Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

An optimal acquisition scheme for Q-band EPR distance measurements using Cu^{2+}-based protein labels

Xiaowei Bogetti,[‡a] Zikri Hasanbasri,[‡a] Hannah R. Hunter [a] and Sunil Saxena [a]

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, 15260, USA

* To whom correspondence should be addressed: ORCID: 0000-0001-9098-6114, Phone (412) 624-8680. Email: sksaxena@pitt.edu

To determine the coordination of Cu^{2+}-NTA to dHis sites, three-pulse electron-spin echo envelope modulation (ESEEM) experiments[1,2] were performed at 20 K with a Bruker ElexSys E680 X-band FT/CW spectrometer with a Bruker EN4118X-MD4 resonator. The pulse sequence was $(\pi/2)-\tau-(\pi/2)-t-(\pi/2)-\tau-$echo. The first pulse delay time, τ, was 140 ns. The second pulse delay time, t, was 288 ns and lengthened by 16 ns for each iteration. The experiments were performed at the magnetic field with the most intense echo based on the echo-detected field sweep. Four-step phase cycling was used to eliminate the undesired echoes. The data were acquired for ca. 20 min. A stretched exponential decay was fitted to and subtracted from the time-domain raw ESEEM signal. Hamming was applied to filter the background noise, followed by zero filling with 2048 points. Fast Fourier Transform was employed and the absolute value was taken to form the ESEEM spectra, which has been shown in Fig. S1B.

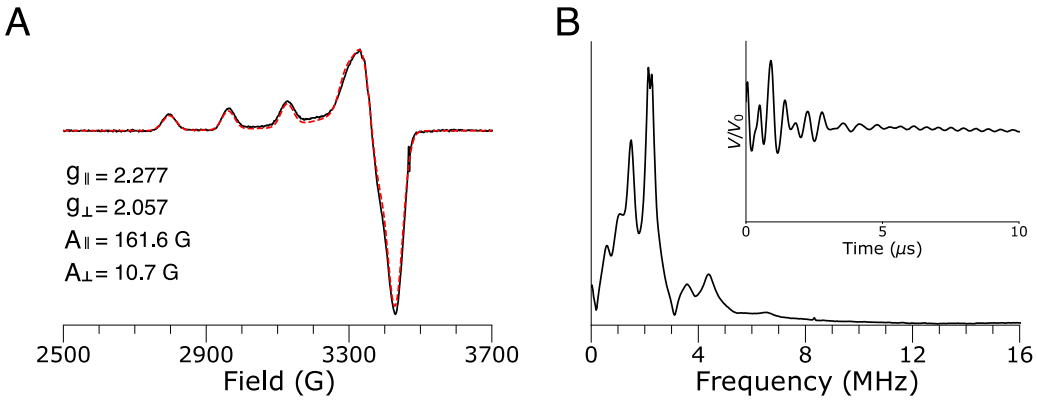

Fig. S1: (A) CW simulations of dHis-Cu-NTA labeled hGSTA1-1 with GSHex. The simulation was performed using EasySpin.[3] The black solid is the experimental data, and the red dashed is the best fit simulation. One-component fit was obtained with $g_{\parallel} = 2.277$, $g_{\perp} = 2.057$, $A_{\parallel} = 161.6$ G and $A_{\perp} = 10.7$ G. The $g_{\parallel}$ and $A_{\parallel}$ values are consistent with multi-nitrogen coordination to Cu^{2+}.[4,5] (B) ESEEM experiment was performed on the same sample. Characteristic peaks around 0.5, 1, 1.5 MHz, and ca. 4 MHz suggests histidine coordination to Cu^{2+}. The double peaks around 2 MHz indicate solvent deuterium. The raw time-domain of the ESEEM is shown as inset.

