Supporting Information

The double histidine Cu^{2+} binding motif – A highly rigid, site-specific spin probe for electron spin resonance distance measurements

Timothy F. Cunningham, Miriam R. Putterman, Astha Desai, W. Seth Horne, and Sunil Saxena

Department of Chemistry, University of Pittsburgh, 219 Parkman Ave., Pittsburgh, PA 15260

Materials and Methods

Construction of GB1 Mutants. The plasmid encoding for wild type GB1 was kindly provided by Prof. Angela Gronenborn (University of Pittsburgh). The histidine mutations were chosen to incorporate the double histidine motif into the α-helix (28H/32H) and the β-sheet (6H/8H). In order to compare Cu^{2+}-Cu^{2+} distance with MTSL based labels, a homologous double cysteine mutant was created with an α-helix site (28C) and a β-sheet site (6C) to be converted to R1. Mutations were performed one at a time and mutant plasmid was used to create the final desired mutants (6H/8H, 28H/32H, 6H/8H/28H/32H, and 6C/28C). Each mutant was created using the appropriate plasmid DNA, the primers encoding for the desired mutation (Invitrogen, Carlsbad, CA), and the KAPA Hifi Hotstart Ready Mix (Kapa Biosystems, Cape Town, South Africa). Resultant PCR reaction mixtures were treated with DpnI (New England Biolabs, Boston, MA), transformed into XL1-Blue Supercompetent cells (New England Biolabs), and grown overnight on culture plates containing Luria-Bertani (LB) broth with 100 mg/mL ampicillin. Colonies were picked and grown overnight in 50 mL of LB with 100 mg/mL ampicillin and the plasmid DNA was purified using the PureYield Plasmid Midiprep System (Promega, Madison, WI). All mutations were confirmed by DNA sequencing (Genomic and Proteomics Core Laboratories, University of Pittsburgh, Pittsburgh, PA) and subsequently transformed into BL21(DE3) competent cells (New England Biolabs) for expression.

Protein Expression, Purification, and Labeling. Expression/purification of all GB1 mutants and the labeling of the 6C/28C mutant with MTSL was performed as was described previously.[1]

Crystallization, Data Collection, and Structure Determination. Crystals of the 6H/8H/28H/32H mutant of GB1 were grown by hanging drop vapor diffusion method. 1 μL of an 18 mg/mL solution of protein in water was mixed with 1 μL of a crystallization buffer composed of 0.2 M NaCl, 0.1 M HEPES pH 7.0, and 1.75 M $(NH_4)_2SO_4$. The drop was allowed to equilibrate at room temperature over a well containing the crystallization buffer. A single crystal was flash frozen in liquid nitrogen after cryoprotection in crystallization buffer supplemented with 30% glycerol by volume. Diffraction data were

collected using CuKα radiation on a Rigaku/MSC diffractometer (FR-E generator, Saturn 944 CCD detector) equipped with an X-Stream 2000 low temperature system operated at 100 K. Raw diffraction data were processed with d*TREK. Structure determination and refinement were carried out using the Phenix software suite. The structure was solved by molecular replacement using a published structure of wild-type GB1 (PDB 2QMT) as the search model. Coordinates and structure factors have been deposited in the Protein Data Bank (PDB: 4WH4). Data collection and refinement statistics are given in Table S1.

Data Collection	
Unit cell dimensions (Å, °)	$a = b = 74.3$, $c = 41.3$ $\alpha = \beta = 90$, $\gamma = 120$
Space group	$P3_121$
Resolution (Å)	32.17–2.20 (2.28–2.20)
Total observations	45,773
Unique observations	6,876
Redundancy	6.7 (4.8)
Completeness (%)	99.7 (98.4)
I/σ	14.0 (2.9)
R_{merge} (%)	7.0 (37.3)
Refinement	
Resolution (Å)	32.17–2.20
R (%)	20.7
R_{free} (%)	25.6
Avg. B factor (Å^2)	44.4
RMSD	
Bonds (Å)	0.004
Angles (°)	0.65

Table S1.

Circular Dichroism. CD data were collected using an Olis DSM17 Circular Dichroism Spectrometer. All samples were 40 μM protein in 20 mM sodium phosphate buffer (pH 7.0) and measured in 2 mm

quartz cells. 120 μM of both Cu^{2+} and IDA were included for the appropriate samples. The scans were collected at a temperature of 25 °C from 200 nm to 260 nm with 1 nm increments and a 2 nm bandwidth. The melts were collected by monitoring the signal at 220 nm from 4 °C to 98 °C in 2 °C increments with a 0.5 °C dead band and 2 min equilibration time at each temperature. All measurements were baseline corrected with buffer alone, or buffer including the Cu^{2+} and IDA for the appropriate samples. The melts were fit to a two-state unfolding model to calculate the melting temperatures (Table S2).

