SUPPLEMENTARY MATERIAL

Web server DDfit: a new scheme to process PFG NMR diffusion data with improved precision

Vladislav A. Salikov[1], Olga O. Lebedenko[1], Nikolai R. Skrynnikov[1,2]*, Ivan S. Podkorytov[1]*

[1] Laboratory of Biomolecular NMR, St. Petersburg State University, St. Petersburg 199034, Russia
[2] Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA
* Corresponding authors: n.skrynnikov@spbu.ru, i.podkorytov@spbu.ru

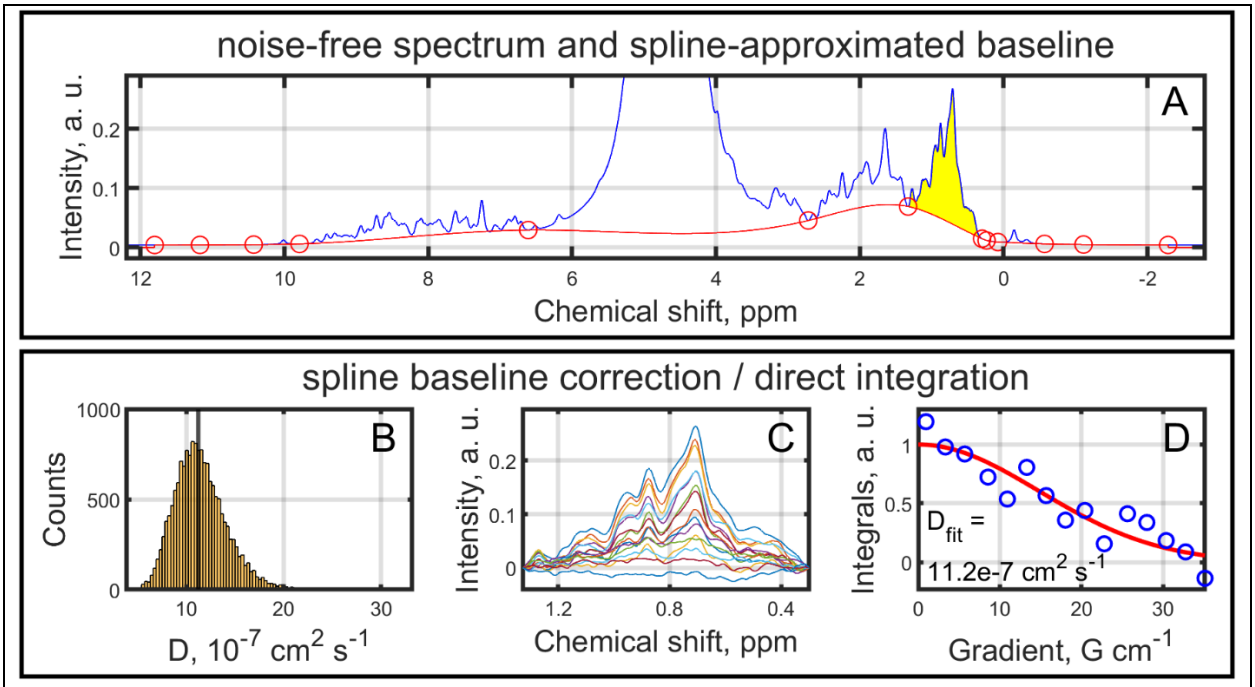

Figure S1. Analyses of the simulated STE data using spline baseline correction and numeric integration. **(A)** Simulated STE spectrum of ubiquitin at zero gradient strength with no added noise. The points to construct spline approximation for spectral baseline were selected manually (red circles); the same points (i.e. frequencies) have been used when processing the simulated spectra $f_k(\omega)$ that contain synthetic noise. The spline-approximated baseline is shown with red line. The portion of the spectrum intended for numeric integration is shaded yellow. (**B**) Histogram of D_{fit} values obtained from Stejskal-Tanner analyses of the N = 16,383 simulated datasets. The horizontal scale is chosen such as to accommodate all of the obtained D_{fit} values. (**C**) Representative dataset corresponding to the median D_{fit} value: a series of simulated STE spectra plotted in the interval from 1.33 to 0.30 ppm. Note one predominantly negative spectrum, which arises from instability of the spline procedure in the presence of random noise. (**D**) Stejskal-Tanner fit of the integral signals obtained from the dataset shown in panel (C). For the asymmetric distribution shown in panel (B), the median value of D_{fit}, 11.19×10^{-7} cm^2 s^{-1}, is somewhat different from the mean, 11.46×10^{-7} cm^2 s^{-1}.

