Association of Biomolecular Systems *via* Pulsed Field Gradient NMR Self-Diffusion Measurements

Amanda S. Altieri,[†] Denise P. Hinton,[‡,§] and R. Andrew Byrd*,[†]

Macromolecular Structure Laboratory
ABL—Basic Research Program
NCI—Frederick Cancer Research and Development Center
Frederick, Maryland 21702
Department of Chemistry, University of Virginia
Charlottesville, Virginia 22901

Received January 26, 1995

Most biochemical interactions involve translational or rotational diffusion of molecules. Self-diffusion is defined as translational motion reflecting the random motions of a molecule in the absence of a concentration gradient. The coefficient of this motion, D_s, is indicative of hydrodynamic properties and molecular size. Numerous proteins are now recognized to depend on oligomerization for biological activity or receptor binding, such as the dimerization of signal transduction proteins including cytokines[1] and SH2 and SH3 domains.[2] Currently, many of these systems are the target of solution-structure determination by NMR. Because of the difficulty and importance of assigning long-range distances in NMR structure determination studies, it is necessary to know the oligomeric state of a biomolecular system prior to structure calculations using NMR data.[3−6]

Traditionally, association of biomolecules has been determined using equilibrium ultracentrifugation or dynamic light scattering. Although effective, these techniques are often carried out at micromolar concentrations which are 3 orders of magnitude less than the millimolar concentrations required by NMR.[7] This discrepancy can be particularly important for concentration dependent association processes. In contrast, D_s measurements of the NMR sample reports on the size of the molecule under conditions identical to those used for structure determination. In this communication, we introduce a modified NMR technique for measuring the D_s and demonstrate that it can be used to determine the aggregation state of protein systems (for sizes up to ~40 kDa) under conditions identical to those used in structure determination by NMR, using unmodified, commercially available, high-resolution spectrometer hardware.

The pioneering work of Stejskal and Tanner[8] showed that the D_s of molecules in solution can be measured using pulsed field gradient (PFG) NMR methods. In the original experiment (90°−τ(PFG)−180°−τ(PFG)−Acq) the strength of the PFG pulse is increased in successive experiments. The echo amplitude [$A(2\tau)$] is differentially attenuated in each spectrum due to translational diffusion, which is related to D_s by[8]

$$A(2\tau) = A(0)\exp[-(\gamma\delta G)^2(\Delta - \delta/3)D_s] \qquad (1)$$

where γ = ^{1}H gyromagnetic ratio, δ = PFG duration (s), G = gradient strength (G/cm), and Δ = time between PFG pulses (s). Accurate D_s measurements require at least 1 order of magnitude signal attentuation, to accurately fit the data to eq 1. The limit on measurable D_s values, therefore, depends on the area of the gradient pulses (δG) and the time allowed for diffusion (Δ). Larger values of D_s ($\geq 10^{-5}$ cm^2/s) can be reliably measured using this experiment. However, as one attempts to measure D_s for large molecules, such as proteins (where $D_s \sim 10^{-6}$ cm^2/s), the time necessary to observe diffusion, Δ, becomes long relative to the transverse relaxation time ($T_2^* \sim 12-40$ ms (see supporting information)), and the signal decays from spin−spin interactions, independent of diffusion.[9] Additionally, homonuclear J coupling during long τ delays causes peak phase distortion and makes accurate amplitude measurements difficult.[9] Increasing the available PFG strength can reduce the τ times and the amount of signal loss from transverse relaxation; however, it can also increase signal distortions due to residual eddy currents.[10]

The stimulated echo experiment[11] (PFGSTE, 90°−τ(PFG)−90°−T−90°−τ(PFG)−Acq) was designed to avoid T_2 relaxation effects by storing the magnetization along the Z axis during T, so that relaxation depends primarily on T_1, which is usually much longer than T_2 for proteins. Longer diffusion times (Δ) can then be used for larger proteins without significant signal loss from relaxation. However, this sequence does not compensate for residual eddy currents, which are present during acquisition and cause amplitude distortions that complicate accurate measurement of D_s. This problem has been discussed in detail by Gibbs and Johnson,[12] who proposed a longitudinal encode−decode (LED) experiment. In this method, magnetization is stored along the Z axis during T (as in PFGSTE), and an additional Z storage period is included before acquisition, to allow the residual eddy currents to decay. The net effect is the ability to use long diffusion times, such that the slower motion