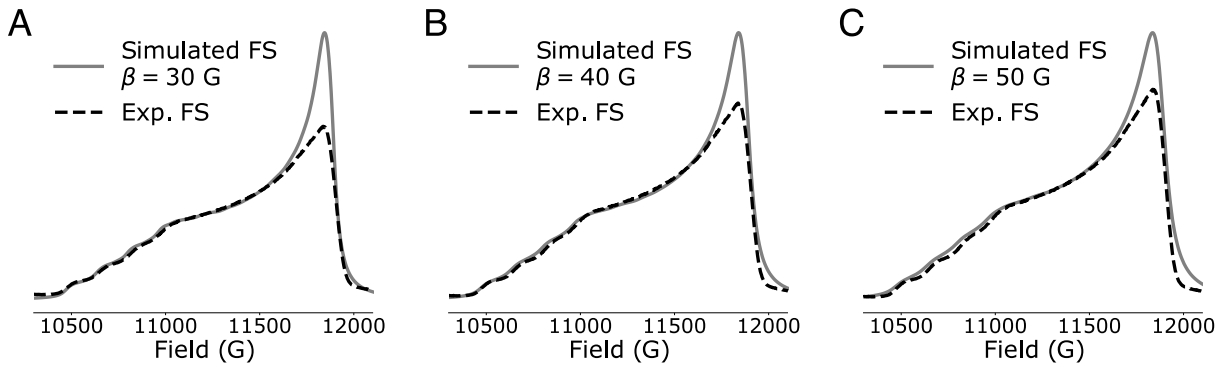

Fig. S2: Fitting of experimental FS-ESE spectrum of dHis-Cu^{2+} labeled hGSTA1-1. The simulated FS-ESE is systematically tested with different Lorentzian broadening parameter, β, values of (A) 30 G, (B) 40 G, and (C) 50 G. We note that phase memory times vary across the magnetic field,[6] which was not accounted for in our FS-ESE simulations. Hence, there may exist some difference in the intensity of the experimental versus simulated spectrum. Nevertheless, with a β of 30 G, the intensity at the $g_\perp$ region of ca.11800 G is not properly fitted. However, a β of 50 G leads to a loss of features at the $g_\parallel$ region of ca.10500 G to ca.11000 G. Therefore, β of 40 G is a good compromise between the features at $g_\parallel$ and the intensity at $g_\perp$.

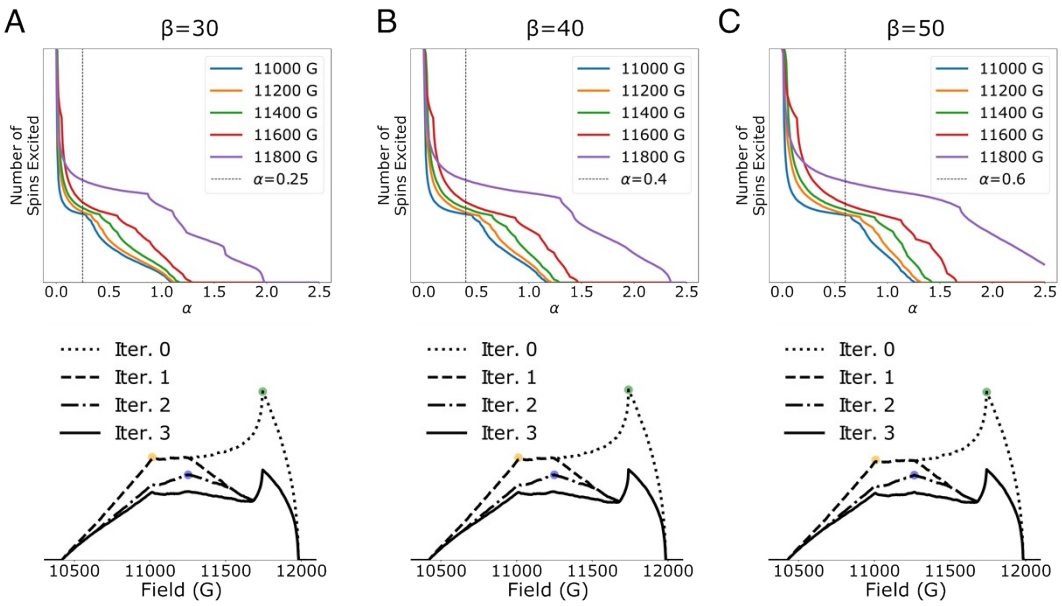

Fig. S3: Plots of the number of excited spins as a function of threshold parameter, α, at different fields across the simulated FS-ESE spectrum for β of (A) 30, (B) 40, and (C) 50. At each field, the plot contains three regions. The first region contains a sharp decay of the number of the excited spins at low α values. This region indicates that low α leads to oversampling of spins. This is a result of the slow decay of I(B) since each Lorentzian goes to 0 only at a magnetic field of infinity. The second region shows a plateau where the change in α does not significantly change the number of excited spins. The third region shows a further decay of number of spins at high α values. In this region the choice of α leads to a undersampling. The optimal α value for each in β was chosen using the criteria that this value provides a stable count (i.e. is in the flat region) for all fields. This α value is shown by dashed line in the top panel. The bottom panel shows the Φ curve and the three identified fields for each β value with the corresponding optimal α value shown in the top panel.

