Electronic Supplementary Material (ESI) for ChemComm.

SUPPORTING INFORMATION

T_m filtering by ^{1}H-methyl labeling in a deuterated protein for pulsed double electron-electron resonance EPR

Thomas Schmidt and G. Marius Clore*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, U.S.A.

Experimental Procedures

Expression, purification and labeling of Protein A. Fully deuterated AviTag-protein A, with two surface exposed, engineered cysteine residues (Q39C and K88C) was expressed in E. coli and purified as described previously.[1] The AviTag extends from residues 1-29, and protein A from residues 30-90; residues 1-38 are disordered in solution. Incorporation of protonated methyl groups of Leu ($C^{\delta}H_3$ and $C^{\delta}H$) and Ile($C^{\delta}H$) in a fully deuterated background was carried out as described previously[2] by growing the bacteria in minimal D_2O medium with ammonium chloride as the sole nitrogen source, U-[^{2}H]D-glucose as the main carbon source, and the appropriate α-keto acid precursor for $^{13}CH_3$ labeling: α-ketoisovaleric acid (^{13}C5, 98%, 3-D1, 98%) for Leu and Val (Cambridge Isotope Laboratories CDLM-4418-PK) and α-ketobutyric acid (methyl-^{13}C, 99%; 3,3-D2, 98%) for Ile (Cambridge Isotope Laboratories CDLM-7318-PK). Note there are no valines in AviTag-Protein A. Nitroxide (R1) spin-labeling was carried out with S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL; Toronto Research Chemicals) as described previously.[1] Nitroxide-labeling was verified by electrospray ionization mass spectrometry. The sample for NMR comprised 50 μM AviTag-Protein A, 0.85 mM KH_2PO_4, 25 mM Na_2HPO_4, pH 7.4, 75 mM NaCl, and d8-glycerol(30% v/v)/D_2O (70% v/v).

Pulsed Q-band EPR spectroscopy. Pulsed EPR data were collected at Q-band (33.8 GHz) and 50 K on a Bruker E-580 spectrometer equipped with a 150 W traveling-wave tube amplifier, a model ER5107D2 resonator, and a cryofree cooling unit, as described previously. DEER experiments were acquired using a conventional four-pulse sequence. The observer and ELDOR pump pulses were separated by ca. 90 MHz with the observer π/2 and π pulses set to 12 and 24 ns, respectively, and the ELDOR π pulse to 10 ns. The pump frequency was centered at the Q-band nitroxide spectrum located at +40 MHz from the center of the resonator frequency. The τ_1 value of 400 ns for the first echo-period time was incremented eight times in 16 ns steps to average ^{2}H modulation; the position of the ELDOR pump pulse was incremented in steps of Δt = 8 ns. The bandwidth of the overcoupled resonator was 120 MHz. All DEER echo curve were acquired for τ_2 = 4 μs, with the exception of the DEER echo curve for τ_2 = 4 μs (T = 8 μs), where t_2 was set to 3 μs to avoid the persistent "2+1" echo perturbation of the DEER echo curves at a time of about τ_1 from the final observe π pulse. DEER data were recorded with values of the dipolar evolution time T (= $2\tau_1$) ranging from 8 to 80 μs for [Leu-$C^{\delta}H_3$/^{2}H] AviTag-Protein A, and from 10 to 60 μs for [Ile-$C^{\delta}H$/^{2}H] aviTag-Protein A. Measurement times were approximately as follows: for T = 8 - 20 μs, 1 hr; 24 - 32 μs, 6 hrs; 36 - 50 μs, 12 hrs; 60 - 80 μs, 24 hrs.

***P*(*r*) distributions from DEER data.** Four different approaches were used to obtain *P*(*r*) distributions from the DEER echo curves: (1) model free Tikhonov regularization using the program DeerAnalysis 2016[3] (with the Tikhonov regularization parameter α automatically determined at each iteration) ; (2) the program DD[4] that uses a sum of Gaussians to directly fit the experimental DEER data (including automated background correction with a best-fit exponential decay); (3) the program WavPDS[5] that makes use of the Daubechis 6 wavelet family to filter out noise, followed by Tikhonov regularization with DeerAnalysis;[3] (4) the program WavPDS followed by singular value decomposition (SVD).[5] Validated Tikhonov regularization was carried out in DeerAnalysis varying the modulation depth (11 steps from 0.25 to 0.5), background density (11 steps from 0.04 to 0.15) and background start (6 steps from 500 to 1800 ns) for a total of 726 permutations. The optimal number of Gaussians (two for the current data) used by the program DD to represent the DEER data, was assessed using the Akaike information criterion corrected for finite sample size (AICc) and the Bayesian information criterion (BIC).[4]

Table S1 Optimized values of the scale factors λ_1 and λ_2 describing the amplitudes of the contributions of the ESSEM effect and phase memory relaxation, respectively, to the Hahn spin echo curve for [Leu-CH_3/^{2}H] aviTag-Protein A (Q39C-R1).[a]

Method[b]	λ_1	λ_2
DA[c]	7000 ± 100	24,000 ± 100
DA + validation[d]	7500 ± 100	23,500 ± 100
DD	7300 ± 200	23,600 ± 200
WavPDS + DA	7200 ± 100	23,800 ± 100
WavPDS + SVD	7400 ± 100	23,600 ± 100

[a]See Table 1 of the main text for the optimized values for the phase memory relaxation rates (R_m) for the *a* and *b* populations of the Q39C-R1 spin label in [Leu-CH_3/^{2}H]-AviTag-Protein A, the occupancy of the *b* state (p_B), and the apparent relaxation rate, R_{ESEEM}, for the initial fast decay of the Hahn spin echo curve due to the ESEEM effect. [b]Method used to derive the $P(r)$ distributions from the DEER echo curves. [c]DA, DeerAnalysis with Tikhonov regularization. [d]Validated Tikhonov regularization was carried out in DeerAnalysis varying the modulation depth (11 steps from 0.25 to 0.5), background density (11 steps from 0.04 to 0.15) and background start (6 steps from 500 to 1800 ns) for a total of 726 permutations.