ESR Measurements. All ESR experiments were performed on a Bruker Elexsys 580 spectrometer equipped with a Bruker ER4118X-MD5 resonator or a Bruker ElexSys E680 X-band FT/CW spectrometer equipped with a Bruker EN4118X-MD4 resonator. The temperature for all experiments was controlled using an Oxford ITC503 temperature controller with an Oxford ER 4118CF gas flow cryostat. All samples were 150 μL with protein in 50 mM N-ethyl morpholine (NEM) at pH 7.4, 25% glycerol, isotopic $^{65}CuCl_2$ (Cambridge Isotopes, Tewksbury, MA), and iminodiacetic acid (IDA) if the IDA was specified as part of the sample. Protein, Cu^{2+}, and IDA concentrations are dependent on the experiment performed.

Continuous wave experiments were carried out at 80 K at X-band frequencies at a 1:1:1 ratio of Cu^{2+}:IDA:GB1 with a GB1 concentration of 500 μM. The field was swept from 2200 G to 4200 G for 1024 points with a modulation amplitude of 4 G, a modulation frequency of 100 kHz, and an incident power of 0.1992 mW. All CW spectra were fit with the Bruker Simfonia software.

Electron spin echo envelope modulation experiments were performed at 80 K at a 2:2:1 ratio of Cu^{2+}:IDA:GB1 with a GB1 concentration of 500 μM. A $\pi/2–\tau–\pi/2–T–\pi/2–$echo pulse sequence was used with a $\pi/2$ pulse length of 16 ns and the magnetic field fixed at the maximum of 3350 G. The first time delay, τ, was 144 ns and the second time delay, T, started at 288 ns and was incrementally increased at steps of 16 ns. The resultant signal was phase corrected and the real part was fast Fourier transformed using the Bruker WinEPR software.

The DEER distance measurement was performed on the MTSSL labeled mutant as described previously.[1] The Cu^{2+} distance measurement was performed at 20K. The four pulse sequence utilized for the measurements was $(\pi/2)\nu_1$-τ_1-$(\pi)\nu_1$-T-$(\pi)\nu_2$-τ_2-$(\pi)\nu_1$-τ_2-echo. For the distance measurements displayed in Figure 2c, the pump pulse (ν_2) was placed at the maximum of the Cu^{2+} spectrum and the observer pulses (ν1) are placed 150 MHz downfield from the maximum. The observer pulse (ν1) lengths were 16ns and 32ns for the $\pi/2$ and π pulses, and the pump pulse (ν_2) length was 16ns. The parameter τ_1 was set to 200 ns and T to 160 ns initially with T being increased by 10ns for 128 steps. τ_2 was adjusted such that τ_2 + T = 1300 ns. Shot repetition time was set to 2000 μs and collection time was approximately 10 hours for the R1 measurement, and for the Cu^{2+} measurement, shot repetition time was set to 300 μs and the collection time was approximately 16 hours.

Results and Discussion

Circular Dichroism Scans and Melts. Circular dichroism was used to compare the global structure and stability among wild-type GB1 and its various mutants. In CD scans (Figure S1), the signatures of all mutants are similar in shape and magnitude to WT indicating that the fold of GB1 is maintained, including all histidine mutants in the presence of Cu^{2+} and IDA. Analysis of temperature-induced unfolding (Figure S2) reveals cooperative two-state unfolding behavior in all samples, consistent with wild-type GB1. The melting temperatures, T_m, are somewhat decreased in the mutants indicating a loss in stability regardless of cysteine or histidine modifications (Table S2). Interestingly, the 28H/32H and 6H/8H/28H/32H mutants show increased T_m in the presence of Cu^{2+}-IDA. This observation is consistent with similar stabilization observed previously for a dHis site in an α-helix.[2] The destabilization of the protein upon His incorporation in the β-sheet is consistent with the low sheet folding propensity of the residue.[3]

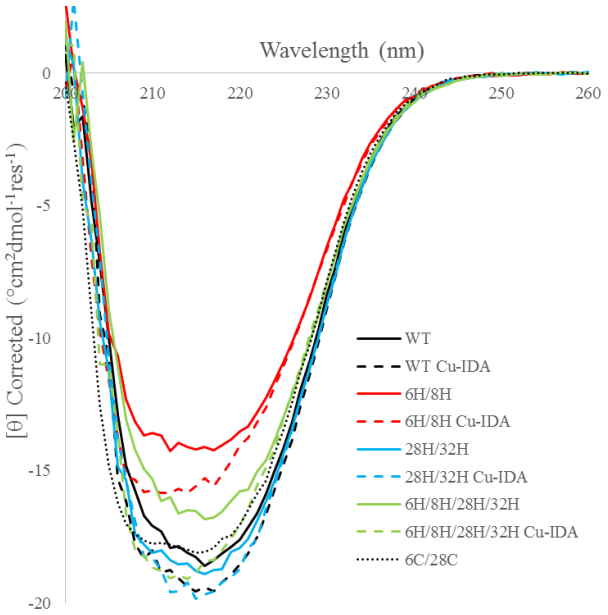

Figure S1: The CD scans of WT, 6H/8H, 28H/32H, and 6H/8H/28H/32H GB1 with and without Cu^{2+}-IDA as well as 6C/28C.

