

**SITE-DIRECTED Cu²⁺ LABELING METHODOLOGIES FOR OBTAINING DISTANCE
CONSTRAINTS IN PROTEINS AND DNA**

by

Shreya Ghosh

Bachelor of Science, Visva-Bharati University, 2012

Master of Science, Visva-Bharati University, 2014

Submitted to the Graduate Faculty of the
Dietrich School of Arts and Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Shreya Ghosh

It was defended on

August 19, 2020

and approved by

Haitao Liu, Professor, Department of Chemistry

Sean Garrett-Roe, Professor, Department of Chemistry

Junmei Wang, Professor, Department of Pharmaceutical Sciences

Dissertation Director: Sunil Saxena, Professor, Department of Chemistry

Copyright © by Shreya Ghosh

2020

SITE-DIRECTED Cu²⁺ LABELING METHODOLOGIES FOR OBTAINING DISTANCE CONSTRAINTS IN PROTEINS AND DNA

Shreya Ghosh, PhD

University of Pittsburgh, 2020

Electron spin resonance (ESR) spectroscopy along with site-specific introduction of spin labels into macromolecules is a powerful tool to obtain a multifaceted view of macromolecule structure, flexibility, and dynamics. In this dissertation, we develop Cu²⁺-based ESR reporters for proteins and DNA that are small, rigid and can provide precise information on the backbone fluctuations without requiring additional modeling. First, we demonstrate a nucleotide and structure-independent Cu²⁺-based label for DNA, that can be incorporated anywhere in the DNA duplex. We perform pulsed ESR based distance measurements on several duplexes with varying base pair separation between the labels. Using the distance measurements along with modeling, we illustrate that this methodology is capable of directly reporting on DNA backbone conformations in solution. Additionally, we perform molecular dynamics simulations using high-quality force field parameters developed for the Cu²⁺-label in DNA. Such analysis provides detailed atomic insights into the conformational fluctuations of the label and a more nuanced picture of the ESR distance measurements.

For proteins, we further develop a labeling strategy where Cu²⁺ site-specifically binds to two strategically placed histidine residues in a protein. Systematic analysis confirms that the Cu²⁺-complex binds specifically to the engineered binding sites. Distance measurements using this label show an improved resolution with a two-fold increase in the signal-to-noise ratio. The small size and the rigidity of the Cu²⁺-label shows promise over traditional labels and will allow for readily

elucidating protein backbone flexibility, distinguishing between different protein conformations in solution and determining relative orientations of different protein subunits. In addition, we also apply Cu^{2+} -based ESR measurements to understand the influence of metal binding on the homodimeric antimicrobial protein, Calgranulin C. We demonstrate that despite being homodimeric, the two native Cu^{2+} binding sites in Calgranulin C have different coordination environments in solution, where only one of the Cu^{2+} shows backbone coordination. Finally, ESR distance measurements on the Cu^{2+} -bound protein provide multiple distances indicating metal-induced protein oligomerization in solution. Overall, this dissertation highlights the advantages of Cu^{2+} -based labeling strategies in proteins and DNA that can find application for understanding conformational changes associated with protein-DNA interactions in important physiological functions.

TABLE OF CONTENTS

PREFACE.....	xxiv
1.0 INTRODUCTION.....	1
1.1 ELECTRON SPIN RESONANCE SPECTROSCOPY	1
1.2 SPIN HAMILTONIAN	8
1.3 CONTINUOUS WAVE (CW) ESR	9
1.3.1 CW-ESR spectrum for Cu²⁺	9
1.3.2 Characterization of Cu²⁺ coordination environment using CW-ESR spectrum	13
1.3.3 Binding affinity of Cu²⁺ to its binding site	16
1.4 ELECTRON SPIN ECHO ENVELOPE MODULATION (ESEEM).....	17
1.5 DOUBLE ELECTRON ELECTRON RESONANCE	22
1.5.1 Determining the population of different labeled species and oligomeric states	25
1.5.2 Obtaining distance distribution	26
2.0 Cu²⁺-BASED DISTANCE MEASUREMENTS BY PULSED ESR PROVIDE	
DISTANCE CONSTRAINTS FOR DNA BACKBONE.....	28
2.1 INTRODUCTION	28
2.2 MATERIALS AND METHODS.....	31
2.2.1 Oligonucleotides	31
2.2.2 ESR measurements	32
2.2.3 MD simulations	34