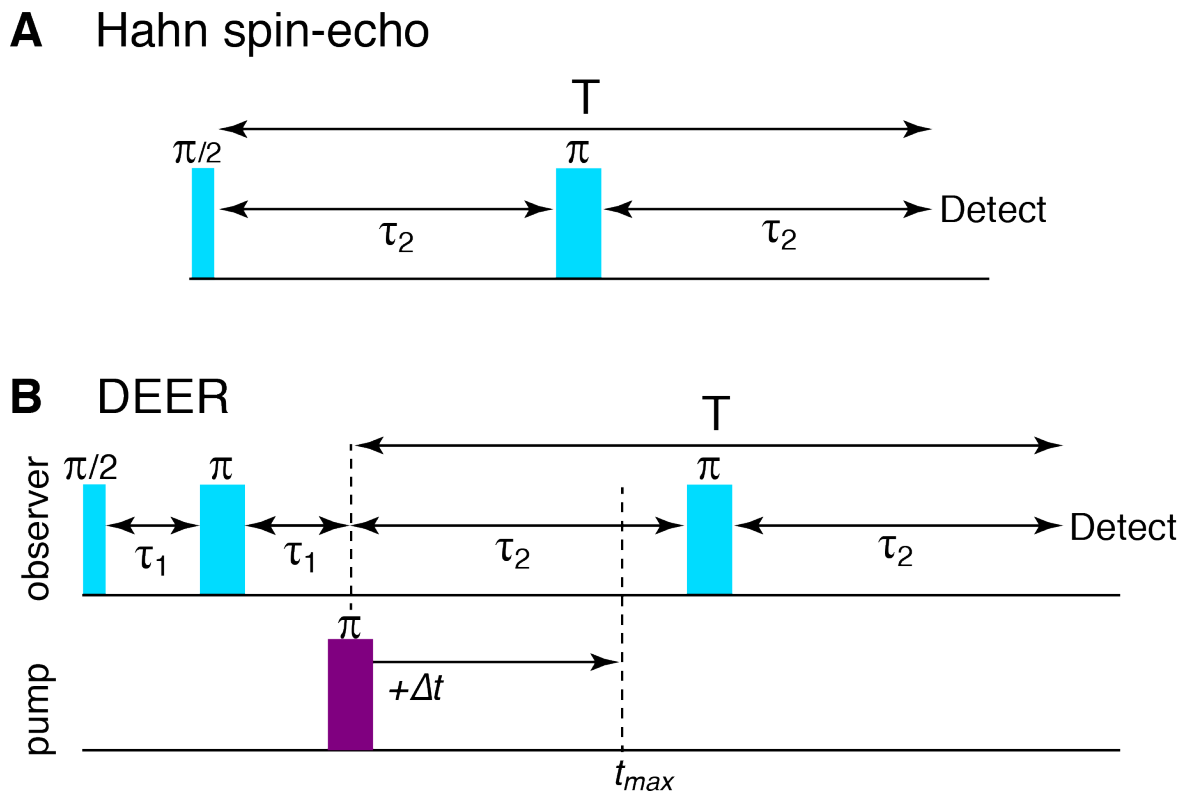

Figure S1. Pulse schemes for (A) two-pulse Hahn spin echo and (B) four-pulse DEER[8] experiments. In the current work, the duration of the blue $\pi/2$ and π pulses are 12 and 24 ns, respectively, and that of the purple π pulse is 10 ns.