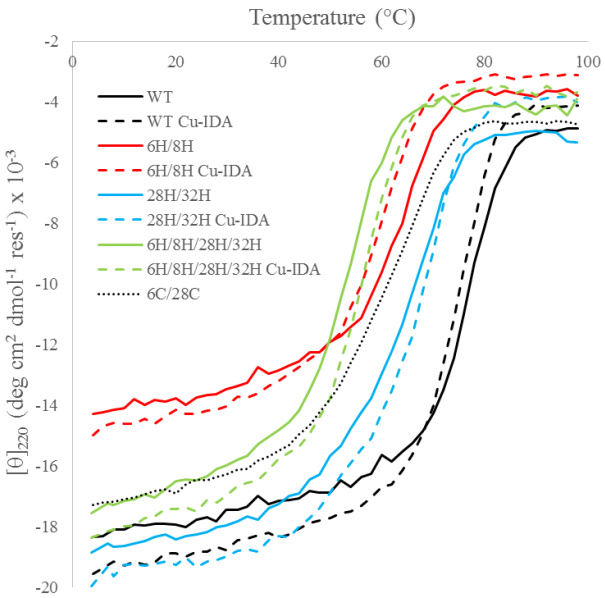

Figure S2: The CD thermal melts of WT, 6H/8H, 28H/32H, and 6H/8H/28H/32H GB1 with and without 3:1 Cu^{2+}-IDA, in addition to 6C/28C, monitored at 220 nm.

	No Cu^{2+}	**Cu^{2+}-IDA**
WT	**77.2±0.3**	**75.3±0.2**
6H/8H	**64.4±0.2**	**60.1±0.2**
28H/32H	**66.2±0.3**	**68.1±0.5**
6H/8H/28H/32H	**54.0±0.2**	**56.8±0.1**
6C/28C	**62.1±0.3**	**-**

Table S2: The melting temperatures (°C) of each sample calculated from fits of the CD melts.

Cu^{2+}-IDA, Cu^{2+}-WT GB1, and Cu^{2+}-IDA-WT GB1 ESR Measurements. As a means to assess binding of the Cu^{2+} and IDA, ESR data was collected on the complex alone as a control. In NEM buffer, free Cu^{2+} is ESR silent[4] and thus the Cu^{2+}-IDA CW signal as seen in Figure S3 is arises from the binding of the Cu^{2+} to the IDA. Additionally, ESEEM data was collected on the complex and the observed signal is shown in Figure S4. As ESEEM only reports on atoms which are remotely coordinating (5 Å to 8 Å away), only ^{1}H ESEEM from solvent waters is expected and indeed observed given the binding environment shown in Figure 1b.

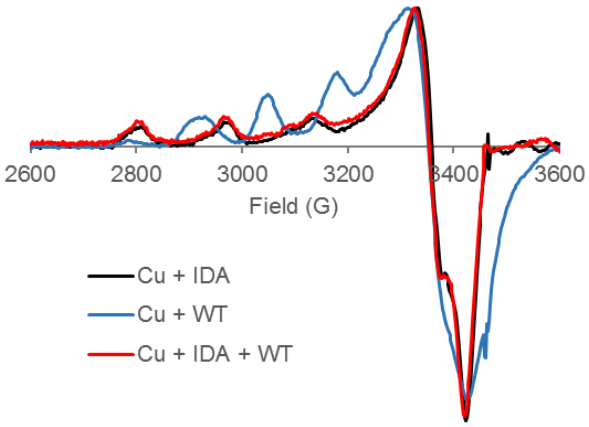

Figure S3: Overlay of the continuous wave spectra for Cu^{2+}-IDA, Cu^{2+}-WT GB1, Cu^{2+}-IDA-WT GB1.

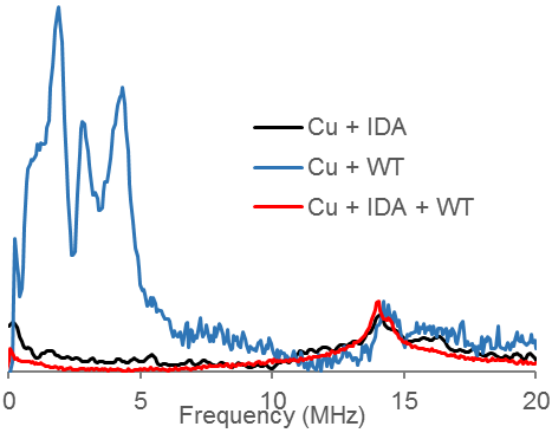

Figure S4: Overlay of the ESEEM spectra for Cu^{2+}-IDA, Cu^{2+}-WT GB1, Cu^{2+}-IDA-WT GB1. The spectra have been normalized according to ^{1}H ESEEM intensity.