2.3 RESULTS AND DISCUSSION.....	35
2.3.1 Characterization of Cu²⁺ bound to DPA-DNA.....	36
2.3.2 DEER measurements show no observable orientational selectivity at both X and Q-band frequencies.....	39
2.3.3 DEER measurements using Cu²⁺-DPA motif can probe the DNA backbone.....	44
2.3.4 MD simulations show good agreement with the most probable distance from DEER.....	46
2.3.5 DNA twist-stretch model enables estimation of label flexibility.....	49
2.3.6 Twist-stretch and in-silico models show that Cu²⁺-Cu²⁺ distance can predict on DNA backbone despite having a radial offset.....	55
2.3.7 In-silico model of non-linear DNA shows that Cu²⁺-DPA can report on DNA backbone conformations of bent/kinked DNA.....	57
2.4 CONCLUSION.....	58
2.5 ACKNOWLEDGEMENTS.....	59
3.0 MOLECULAR DYNAMICS SIMULATIONS USING FORCE FIELD PARAMETERS OF Cu²⁺-BASED DNA LABELS PROVIDE ATOMIC INSIGHTS INTO PULSED ESR BASED DISTANCE CONSTRAINTS.....	60
3.1 INTRODUCTION.....	60
3.2 MATERIALS AND METHODS.....	62
3.2.1 HYSORE.....	62
3.2.2 DFT optimization.....	63
3.2.3 MD simulations.....	64

3.3 RESULTS AND DISCUSSION.....	65
3.3.1 DFT-optimized structure agrees well with crystal structure	66
3.3.2 HYSCORE indicates the presence of the fourth equatorially coordinating atom	67
3.3.3 MD simulations show that the Cu²⁺-Cu²⁺ distance reasonably agrees with the backbone distance	69
3.3.4 DPA-DNA based MD simulations capture the experimental most probable distance.....	71
3.3.5 Orientation of Cu²⁺-DPA helps reduce the effect of linker offset.....	73
3.3.6 MD simulations show that Cu²⁺-DPA have varied mobility that depends on position	75
3.3.7 MD simulations provide insight into the use of DEER for such measurements	77
3.3.8 Many orientations are excited even at a single magnetic field.....	81
3.4 CONCLUSION	82
3.5 ACKNOWLEDGEMENTS	83
4.0 THE Cu²⁺-NITRILOTRIACETIC ACID COMPLEX IMPROVES LOADING OF α-HELICAL DOUBLE HISTIDINE SITE FOR PRECISE DISTANCE MEASUREMENTS BY PULSED ESR	84
4.1 INTRODCUTION	84
4.2 MATERIALS AND METHODS.....	88
4.2.1 Protein expression, purification and labeling.....	88
4.2.2 ESR measurements	88

4.2.3 Analysis of the modulation depth of DEER.....	90
4.3 RESULTS AND DISCUSSION.....	92
4.3.1 CW-ESR spectroscopy to show complexation between Cu ²⁺ and NTA	93
4.3.2 Binding affinity of Cu ²⁺ -NTA to dHis site using CW-ESR spectroscopy	94
4.3.3 Cu ²⁺ -NTA further improves dHis-based DEER signal	103
4.4 CONCLUSION.....	107
4.5 ACKNOWLEDGEMENTS	108
5.0 Cu²⁺ ESR REVEALS TWO DISTINCT BINDING SITES AND	
OLIGOMERIZATION OF INNATE IMMUNE PROTEIN CALGRANULIN C	109
5.1 INTRODUCTION	109
5.2 MATERIALS AND METHODS.....	112
5.2.1 Sample preparation.....	112
5.2.2 ESR measurements	112
5.3 RESULTS AND DISCUSSION.....	114
5.4 CONCLUSION.....	122
5.5 ACKNOWLEDGEMENTS	123
APPENDIX A: RESIDUE TOPOLOGY AND KEY FORCE FIELD PARAMETERS	
.....	124
APPENDIX A.1 RESIDUE TOPOLOGY OF Cu²⁺-DPA-WAT₁	124
APPENDIX A.2 RESIDUE TOPOLOGY OF DSPACER	126
APPENDIX A.3 KEY FORCE FIELD PARAMETERS OF Cu²⁺-DPA-WAT₁	127
BIBLIOGRAPHY	131

