Supporting Information

Measuring Protein Concentration by Diffusion-Filtered Quantitative Nuclear Magnetic Resonance Spectroscopy

Scott A. Bradley, Wesley C. Jackson, Jr. and Patrick P. Mahoney

Eli Lilly and Company, Indianapolis, IN 46285

Contents

List of Figures

Measuring Diffusion

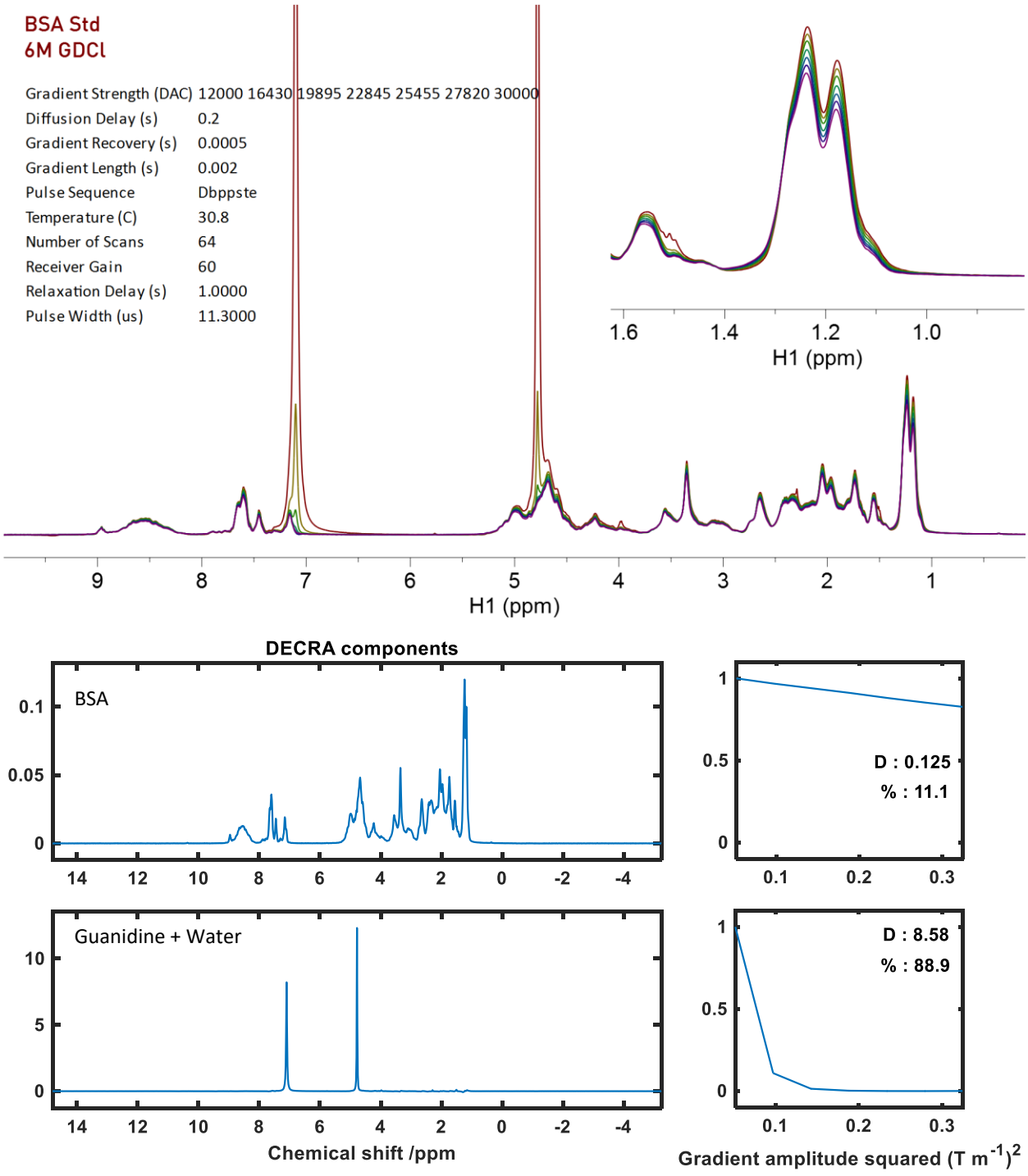


Figure S1. Representative data from the diffusion coefficient measurement. Top: NMR spectra of the denature BSA standard acquired with the bppste pulse sequence. Inset: expansion of the peak for the I/L/V resonances. Bottom Left: the pure NMR spectra for the different diffusing species as determined by the DECRA algorithm. Bottom Right: Stejkal-Tanner plots and diffusion coefficients from the DECRA algorithm.

Measuring T_1 and T_2 with the Agilent Dbppste Pulse Sequence

The molar concentration of the protein in the NMR tube, c_P^{NMR}, can be calculated from the known molar concentration of the external reference standard, c_{RS}, using PULCON

$$c_P^{NMR} = c_{RS} \frac{A_P\, T_P\, pw_P\, S_{RS}\, H_{RS}}{A_{RS}\, T_{RS}\, pw_{RS}\, S_P\, H_P} f_{bppste} \qquad \text{(eq. S1)}$$

where A_P and A_{RS} are the peak area of the protein and reference standard peaks, T is the temperature at which the NMR spectra were acquired, pw is the 90 ^{o} pulse width and S is the number of scans used during the acquisition of the respective spectra, and H is the number of protons that constitutes the respective peaks. The correction factor f_{bppste} compensates for peak attenuation in the bppste pulse sequence (Figure S2) due to diffusion, T_1 relaxation, and T_2 relaxation. It is given by