One potential hindrance of using the dHis motif is that the concentration of Cu^{2+} necessary to fully load all sites could potentially lead to nonspecific binding of the Cu^{2+} to undesired locations of the protein. This tends to be common for many proteins given the prevalence of Cu^{2+} binding residues and this was indeed the case for GB1 as nonspecific binding had been observed previously through NMR.[5] CW and ESEEM spectra were collected on WT GB1 and Cu^{2+}, as shown in Figures S3 and S4 respectively. The CW spectrum shows a signal which in NEM indicates that Cu^{2+} is indeed binding. Additionally, when compared to the signal of Cu^{2+} and the dHis mutants in Figure 2b, the signals are distinct, which indicates different binding environments. The ESEEM spectrum shows the presence of remotely coordinating nitrogen atoms, most likely a backbone nitrogen given the peak observed at 2.8 MHz.[6]

Similar experiments were performed with WT GB1, Cu^{2+}, and IDA. The CW spectrum (Figure S3) displays a signal very similar to Cu^{2+}-IDA alone. Furthermore, the ESEEM spectrum displays the removal of the nitrogen signal observed for the WT GB1 and Cu^{2+} alone and is similar to the minimal ESEEM spectrum observed for Cu^{2+}-IDA alone (Figure S4). Thus as intended, the IDA is preventing the Cu^{2+} from binding to the nonspecific site. Additional experiments displayed that IDA added to WT GB1 already coordinating Cu^{2+} was able to entirely remove the nonspecific binding.

CW of dHis Mutants and Fits. As a means to assess the directly coordinating atoms in the dHis mutants, CW spectra were collected for both 6H/8H-Cu^{2+}-IDA and 28H/32H-Cu^{2+}-IDA. Both spectra and subsequent fits are displayed in Figure S5 and the fitting parameters are in the main text. The parameters of $A_{\parallel}$ and $g_{\parallel}$ report directly on the number of coordinating nitrogens and oxygens and values required for the fit match well with the expected binding environment shown in Figure 2b.

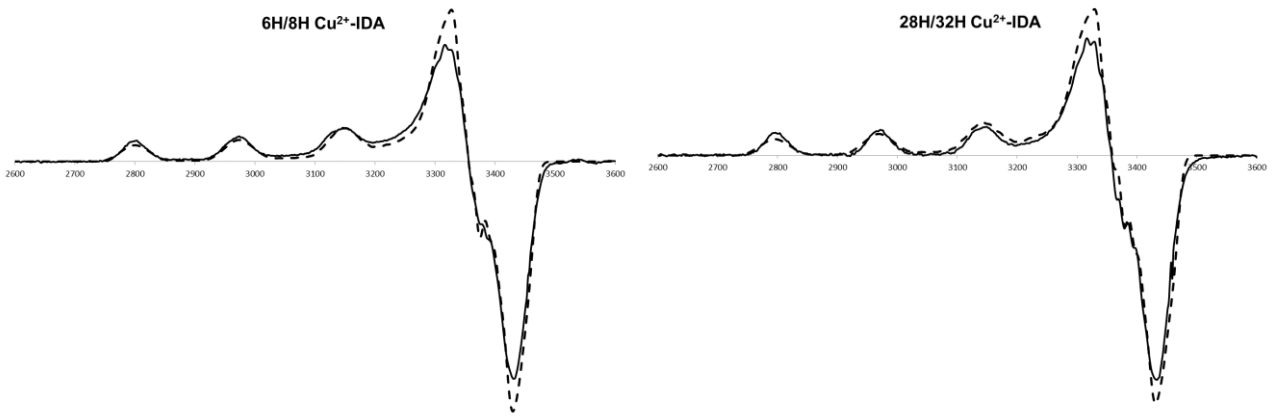

Figure S5: CW spectra of 6H/8H and 28H/32H GB1 with Cu^{2+}-IDA (solid) and the corresponding fits (dashed) from Bruker Simfonia software. Fittings parameters are in the main text.

Comparison of dHis Cu^{2+} ESEEM to Model Complexes. The ESEEM spectra of 6H/8H-Cu^{2+}-IDA and 28H/32H-Cu^{2+}-IDA are shown in Figure 1c and the relative similarity of the spectra indicates similar binding environments at each site despite the differences in local environment. As a means to establish the simultaneous coordination of both histidines at both sites, the ESEEM data was compared to the resultant ESEEM spectra of model histidine complexes (Figure S6b) that we have investigated previously.[7] As discussed in the main text, the ratio of NQI peaks to the DQ peaks will decrease as the number of coordinating histidines increases. Accordingly, all data including model complexes and dHis samples were normalized by max intensity of the NQI feature and the relative intensity of the DQ feature was assessed. As can be readily observed qualitatively in Figure S6, the relative DQ intensity of both dHis mutants matches closest to that of the two histidine complex. While not an exact match, this comparison in addition to the small 8 MHz feature indicated in the main text support that at both dHis sites, both histidines are involved in the coordination of the Cu^{2+} in the presence of IDA.