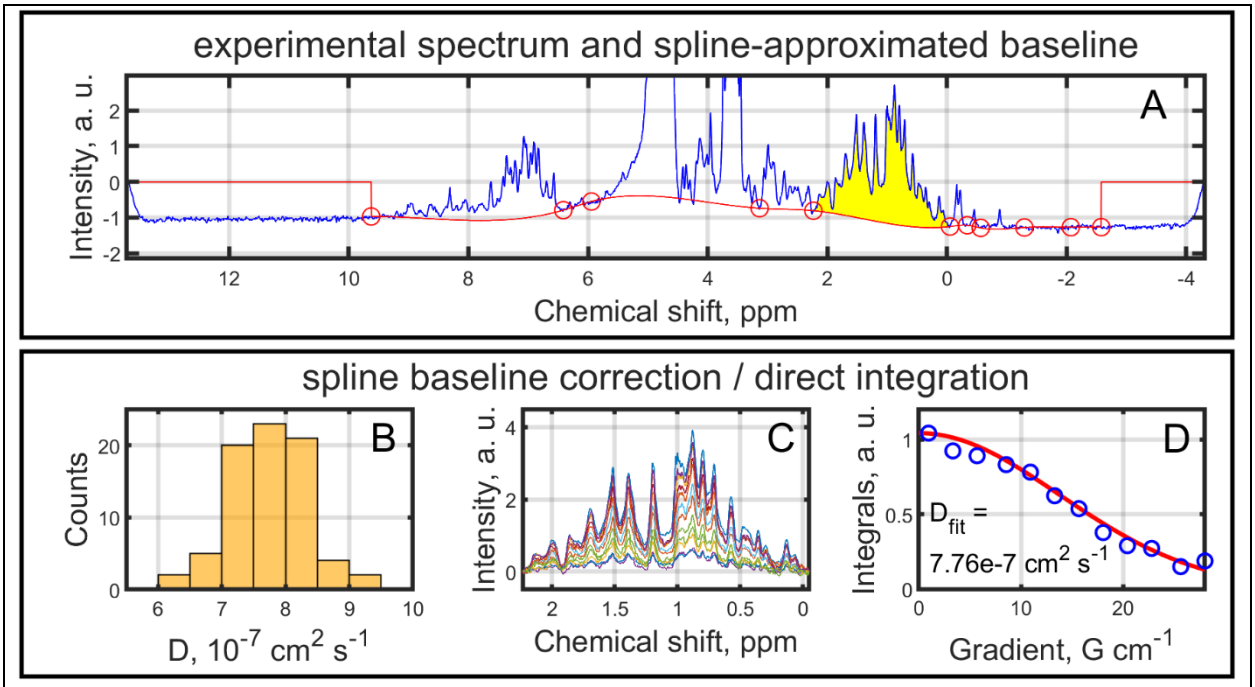

Figure S2. Analyses of the experimental DSTE data using spline baseline correction and numeric integration. **(A)** DSTE spectrum of HEWL at lowest gradient strength 0.95 G/cm. The points to construct spline approximation for spectral baseline were selected manually (red circles); the same points (i.e. frequencies) have been used when processing all of the experimental spectra $f_k(\omega)$. The spline-approximated baseline is shown with red line. The portion of the spectrum intended for numeric integration (from 2.23 to -0.05 ppm, region I) is shaded yellow. (**B**) Histogram of D_{fit} values obtained from Stejskal-Tanner-Jerschow-Muller analyses of the 77 replicate datasets. (**C**) Representative dataset corresponding to the median D_{fit} value. (**D**) Stejskal-Tanner-Jerschow-Muller fit of the integral signals obtained from the dataset shown in panel (C).

