Journal of Biomolecular NMR, **13:** 223–232, 1999.
KLUWER/ESCOM

Protein-ligand interactions measured by ^{15}N-filtered diffusion experiments

Marcus L. Tillett[a], Mark A. Horsfield[b], Lu-Yun Lian[a,*] & Timothy J. Norwood[c]
[a]*Biological NMR Centre, Department of Biochemistry, Leicester University, University Road, Leicester LE1 7RH, U.K.*
[b]*Division of Medical Physics, Leicester University, Leicester Royal Infirmary, Leicester LE1 5WW, U.K.*
[c]*Department of Chemistry, Leicester University, University Road, Leicester LE1 7RH, U.K.*

Received 30 July 1998; Accepted 9 December 1998

Key words: diffusion, heteronuclear filtration, peptide binding, protein, SH3

Abstract

NMR diffusion coefficient measurements have been shown to be sensitive to the conformational and oligomeric states of proteins. Recently, heteronuclear-filtered diffusion experiments have been proposed [Dingley et al. (1997) *J. Biomol. NMR*, **10**, 1–8]. Several new heteronuclear-filtered diffusion pulse sequences are proposed which are shown to have superior sensitivity to those previously proposed. One of these new heteronuclear-filtered diffusion experiments has been used to study the binding of an SH3 domain to a peptide. Using this system, we show that it is possible to measure binding constants from diffusion coefficient measurements.

Introduction

Diffusion measurements have proved to be a useful tool for detecting aggregation in protein solutions (Altieri et al., 1995; Dingley et al., 1995; Lin et al., 1995; Ilyina et al., 1997; Krishnan, 1997; Price et al., 1997). This information is important for both structural and dynamic studies of proteins by NMR. In addition, diffusion measurements can be used to investigate hydration (Baranowska et al., 1996) and to detect conformation changes (Jones et al., 1997; Pan et al., 1997) in macromolecules. NMR diffusion measurements provide a useful alternative to methods such as ultracentrifugation or light scattering which are not always suitable for use with samples prepared for NMR studies.

Diffusion is usually measured by NMR by first dephasing and then rephasing the transverse magnetization in the presence of a linear magnetic field gradient pulse (Stejskal et al., 1965; Tanner, 1970; Gibbs et al., 1991; Norwood, 1993). During each magnetic field gradient pulse the magnetization accumulates a phase that depends upon its spatial location in the direction of the magnetic field gradient. The phases accumulated during the two gradient pulses have opposite signs and will thus cancel out, giving rise to an echo, provided the pulses have the same area. However, any motion of the spins, such as that resulting from diffusion, in the direction of the magnetic field gradient between the dephasing and rephasing periods will result in a difference in the phases accumulated during the two gradient pulses and hence incomplete refocusing of the magnetization. The diffusion coefficient can be calculated from the resulting reduction in amplitude of the echo. In practice, experiments are usually performed with a number of different magnetic field gradient strengths. While most measurements are made with pulse sequences incorporating pulsed magnetic field gradients, static gradients can also be used.

The diffusion coefficient depends on a number of factors, including the shape and size of the diffusing molecule and the viscosity of the medium through which it is diffusing (Stilbs, 1989). While attempts have been made to quantify the diffusion coefficients of proteins in terms of both molecular weight (Dingley et al., 1995) and solvent accessible surface area (Krishnan, 1997), the results have only been semi-

*To whom correspondence should be addressed. E-mail: yun@le.ac.uk

quantitative, in part due to the approximations that have to be made about the quantity and location of hydrated water molecules.

The earliest pulse sequences for measuring diffusion, such as the pulsed gradient spin echo (PGSE) (Stejskal et al., 1965) were based on a spin echo. Subsequently sequences based on stimulated echoes (Tanner, 1970) were also developed. Since, for biological macromolecules, $T_1 \gg T_2$, the latter have been most widely used for measurements of proteins since these enable the magnetization to be stored along the z-axis whenever it is not being dephased or rephased by a magnetic field gradient pulse. While most pulse sequences only work well when shielded magnetic field gradients are used, variants have been suggested that do not require them (Gibbs et al., 1991; Norwood, 1993). In addition, we have recently identified and proposed solutions for some problems which may result in erroneous measurements, including nonlinearity of the gradient, restricted diffusion and convection within the sample (Tillett et al., 1998). A pulse sequence has recently been proposed to suppress convection artifacts (Jerschow et al., 1997), though its length may make it unsuitable for use with proteins.

Most structural studies by NMR, of all but the smallest protein molecules, utilize labelled samples. Thus, by incorporating heteronuclear filtration into the diffusion pulse sequence it is possible to eliminate the signals arising from any unlabelled components, distinguishing between the labelled protein and any unlabelled ligands in the solution. Pulse sequences for use with labelled samples have recently been proposed (Dingley et al., 1997).

Here we propose a number of new heteronuclear-filtered diffusion experiments. These are compared to those previously proposed and are shown to exhibit superior sensitivity. They are used to study the binding of an Src homology 3 (SH3) domain to a ligand. The SH3 domains consist of 60–70 amino acids and mediate protein-protein interactions by binding to proline-rich sequences. The ligand used in this study was a target peptide from the cytoplasmic domain of p22phox. The binding affinities between SH3 domains and their ligands are low, of the order of 5–100 μM. The methods described here offer the possibility to screen for and establish binding affinities of a range of ligands for these SH3 domains under NMR sample conditions; at the same time, the recognition site can be located.

Experimental

Materials

NdeI and BamHI restriction sites were introduced at the 5′ and 3′ ends respectively of the protein G domain II gene. The PCR product was sequenced before subcloning into pET11a. For protein expression, E. coli BL21(DE3) cells were transformed with the resulting construct. Uniformly ^{15}N-labelled protein G domain II was obtained by growing the transformed E. coli in 2× M9 medium containing 1 g/l of [^{15}N] ammonium chloride as the sole nitrogen source, with 200 μg/ml ampicillin. The protein purification was similar to that described by Lian et al. (1991). A 4 mM solution of protein G was used in 50 mM NaH_2PO_4 buffer at pH 6.2.