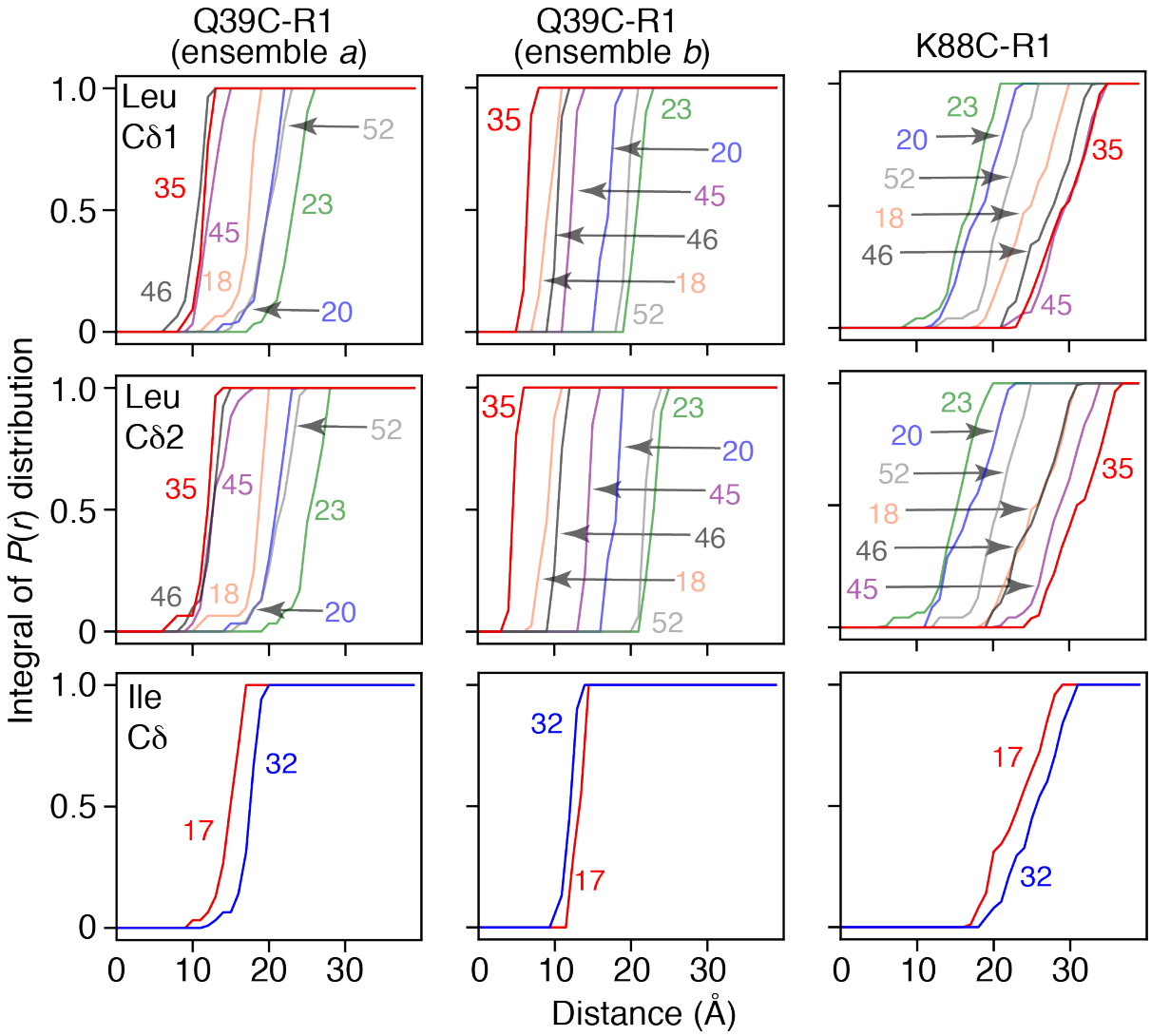

Figure S2. Analysis of distance distributions between the methyl groups of Leu (Cδ1 and Cδ2) and Ile (Cδ) and the nitroxide spin labels at Q39C-R1 and K88C-R1 in AviTag-Protein A. The spin label conformer ensembles were calculated from the coordinates of Protein A (PDB code 1BDD)[9] using the rotamer library program MMM.[10, 11] Q39C-R1 exists in two distinct ensemble populations *a* and *b* (see Fig. 1A of main text). The distances are calculated to the carbon atoms of the methyl groups and displayed as the integral of the *P*(*r*) distributions. Only the methyl groups of Leu35 are within the 4-8 Å range of a nitroxide oxygen, namely that of Q39C-R1 in ensemble *b*, to result in a significant increase in the phase memory relaxation time (T_m) of the unpaired electron.