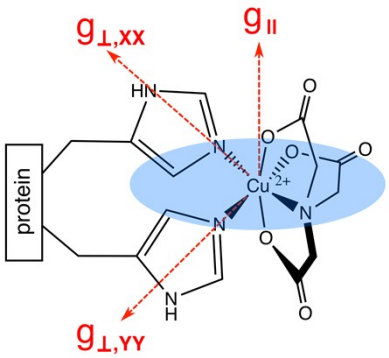

Fig. S4: Directions of $g_{\parallel}$ and $g_{\perp}$ within the dHis-Cu^{2+}-NTA complex. dHis-Cu^{2+}-NTA is an octahedral complex with Cu^{2+} being a d^9 system. Such complexes experience Jahn-Teller effect, leading to the axial elongation. The imidazole nitrogen atoms bind to Cu^{2+} equatorially, leading to the $g_{\parallel}$ perpendicular to the equatorial plane.

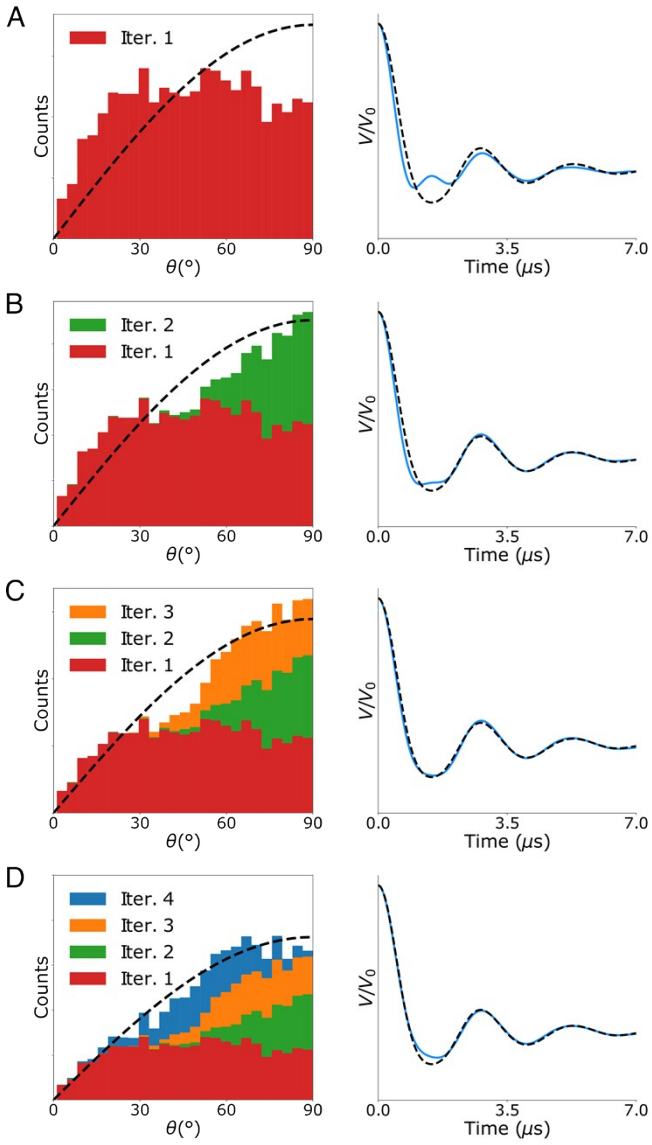

Fig. S5: The left column shows the distribution of probed θ from DEER simulations with a ca.38 MHz observer pulse and 100 MHz pump pulse after iteration 1(A), iteration 2 (B), iteration 3 (C), and iteration 4 (D). Iteration 1, 2, 3 and 4 represent DEER simulations at 100 G, ca.827 G, ca.580 G and ca.340 G lower than the maximum of the ESE-FS spectrum. The first three fields are identified in Fig. 5. The dashed line represents an ideal sinusoidal probability of θ distribution. Overall, only a slight improvement in the sampling of θ, when compared to Fig. 5A in the main text, is observed. The right column of the figure shows the DEER intra-molecular signal from the corresponding probed θ distribution shown in the left. The solid blue line represents the probed DEER signal. The dashed black line represents an ideal DEER signal, assuming all spin-pairs in the *in-silico* model are excited. The fourth DEER does not improve the DEER signal when compared to the DEER from only three fields, shown in Fig. 5B of the main text.