$$f_{bppste} = 2\, e^{\frac{\tau_1}{T_1}}\, e^{\frac{\tau_2}{T_2}}\, e^{D\gamma^2 g^2 \delta^2 \left[\Delta - \frac{\delta}{3} - \frac{\tau_3}{2}\right]} \qquad \text{(eq. S2)}$$

where τ_1 is the time between the second and third 90^o pulses in the bppste pulse sequence, T_1 is the longitudinal relaxation time, τ_2 is the time between the first and second 90^o pulses (and between the third 90^o pulse and acquisition), T_2 is the spin-spin relaxation time, D is the translation diffusion coefficient, γ is the magnetogyric ratio, g is the pulse-field gradient strength, δ is the pulse-field gradient length, Δ is the diffusion delay, and τ_3 is the total time between the pulse-field gradient pulses of the bipolar-pulse pairs in the sequence. In effect, the first exponential term in the equation describes signal attenuation due to T_1 relaxation, the second due to T_2 relaxation, and the third due to diffusion. Of all the terms in the equation, these three quantities (T_1, T_2, and D) are the only unknowns; all others are pulse sequence/instrument parameters or a physical constant (in the case of γ). Therefore, by measuring these three values for a given sample, the concentration of the protein can be calculated using eq S1. The values of T_1, T_2, and D can be experimentally determined with the bppste pulse sequence by acquiring spectra with the spectra with different combinations of pulse sequence parameters as described below.