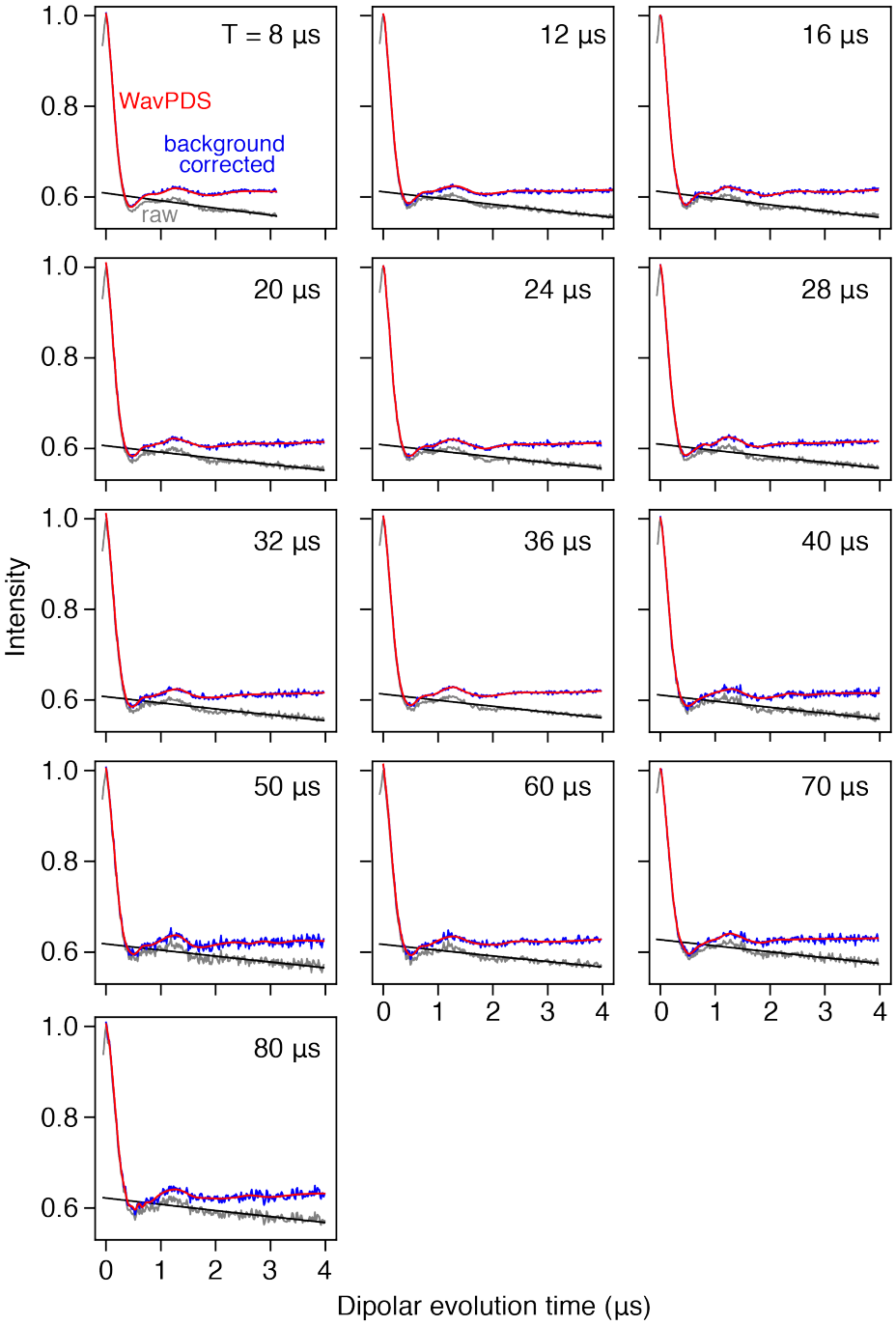

Figure S3. Q-band DEER echo curves for [Leu-$C^{\delta}H_3/^2H$] AviTag-protein A (Q39C-R1/K88C-R1) recorded with different total evolution periods *T*. The color coding is as follows: grey, raw DEER echo curves; blue, DEER echo curves after background correction (black line) obtained using DeerAnalysis;[3] red, denoised and background-corrected DEER echo curves obtained using WavPDS.[5, 7]