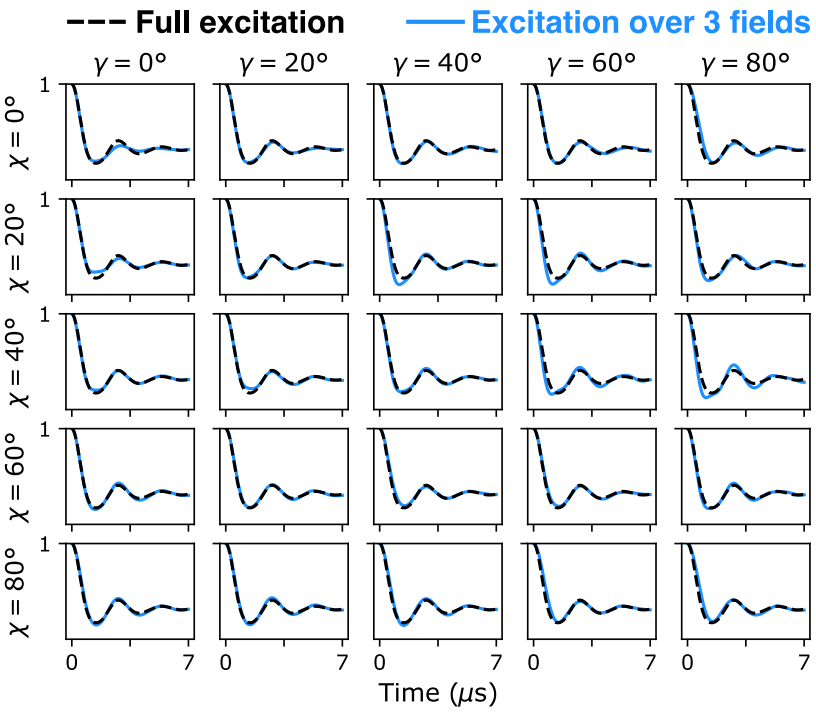

Fig. S6: Averaged DEER time traces simulated with the optimal DEER collection method. The solid blue curves are the simulated DEER time traces, compared to a DEER time trace with all spins excited. The simulation set $\eta = 20°$ and $\sigma_\gamma = \sigma_\chi = \sigma_\eta = 10°$, and iterates through γ and χ from $0°$ to $90°$.

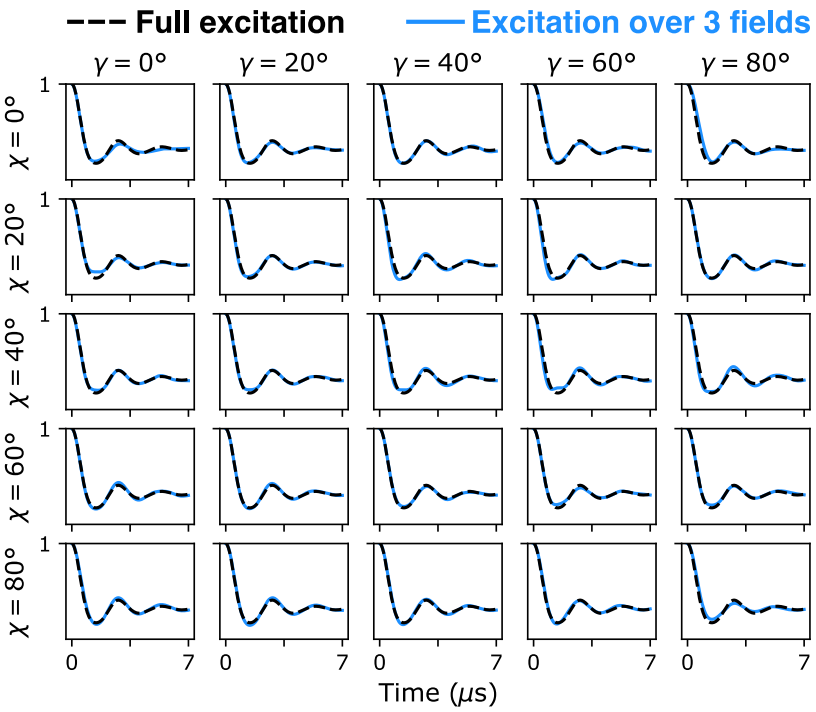

Fig. S7: Averaged DEER time traces simulated with the optimal DEER collection method. The solid blue curves are the simulated DEER time traces, compared to a DEER time trace with all spins excited. The simulation set $\eta = 40°$ and $\sigma_\gamma = \sigma_\chi = \sigma_\eta = 10°$, and iterates through γ and χ from $0°$ to $90°$.