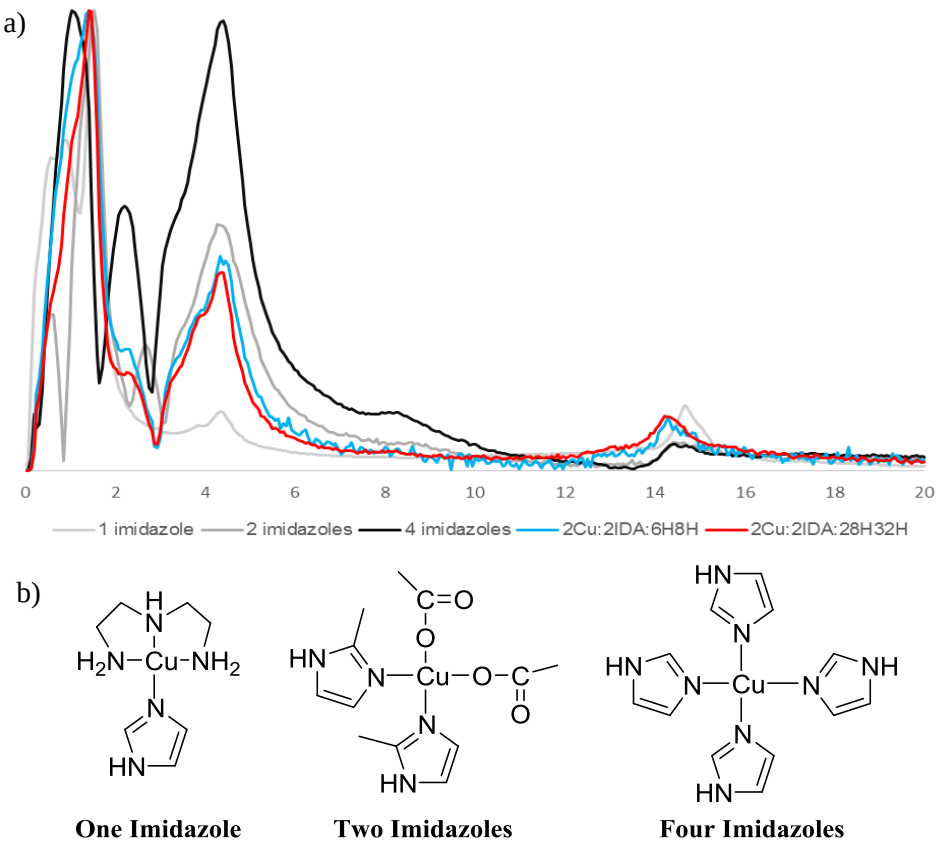

Figure S6: (a) Comparison of the ESEEM spectra of the 6H/8H and 28H/32H mutants with Cu^{2+}-IDA compared to ESEEM spectra for three model imidazole complexes (b). The spectra are normalized to the maximum NQI intensity.

Investigation of the Final Axial Ligand. In Figure 1b, five of the six ligands chelating to the Cu^{2+} ion in a dHis-Cu^{2+}-IDA complex are shown. As a means to investigate the final ligand in the axial position, an ESEEM spectrum was collected for **28H/32H-GB1** with 1 equivalent of Cu^{2+}-IDA in D_2O with all experimental conditions kept identical. The resulting spectrum is displayed in Figure S7. As can be seen, the ^1H ESEEM peak at ~14 MHz has essentially disappeared indicating that solvent H_2O was the primary cause of this signal. The ^2H peak appears at ~2.1 MHz and the breadth of this peak can distinguish between equatorially coordinated, axially coordinated, and ambient D_2O.[8] The peak breadth observed in Figure S7 matches best with axial coordination.

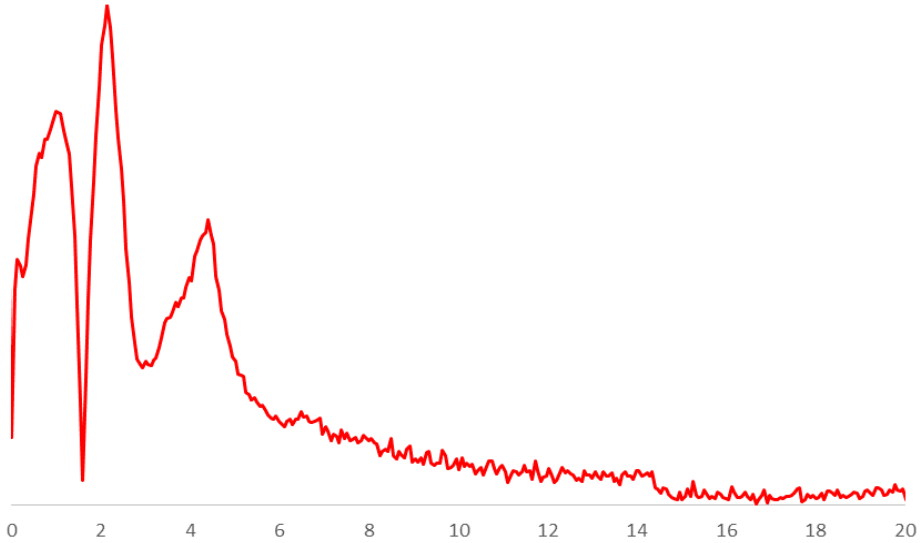

Figure S7: The ESEEM spectra of 1:1 Cu^{2+}-IDA:**28H/32H-GB1** in D_2O. The stark decrease of the ^1H peak at ~14 MHz and the breadth of the ^2H at ~2.1 MHz both suggest axial H_2O coordination (see text).