LIST OF FIGURES

Figure 1-1 Structures of A) MTSSL and B) the Cu²⁺-based label. C) The rotamer of the two labels when incorporated in the same protein and the corresponding distance distributions show that MTSSL is much more flexible than the Cu²⁺ label..... 3

Figure 1-2 The Cu²⁺-based label for DNA shown in red and the nitroxide label shown in green. 6

Figure 1-3 Energy level diagram of Cu²⁺. 10

Figure 1-4 A) Unpaired electron in Cu²⁺ present in the d_{x²-y²} orbital. Type (II) Cu²⁺ shows axial symmetry whereby, $g_{xx}=g_{yy}= g_{\perp}$ and $g_{zz}=g_{\parallel}$. The angle θ is the angle between the applied magnetic field, B_0 and the molecular zz -axis. B) A sample of randomly oriented molecules with random θ values. C) Schematic drawings of an axial g -tensor ellipsoid. For a sample with random orientations, there will be higher probability for the molecules to be at $\theta=90^\circ$ than $\theta=0^\circ$. D) The corresponding ESR spectrum. The intensity corresponding to $\theta=90^\circ$ will be the highest as shown in Trace (I). Trace (II) shows the first derivative of the absorption spectrum, as seen in CW-ESR..... 12

Figure 1-5 CW-ESR spectrum of type-II Cu²⁺. 13

Figure 1-6 A) CW-ESR spectrum of differing Cu²⁺ coordination. The direct coordination environment of Cu²⁺ in each compound is shown by the shaded circles. The lineshape is different between the three spectra due to differing g_{\parallel} and A_{\parallel} values. These results are taken with permission from Ref. 69. B) The Peisach-Blumberg plot depicting the relationship between Cu²⁺ centers with different equatorially coordinated ligands and their respective g_{\parallel} and A_{\parallel} values..... 13

Figure 1-7 A) Jahn-Teller distortion of octahedral geometry leading to a tetragonal distortion B) Energy level diagram of octahedral to square planar geometry. 14

Figure 1-8 A) Background subtracted time domain ESEEM spectrum of Cu²⁺ coordinated to two histidine residues. B) Fourier transformed ESEEM spectrum shows the characteristic NQI peaks below 2 MHz and DQ peak at ~ 4 MHz. 18

Figure 1-9 Interaction of the unpaired electron of Cu²⁺ with the nuclear spin of the remote nitrogen of the Cu²⁺-coordinated imidazole moiety..... 19

Figure 1-10 Raw time-domain ESEEM data showing the calculation of the parameters, a and b and the modulation depth, k. B) An example plot from Ref. 49 showing the plot of k versus the equivalents of a Cu²⁺-complex that binds to histidine residues. The plot shows the contribution of unbound Cu²⁺-complex towards the decrease in k. 21

Figure 1-11 A) Intramolecular interaction (solid line) and intermolecular interaction (dashed line) between two spins in solution. The distance between two spins is represented by the interspin vector, *r*. The angle between *r* and the applied magnetic field B₀ is defined as θ . B) Raw time domain DEER data (solid) and the background decay (dashed). C) Background subtracted time domain DEER data showing intramolecular contributions only. The dashed line shows the modulation depth. 23

Figure 1-12 Dependence of modulation depth, λ , on the population of different labeled spin systems in solution..... 25

Figure 2-1 The mass spectrum of each strand was completed upon synthesis by ATDBio. The mass of each individual strand (in amu) of each duplex is shown above..... 31

Figure 2-2 A) Structures of the Cu^{2+} incorporated DPA phosphoramidite and the dSpacer.