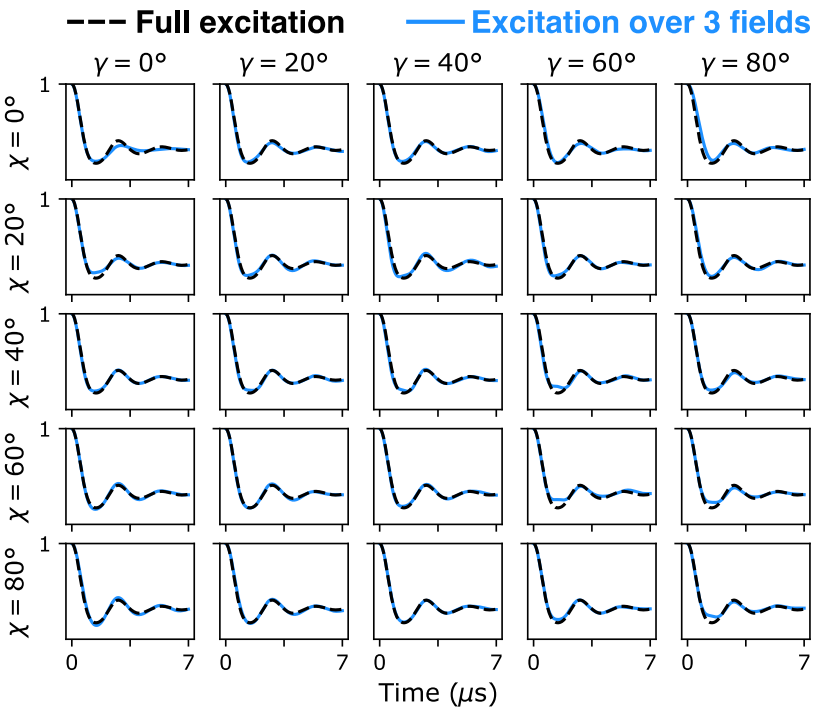

Fig. S8: Averaged DEER time traces simulated with the optimal DEER collection method. The solid blue curves are the simulated DEER time traces, compared to a DEER time trace with all spins excited. The simulation set $\eta = 60°$ and $\sigma_\gamma = \sigma_\chi = \sigma_\eta = 10°$, and iterates through γ and χ from $0°$ to $90°$.

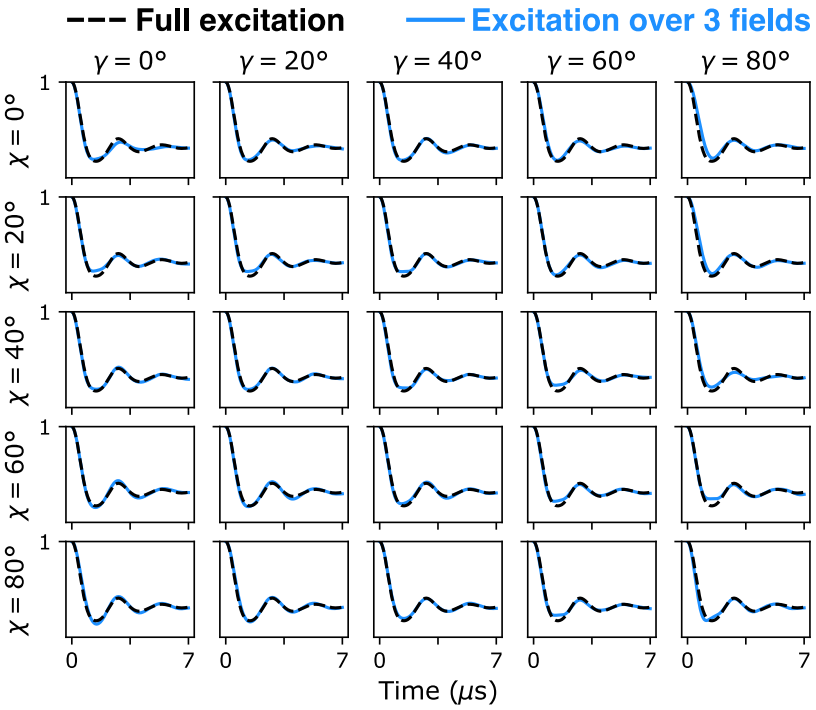

Fig. S9: Averaged DEER time traces simulated with the optimal DEER collection method. The solid blue curves are the simulated DEER time traces, compared to a DEER time trace with all spins excited. The simulation set $\eta = 80°$ and $\sigma_\gamma = \sigma_\chi = \sigma_\eta = 10°$, and iterates through γ and χ from $0°$ to $90°$.