Probing for Orientational Selectivity in the Cu^{2+}-Cu^{2+} DEER Measurement. Orientational selectivity can be reduced in Cu^{2+}-DEER because of a distribution in relative orientations of the g-tensor[9-11] and due to the relative orientations of the g-tensors.[12, 13] Given the expected rigidity of the site, there may be a potential for orientational effects for Cu^{2+} chelated to the dHis sites. Accordingly, the DEER data was acquired at 2 positions in the $g_\perp$ and in the $g_\parallel$ region of the Cu^{2+} spectrum. The locations of the observer and pump pulses are shown in Figure S8a. The baseline subtracted time domain data (Figure S8b) and the resultant Pake patterns (Figure S8c) for each spectrum display no appreciable differences. Additionally, the time domain data is overlaid with the best single Gaussian fits (dotted red in Figure S8b) resulting in the distance distributions shown in Figure S8d. For these distributions, the most probable distance only varies by approximately 0.1 Å with only the breadth of the distributions changing, which is likely due to the lower signal to noise of positions B and C. Taken together, there are negligible orientational effects and as such, only the data from position A was presented in the main text (Figure 3).

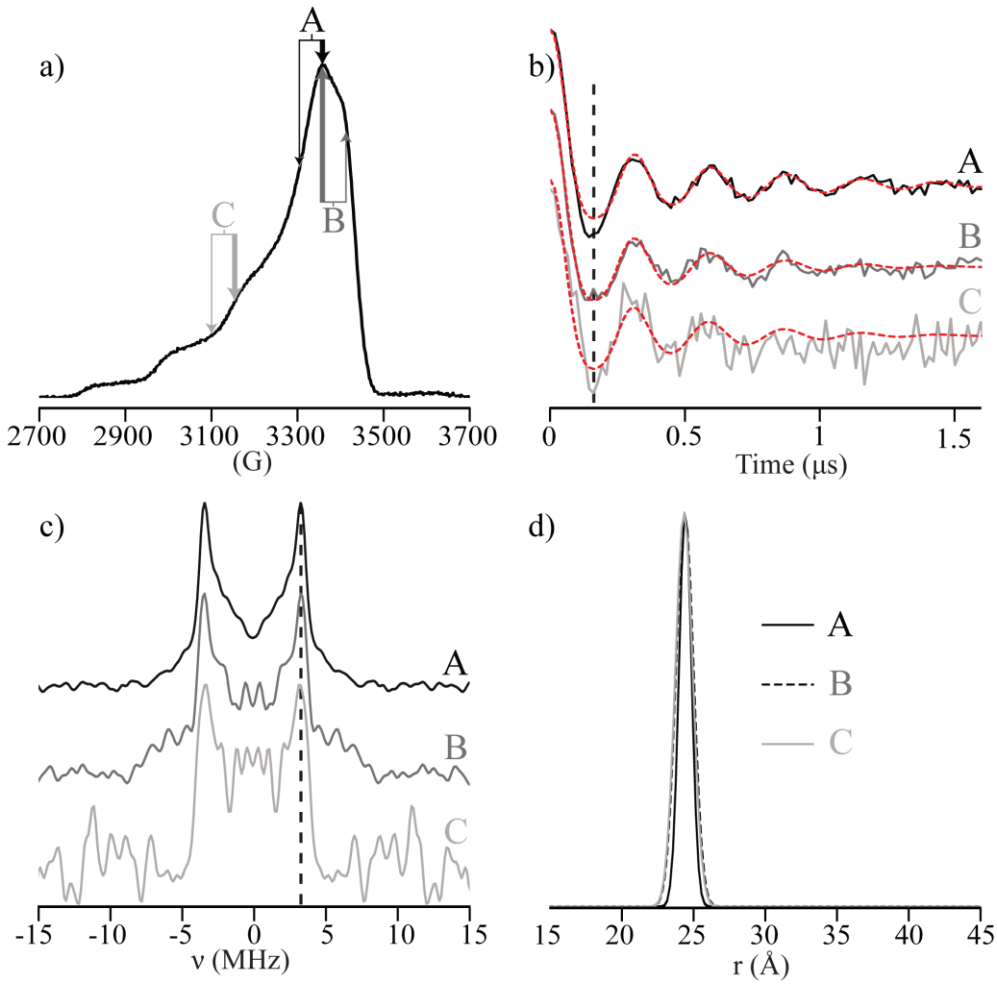

Figure S8: (a) Pulse positions (thin arrow indicate observer pulses and thick arrows indicated pump pulses) probed for orientational effects in the Cu^{2+}-Cu^{2+} DEER measurement. (b) The baseline subtracted time domain data and the best single Gaussian fits (dotted red lines), and (c) the Pake patterns for each of these measurements. (d) The single Gaussian distributions for each of the fits in (b).