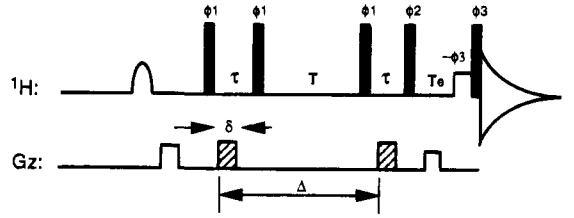

Figure 1. Water-suppressed LED pulse sequence (water-sLED) for diffusion measurements. Solid rectangles are hard 90° pulses; cross-hatched pulses are the PFG encoding and decoding pulses along Z. The low-power ^{1}H pulses are selective pulses on the water resonance. The first sel90°−PFG element dephases the water magnetization,[15] preventing radiation damping during T. The 90°$_{\phi 1}$−PFG−90°$_{\phi 1}$ sequence creates gradient-encoded Z magnetization. Following the T period, a gradient-decoded stimulated echo is created by the 90°$_{\phi 1}$−PFG sequence, which is then stored along the Z axis by the 90°$_{\phi 2}$ pulse. The final PFG−sel90°$_{-\phi 3}$−90°$_{\phi 3}$ sequence eliminates any residual transverse magnetization, maintains water magnetization along the Z axis during acquisition,[16] and reads the stimulated echo in the transverse plane. The pulses were phase cycled as follows: $\phi 1$ = 8(x), 8(−x), 8(y), 8(−y); $\phi 2$ = 4(x), 8(−x), 4(x), 4(y), 8(−y), 4(y); $\phi 3$ = x, y, −x, −y, 2(−x, −y, x, y), x, y, −x, −y, y, −x, −y, x, 2(−y, x, y, −x), y, −x, −y, x; ϕ_{rec} = 2(x, y, −x, −y), 2(−x, −y, x, y), 2(y, −x, −y, x), 2(−y, x, y, −x).

* To whom correspondence should be addressed.
[†] NCI—Frederick Cancer Research and Development Center.
[‡] University of Virginia.
[§] Current address: Department of Radiology, University of Virginia Health Sciences Center, Charlottesville, VA 22908.

(1) Sprang, S. R.; Bazan, J. F. *Curr. Opin. Struct. Biol.* **1993**, *3*, 815−827.
(2) Eck, M. J.; Atwell, S. K.; Shoelson, S. E.; Harrison, S. C. *Nature* **1994**, *368*, 764−769.
(3) Otting, G.; Wuthrich, K. *J. Magn. Reson.* **1989**, *85*, 586−594.
(4) Weiss, M. A. *J. Magn. Reson.* **1990**, *86*, 626−632.
(5) Wider, G.; Weber, C.; Traber, R.; Widmer, H.; Wuthrich, K. *J. Am. Chem. Soc.* **1990**, *112*, 9015−9016.
(6) Folkers, P. J. M.; Folmer, R. H. A.; Konings, R. N. H.; Hilbers, C. W. *J. Am. Chem. Soc.* **1993**, *115*, 3798−3799.
(7) ^{15}N NMR relaxation parameters have also been used to determine rotational correlation times and assess association, however, these experiments require isotopic labeling and a motional model to interpret the data in terms of molecular size. Aggregation state information is often desired prior to isotopic labeling.
(8) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288−292.
(9) Stilbs, P. *Prog. Nucl. Magn. Reson. Spectrosc.* **1987**, *19*, 1−45.
(10) Hrovat, M. I.; Wade, C. G. *J. Magn. Reson.* **1984**, *45*, 67−80.
(11) Tanner, J. E. *J. Chem. Phys.* **1970**, *52*, 2523.
(12) Gibbs, S. J.; Johnson, C. S., Jr. *J. Magn. Reson.* **1991**, *93*, 395−402.