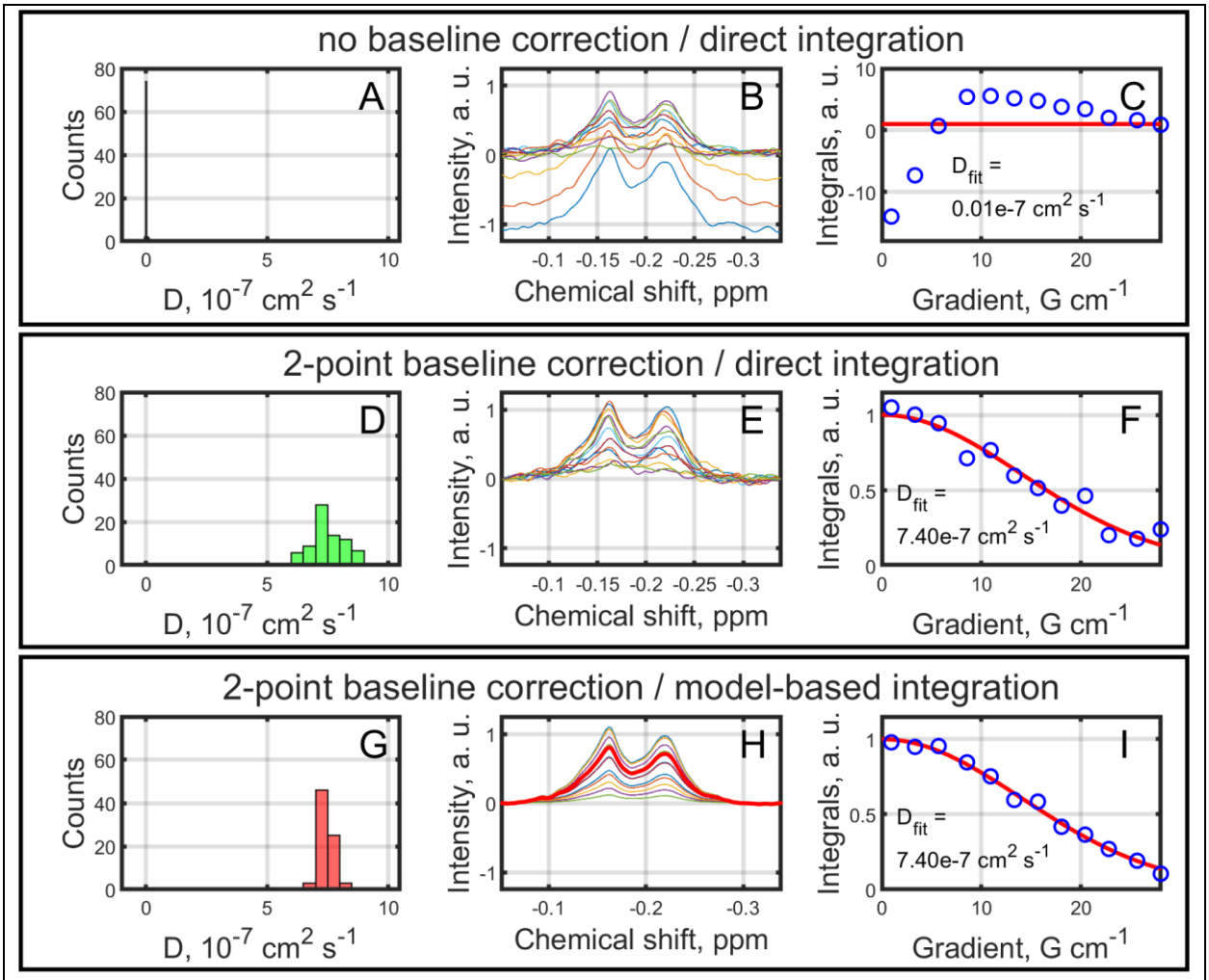

Figure S3. Analyses of the experimental DSTE data from the sample of HEWL using different processing schemes. **Top row**: no baseline correction, direct numeric integration of the spectra in the specified region, from -0.05 to -0.34 ppm (region II). (**A**) Histogram of D_{fit} values obtained from Stejskal-Tanner-Jerschow-Muller analyses of 77 replicate datasets. (**B**) Representative dataset corresponding to the median D_{fit} value. (**C**) Stejskal-Tanner-Jerschow-Muller fit of the integral signals from the dataset shown in panel (B). The analysis is compromised by the baseline shift in some of the spectra (those recorded with weaker gradients), leading to dramatic underestimation of diffusion coefficients. **Middle row**: 2-point baseline correction, direct numeric integration of the spectra. (**D**) Histogram of D_{fit} values. (**E**) Representative dataset corresponding to the median D_{fit} value. (**F**) Stejskal-Tanner-Jerschow-Muller fit of the integral signals $I_k^{(0)}$ from the dataset (E). **Bottom row**: 2-point baseline correction, model-based integration as implemented in DDfit. (**G**) Histogram of D_{fit} values. (**H**) Representative dataset corresponding to the median D_{fit} value, after DDfit treatment. (**I**) Stejskal-Tanner-Jerschow-Muller fit of the integral signals $I_k^{(1)}$ obtained from the dataset (H). The plots similar to panels (H) and (I) are generated as a part of the demo by the DDfit server; these results are from the first DSTE dataset out of seventy seven experimental datasets, yielding D_{fit} = 7.59×10^{-7} cm^2 s^{-1}.