Raw DEER Data and Modulation Depths. Due to the inherently smaller modulation depths of Cu^{2+} DEER measurements when compared to R1 measurements, the data in Figure 3 is displayed in different scales to allow for better visual comparison of the time domain signals. The raw data and baseline subtracted data for both measurements are displayed in Figure S9. As expected, the modulation depth for the Cu^{2+} measurement is indeed much smaller when compared to the R1 signal (~2.2% vs ~33% respectively). Despite the decreased modulation depth, quality signal to noise was achieved in a reasonable collection time (~16 hours).

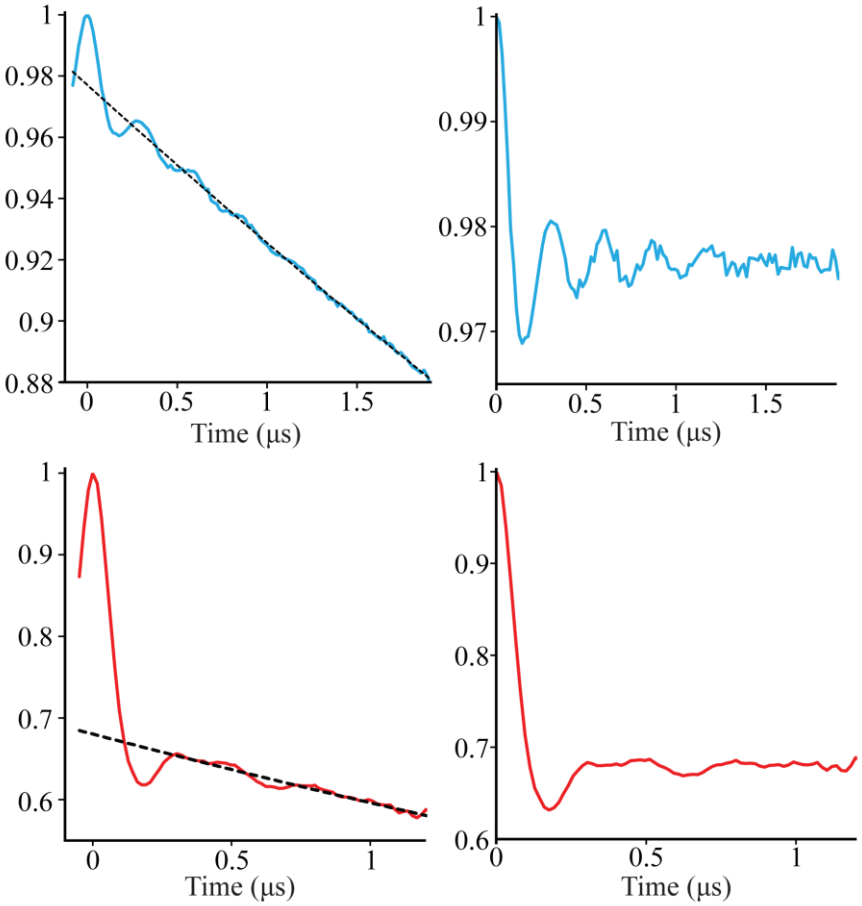

Figure S9: Raw DEER data (left panels) and baseline subtracted DEER data (right panels) for the Cu^{2+} distance measurement (top panels in blue) and the R1 distance measurement (bottom panels in red).

Loading Efficiency. For Cu^{2+} measurements in NEM buffer, free Cu^{2+} forms an insoluble and ESR silent complex[4] and thus simple titrations can be performed to assess binding. In the case of Cu^{2+}-IDA binding to the dHis motif, the added benefit of NEM is lost due to the Cu^{2+}-IDA complex being soluble, and thus ESR active. As a means to assure maximum loading of the dHis sites, 3 equivalents of Cu^{2+}-IDA was added to **6H/8H/28H/32H-GB1** for the DEER measurement. The modulation depth of the DEER measurement is dependent upon the number of spins per species, and with a modulation depth of ~ 2.2%, the calculated loading efficiency is ~ 60%. This number of 60% is an underestimate given that the presence of free spins decreases modulation depth. Regardless, quality signal to noise was achieved for this measurement in a reasonable amount of collection time.

Spin Density Distribution. The spin density of Cu^{2+} chelated to dHis and IDA was investigated with density functional theory calculations. The model constructed based off the crystal structure of Cu^{2+} chelated to two imidazoles and an IDA derivative was also used here with H_2O placed in the unoccupied axial position. Calculations were performed using the B3PW91[14] method with the def2-TZVP[15] basis set in ORCA.[16] The calculated atomic spin densities display 69% of the density on the Cu^{2+}, 6% on one directly coordinating histidine N, 2% on the other coordinating histidine N, 4% on the IDA N, 13% on the equatorial IDA O, 2% on the axial IDA O, and 2% on the water O. The effects of such small delocalization are expected to be minimal in DEER.[17, 18]

References

[1] Cunningham, T., McGoff, M., Sengupta, I., Jaroniec, C., Horne, W., and Saxena, S. (2012) High-resolution structure of a protein spin-label in a solvent-exposed β-sheet and comparison with DEER spectroscopy, *Biochemistry 51*, 6350-6359.