0002-7863/95/1517-7566$09.00/0

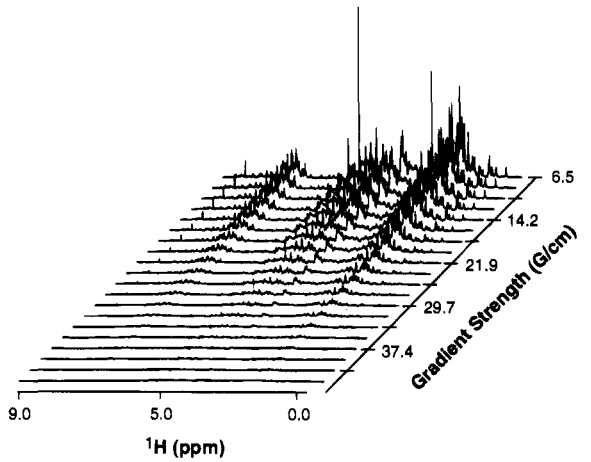

Figure 2. Stack plot of signal attenuation of lysozyme obtained using the water-sLED experiment. The PFG duration (δ) was 7 ms, the diffusion time (Δ) was 132.5 ms, and the PFG strength varied from 6 to 42 G/cm.

of large molecules can be observed with the elimination of eddy current effects.

In order to utilize D_s as a means of assessing the aggregation state of proteins for NMR structural studies, it is necessary that these measurements be made under identical conditions. In terms of hardware, the requirement for short gradient recovery times for high-resolution 3D and 4D experiments has kept the available PFG strength below 40−50 G/cm. Difficulties in measuring D_s for proteins at this PFG strength occur in obtaining sufficient amplitude attenuation as well as in water suppression. Because of the importance of observing exchangeable protons, NMR solution conditions usually require 90% H_2O/10% D_2O as the solvent at 1−2 mM protein concentration, a 10^5-fold excess of water protons. Water suppression is an important element of this technique, since it is essential to achieve the maximum dynamic range to measure protein signal attenuation. Some water suppression is inherent in diffusion measurements of proteins in water.[13,14] However, the difference in diffusion rates does not provide enough suppression to permit operation in 90% H_2O, particularly with the requirement for sufficient signal attenuation within the range of available PFG strength. Consequently, we have modified the LED pulse sequence to incorporate both RAW[15] and flipback[16] sequences for water suppression as shown in Figure 1. The new sequence retains insensitivity to the relaxation and eddy current effects of the LED sequence.[12] In addition, the currently available self-shielded high-resolution PFG probes have sufficiently fast eddy current recovery to eliminate the PFG prepulses and to keep the final eddy current delay short, compared to the original sequence. This modified pulse sequence, water-suppressed LED (water-sLED), provides sufficient water suppression, even at the lower PFG strengths, to observe 100-fold amplitude attenuation in ≈1 mM protein solutions in 90% H_2O.

Using the water-sLED sequence, full amplitude attenuation is acheived for lysozyme (a sample globular protein of 14.5 kDa) as shown in Figure 2. The D_s is obtained by fitting the signal attenuation to eq 1 (after careful calibration of G, see supporting information), which yields a value of 1.08×10^{-6} cm²/s. This result is in good agreement with the light-scattering[17] measurement of 1.06×10^{-6} cm²/s. Using similar conditions, we also observed 100-fold amplitude attenuation for a protein with an aggregate size of ~37 kDa. These results show that accurate measurements of the D_s for biomolecules can be obtained using water-sLED under solution conditions that are identical to those appropriate for NMR structural studies.