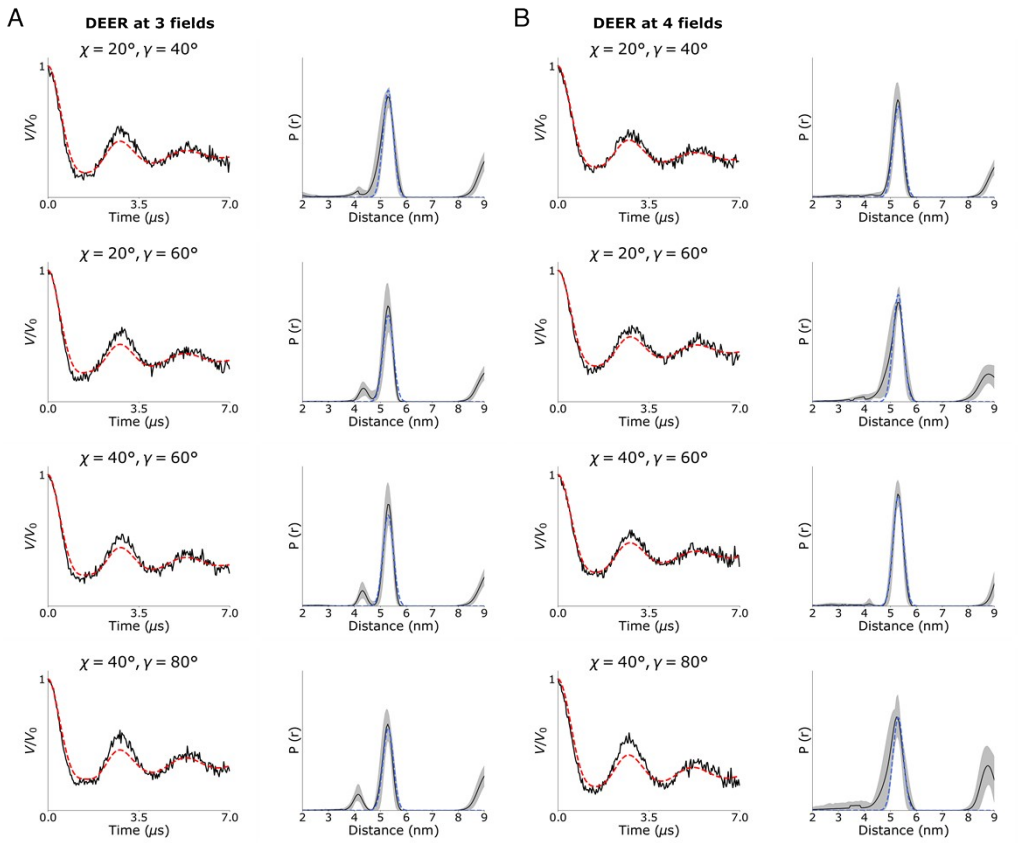

Fig. S10: The left column in (A) and (B) shows the averaged DEER time traces simulated at three and four magnetic fields, respectively. Random noise has been added to the time-domain signals to represent real-life measurements. In addition to the three identified fields, one more field at 338 G lower than the maximum of FS-ESE spectrum make the four-field acquisition scheme. The solid black curves are the simulated DEER time traces, and the red dashed curves are the best fit from DEERAnalysis,[7] for the corresponding angles. The right columns in (A) and (B) show the resulting distance distributions as black solid curves generated by CDA[8,9] for $\eta = 0°$, $\chi = 20°$ and $\gamma = 40°$; $\eta = 0°$, $\chi = 20°$ and $\gamma = 60°$; $\eta = 0°$, $\chi = 40°$ and $\gamma = 60°$; $\eta = 0°$, $\chi = 40°$ and $\gamma = 80°$ respectively, that have shown some deviation from the expected DEER signal in Fig. 8. The uncertainty in the distributions is shown in grey shading. The blue dashed curve shows the expected distance distribution.

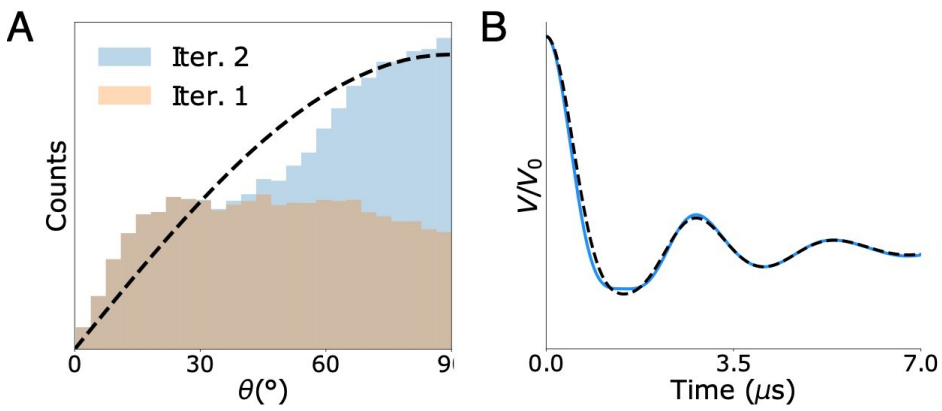

Fig. S11: (A) Distribution of probed θ from two DEER simulations with a ca.38 MHz observer pulse and 300 MHz pump pulse. Iteration 1 and Iteration 2 represents DEER performed at 100 G and ca.740 G lower than the maximum of the ESE-FS spectrum. The dashed line represents an ideal sinusoidal probability of θ distribution. (B) The DEER intra-molecular signal from the DEER simulation is shown as the solid blue line. The black dashed line represents an ideal DEER signal assuming all spin-pairs in the *in-silico* model can be excited.