The SH3(C) domain of phox47 was expressed in E. coli BL21(DE3) pLysS using a plasmid supplied by Prof. A. Segal, University College, London. Uniformly ^{15}N-labelled SH3(C) domain was obtained by growing the cells in a similar medium as that described for protein G. The E. coli cultures were harvested by centrifugation after growth and induction, and the cells were lysed by sonication. Protein purification was performed using a Q-sepharose soft-gel anion-exchange column, followed by a HiLoad Superdex-75 gel filtration FLPC column. The SH3(C) sample was made up in 50 mM NaH_2PO_4 buffer with 11% D_2O at a concentration of 0.46 mM, calculated from the UV extinction coefficient, and the pH was adjusted to 6.0.

p22phox peptide was made by solid phase synthesis (by the Protein and Nucleic Acid Laboratory, Leicester University). p22phox peptide solutions were made up in 20 mM NaH_2PO_4 buffer with 11% D_2O and the pH was adjusted to 6.0, the concentration was calculated from the mass of solid p22phox peptide used.

All samples were placed in 5 mm susceptibility-matched NMR microtubes (Shigemi, Japan) and the sample height was adjusted to extend beyond the region excited by the slice selection procedure (see section NMR experiments).

Simulated diffusion data

Data was simulated for both ‘small’ and ‘medium’ sized proteins, with molecular weights of 8 kDa and 14 kDa respectively, at 500 MHz and a gradient strength of 0.5 Tm^{-1}. Relaxation rates for eglin c were used as a model for the 8 kDa protein, $R(2H_zN_z) = 6.45\ s^{-1}$, $R(H_z^N) = 4.62\ s^{-1}$, $R(H_x^N) = 10.0\ s^{-1}$ and $R(2H_xN_x) = 25.0\ s^{-1}$ (Peng et al., 1995). The dif-

fusion coefficient for an 8 kDa protein was assumed to be 1.3×10^{-10} m^2 s^{-1}, which is the diffusion coefficient of protein G at 298 K. Relaxation rates for Villin 14T were used as the model for a 14 kDa protein, $R(2H_zN_z) = 11.4$ s^{-1}, $R(H_z^N) = 7.75$ s^{-1} (Markus et al., 1996) and $R(H_x^N) = 46.2$ s^{-1} (Dayie et al., 1994). $R(2H_xN_x)$ was estimated to be 92.4 s^{-1}. The diffusion coefficient for a 14 kDa protein was assumed to be 1.1×10^{-10} m^2 s^{-1}, which is the diffusion coefficient for lysozyme at 298 K (Ilyina et al., 1997). Simulation parameters were calculated to give 90% attenuation of the initial intensity over a gradient value range of 0–0.5 Tm^{-1} while achieving maximum signal-to-noise. A delay of 400 μs to allow eddy current to die away was included after each gradient pulse. The parameters calculated for the pulse sequences in Figure 1A–E were as follows: Figure 1A, $\Delta = 32.4$ ms and $\delta = 9$ ms for 8 kDa; $\Delta = 53.2$ ms and $\delta = 7.5$ ms for 14 kDa. Figure 1B, $\Delta = 36.3$ ms and $\delta = 8.5$ ms for 8 kDa; $\Delta = 53.6$ ms and $\delta = 7.5$ ms for 14 kDa, with the ratio of $G{:}G' = \gamma_H{:}\gamma_H{-}\gamma_N$. Figure 1C, $\Delta = 69.4$ ms and $\delta = 6$ ms for 8 kDa; $\Delta = 81.7$ ms and $\delta = 6$ ms for 14 kDa. Figure 1D, $\Delta = 27.5$ ms, $\delta = 7.5$ ms and $\delta' = 2.5$ ms for 8 kDa; $\Delta = 38.4$ ms, $\delta = 7$ ms and $\delta' = 2$ ms for 14 kDa. Figure 1E, $\Delta = 27.7$ ms and $\delta_1 = \delta_2 = 5$ ms for 8 kDa; $\Delta = 32.2$ ms and $\delta_1 = \delta_2 = 5$ ms for 14 kDa.

NMR experiments

NMR data were acquired on a Bruker DMX 500 MHz spectrometer operating at 500.1 MHz for ^1H, using a 5 mm triple resonance ($^{15}N/^{13}C/^1H$) probe equipped with a single axis actively shielded gradient. The linearity of the gradient was tested and a region of the sample over which the gradient was linear was selected using slice selection, as described previously (Tillett et al., 1998). The gradient strength was found to be 0.517 Tm^{-1} by calibration using a standard sample of HDO at 298 K (Mills, 1965).

All experiments presented in Figure 3 were conducted at 298 K and 128 transients were averaged for each FID. The length of the gradient pulse was kept constant and the amplitude was incremented in each case. For the sequences in Figure 1A, C–E the gradient was stepped from 10% to 100% of the maximum gradient strength, in step sizes of 10%. For the sequence in Figure 1B, where $G{:}G' = \gamma_H{:}\gamma_H{-}\gamma_N$, G' was stepped from 10% to 100% of the maximum gradient strength, in step sizes of 10%. Other parameters for the pulse sequences in Figure 1 were as follows: Figure 1A, $\Delta = 23.4$ ms and $\delta = 10.5$ ms. Fig-