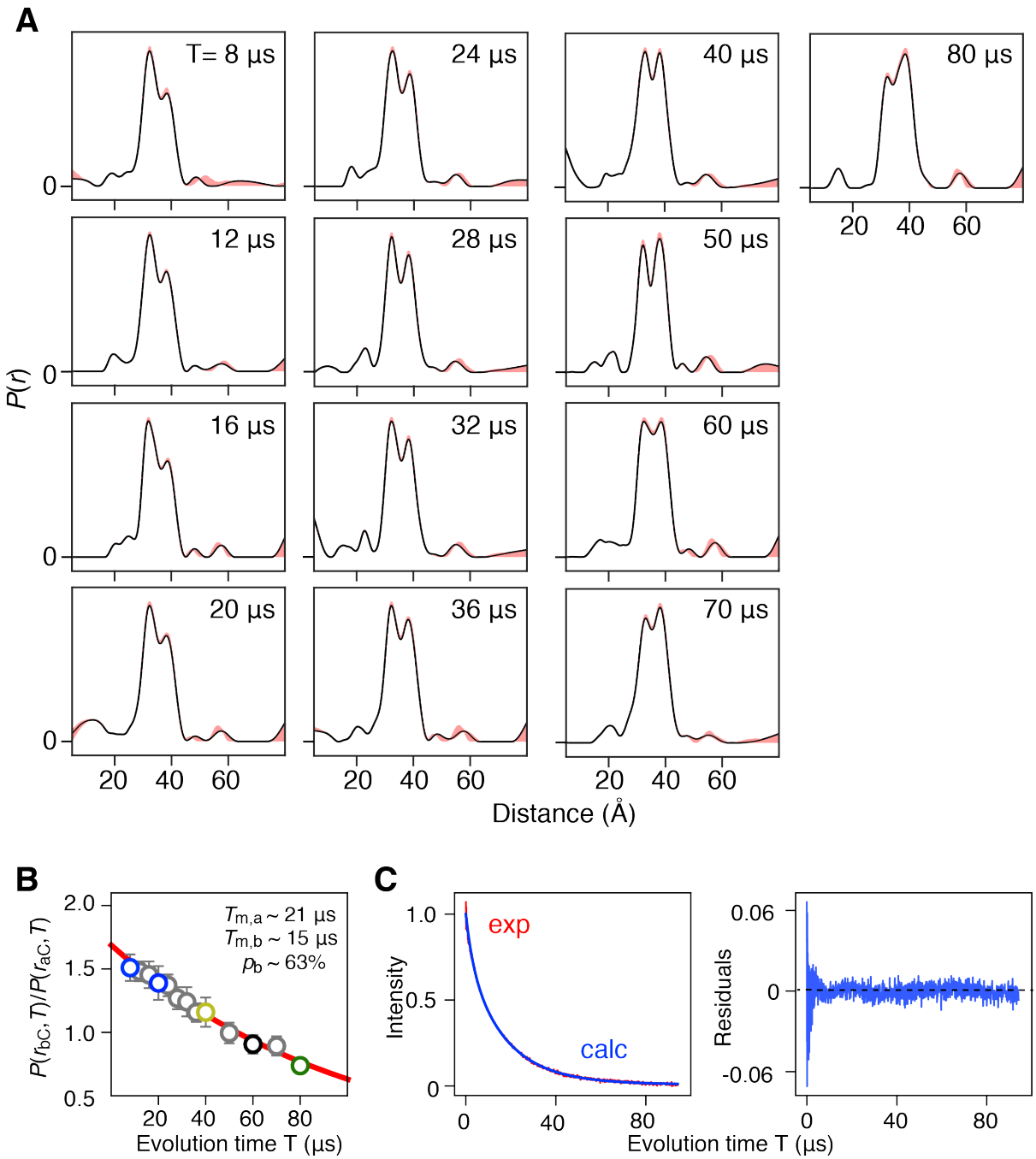

Figure S4. Analysis of the evolution time *T* dependence of the *P*(*r*) distributions for [Leu-CH_3/^{2}H] AviTag-Protein A (Q39C-R1/K88C-R1) obtained from the DEER echo curves using DEERAnalysis with validated Tikhonov regulation.[3] (**A**) *P*(*r*) distributions for values of *T* (= 2τ_2) ranging from 4 to 80 µs. The *P*(*r*) distributions with the best r.m.s.d. are displayed in black; the lilac regions indicate the full *P*(*r*) variation over all 726 trials varying the modulation depth (11 steps), background density (11 steps) and background start (6 steps). Results of global best-fitting Eqs. (1) and (2) to the experimental evolution time *T* (= 2τ_2) dependence of (**B**) the peak intensity ratio, $P(r_{bC},T)/P(r_{aC},T)$, and (**C**) the Hahn spin echo curve. The experimental data and best-fits obtained with Eqs. (1) and (2) (of the main text) are displayed by the circles and blue line, respectively in (B) and by the red and blue curves, respectively, in (C). Normalization in (C) was carried out by dividing the experimental and calculated curves by the optimized value of the sum of the two scale factors ($\lambda_1 + \lambda_2$). The right panel in (**C**) is a plot of the distribution of residuals between experimental and calculated Hahn spin echo curves.

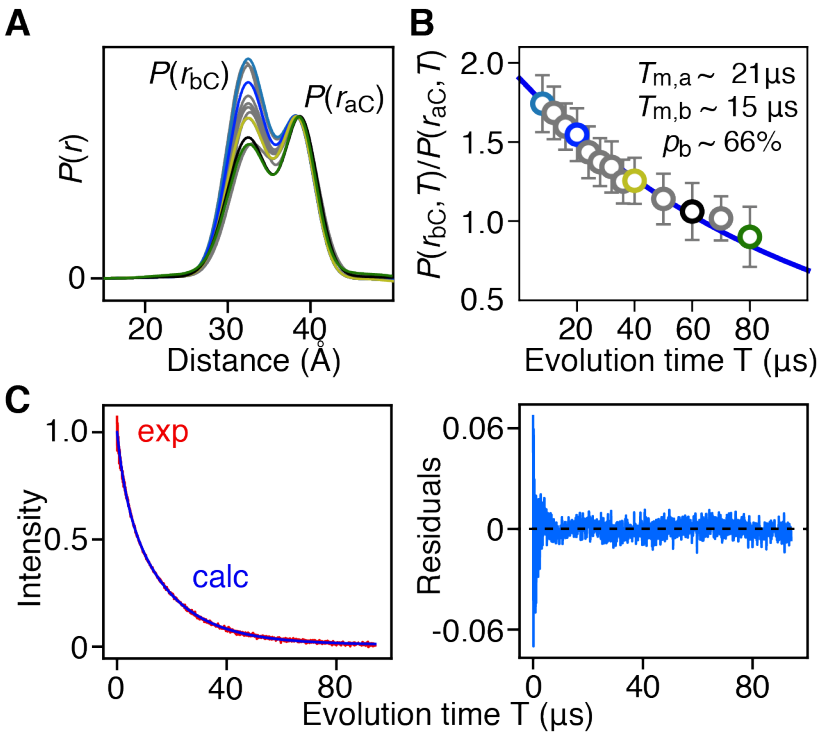

Figure S5. Analysis of the evolution time *T* dependence of the *P(r)* distributions for [Leu-CH_3/^{2}H] AviTag-Protein A (Q39C-R1/K88C-R1) obtained from the DEER echo curves using the program DD[4] with a sum of two Gaussians. (**A**) *P*(*r*) distributions for values of T (= $2\tau_2$) ranging from 4 to 80 µs, normalized with regard to the *a* peak at 38 Å. The color coding corresponds to that of the experimental points in panel B. Results of global best-fitting Eqs. (1) and (2) to the experimental evolution time T (= $2\tau_2$) dependence of (**B**) the peak intensity ratio, $P(r_{bC},T)/P(r_{aC},T)$, and (**C**) the Hahn spin echo curve. The experimental data and best-fits obtained with Eqs. (1) and (2) (of the main text) are displayed by the circles and blue line, respectively in (B) and by the red and blue curves, respectively, in (C). Normalization in (C) was carried out by dividing the experimental and calculated curves by the optimized value of the sum of the two scale factors ($\lambda_1 + \lambda_2$). The right panel in (**C**) is a plot of the distribution of residuals between experimental and calculated Hahn spin echo curves.

