPSRC for making its purchase possible. One of us (M.L.T.) thanks the MRC for a Research Studentship.

REFERENCES

1. P. Stilbs, Fourier transform pulsed-gradient spin-echo studies of molecular diffusion, *Prog. NMR Spectrosc.* **19,** 1–45 (1989).
2. J. E. Tanner, Transient diffusion in a system partitioned by permeable barriers. Application to NMR measurements with a pulsed field gradient, *J. Chem. Phys.* **69,** 1748–1754 (1978).
3. A. J. Dingley, J. P. Mackay, B. E. Champan, M. B. Morris, P. W. Kuchel, B. D. Hambly, and G. F. King, Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: Application to myosin light chain 2, *J. Biomol. NMR* **6,** 321–328 (1995).
4. A. S. Altieri, D. P. Hinton, and R. A. Byrd, Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements, *J. Am. Chem. Soc.* **117,** 7566–7567 (1995).
5. H. M. Baranowska and K. J. Olszewski, The hydration of proteins in solution by self-diffusion coefficients NMR study, *Biochim. Biophys. Acta* **1289,** 312–314 (1996).
6. V. V. Krishnan, Determination of oligomeric state of proteins in solution from pulsed-field-gradient self-diffusion coefficient measurements. A comparison of experimental, theoretical and hard sphere approximated values, *J. Magn. Reson.* **124,** 468–473 (1997).
7. E. Ilyina, V. Roongta, H. Pan, C. Woodward, and K. H. Mayo, A pulsed-field gradient NMR study of bovine pancreatic trypsin inhibitor self-association, *Biochemistry* **36,** 3383–3388 (1997).
8. M. Lin and C. K. Larive, Detection of insulin aggregates with pulsed-field gradient nuclear magnetic resonance spectroscopy, *Anal. Biochem.* **229,** 214–220 (1995).
9. E. O. Stejskal and J. E. Tanner, Spin diffusion measurements: Spin echoes in the presence of time-dependent magnetic field gradients, *J. Chem. Phys.* **42,** 288–292 (1965).
10. J. E. Tanner, Use of stimulated echoes in NMR diffusion studies, *J. Chem. Phys.* **52,** 2523–2526 (1970).
11. S. J. Gibbs and C. S. Johnson, Jr., A PFG experiment for accurate diffusion and flow studies in the presence of eddy currents, *J. Magn. Reson.* **93,** 395–402 (1991).
12. T. J. Norwood, New NMR methods for measuring diffusion, *J. Magn. Reson. A* **103,** 258–267 (1993).
13. D. Wu, A. Chen, and C. S. Johnson, Jr., An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses, *J. Magn. Reson. A* **115,** 260–264 (1995).
14. P. Le Doussal and P. N. Sen, Decay of nuclear magnetization by diffusion in a parabolic magnetic field: An exactly solvable model, *Phys. Rev. B* **46,** 3465–3485 (1992).
15. P. Mansfield and P. G. Morris, ''NMR Imaging in Biomedicine,'' Academic Press, New York (1982).
16. R. E. Hurd, B. K. John, P. Webb, and D. Plant, Spatially selective suppression of B_1 inhomogeneity, *J. Magn. Reson.* **99,** 632–637 (1992).
17. R. E. Hurd, A. Deese, M. O'Neil Johnson, S. Sukumar, and P. C. M. Van Zijl, Impact of differential linearity in gradient-enhanced NMR, *J. Magn. Reson. A* **119,** 285–288 (1996).
18. W. B. Hyslop and P. C. Lauterbur, Effects of restricted diffusion on microscopic NMR imaging, *J. Magn. Reson.* **94,** 510–510 (1991).
19. M. L. Tillett, 13th International Meeting on NMR Spectroscopy, Exeter, United Kingdom, July 6–11, 1997.
20. A. Jerschow and N. Muller, Suppression of convection artifacts in stimulated-echo diffusion experiments. Double-stimulated echo experiments, *J. Magn. Reson.* **125,** 372–375 (1997).