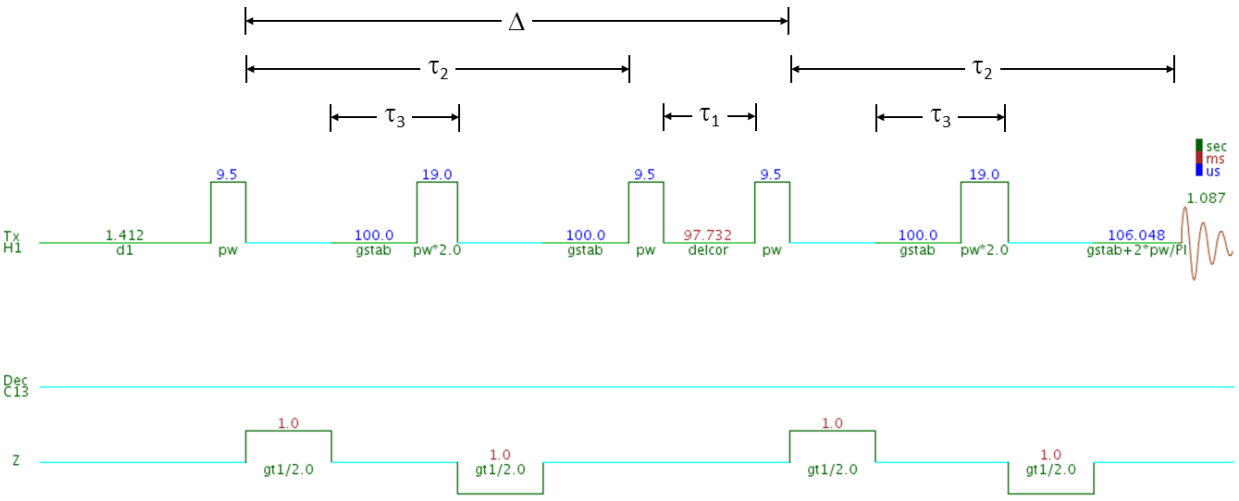


Figure S2. The bppste pulse sequence for Agilent (a.k.a., Dbppste) spectrometers along with the definition of the various delays discussed in the text.

First, rewriting the eq S2 with the Agilent Dbppste pulse sequence acquisition parameters yields

$$f_{BPPSTE} = 2\, e^{\frac{(del-gt1-2\cdot gstab-4\cdot pw90-3\cdot rof1)}{T_1}} e^{\frac{2\cdot(gt1+2\cdot gstab+2\cdot pw90+2\cdot rof1)}{T_2}}$$

$$e^{D\gamma^2 gzlvl1^2 gcal^2 \delta^2\left[del-\frac{gt1}{3}-\frac{(gstab+2\cdot pw90+rof1)}{2}\right]} \qquad \text{(eq. S3)}$$

where *del* is the diffusion delay (Δ), *gt1* is the gradient length (δ), *gzlvl1* is the gradient strength, *gcal* is the multiplier to convert the *gzlvl1* to Gauss, *gstab* is the gradient recovery delay, *pw90* is the 90°-pulse length (*PW*), and *rof1* is the short delay prior to an RF pulse (Figure S2).

For the T_1 experiment, five spectra were acquired with the Dbppste pulse sequence using combinations of *del* and *gzlvl1*, such that

$$gzlvl1_n = gzlvl1_1 \times \sqrt{\frac{\Delta\prime_1}{\Delta\prime_n}} \qquad \text{(eq. S4)}$$

where Δ'is the correct diffusion delay given by

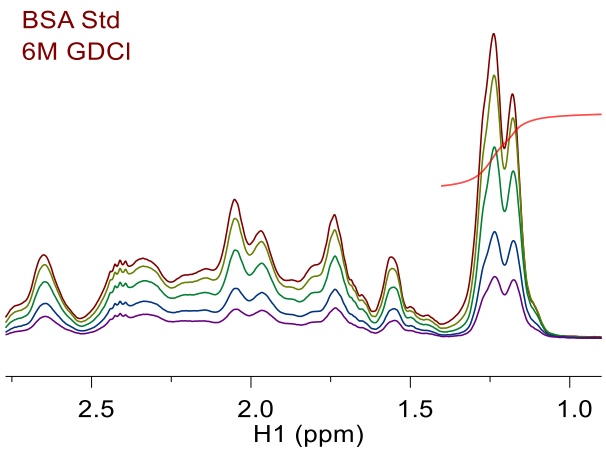
$$\Delta' = del - \frac{gt1}{3} - \frac{(gstab+2\cdot pw90+rof1)}{2} \qquad \text{(eq. S5)}$$