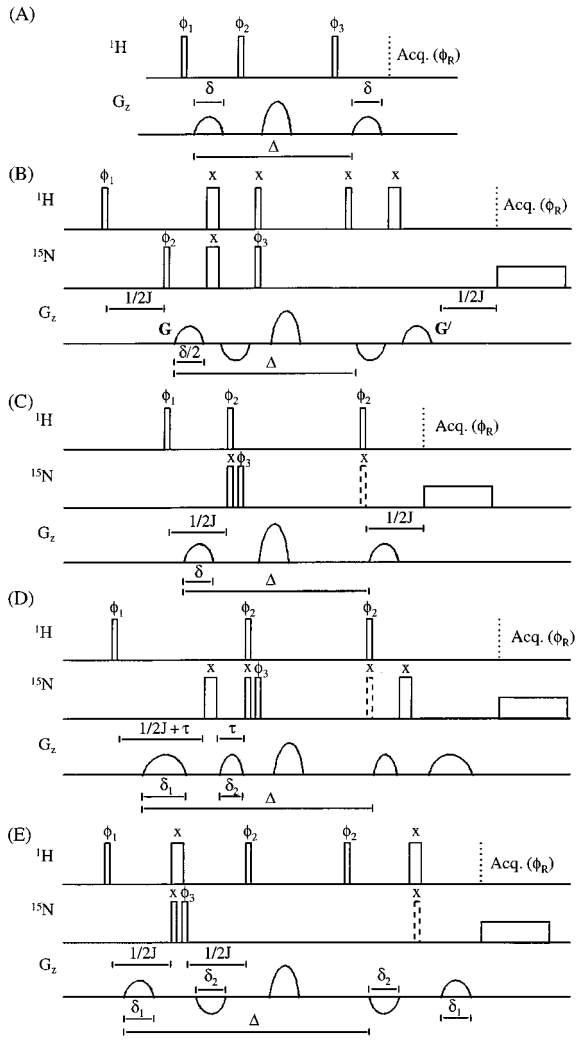

Figure 1. Five pulse sequences for measuring diffusion coefficients: (A) stimulated-echo PGSE, (B)–(E) incorporating heteronuclear filtration. Sequences (C)–(E) can be used to edit signals of either unlabelled or labelled components. Thin rectangular bars represent π/2 pulses, thick rectangular bars represent π pulses and dashed rectangular bars represent π/2 pulses to be included when editing out ^{15}N-labelled components. The small sinusoidal gradients indicate pulses that are for diffusion encoding and the large sinusoidal gradient pulses indicate 'crusher' gradients to dephase unwanted magnetisation. Phase cycling for (A): $\phi_1 = 8(x), 8(-x)$; $\phi_2 = x, y, -x, -y$; $\phi_3 = x, y, -x, -y, -x, -y, x, y$; $\phi_R = 2(x, -x), 4(-x, x), 2(x, -x)$. Phase cycling for (B) (Dingley et al., 1997): $\phi_1 = 4(x), 4(-x)$; $\phi_2 = x, -x$; $\phi_3 = 2(x), 2(-x)$; $\phi_R = x, 2(-x), x, -x, 2(x), -x$. The ratio of gradient amplitudes $G{:}G'$ is set to select the desired multiple-quantum coherence. Phase cycling for (C), (D) and (E): $\phi_1 = 2(x), 2(-x)$; $\phi_2 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_3 = x, -x$; for selection of ^{15}N-labelled components $\phi_R = x, 2(-x), x, -x, 2(x), -x.$; for editing out ^{15}N-labelled components $\phi_R = 2(x), 4(-x), 2(x)$.

ure 1B, $\Delta = 23.76$ ms and $\delta = 9$ ms. Figure 1C, $\Delta = 150$ ms and $\delta = 4$ ms. Figure 1D, $\Delta = 20.84$ ms, $\delta_1 = 8$ ms and $\delta_2 = 3.5$ ms. Figure 1E, $\Delta = 38.01$ ms, $\delta_1 = 3.5$ ms and $\delta_2 = 4.4$ ms.

Data presented in Figure 4 were collected at 285 K using the pulse sequence in Figure 1D, phase cycling was used to edit either ^{15}N-labelled or unlabelled components (given in Figure 1). For SH3(C), 20 spectra were obtained by incrementing the gradient amplitude in steps of 5% from 5% to 100% of the maximum gradient strength, $\Delta = 26.22$ ms, $\delta_1 = 8$ ms, $\delta_2 = 3.5$ ms and 32 or 64 scans. For 'bound' p22phox peptide, $\Delta = 20.82$ ms, $\delta_1 = 8$ ms, $\delta_2 = 3.5$ ms and 256 scans for the lowest concentration and $\Delta = 18.82$ ms, $\delta_1 = 7$ ms, $\delta_2 = 2.5$ ms and 32–128 scans for the remaining concentrations. For free p22phox peptide, $\Delta = 18.82$ ms, $\delta_1 = 7$ ms, $\delta_2 = 2.5$ ms and 32–256 scans for all the concentrations. For all p22phox peptide measurements 10 spectra were obtained as described for Figure 3.

All FIDs consisted of 4096 points, the total relaxation delay was 3.5 s including presaturation for 1.5 s, all shaped rf pulses were Gaussian pulses 2 ms in duration and nitrogen decoupling during acquisition was achieved using GARP-1 (Shaka et al., 1985).

Results and discussion

Four pulse sequences designed to produce heteronuclear-filtered diffusion measurements are given in Figure 1B–E. The stimulated echo PGSE pulse sequence is given in Figure 1A for comparison, and a previously proposed heteronuclear sequence in Figure 1B (Dingley et al., 1997). We will firstly discuss critically the experiment proposed by Dingley et al.; methods for improving the sensitivity of heteronuclear-filtration diffusion experiment will then be proposed.