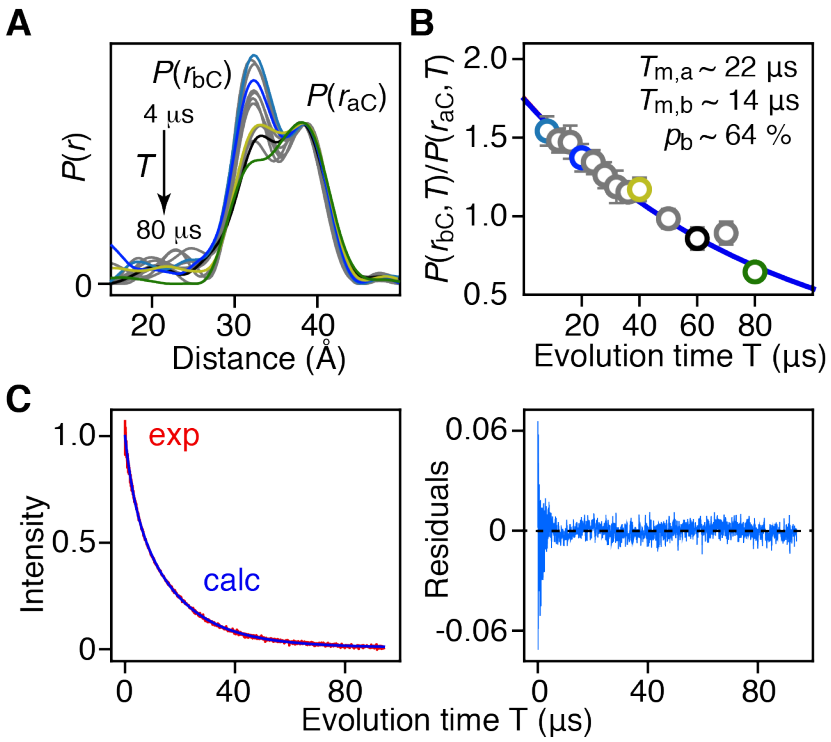

Figure S6. Analysis of the evolution time *T* dependence of the *P*(*r*) distributions for [Leu-CH_3/^{2}H] AviTag-Protein A (Q39C-R1/K88C-R1) obtained from the DEER echo curves using the program WavPDS[5] for denoising followed by Tikhonov regularization.[3] (**A**) *P*(*r*) distributions for values of *T* (= $2\tau_2$) ranging from 4 to 80 µs, normalized with regard to the *a* peak at 38 Å. The color coding corresponds to that of the experimental points in panel B. Results of global best-fitting Eqs. (1) and (2) to the experimental evolution time *T* (= $2\tau_2$) dependence of (**B**) the peak intensity ratio, $P(r_{bC},T)/P(r_{aC},T)$, and (**C**) the Hahn spin echo curve. The experimental ratios of the *b* to *a* peak intensities, $P(r_{bC},T)/P(r_{aC},T)$, are obtained from the integrated intensities of two Gaussians fitted to the bimodal *P*(*r*) distributions. The experimental data and best-fits obtained with Eqs. (1) and (2) (of the main text) are displayed by the circles and blue line, respectively in (B) and by the red and blue curves, respectively, in (C). Normalization in (C) was carried out by dividing the experimental and calculated curves by the optimized value of the sum of the two scale factors ($\lambda_1 + \lambda_2$). The right panel in (**C**) is a plot of the distribution of residuals between experimental and calculated Hahn spin echo curves.

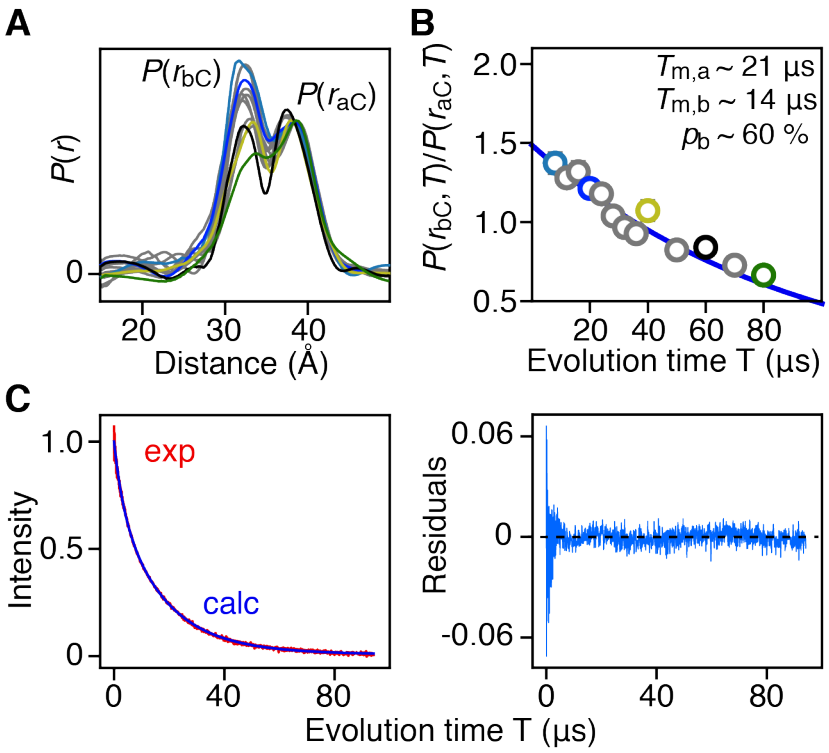

Figure S7. Analysis of the evolution time *T* dependence of the *P(r)* distributions for [Leu-CH_3/^{2}H] AviTag-Protein A (Q39C-R1/K88C-R1) obtained from the DEER echo curves using the program WavPDS for denoising followed by SVD.[5-7] (**A**) *P*(*r*) distributions for values of *T* (= $2\tau_2$) ranging from 4 to 80 µs, normalized with regard to the *a* peak at 38 Å. The color coding corresponds to that of the experimental points in panel B. Results of global best-fitting Eqs. (1) and (2) to the experimental evolution time *T* (= $2\tau_2$) dependence of (**B**) the peak intensity ratio, $P(r_{bC},T)/P(r_{aC},T)$, and (**C**) the Hahn spin echo curve. The experimental ratios of the *b* to *a* peak intensities, $P(r_{bC},T)/P(r_{aC},T)$, are obtained from the integrated intensities of two Gaussians fitted to the bimodal *P*(*r*) distributions. The experimental data and best-fits obtained with Eqs. (1) and (2) (of the main text) are displayed by the circles and blue line, respectively in (B) and by the red and blue curves, respectively, in (C). Normalization in (C) was carried out by dividing the experimental and calculated curves by the optimized value of the sum of the two scale factors ($\lambda_1 + \lambda_2$). The right panel in (**C**) is a plot of the distribution of residuals between experimental and calculated Hahn spin echo curves.