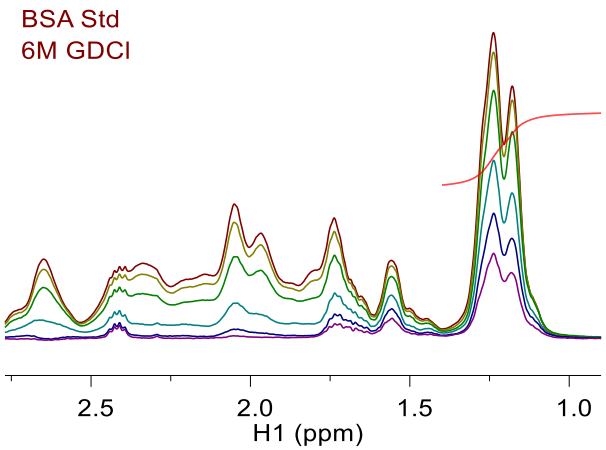
This keeps the T_2 and D portions of the equation constant, making signal attenuation governed solely by T_1 nuclear relaxation. The effective T_1 for the leucine, isoleucine, and valine methyls can be calculated from a mono-exponential fit of A_P^{LIV} vs. τ_1 ($\tau_1 = del - gt1 - 2\,gstab - 4\,pw90 - 3\,rof1$). An example of the spectra and the corresponding plot are shown Figure S3.

For the T_2 experiment, six spectra were acquired with combinations of *gstab*, *del*, and *gzlvl1*, such that

$$del_n = del_1 + 2\cdot(gstab_n - gstab_1) \qquad \text{(Eq S6)}$$

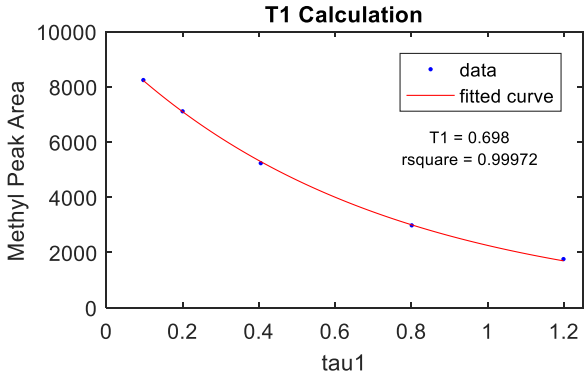
and $gzlvl_n$ is determined as shown above. This keeps the T_1 and D portions of the equation constant, making signal attenuation governed solely by T_2. The effective T_2 for the leucine, isoleucine, and valine methyls can be calculated from a mono-exponential fit of A_P^{LIV} vs. τ_2 (τ_2 can be reduced to 4 *gstab*). An example of the spectra and the corresponding plot are shown Figure S3.





Pulse Sequence	Dbppste
Number of Scans	16
Receiver Gain	60
Relaxation Delay	3.6369
Pulse Width	11.1500
Gradient Strength	30000 21006 14800 10537 8616
Gradient Length	0.002
Gradient Recovery	0.0005
Diffusion Delay	0.1 0.203 0.408 0.804 1.202

Pulse Sequence	Dbppste
Number of Scans	16
Receiver Gain	60
Relaxation Delay	1.0000
Pulse Width	11.1500
Gradient Strength	30025 30076 30177 30383 30594 30809
Gradient Length	0.0014
Gradient Recovery	0.001 0.002 0.004 0.008 0.012 0.016
Diffusion Delay	0.1500005 0.1500015 0.1500035 0.1500075 0.1500115 0.1500155



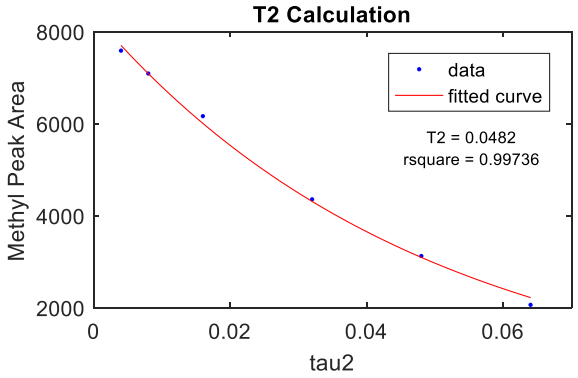


Figure S3. Representative data from the T_1 and T_2 measurements using the bppste pulse sequence. Top Left: NMR spectra from and parameters for the bppste pulse sequence setup to measuring T_1. Top Right: NMR spectra from and parameters for the bppste pulse sequence setup to measuring T_2. Bottom Left: Plot for T_1 calculation. Bottom Right: Plot for T_2 calculation.