B) Sequences of the control DNA and the DPA-DNA duplexes with different base pair separation, n. 35

Figure 2-3 A) CW-ESR spectrum of Cu^{2+} bound DPA-DNA duplex (solid line). The single component fit, with g_{\parallel} and A_{\parallel} values of 2.240 and 170 G, respectively is overlaid (black dashed). B) Comparison of CW-ESR spectra between Cu^{2+} bound to control DNA (grey solid) and Cu^{2+} bound to DPA-DNA (red solid)..... 36

Figure 2-4 Simulation of the CW-ESR spectrum of the control DNA shows a single component fit with g_{\parallel} and A_{\parallel} values of 2.280 and 163 G, respectively. 38

Figure 2-5 DEER signals of Cu^{2+} -DPA-DNA (n=11) performed at Q-band frequency. DEER performed at A) g_{\parallel} region (11220 G) B) g_{\perp} region (11820 G) C) 8 different magnetic fields. 39

Figure 2-6 A) Field-swept electron spin echo spectrum of Cu^{2+} -DPA-DNA (n=11) at Q-band frequency. The lines show the different magnetic fields at which DEER was performed. B) Background subtracted time domain DEER signals at the different magnetic fields. The y-axis has been offset for ease of visualization. The data show minimal difference in dipolar frequency at the different fields. C) Background subtracted time domain data at g_{\parallel} (dashed, 11220 G) and g_{\perp} (solid, 11820 G) regions. These two regions also do not show any distinct difference in modulation frequency, suggesting Cu^{2+} -DPA shows minimal orientational effects. D) Distance distribution obtained via Tikhonov regularization, for both g_{\parallel} (dashed, 11220 G) and g_{\perp} (solid, 11820 G) regions. The similarity in the distributions further confirms that Cu^{2+} -DPA is not orientation selective for non-complexed DNA..... 40

Figure 2-7 A) Background subtracted time domain DEER signal of Cu²⁺-DPA-DNA (n=11) at X-band frequency at g_{||} (grey) and g_⊥ (black) regions. The field corresponding to the g_{||} and g_⊥ regions are shown in the field-swept electron spin echo spectrum in the inset. There is no distinguishable difference in the dipolar frequency between the two data. B) Distance distributions obtained as a sum of DEER time traces collected at different fields, ranging from g_{||} and g_⊥ regions at Q- band frequency (dashed red) and at g_⊥ region at X-band frequency (solid black). Both distributions, analysed via Tikhonov regularization, show similar most probable distance and width..... 42

Figure 2-8 Raw time domain DEER signals for DNA duplexes with n ranging from 9 to 12, performed at fields corresponding to g_{||} (grey) and g_⊥ (blue) at X-band frequency and the corresponding background subtracted time domain DEER data..... 43

Figure 2-9 A) Background subtracted time domain DEER data of the DPA-DNA duplexes ranging from n=8 to 12 base pairs. The modulation frequency increases as the base pair separation increases. The y-axis has been offset for ease of visualization. B) Duplexes showing the Cu²⁺-DPA and dSpacer positions for varying n. (C) Pake pattern for duplexes n=8 to 12..... 44

Figure 2-10 A) Area-normalized distance distributions obtained for duplexes n=8 to 12 via Tikhonov regularization. The distributions show an increase in the most probable distance with increasing n. B) Plot of most probable distance versus base pair separation, n. On fitting to a linear trend with a y-intercept of 0, we obtained a slope of 0.35 nm with an error of 0.003 nm. The slope is in reasonable agreement with the reported base pair separation for a B-DNA..... 45

Figure 2-11 Backbone carbon atoms of an unmodified nucleotide compared to the DPA phosphoramidite. The backbone carbon atom in DPA (grey circle) can resemble either A) a C3' atom when viewed from the 3' end or B) a C4' atom when viewed from the 5' end. 46

Figure 2-12 The first panel shows the DNA duplex inside a water box, used in our simulations. C3'-C3' (blue dashed) and C4'-C4' (red dotted) distance distributions obtained from MD simulations and experimental distance distributions (black solid). MD simulated distributions show a good agreement with the experimental most probable distance, within 1-2 Å. 47

Figure 2-13 Comparison of the most probable distance from the experimental results with A) C3'-C3' distances and B) C4'-C4' distances from MD simulations. The data show good agreement for different base pair separations, n. 48

Figure 2-14 A) DPA structure, built on available crystal structure parameters, is incorporated in a DNA duplex and the approximate length of the label is calculated to be 3.3 Å. Geometric model of the B-DNA helix. The model comprises of two components B) axial and C) transverse. 49