The experiment proposed by Dingley et al. (Figure 1B) filters the magnetization through heteronuclear multiple-quantum coherence, zero-quantum coherence for ^{1}H-^{15}N and double-quantum coherence for ^{1}H-^{13}C. This sequence also applies the dephasing gradient to the heteronuclear multiple-quantum coherence in an attempt to maximise sensitivity since the multiple-quantum coherence present during the dephasing period in each case is more sensitive to the magnetic field gradient than ^{1}H single-quantum coherence. However, more is lost than is gained since the structure of the experiment results in the inevitable loss of 50% more signal than in a conventional stimulated echo PGSE experiment: only zero or double-quantum coherence is selected, thus sacrificing half of the signal. In addition, signal is also lost due to transverse relaxation during the two $1/2J_{IS}$ periods. These periods could have been used to encode diffusion, thus reducing the lengths of the remaining transverse evolution periods, but for reasons which are unclear, were not. The intensity of the observed signal at the start of acquisition, *I(t)*, for this experiment calculated according to the method given in (Stejskal et al., 1965) is given by:

$$I(t) = (\frac{1}{4})I(0)\exp -\{R(t) + \frac{(\gamma_I + \gamma_S)^2}{\gamma_S} D\frac{G^2}{6}\delta^2(6\Delta\gamma_I + 2\delta\gamma_S - 2\delta\gamma_I - 3t'\gamma_I)\} \quad (1)$$

where t′ is the time between each pair of gradient pulses and γ_I and γ_S are the gyromagnetic ratios of the I-spin and the S-spin, respectively. *R(t)* is given by:

$$R(t) = \frac{R(I_x)}{J_{IS}} + \delta R(I_x) + \delta R(2I_xS_x) + (\Delta - \delta)R(2I_zS_z) \quad (2)$$

where $R(I_x)$ and $R(I_xS_x)$ are the relaxation rates of the I-spin single-quantum and IS multiple-quantum coherences, respectively, and $R(I_z)$ and $R(I_zS_z)$ are the relaxation rates of the I-spin longitudinal magnetization and the IS longitudinal two spin order. It should be noted that the expression for diffusion attenuation in the experiment given by Dingley et al. (1997) neglects the difference in length of the dephasing and rephasing gradient pulses and ignores the delay between each pair of gradient pulses. If sinusoidally shaped gradient pulses are used, as is more usually the case, this expression becomes:

$$I(t) = (\frac{1}{4})I(0)\exp -\{R(t) + D\frac{G^2\delta^2}{8\gamma_I\pi^2} [32\Delta\gamma_I(\gamma_I + \gamma_S)^2 - 16t'\gamma_I(\gamma_I + \gamma_S)^2 + \delta(11\gamma_S^3 + 12\gamma_I\gamma_S^2 - 9\gamma_I^2\gamma_S - 10\gamma_I^3)]\} \quad (3)$$

where *R(t)* is defined as before (Equation 2).

A more efficient way to implement heteronuclear filtration is to filter the magnetization through both heteronuclear zero- and double-quantum coherence and to use the $1/2J_{IS}$ periods to encode diffusion, shown in Figure 1C. This pulse sequence has an efficiency comparable to homonuclear stimulated echo-based diffusion experiments. Assuming sinusoidal

gradient pulses, the observed signal intensity is given by:

$$I(t) = (\frac{1}{2}) I(0) \exp -\{R(t) + DG^2\gamma_I^2 \left(\frac{2\delta}{\pi}\right)^2 (\Delta - \frac{\delta}{4})\} \quad (4)$$

where *R(t)* is given by:

$$R(t) = \frac{R(I_x)}{J_{IS}} + (\Delta - \delta) R(2I_z S_z) \quad (5)$$

The major disadvantage of this sequence over the previous one is that the lengths of the magnetic field gradient pulses are restricted to $1/2J_{IS}$ which will typically be 5.4 ms for ^{1}H-^{15}N and 3.6 ms for ^{1}H-^{13}C (Bax et al., 1993). A variation of this pulse sequence that allows diffusion encoding gradient pulses of arbitrary length to be used is given in Figure 1D. The two additional 180°(S) pulses serve to restrict heteronuclear scalar coupling evolution to a time of $1/2J_{IS}$ in each case. The observed signal intensity for this experiment, again assuming sinusoidal gradient pulses, is given by:

$$I(t) = \left(\frac{1}{2}\right) I(0) \exp -\{R(t) + D\gamma_I^2 \left(\frac{G}{\pi}\right)^2 [4\Delta(\delta_1 + \delta_2)^2 - \delta_1^3 - \delta_1^2(4\delta_2 + t') - \delta_1(4\delta_2^2 + 8\delta_2 t') - \delta_2^3]\} \quad (6)$$

where *R(t)* is given by:

$$R(t) = \frac{R(I_x)}{J_{IS}} + 2\delta_2 R(I_x) + (\Delta - \delta_1 - \delta_2) R(2I_z S_z) \quad (7)$$

The above pulse sequences (Figure 1B–D) all store the magnetization as heteronuclear two-spin order between the dephasing and rephasing periods. An alternative pulse sequence that utilises ^{1}H longitudinal magnetization during this period is given in Figure 1E. This may be the preferred experiment in some instances since usually $R(H_z) < R(2H_zN_z)$, $R(2H_zC_z)$. Its main disadvantage is that the minimum time spent in the transverse plane by the magnetization is increased to $(1/J_{IS} + \delta)$, although this is still shorter than for Figure 1B. With sinusoidal gradients, the observed signal intensity for this experiment, where $\delta_1 = \delta_2$, is given by:

$$I(t) = \left(\frac{1}{2}\right) I(0) \exp -\{R(t) + D\left(\frac{G\gamma_I\delta}{2\pi}\right)^2 (16\Delta - 5\delta - 8t')\} \quad (8)$$

where $\delta = \delta_1 + \delta_2$ and *R(t)* is given by:

$$R(t) = \frac{R(I_x)}{J_{IS}} + \delta R(I_x) + (\Delta - \delta) R(I_z) \quad (9)$$

where $\delta_1 \neq \delta_2$, the signal intensity is given by Equation 6. By using the alternative phase cycles given in the figure caption and incorporating the pulses indicated by dashed lines, these pulse sequences (Figure 1C–E) can also be used to edit out the signal from labelled molecules from the measured spectrum. We also note that while the magnetization is stored along the z-axis it is possible to suppress signal loss due to cross-relaxation by utilising the appropriate pulse sequence (Dingley et al., 1997).