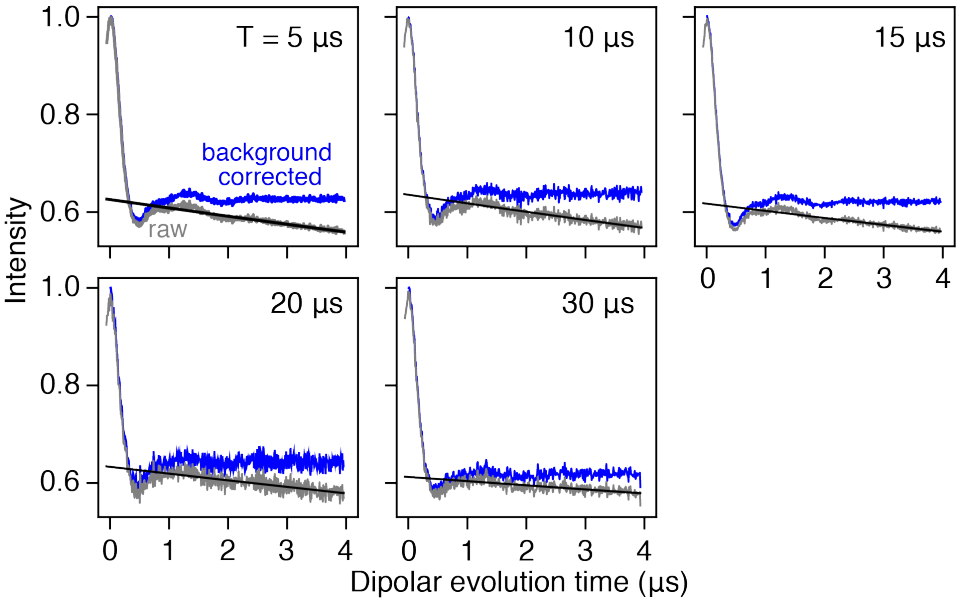

Figure S8. Q-band DEER echo curves for [Ile-$C^{\delta}H_3/^2H$] AviTag-protein A (Q39C-R1/K88C-R1) recorded with different total evolution periods *T*. The color coding is as follows: grey, raw DEER echo curves; blue, DEER echo curves after background correction (black line) obtained using DeerAnalysis.[3]