Figure 2-15 Distance distributions obtained from the twist-stretch model of the DNA compared to the experimental results for different base pair separations between the spin labels. 51

Figure 2-16 Background subtracted time-domain DEER data for the different DNA duplexes with variable α -values (1, 10 and 100). There is a good agreement of the experiment with the fit, over the different α -values used. 52

Figure 2-17 A) Comparison of standard deviations of distance distributions for varying α -values. B) Comparison of the average of the standard deviations for different α with that obtained from the optimum α -value..... 53

Figure 2-18 A) Plot of most probable distance versus base pair separation for the geometric model of DNA's twist-stretch motion (blue) and experiment (orange). The results agree within the error of the experiment. B) Plot of standard deviation versus base pair separation for model (blue) and experimental results (orange). The model considers an additional ~ 1.6 Å flexibility of the Cu^{2+} -DPA label and agrees with the experimental results. 54

Figure 2-19 A) Two different models of the straight DNA where the Cu^{2+} -DPA motifs are a) on top of each other and b) opposing one another. The Cu^{2+} is modeled to be at 3.3 Å from the DNA backbone. In the first case, the Cu^{2+} - Cu^{2+} distance can be directly translated to DNA backbone distance as the 3.3 Å deviation does not play any role here. In the second case, the deviation, being significantly smaller than the axial distance, also does not contribute to the resultant distance. B) Plot of most probable distance of the DNA backbone and Cu^{2+} -DPA-DNA, obtained from the twist-stretch model, for different base pair separations..... 55

Figure 2-20 Comparison of distance from in-silico model of Cu^{2+} -DPA-DNA with most probable distance from MD simulations for A) $\text{C3}'$ - $\text{C3}'$ and B) $\text{C4}'$ - $\text{C4}'$. We see a slightly better agreement of the Cu^{2+} - Cu^{2+} distance with the $\text{C4}'$ - $\text{C4}'$ backbone distance. 56

Figure 2-21 A) Model of a bent DNA showing the backbone distance and the Cu^{2+} - Cu^{2+} distance. In the first case, the position of the two labels being almost parallel causes the Cu^{2+} - Cu^{2+} distance to directly translate into the DNA backbone distance. In the second

case, the labels are anti-parallel. However, the offset of Cu^{2+} from the backbone does not significantly contribute to the measured distance. B) and C) Modeled DNA structures with a bent angle of 90° and 150° , respectively. D) PDB of a bent DNA (PDB: 1A73)¹⁸³. C3' (orange squares) and C4' (green triangles) DNA backbone distances were measured. In-silico modeling was performed incorporating Cu^{2+} -DPA phosphoramidite into the DNA duplex for different base pair separations and the corresponding Cu^{2+} - Cu^{2+} distance was measured (black circles). All the three distances show reasonable agreement within the caveats of the model..... 57

Figure 3-1 DPA-DNA duplex with the Cu^{2+} -DPA motif (represented by rectangle) and the dSpacer (represented by circle). The base pair separation between the two Cu^{2+} -DPA motifs is denoted by n. 65

Figure 3-2 Cu^{2+} -DPA structures coordinated with A) three B) two and C) one water molecule. 66

Figure 3-3 A) DFT optimized structure of Cu^{2+} -DPA. B) Comparison of bond angles and bond lengths with crystal structure¹⁷⁵ of Cu^{2+} -DPA show good agreement with the DFT-optimized structure. 67

Figure 3-4 ¹H HYSCORE spectra of A) Cu^{2+} -DPA B) Cu^{2+} -DPA-DNA and C) Cu^{2+} -control DNA, analyzed to the same contour levels. In Cu^{2+} -DPA and Cu^{2+} -DPA-DNA, the proton signal results from equatorially coordinated water molecule (rectangle) and weakly coupled solvent (circle). In control DNA, the proton signal results from only weakly coupled solvent molecules..... 68

Figure 3-5 A) Cu^{2+} -DPA phosphoramidite with the backbone carbon atom, C' marked in red. The Cu^{2+} - Cu^{2+} distance is denoted by the black solid line and the backbone C'-C'

distance is represented by the red dashed line in a DPA-DNA duplex. The Cu^{2+} - Cu^{2+} distance (black solid) is compared with the backbone distance (red dashed) for B) $n=9$ C) $n=10$ D) $n=11$ and E) $n=12$ base pair separations. F) Plot showing the experimental and MD Cu^{2+} - Cu^{2+} and MD C'-C' most probable distances. All the distance distributions agree reasonably well..... 69