To compare the efficiencies of these pulse sequences, diffusion data was simulated. The range of diffusion attenuations to be spanned in each case was fixed at 90%, and for a constant range of gradient strengths (0–0.5 Tm^{-1}) the values of the delays that gave the maximum signal-to-noise ratio in each case were determined. The X nucleus was taken as ^{15}N and the 1/2J period was allowed to exceed 5.4 ms if the loss in signal because of incomplete coherence transfer was more than compensated for by an increase in the overall signal-to-noise ratio. Homonuclear coupling between the amide proton and the alpha proton was also included, and was set at 9 Hz, which is the average value for a parallel beta-sheet (Pardi et al., 1984) and represents the worst case. For the sequence in Figure 1B, the ratio of G:G′ was set to γ_H:γ_H–γ_N to select zero-quantum coherence with the maximum value for G′ set to 100%, and the corresponding value of G was calculated to be 82.44%. Optimum conditions were calculated for both ‘small’ and ‘medium’ sized proteins, of molecular weights 8 and 14 kDa with correlation times of 4.15 and 10.5 ns respectively, using the conditions outlined above. The results are shown in Figure 2A–B.

Figure 2A shows the optimised decay curves for a ‘small’ protein. The decay curves for the pulse sequences in Figure 1A, D and E have very similar intensities. The decay curve for the pulse sequence in Figure 1C has a lower intensity than these previous three and the decay curve for the pulse sequence in Figure 1B has a lower intensity still. There are

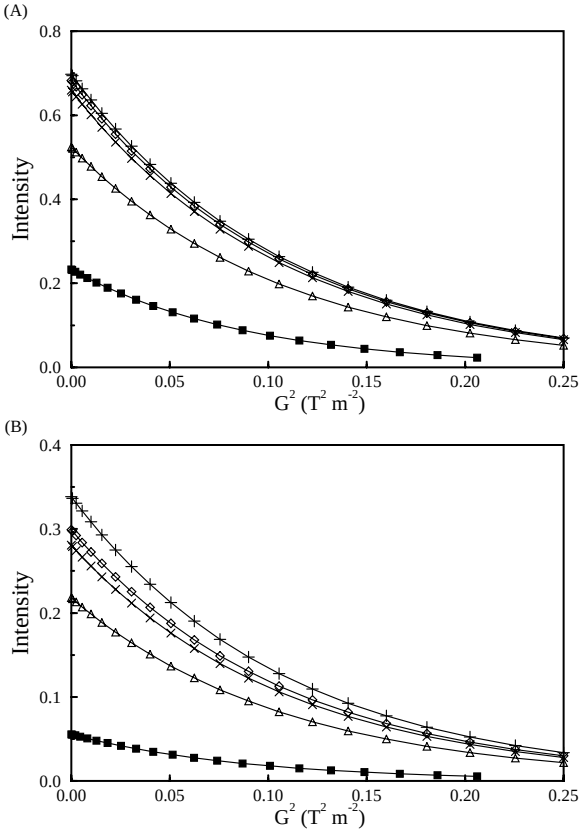

Figure 2. Simulated diffusion data for 'small' (A) and 'medium' (B) sized proteins. Data for the pulse sequences, in Figure 1, are denoted by the following symbols: Figure 1A (+), 1B (■), 1C (Δ), 1D (X), 1E (◊).

three factors that contribute to these differences. First, the observed coherence transfer pathway(s) are determined by the structure of the pulse sequence; for example, selecting heteronuclear zero-quantum coherence in a pulse sequence where both zero- and double-quantum coherence are created in equal amounts will result in an inevitable loss of 50% of the signal. Second, the pulse sequences have different relaxation dependencies; for example, $R(2H_zN_z)$ decays more rapidly than $R(H_z^N)$ which means that experiments that store magnetisation as $2H_zN_z$ during much of Δ will lose more signal than those which store it as H_z^N. Last, the sensitivity of a pulse sequence to diffusion depends on the coherences present during the diffusion-encoding gradient pulses; for example, homonuclear double-quantum coherence is twice as sensitive to a magnetic field gradient as the corresponding single-quantum coherence.

The pulse sequences in Figure 1A, D and E have different structures but for a 'small' protein the optimum delays for each are such that the times spent in the xy-plane as ^{1}H single-quantum coherence, which is the dominant source of non-diffusion-related attenuation, are similar. Consequently their decay curves are of similar intensity. The decay curve for the sequence in Figure 1C is of lower intensity than the previous three because of the structure of the sequence. The period during which the diffusion-encoding gradient can be applied has an optimal length of 1/2J from the point of view of coherence transfer efficiency, which is too short to allow the optimum gradient length to be used. This problem is overcome by the more flexible version of the pulse sequence presented in Figure 1D. The sequence in Figure 1B, proposed by Dingley et al., suffers far more severely than the other heteronuclear-filtered experiments from loss of signal, by a factor of 2.9 when compared to the best heteronuclear-filtered experiment (Figure 1E). There are two reasons for this loss of signal. First, the sequence in Figure 1B selects heteronuclear zero-quantum coherence from an equal mixture of both zero- and double-quantum coherences which were generated, resulting in a loss of 50% of the signal. Second, the time spent in the xy-plane is longer than in any other experiment, resulting in greater transverse relaxation losses. The latter is due to two 1/2J periods which are not used for encoding diffusion; these could have been incorporated into the diffusion-encoding periods, thus shortening their length. While using heteronuclear zero-quantum coherence (or double-quantum coherence for ^{1}H-^{13}C) to encode diffusion does result in an enhancement of sensitivity to the magnetic field gradient of $[(\gamma_H - \gamma_N)/\gamma_H] = 1.1$, this gain is more than offset by the sources of signal loss noted above.