[2] Nicoll, A. J., Miller, D. J., Futterer, K., Ravelli, R., and Allemann, R. K. (2006) Designed high affinity Cu2+-binding alpha-helical foldamer, *J. Am. Chem. Soc. 128*, 9187-9193.

[3] Minor, D. L., Jr., and Kim, P. S. (1994) Measurement of the beta-sheet-forming propensities of amino acids, *Nature 367*, 660-663.

[4] Aronoff-Spencer, E., Burns, C. S., Avdievich, N. I., Gerfen, G. J., Peisach, J., Antholine, W. E., Ball, H. L., Cohen, F. E., Prusiner, S. B., and Millhauser, G. L. (2000) Identification of the Cu2+ binding sites in the N-terminal domain of the prion protein by EPR and CD spectroscopy, *Biochemistry 39*, 13760-13771.

[5] Nadaud, P. S., Sengupta, I., Helmus, J. J., and Jaroniec, C. P. (2011) Evaluation of the Influence of Intermolecular Electron-Nucleus Couplings and Intrinsic Metal Binding Sites on the Measurement of ^{15}N Longitudinal Paramagnetic Relaxation Enhancements in Proteins by Solid-State NMR, *J. Biomol. NMR 51*, 293-302.

[6] Burns, C. S., Aronoff-Spencer, E., Dunham, C. M., Lario, P., Avdievich, N. I., Antholine, W. E., Olmstead, M. M., Vrielink, A., Gerfen, G. J., Peisach, J., Scott, W. G., and Millhauser, G. L. (2002) Molecular features of the copper binding sites in the octarepeat domain of the prion protein, *Biochemistry 41*, 3991-4001.

[7] Silva, K. I., Michael, B. C., Geib, S. J., and Saxena, S. (2014) ESEEM analysis of multi-histidine Cu(II)-coordination in model complexes, peptides, and amyloid-beta, *J. Phys. Chem. B 118*, 8935-8944.

[8] Lu, J., Bender, C. J., McCracken, J., Peisach, J., Severns, J. C., and McMillin, D. R. (1992) Pulsed EPR studies of the type 2 copper binding site in the mercury derivative of laccase, *Biochemistry 31*, 6265-6272.

[9] Yang, Z., Kise, D., and Saxena, S. (2010) An approach towards the measurement of nanometer range distances based on Cu2+ ions and ESR, *J. Phys. Chem. B 114*, 6165-6174.

[10] Yang, Z., Kurpiewski, M. R., Ji, M., Townsend, J. E., Mehta, P., Jen-Jacobson, L., and Saxena, S. (2012) ESR spectroscopy identifies inhibitory Cu2+ sites in a DNA-modifying enzyme to reveal determinants of catalytic specificity, *Proc. Natl. Acad. Sci. U. S. A. 109*, E993-1000.
[11] Ji, M., Ruthstein, S., and Saxena, S. (2013) Paramagnetic Metal Ions in Pulsed ESR Distance Distribution Measurements, *Acc. Chem. Res.*
[12] Yang, Z., Becker, J., and Saxena, S. (2007) On Cu(II)-Cu(II) distance measurements using pulsed electron electron double resonance, *J. Magn. Reson. 188*, 337-343.
[13] Merz, G. E., Borbat, P. P., Pratt, A. J., Getzoff, E. D., Freed, J. H., and Crane, B. R. (2014) Copper-Based Pulsed Dipolar ESR Spectroscopy as a Probe of Protein Conformation Linked to Disease States, *Biophys. J. 107*, 1669-1674.
[14] Perdew, J. P., Burke, K., and Wang, Y. (1996) Generalized gradient approximation for the exchange-correlation hole of a many-electron system, *Phys. Rev. B: Condens. Matter 54*, 16533-16539.
[15] Weigend, F., and Ahlrichs, R. (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, *PCCP 7*, 3297-3305.
[16] Neese, F. (2012) The ORCA Program System, *WIREs Comput. Mol. Sci. 2*, 73-78.
[17] Bode, B. E., Plackmeyer, J., Prisner, T. F., and Schiemann, O. (2008) PELDOR measurements on a nitroxide-labeled Cu(II) porphyrin: orientation selection, spin-density distribution, and conformational flexibility, *J. Phys. Chem. A 112*, 5064-5073.
[18] Lovett, J. E., Bowen, A. M., Timmel, C. R., Jones, M. W., Dilworth, J. R., Caprotti, D., Bell, S. G., Wong, L. L., and Harmer, J. (2009) Structural information from orientationally selective DEER spectroscopy, *PCCP 11*, 6840-6848.