Figure 3-6 Plot of A) most probable distances and B) standard deviation of distance distributions from ESR (black), 100 ns MD (blue) and 1 μs MD (red) against base pair separation (n)..... 71

Figure 3-7 A) The frames from MD trajectories considered for obtaining the spatial distribution of Cu^{2+} are within $\pm 1 \text{ \AA}$ of the Cu^{2+} - Cu^{2+} most probable distance and shown by the shaded region. B) The frames are aligned with respect to the DPA backbone atoms marked in red. C) The distribution of the Cu^{2+} centers for the DPA site 1 (blue, circles) and site 2 (orange, rectangles) for duplexes $n= 9-12$. D) The top-down view of the DNA where the blue and orange spheres represent the Cu^{2+} at DPA sites 1 and 2, respectively. The grey sphere represents the backbone carbon atom, C' 73

Figure 3-8 A) Angles between the C4' atom of the adjacent base to the DPA (blue sphere), the C' backbone atom of DPA (grey sphere) and the Cu^{2+} (cyan and orange for DPA site 1 and 2, respectively. B) Table showing the values for the two DPA sites C) Plot of the angles versus base pair separation for the two DPA sites. 74

Figure 3-9 The root mean square fluctuations (RMSF) of all the bases in the DPA-DNA duplexes with A) $n=9$ B) $n=10$ C) $n=11$ and D) $n=12$. The grey and black denote the two strands in the DNA and the red and blue denote Cu^{2+} -DPA and dSpacer, respectively. As can be seen, the RMSF is high for the terminal bases, as expected. The RMSF for Cu^{2+} -

DPA and the abasic dSpacer positions are generally high, indicating more flexibility than the natural bases. 75

Figure 3-10 Fluctuations in bond length between Cu^{2+} and two coordinating nitrogen atoms – A) Bond between Cu^{2+} and the backbone N1 atom. B) Fluctuations in Cu^{2+} -N1 bond length in the two DPA sites, sampled over 1 μs of MD for the duplex $n=11$ (top chart) and the probability distribution (bottom chart). C) and D) Same analysis repeated for bond length between Cu^{2+} and the pyridine N2 atom. A change of $\sim 25\%$ in the bond length is observed for Cu^{2+} -N1 and $\sim 30\%$ for Cu^{2+} -N2. 77

Figure 3-11 Dihedral angles measured between Cu^{2+} and its coordinating atoms for 1 μs of MD run on DPA-DNA duplex with $n=11$. The dihedral angle is denoted by the red arrow (top chart). Bottom chart shows the probability distribution of the dihedrals between the two DPA sites, DPA1 (blue) and DPA2 (orange). 78

Figure 3-12 The distribution of g_{\parallel} directions (cyan) in the Cu^{2+} -DPA system, sampled every 10 ns of the 1 μs MD run. The blue and red bonds show the Cu^{2+} coordination with nitrogen and oxygen, respectively. The cyan represents the g_{\parallel} . The g -tensor calculations were performed with ORCA^{214,215}. 79

Figure 3-13 The relative orientations between the two spins, A and B, are indicated by three angles, χ , γ and η 80

Figure 3-14 A) DEER performed on DPA-DNA duplex with $n=11$ at Q-band frequency. The red lines show the various fields at which DEER was performed. B) The simulated DEER time traces at each field at the Q-band frequency. The y-axis is offset for ease of visualization. The red dashed line represents the first period of the modulations for all fields. C) Background subtracted time domain data at g_{\parallel} (11220 G, dashed) and g_{\perp} (11820

G, solid) regions. The figure shows the lack of orientational selectivity effects at Q-band frequency for Cu²⁺-DPA..... 80

Figure 3-15 MD frames of DPA-DNA duplex (n=11) sampled every 10 ns for a total 1 μs. The black sphere in the center is the reference Cu²⁺. All frames are aligned to the g_{||} axis of the reference Cu²⁺. The second Cu²⁺ position is represented as orange spheres. The g_{||} direction, marked in the figure, is calculated from ORCA^{214,215}. B) The entire DNA duplex is shown for 5 snapshots for simple viewing. The figure suggests that there is a wide range of molecular orientations at a single g_{||} orientation. 81