The decay curves for a 'medium' sized protein are shown in Figure 2B. The three pulse sequences given in Figure 1A, D and E no longer produce similar intensities, though the same trend as before is still observed for the pulse sequences in Figure 1B and C. The simulations for the 'medium' sized protein generally give less than half the signal obtained for the 'small' protein. This is due to the increased relaxation rate of ^{1}H transverse magnetisation from 10.0 s^{-1} to 46.2 s^{-1} and the longer diffusion-encoding intervals required to obtain the same range of diffusion attenuations (since the protein diffuses more slowly). All the heteronuclear-filtered experiments have a lower signal intensity than stimulated-echo PGSE (Figure 1A). For the sequence given in Figure 1E, this is because the initial transverse evolution period is fixed at 1/J, which is longer than is required for diffusion encoding alone.

The sequence in Figure 1D is the next most efficient experiment, but loses out because magnetization is stored as $2H_zN_z$ instead of H_z^N (which relaxes more slowly) during much of Δ. The pulse sequence in Figure 1C suffers again from the same problems described for a 'small' protein. The problems associated with the sequence in Figure 1B have been discussed above, though for a 'medium' sized protein the loss of signal due to transverse relaxation is far worse, resulting in a factor of 5.4 reduction in signal intensity when compared to the best heteronuclear-filtered experiment (Figure 1E).

For larger proteins, transverse relaxation will clearly have a much more deleterious effect on the signal intensity for the pulse sequence given in Figure 1B than for the most efficient heteronuclear-filtered experiments, largely because of the time the magnetisation has to spend in the xy-plane. This is contrary to the expectations raised by Dingley et al. (1997) who suggest that their experiment is more likely to be useful for studying larger proteins.

For a homonuclear coupling of 4 Hz, which is the average value for an α-helix (Pardi et al., 1984), the signal intensity for all the sequences would increase by between 3–14%. The pulse sequence given in Figure 1B is the one most affected by signal intensity loss due to homonuclear scalar coupling. However, this has only a slight effect on the relative efficiency of the sequence in Figure 1B; when compared to the best heteronuclear-filtered experiment (Figure 1E) a factor of 2.7 and 5.1 reduction in signal intensity is observed for the 'small' and 'medium' sized proteins, respectively.

Four of the pulse sequences given in Figure 1 were tested experimentally on the amide region of the protein G spectrum, Figure 3. In each case, the first proton pulse was made slice-selective. This was done to ensure that the experiments were performed on a region of the sample over which the magnetic field gradient is linear; this was necessary because, as previously shown (Tillett et al., 1998), not all commercial gradients are linear and gradient non-linearity can cause distortions in the resulting data. The subsequent proton 90° pulses were made frequency selective to alleviate any exchange effects. This precaution is not necessary if only non-exchanging protons are measured. Water suppression was afforded by presaturation of the water resonance, though this could also have been achieved by the use of WATERGATE (Piotto et al., 1992). The pulse sequences were optimised with the criteria outlined above for the simulations. The intensities

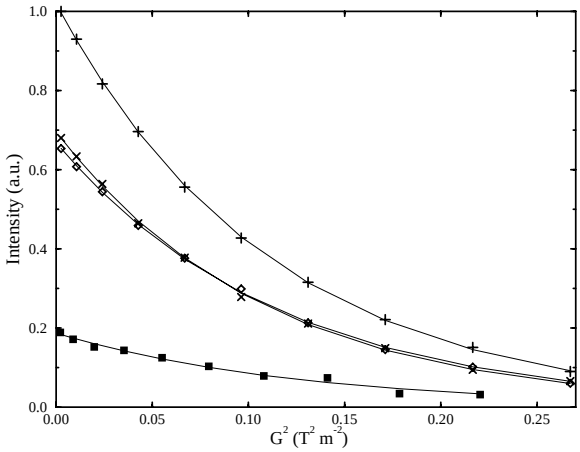

Figure 3. Experimental diffusion data for protein G obtained with the pulse sequences in Figure 1A, B, D and E. The symbols and the diffusion coefficients for each pulse sequence are: Figure 1A (+) $1.35 \pm 0.01 \times 10^{-10}$ m^2 s^{-1}, 1B (■)$1.30 \pm 0.08 \times 10^{-10}$ m^2 s^{-1}, 1D (X) $1.38 \pm 0.02 \times 10^{-10}$ m^2 s^{-1}, 1E (◊)$1.36 \pm 0.01 \times 10^{-10}$ m^{2}s^{-1}. The noise level was used to estimate a standard deviation of 0.02 for all data points.

have been normalized to the initial intensity of the stimulated-echo PGSE pulse sequence (Figure 1A). The trend in the experimental decay curves for protein G is similar to that observed in the simulations given in Figure 2. The decay curve for the pulse sequence in Figure 1C is not shown since it is superseded by the improved version shown in Figure 1D. In this example, the enhancement in signal intensity achieved by use of the sequence in Figure 1D over the sequence in Figure 1B is 3.6-fold. The only substantial deviation from the trend predicted by the simulated diffusion data is the decrease in signal of all the heteronuclear filter sequences when compared to the stimulated-echo PGSE (Figure 1A). For all the pulse sequences, the signal reduction is most likely due to a mis-estimation of the relaxation rates used to calculate delay values when the sequences were optimised and the non-ideal behaviour of the pulses. For the sequence in Figure 1E, a further contributing factor was the time needed to implement the slice-selection procedure in the first 1/2J period, which reduced the time available for the diffusion-encoding gradients. If a linear magnetic field gradient had been available, slice selection would not have been necessary.