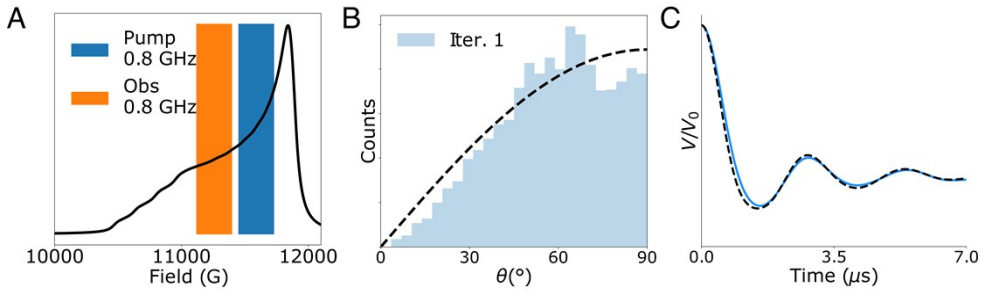

Fig. S12: (A) Pump and observer pulses exciting regions of the FS-ESE spectrum, depicted by the blue and orange region, respectively. The pump excitation is centered at 11592 G, while the observer excitation is centered at 11259 G. These pulses represent a total excitation bandwidth of 1.6 GHz. (B) Distribution of probed θ from a single DEER simulation with a 0.8 GHz observer pulse and 0.8 GHz pump pulse at regions depicted by (A). The dashed line represents an ideal sinusoidal probability of θ distribution. (C) The DEER intra-molecular signal from the DEER simulation is shown as the solid blue line. The dashed black line represents an ideal DEER signal assuming all spin-pairs in the in-silico model can be excited.

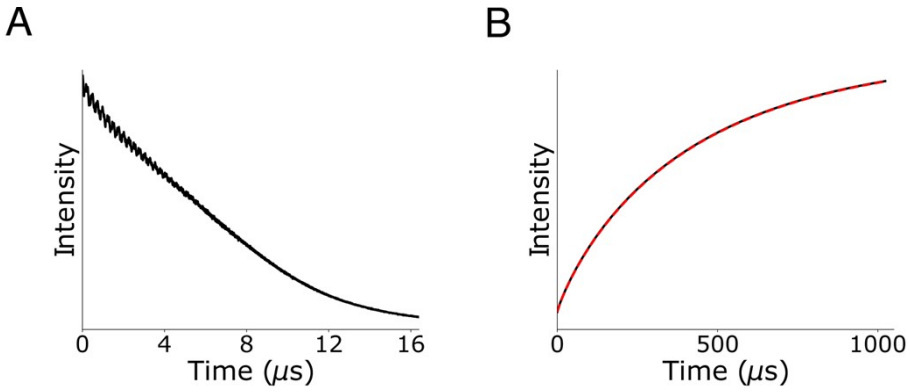

Fig. S13: (A) Primary Hahn echo decay data. Due to the TWT gate time limit, the full two-pulse decay could not be collected. The best fit using the stretched exponential decay shows a phase memory time, T_m, of over 9 μs. (B) Primary inversion recovery data and fit of the stretched exponential decay with the form $I(t) = c * [1 - 2 * e^{-(\frac{t}{T_1})^y}]$ for the 800 μM Cu^{2+}-NTA labeled hGSTA1-1 mutant. The best fit, shown as the dashed red line, suggests that $T_1 = 418 \pm 84\ \mu s$.

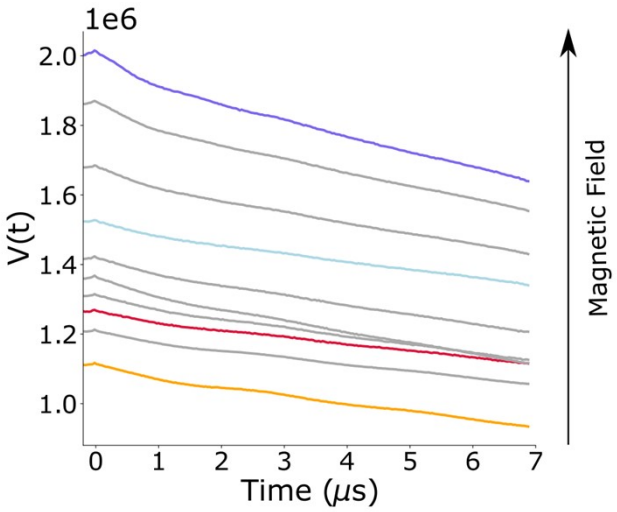

Fig. S14: The primary DEER time domain collected at each magnetic field shown in Fig. 10A, normalized to the intensity of the field-swept spectrum.