Integration of Protein Peak

The area of the peaks for the isoleucine, leucine, and valine methyl groups was obtained by deconvolution of the spectra. In MNova, the best results were acheived using a generalized Lorentzian shape with a Lorentzian/Gaussian ratio of 1.5, and allowing the software to optimize the chemical shift, height, and linewidth (see Figure S4). The areas of the resulting peaks were summed in order to obtain the overall area used to calculate the concentration.

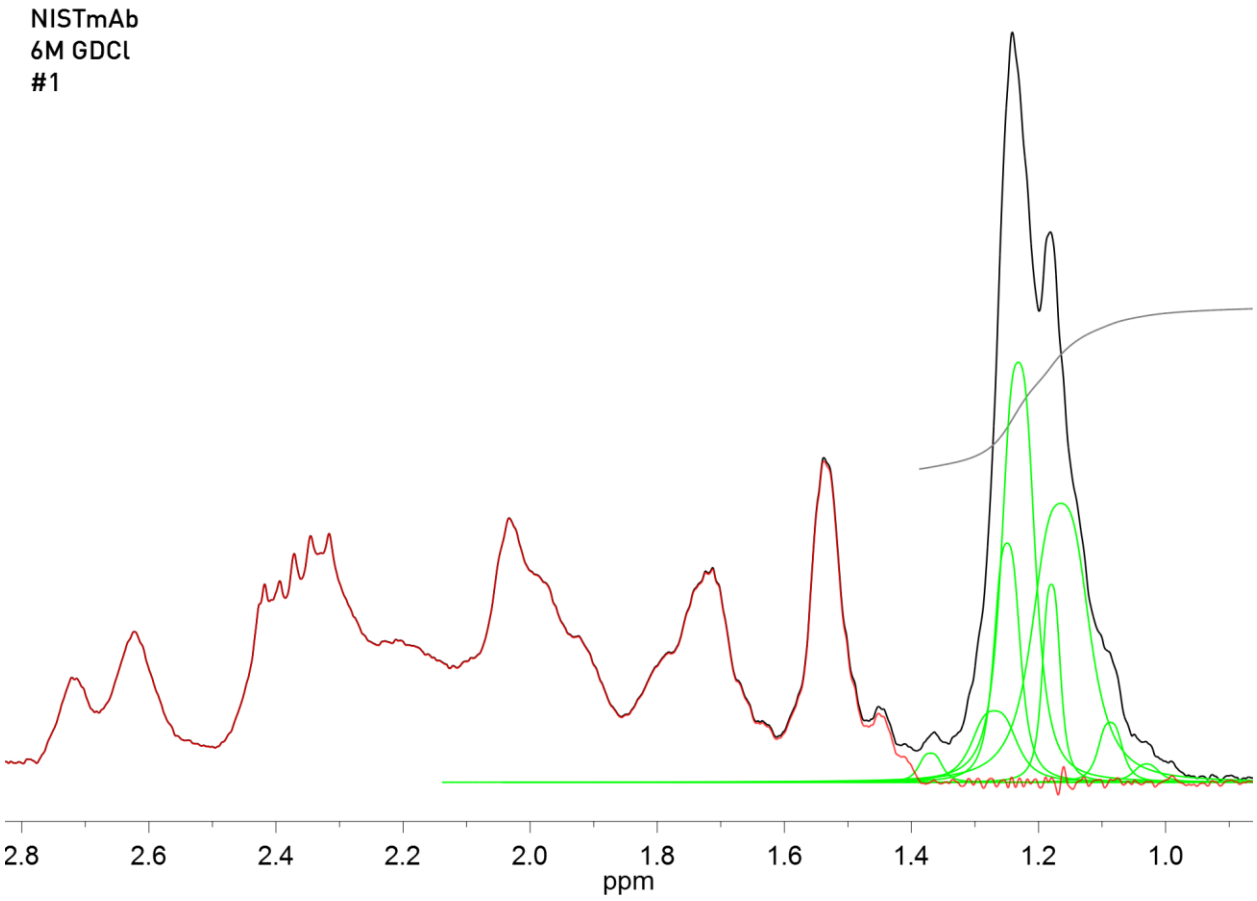