One of the potentially important uses of NMR diffusion measurements is in the area of protein-ligand binding. In order to study the binding of a ligand to a protein it is necessary to measure a property of the ligand or the protein which changes as a function of

the ratio of their free and bound forms. For diffusion studies this can be conveniently achieved by measuring the diffusion coefficient of the ligand as a function of its concentration while keeping the protein concentration constant. To measure the diffusion coefficient of the ligand in a solution that contains both protein and ligand, there are four possible options. (1) Measure a resonance of the ligand which does not overlap with those of the protein. (2) Use a multiexponential analysis to resolve the decay curves of the overlapping protein and ligand resonances as in the DOSY experiment (Morris et al., 1992; Chen et al., 1995). (3) Use relaxation or diffusion filtering in order to remove signals of the protein from the spectrum and leave only signals from the ligand (Hajduk et al., 1997; Ponstingl et al., 1997). (4) Use isotopic labelling of either the protein or the ligand so that heteronuclear filtration can be used to separate the two (Dalvit et al., 1998). Option (1) may be possible in a limited number of cases such as the binding of DNA or RNA to a protein, but in most cases where the ligand is a peptide or other small molecule it is not. Option (2), in common with most procedures utilising multiexponential fitting, is susceptible to relatively large errors. Option (3) has already been demonstrated for a protein and small organic molecules (Hajduk et al., 1997; Ponstingl et al., 1997), but relaxation or diffusion filtering relies on the sizes of the two molecules, and thus their relaxation properties or diffusion coefficients, being substantially different. This is not the case for many protein-peptide and protein-DNA/RNA interactions and may also result in an unnecessary loss of signal. Option (4) involves the moderately expensive procedure of isotopic labelling; however, an isotopically labelled protein is often available from its use in structure determination. By use of a heteronuclear-filtered pulse sequence, such as that in Figure 1D with inclusion of the dashed 90° pulse and the alternative phase cycle given in the figure caption, ligand signals can be separated from those of an isotopically labelled protein by editing out the latter. This is accomplished without the loss of signal associated with option (3) and no restrictions on the relative sizes of the protein and the ligand are imposed. With the sequences in Figure 1C–E, measurement of the diffusion coefficient of both the labelled protein and the unlabelled ligand can be made by using the appropriate phase cycles. This is necessary if quantitative results rather than purely qualitative observations are to be made.

The binding of the SH3(C) protein to a p22phox peptide has been studied by measuring the diffusion coefficient of both the protein and the peptide as a function of peptide concentration. The results are shown in Figure 4. The peptide in the solution containing both protein and peptide is referred to as 'bound' peptide. The diffusion coefficient of the free p22phox peptide was measured in the same range of concentrations, as also shown in Figure 4. The concentration of SH3(C) was fixed at 0.46 mM and the peptide was added to make the total concentration of peptide between 0.23 mM and 10.5 mM, i.e., in a protein:peptide ratio ranging from 1:0.5 to 1:22.8. In order to measure the diffusion coefficient of two components of this system, SH3(C) was isotopically labelled with ^{15}N and p22phox peptide was unlabeled. The diffusion coefficient of the 'bound' peptide at 0.23 mM is approximately equal to that of the protein, and at 10.5 mM it is close to that of the free peptide at the same concentration. The diffusion coefficients of both the protein and the 'bound' peptide for 3.7 and 7.5 mM deviate noticeably from the fitted lines. This is attributed to the fact that these two solutions were also found to have different pH values from the samples at other concentrations. It is known that the diffusion coefficient is dependent on the pH due to the change in the charged state of the ionisable side chains of the protein that a change in pH results in (Cussler, 1997; Pan et al., 1997). Alternatively this deviation may be due to aggregation of SH3(C) which could be pH dependent. These effects are currently under investigation.

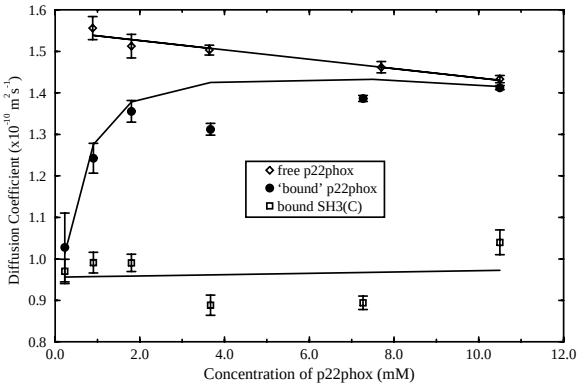

Figure 4. Plot of diffusion coefficient versus concentration of p22phox peptide, for p22phox peptide and SH3(C) in a solution containing both and a solution of p22phox peptide. Data for 'bound' p22phox peptide was fitted with Equation 12, free p22phox peptide and bound SH3(C) data was fitted linearly.

An expression for the diffusion coefficient of the 'bound' peptide (D_{obs}) is given by Equation 10 (Johnson, 1993). D_{bound} is the diffusion coefficient of the

protein, D_{free} is the diffusion coefficient of the free peptide obtained from a solution of peptide without protein and **a** is the fraction of peptide bound to the protein.

$$D_{obs} = a D_{bound} + (1 - a) D_{\text{free}} \quad (10)$$

An expression for K_d is given in Equation 11, where the subscript '**tot**' refers to the total concentration of protein or peptide.

$$K_d = \frac{([Pro]_{tot} - [Complex])([Pep]_{tot} - [Complex])}{[Complex]} \quad (11)$$

By analogy with the equation used in fluorescence spectrophotometry to measure protein-ligand binding (Nomanbhoy et al., 1996), rearranging Equations 10 and 11 gives an expression from which K_d can be calculated:

$$\frac{D_{free} - D_{obs}}{D_{free} - D_{bound}} = \frac{([Pro]_{tot} + [Pep]_{tot} + K_d)}{2[Pep]_{tot}} - \frac{\sqrt{([Pro]_{tot} + [Pep]_{tot} + K_d)^2 - 4[Pro]_{tot}[Pep]_{tot}}}{2[Pep]_{tot}} \quad (12)$$

Equation 12 can also be used to take into account changes in viscosity, if D_{bound} and D_{free} are known for each peptide concentration. Equations for measuring the binding constant of peptide dimerisation and association have also been proposed (Mansfield et al., 1998; Orfi et al., 1998).