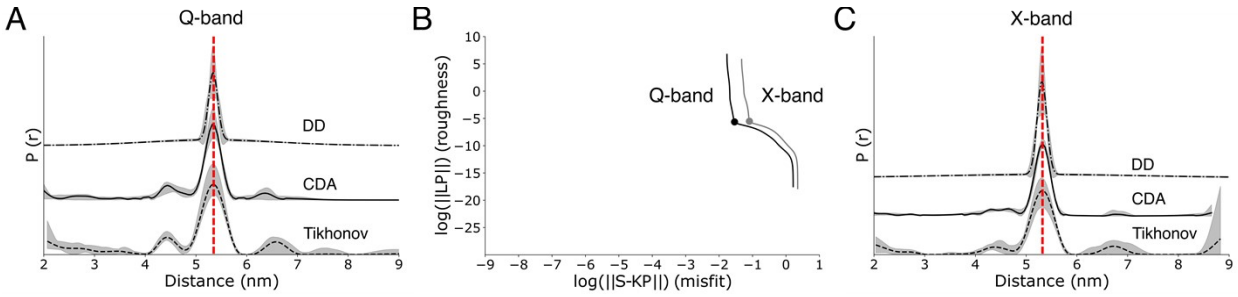

Fig. S15: (A) The Q-band DEER data collected at three magnetic fields was analyzed using Tikhonov regularization in black dashed curve,[10] comparative DEER Analyzer (CDA) in black solid curve,[8,9] as well as DD[11–13] in black dash-dot curve. The uncertainty in the distributions is presented in grey shade. Only the 5.3 nm distance, labeled by the vertical red dashed line, is consistently shown by all analysis tools. (B) L-curves by Tikhonov regularization[10] and α-values in full circles are for the summed Q-band DEER using the minimal DEER collection method (shown in black) and X-band DEER (shown in grey). (C) The X-band DEER data was analyzed using Tikhonov regularization in black dashed curve,[10] CDA in solid black curve,[8,9] as well as DD[11–13] in black dash-dot curve. The uncertainty in the distributions is presented in grey shade. Only the 5.3 nm distance, labeled by the vertical red dashed line, is consistently shown by all analysis tools. The X-band distance distribution agrees with the Q-band DEER results using the $g_{\parallel}$ skewed method.

References

1 W. B. Mims, Envelope Modulation in Spin-Echo Experiments, *Phys Rev B*, 1972, **5**, 2409–2419.
2 W. B. Mims, Amplitudes of Superhyperfine Frequencies Displayed in the Electron-Spin-Echo Envelope, *Phys. Rev. B*, 1972, **6**, 3543–3545.
3 S. Stoll and A. Schweiger, EasySpin, a comprehensive software package for spectral simulation and analysis in EPR, *J. Magn. Reson.*, 2006, **178**, 42–55.
4 J. Peisach and W. E. Blumberg, Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins, *Arch. Biochem. Biophys.*, 1974, **165**, 691–708.
5 K. I. Silva, B. C. Michael, S. J. Geib and S. Saxena, ESEEM Analysis of Multi-Histidine Cu(II)-Coordination in Model Complexes, Peptides, and Amyloid-β, *J. Phys. Chem. B*, 2014, **118**, 8935–8944.
6 S. Saxena and J. H. Freed, Two-Dimensional Electron Spin Resonance and Slow Motions, *J. Phys. Chem. A*, 1997, **101**, 7998–8008.
7 G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger and H. Jung, DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data, *Appl. Magn. Reson.*, 2006, **30**, 473–498.
8 S. G. Worswick, J. A. Spencer, G. Jeschke and I. Kuprov, Deep neural network processing of DEER data, *Sci. Adv.*, 2018, **4**, eaat5218.
9 L. Fábregas Ibáñez, G. Jeschke and S. Stoll, DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data, *Magn. Reson.*, 2020, **1**, 209–224.
10 G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger and H. Jung, DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data, *Appl. Magn. Reson.*, 2006, **30**, 473–498.
11 S. Brandon, A. H. Beth and E. J. Hustedt, The global analysis of DEER data, *J. Magn. Reson.*, 2012, **218**, 93–104.
12 P. Z. Qin and K. Warncke, *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions Part B*, Academic Press, 2015.
13 E. J. Hustedt, F. Marinelli, R. A. Stein, J. D. Faraldo-Gómez and H. S. Mchaourab, Confidence Analysis of DEER Data and Its Structural Interpretation with Ensemble-Biased Metadynamics, *Biophys. J.*, 2018, **115**, 1200–1216.