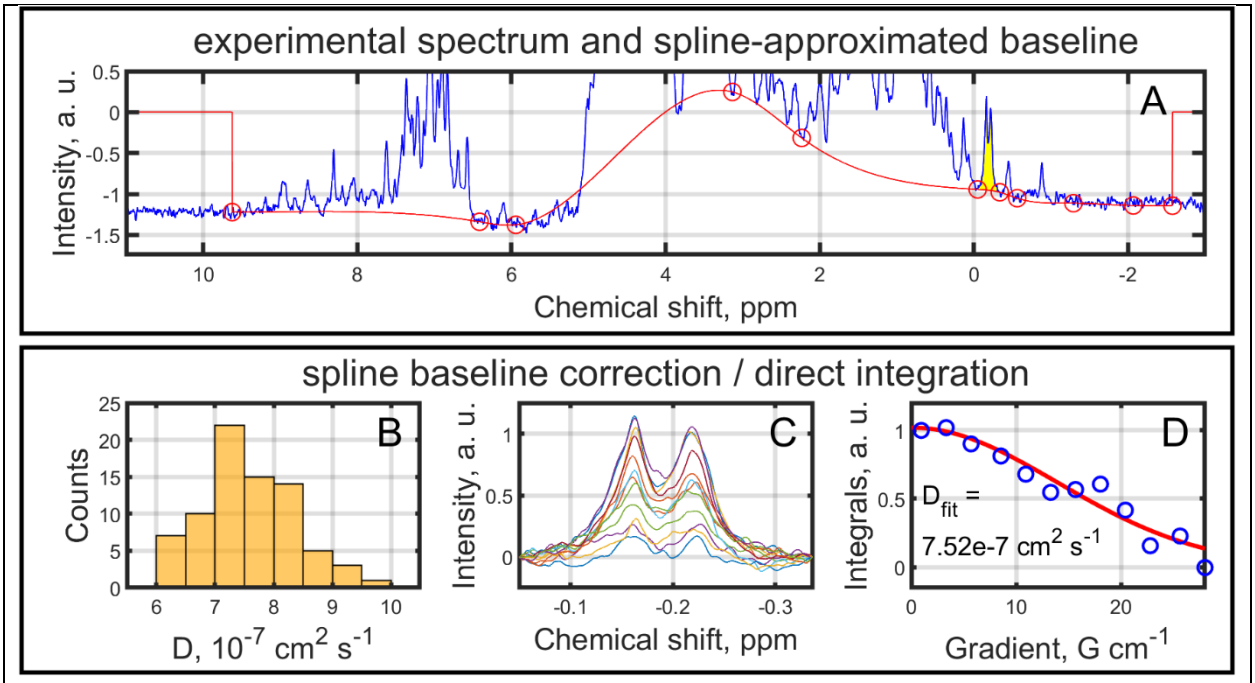

Figure S4. Analyses of the experimental DSTE data using spline baseline correction and numeric integration. **(A)** DSTE spectrum of HEWL at lowest gradient strength 0.95 G/cm. The points to construct spline approximation for spectral baseline were selected manually (red circles); the same points (i.e. frequencies) have been used when processing all of the experimental spectra $f_k(\omega)$. The spline-approximated baseline is shown with red line. The portion of the spectrum intended for numeric integration (from -0.05 to -0.34 ppm, region II) is shaded yellow. (**B**) Histogram of D_{fit} values obtained from Stejskal-Tanner-Jerschow-Muller analyses of the 77 replicate datasets. (**C**) Representative dataset corresponding to the median D_{fit} value. (**D**) Stejskal-Tanner-Jerschow-Muller fit of the integral signals obtained from the dataset shown in panel (C).