Equation 12 was used to calculate K_d for SH3(C) and p22phox peptide. If the two deviant points are omitted, K_d is calculated to be 21 ± 14 μM; this value was used to calculated the fitted line in Figure 4. This is comparable to a value of 29 ± 3 μM obtained from fluorescence spectrophotometry (Galbraith, 1998). It can be seen from these results that diffusion measurements offer the opportunity to make quantitative measurements of K_d for most protein-ligand systems where one of the components is available in an isotopically labelled form. However, the K_d calculated from the diffusion data was found to rely heavily on the early data points.

Conclusions

New heteronuclear-filtered experiments have been proposed which show superior sensitivity to those previously reported (Dingley et al., 1997). This has been demonstrated for both simulated diffusion data and diffusion data obtained experimentally for protein G. This relative improvement in sensitivity is likely to increase for larger proteins, since the time in the transverse plane is kept to a minimum in these new experiments. Using one of these new heteronuclear-filtered experiments, the binding of SH3(C) to a p22phox peptide was studied. By the use of two alternate phase cycling schemes it is possible to measure the diffusion coefficients of both the isotopically labelled and unlabelled components of the same solution. This leads to the possibility of measuring the binding constant, K_d, for many protein-ligand systems where one component is isotopically labelled.

Acknowledgements

We thank Vicky Galbraith for production of the ^{15}N-labelled SH3(C) and helpful discussions of the work produced for her Ph.D Thesis. One of us (M.L.T.) thanks the MRC for a Research Studentship. We also acknowledge the Wellcome Trust and the U.K. Biotechnology and Biological Sciences Research Council (BBSRC) for financial support.

References

Altieri, A.S., Hinton, D.P. and Byrd, R.A. (1995) *J. Am. Chem. Soc.*, **117**, 7566–7567.

Baranowska, H.M. and Olszewski, K.J. (1996) *Biochim. Biophys. Acta*, **1289**, 312–314.

Bax, A. and Grzesiek, S. (1993) *Acc. Chem. Res.*, **26**, 131–138.

Chen, A.D., Wu, D.H. and Johnson, C.S. (1995) *J. Phys. Chem.*, **99**, 828–834.

Cussler, E.L. (1997) *Diffusion: Mass Transfer in Fluid Systems*, Cambridge University Press, Cambridge.

Dalvit, C., Ramage, P. and Hommel, U. (1998) *J. Magn. Reson.*, **131**, 148–153.

Dayie, K.T. and Wagner, G. (1994) *J. Magn. Reson.*, **A111**, 121–126.

Dingley, A.J., Mackay, J.P., Chapman, B.E., Morris, M.B., Kuchel, P.W., Hambly, B.D. and King, G.F. (1995) *J. Biomol. NMR*, **6**, 321–328.

Dingley, A.J., Mackay, J.P., Shaw, G.L., Hambly, B.D. and King, G.F. (1997) *J. Biomol. NMR*, **10**, 1–8.

Galbraith, V.A. (1999) Ph. D. Thesis (in preparation), Department of Biochemistry, University of Leicester, U.K.

Gibbs, S.J. and Johnson, C.S. (1991) *J. Magn. Reson.*, **93**, 395–402.

Hajduk, P.J., Olejniczak, E.T. and Fesik, S.W. (1997) *J. Am. Chem. Soc.*, **119**, 12257–12261.

Ilyina, E., Roongta, V., Pan, H., Woodward, C. and Mayo, K.H. (1997) *Biochemistry*, **36**, 3383–3388.

Jerschow, A. and Muller, N. (1997) *J. Magn. Reson.*, **125**, 372–375.

Johnson, C.S. (1993) *J. Magn. Reson.*, **A102**, 214–218.

Jones, J.A., Wilkins, D.K., Smith, L.J. and Dobson, C.M. (1997) *J. Biomol. NMR*, **10**, 199–203.
Krishnan, V.V. (1997) *J. Magn. Reson.*, **124**, 468–473.
Lian, L.Y., Yang, J.C., Derrick, J.P., Sutcliffe, M.J., Roberts, G.C.K., Murphy, J.P., Goward, C.R. and Atkinson, T. (1991) *Biochemistry*, **30**, 5335–5340.
Lin, M.F. and Larive, C.K. (1995) *Anal. Biochem.*, **229**, 214–220.
Mansfield, S.L., Jayawickrama, D.A., Timmons, J.S. and Larive, C.K. (1998) *Biochim. Biophys. Acta*, **1382**, 257–265.
Markus, M.A., Dayie, K.T., Matsudaira, P. and Wagner, G. (1996) *Biochemistry*, **35**, 1722–1732.
Mills, R. (1965) *J. Phys. Chem.*, **69**, 3116–3119.
Morris, K.F. and Johnson, C.S. (1992) *J. Am. Chem. Soc.*, **114**, 3139–3141.
Nomanbhoy, T.K. and Cerione, R.A. (1996) *J. Biol. Chem.*, **271**, 10004–10009.
Norwood, T.J. (1993) *J. Magn. Reson.*, **A103**, 258–267.
Orfi, L., Lin, M.F. and Larive, C.K. (1998) *Anal. Chem.*, **70**, 1339–1345.
Pan, H., Barany, G. and Woodward, G. (1997) *Protein Sci.*, **6**, 1985–1992.
Pardi, A., Billeter, M. and Wüthrich, K. (1984) *J. Mol. Biol.*, **180**, 741–751.
Peng, J.W. and Wagner, G. (1995) *Biochemistry*, **34**, 16733–16752.
Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
Ponstingl, H. and Otting, G. (1997) *J. Biomol. NMR*, **9**, 441–444.
Price, W.S., Nara, M. and Arata, Y. (1997) *Biophys. Chem.*, **65**, 179–187.
Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.
Stejskal, E.O. and Tanner, J.E. (1965) *J. Chem. Phys.*, **42**, 288–292.
Stilbs, P. (1989) *Progr. NMR Spectrosc.*, **19**, 1–45.
Tanner, J.E. (1970) *J. Chem. Phys.*, **52**, 2523–2526.
Tillett, M.L., Lian, L.Y. and Norwood, T.J. (1998) *J. Magn. Reson.*, **133**, 379–384.