Figure S4. Representative data showing the DF-qNMR spectrum with line fitting of the leucine, isoleucine, and valine signals for integration.

Linearity of DF-qNMR Results

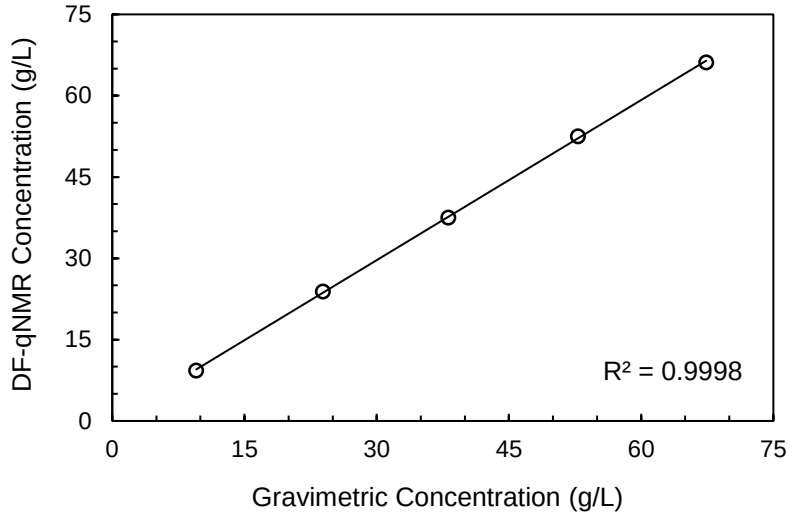


Figure S5. Plot of the DF-qNMR concentrations versus the gravimetric concentrations for NIST BSA 927e SRM. The solid line is the result of linear regression.

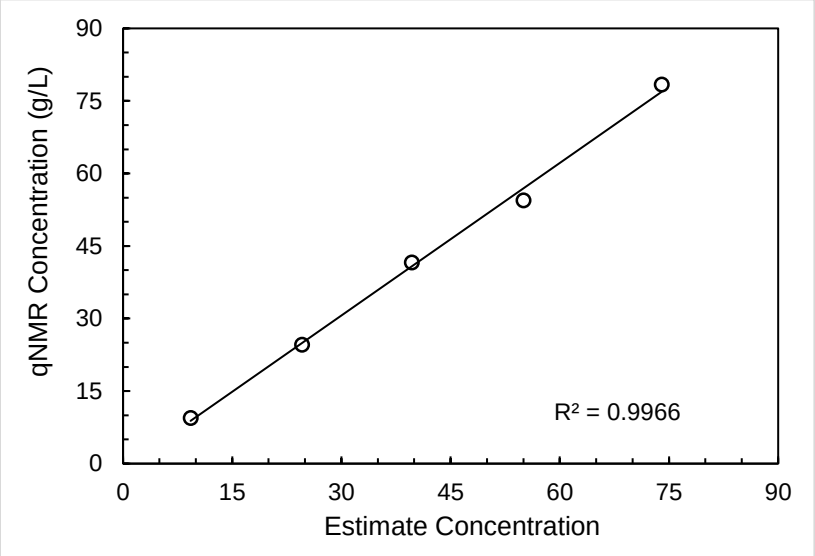


Figure S6. Plot of the DF-qNMR concentrations versus the gravimetric concentrations for the bispecific antibody. The solid line is the result of linear regression.