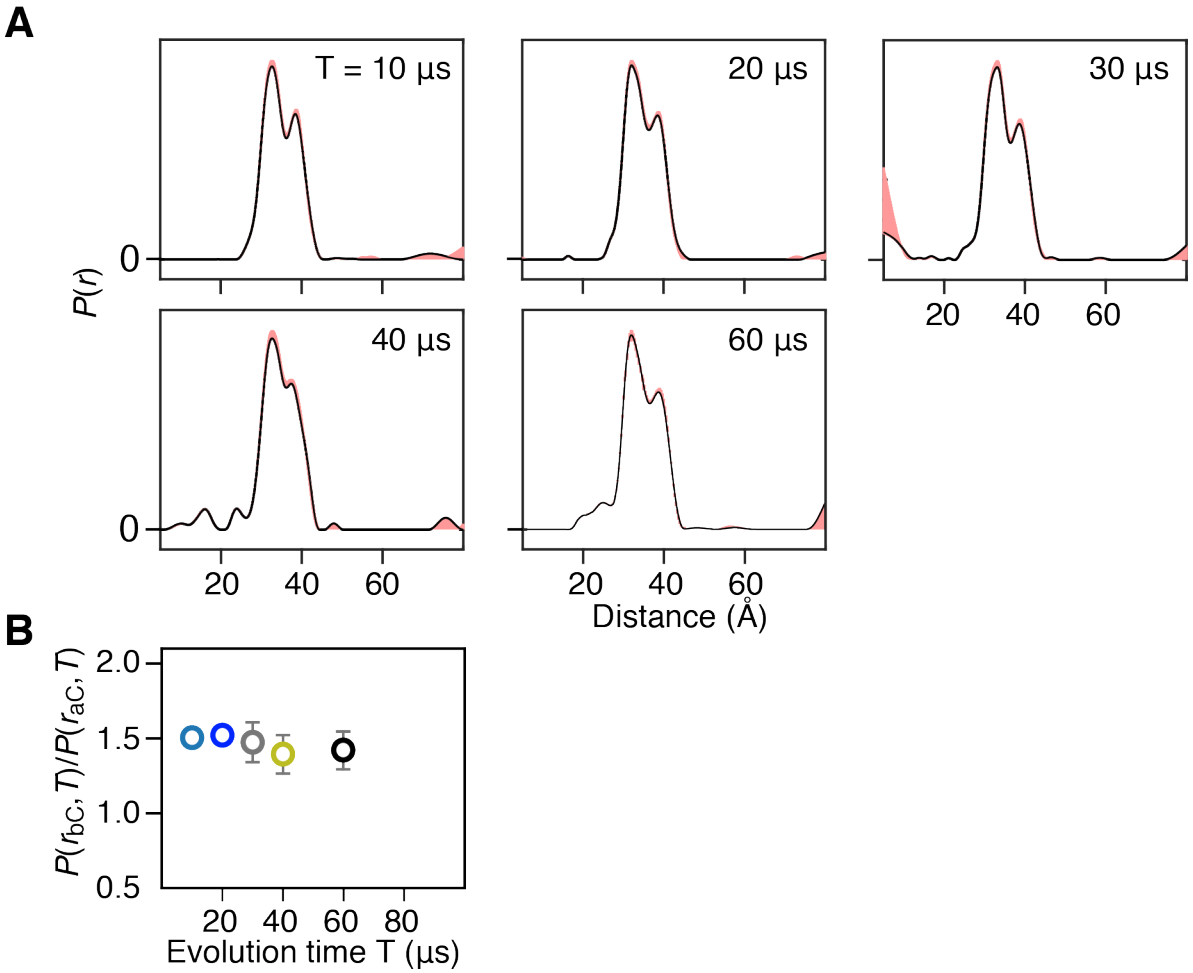

Figure S9. Analysis of the evolution time *T* dependence of the *P(r)* distributions for [Ile-C$^{\delta}$H$_3$/^{2}H] AviTag-Protein A (Q39C-R1/K88C-R1) obtained from the DEER echo curves using the program DeerAnalysis[3] with validated Tikhonov regularization. (**A**) *P*(*r*) distributions for values of *T* (= 2τ_2) ranging from 10 to 60 µs. The *P*(*r*) distributions with the best r.m.s.d. are displayed in black; the lilac regions indicate the full *P*(*r*) variation over all 726 trials varying the modulation depth (11 steps), background density (11 steps) and background start (6 steps). (**B**) Evolution time *T* (= 2τ_2) of the peak intensity ratio, $P(r_{bC},T)/P(r_{aC},T)$. The ratios of *b* to *a* peak intensities, $P(r_{bC},T)/P(r_{aC},T)$, are obtained from the integrated intensities of two Gaussians fitted to the bimodal *P*(*r*) distributions.

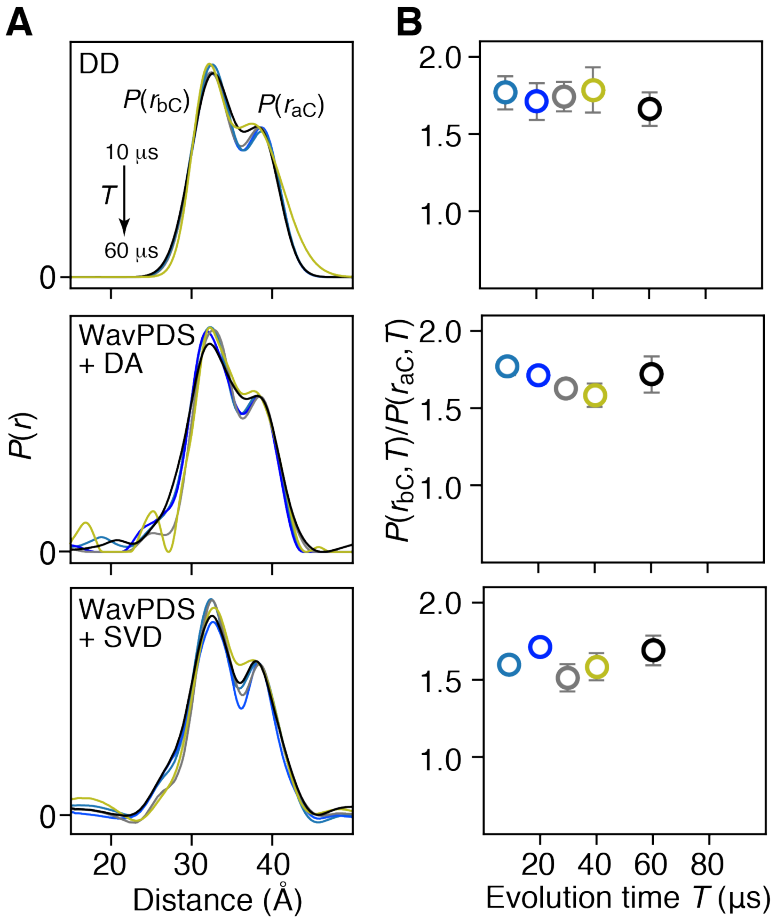

Figure S10. Analysis of the evolution time *T* dependence of the *P(r)* distributions for [Ile-C$^{\delta}$H$_3$/^{2}H] AviTag-Protein A (Q39C-R1/K88C-R1) obtained from the DEER echo curves using the programs DD[4] (top), WavPDS[5] for denoising followed by Tikhonov regularization with DeerAnalysis (DA)[3] (middle), and WavPDS followed by SVD[5-7] (bottom). (**A**) *P*(*r*) distributions for values of *T* (= 2τ_2) ranging from 10 to 60 µs, normalized with regard to the *a* peak at 38 Å. The color coding corresponds to that of the experimental points in panel B. (**B**) Evolution time *T* (= 2τ_2) of the peak intensity ratio, $P(r_{bC},T)/P(r_{aC},T)$. The ratios of *b* to *a* peak intensities, $P(r_{bC},T)/P(r_{aC},T)$, are obtained from the integrated intensities of two Gaussians fitted to the bimodal *P*(*r*) distributions.

Supplementary References

1. J. L. Baber, J. M. Louis and G. M. Clore, *Angew. Chem. Int. Ed. Engl.*, 2015, **54**, 5336-5339.
2. V. Tugarinov, V. Kanelis and L. E. Kay, *Nat. Protoc.*, 2006, **1**, 749-754.
3. G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmerman, J. E. Banham, C. R. Timmel, D. Hilger and H. Jung, *Appl. Magn. Reson.*, 2006, **30**, 473-498.
4. S. Brandon, A. H. Beth and E. J. Hustedt, *J. Magn. Reson.*, 2012, **218**, 93-104.
5. M. Srivastava, E. R. Georgieva and J. H. Freed, *J. Phys. Chem. A*, 2017, **121**, 2452-2465.
6. M. Srivastava and J. H. Freed, *J. Phys. Chem. Lett.*, 2017, **8**, 5648-5655.
7. M. Srivastava and J. H. Freed, *J. Phys. Chem. A*, 2019, **123**, 359-370.
8. M. Pannier, S. Veit, A. Godt, G. Jeschke and H. W. Spiess, *J. Magn. Reson.*, 2000, **142**, 331-340.
9. H. Gouda, H. Torigoe, A. Saito, M. Sato, Y. Arata and I. Shimada, *Biochemistry*, 1992, **31**, 9665-9672.
10. Y. Polyhach, E. Bordignon and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2356-2366.
11. G. Jeschke, *Protein Sci.*, 2018, **27**, 76-85.