An important goal toward using D_s to optimize protein solution conditions is the ability to identify and characterize oligomerization. We have examined two different cytokine systems in order to demonstrate the viability of the water-sLED technique for obtaining this important information. For the simplest case of dimerization, the expected change in D_s may be estimated by approximating the monomer−monomer interaction as a hard-sphere molecular contact. Using this approximation, the expected change in the frictional coefficient, f, upon oligomerization for various geometries of n-mers has been calculated and tabulated by Teller *et al.*[18] The Stokes−Einstein equation, $D = k_BT/f$, is then used to calculate the expected decrease in D_s upon dimerization. The resulting ratio $D_{s,dimer}:D_{s,monomer}$ is 0.75. We tested this model using ubiquitin as a model monomeric protein (76 amino acid residues, 8.6 kDa) of about the same size as the monomeric subunit of the cytokine, monocyte chemotactic protein-1 (MCP-1, 76 amino acid residues, 8.7 kDa). Both proteins are globular and well-behaved in solution. The experimental measurements of D_s are 1.49×10^{-6} cm²/s and 1.08×10^{-6} cm²/s for ubiquitin and MCP-1, respectively. The ratio $D_{s,MCP\text{-}1}:D_{s,ubiquitin} = 0.72$ is in reasonable agreement with the monomer−dimer model described above and confirms the dimeric nature of MCP-1. The solution structure of MCP-1, determined by NMR methods, is in agreement with these findings.[19] Similarly, D_s measurements of the 18.6 kDa cytokine interleukin-10 (IL-10) yield a value of 0.82×10^{-6} cm²/s, which, in comparison to either lysozyme or MCP-1, indicates that IL-10 is also a dimer ($D_{s,IL\text{-}10}:D_{s,lysozyme/MCP\text{-}1} = 0.76$) in solution. The measured value of 0.82×10^{-6} cm²/s is in agreement with proteins of similar aggregate size.[20] The dimerization of IL-10 has been confirmed by Zdanov *et al.*[21] Clearly, unambiguous determination of oligomerization states in solution will contribute to both the procedural aspects of solution structure calculation and to the characterization of functional states of proteins. Additional applications of this method to structure studies in general, and particularly to NMR, include finding optimal conditions[22] for systems that undergo nonspecific aggregation and for studies of protein−protein or protein−nucleic acid complexes.

Acknowledgment. The authors thank Paul J. Dowling, Jr., Donna M. Baldisseri, Tracy M. Handel, Peter J. Domaille, Steven J. Gibbs, Donghui Wu, and Charles S. Johnson, Jr., for helpful discussions. Research sponsored in part by the National Cancer Institute, DHHS, under contract with ABL. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Supporting Information Available: Table of measured relaxation times and diffusion coefficients for four proteins, experimental details, and a more complete description of PFGSE, PFGSTE, and water-sLED pulse sequences (4 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA950259V

(13) Everhart, C. H.; Johnson, C. S., Jr. *J. Magn. Reson.* **1982**, *48*, 466.
(14) van Zijl, P. C. M.; Moonen, C. T. W. *J. Magn. Reson.* **1990**, *87*, 18−25.
(15) Altieri, A. S.; Byrd, R. A. *J. Magn. Reson., Ser. B* **1995**, *107*, 260−266.
(16) Grzesiek, S. Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593−12594.
(17) Dubin, S. B.; Clark, N. A.; Benedek, G. B. *J. Chem. Phys.* **1971**, *54*, 5158−5164.
(18) Teller, D. C.; Swanson, E.; DeHaen, C. *Methods Enzymol.* **1979**, *61*, 103−124.
(19) Handel, T. M.; Domaille, P. J. Private communication.
(20) Kuntz, I. D.; Kauzmann, W. *Adv. Protein Chem.* **1972**, *28*, 236−345.
(21) Zdanov, A.; Schalk-Hihi, C.; Gustchina, A.; Tsang, M.; Weatherbee, J.; Wlodawer, A. *Structure* **1995**, *3*, 591−601.
(22) Wagner, G. *J. Biomol. NMR* **1993**, *3*, 375−385.