Figure 4-1 The dHis motif and the two Cu²⁺-complexes, NTA and IDA (top and bottom respectively) that create the dHis-Cu²⁺-NTA spin label (top) and the dHis-Cu²⁺-IDA spin label (bottom). 86

Figure 4-2 The spectrum of CuCl₂ in water (gray solid line) is markedly different from that of Cu²⁺-NTA in water. 94

Figure 4-3 A) CW-ESR spectrum of Cu²⁺-NTA bound to the α-helical dHis site in GB1, where dHis:Cu²⁺-NTA = 1:0.25. The spectrum shows a single component fit indicating that Cu²⁺-NTA is completely bound to the protein. B) CW-ESR spectrum of Cu²⁺-NTA bound to the α-helical dHis site in GB1, where dHis:Cu²⁺-NTA = 1:2. This spectrum has contributions from both Cu²⁺-NTA bound to the dHis as well as free Cu²⁺-NTA in solution..... 95

Figure 4-4 CW-ESR titrations for 28H/32H GB1 at different equivalents of Cu²⁺-NTA. The protein concentration is 400 μM..... 96

Figure 4-5 A) Plot of bound Cu²⁺-NTA per protein versus Cu²⁺-NTA equivalents gives the apparent binding affinity of Cu²⁺-NTA to the dHis site in the α-helix, 28H/32H (K_d = 285 μM) in GB1. B) Plot showing the amount of bound Cu²⁺-NTA per unbound Cu²⁺-NTA at

different Cu²⁺-NTA equivalents for 28H/32H-GB1. C) Plot showing the apparent binding affinity of Cu²⁺-NTA to the β-sheet, 6H/8H (K_d = 70 μM). D) Plot showing the amount of bound Cu²⁺-NTA per unbound Cu²⁺-NTA at different Cu²⁺-NTA equivalents for 6H/8H. 97

Figure 4-6 A) The CW-ESR spectrum of 6H/8H-GB1 (grey line) containing a 1:0.25 ratio of dHis:Cu²⁺-NTA. Simulation shows a single component fit, indicating that Cu²⁺-NTA is entirely bound to the protein. B) CW-ESR spectrum of 6H/8H-GB1 at a dHis:Cu²⁺-NTA ration of 1:2 (grey line) with its respective simulation (black dashed). The simulation was generated by the addition of two individual spectra: 1) free Cu²⁺ NTA (inset, black) and 2) Cu²⁺-NTA bound to its respective α-helical or β-sheet dHis site (inset, grey). We call the free Cu²⁺-NTA Component 1 and the bound Cu²⁺-NTA Component 2 as seen in the inset. C) CW-ESR titration of 6H/8H-GB1 at different equivalents of Cu²⁺-NTA. The protein concentration is 400 μM. 98

Figure 4-7 Both β-sheet dHis sites used; 6H/8H which positions the dHis binding site within a typical β-sheet location and 15H/17H which positions the dHis site within an edge-strand β-sheet site. 99

Figure 4-8 CW-ESR data for the 15H/17H-GB1 construct which places the dHis Cu²⁺-binding site with an edge-strand β-sheet. A) CW-ESR spectra of 15H/17H-GB1 with varying equivalents of Cu²⁺-NTA incrementally added. B) Example of two component simulation of 15H/17H-GB1 CW-ESR spectrum. C) Plot used to determine the apparent dissociation constant, K_d = 100 μM. D) Plot illustrating the ratio of bound Cu²⁺-NTA per unbound Cu²⁺-NTA in solution as different equivalents of Cu²⁺-NTA are added. 100

Figure 4-9 Background corrected ESEEM signal of 28H/32H-GB1 (A), 6H/8H-GB1 (B) and 15H/17H-GB1 (C) with varying equivalents of Cu²⁺-NTA. As the amount of unbound Cu²⁺-NTA in solution increases, the ESEEM modulations associated with dHis-Cu²⁺-NTA coordination become diminished..... 101

Figure 4-10 A) Background corrected ESEEM signal of 28H/32H-GB1 with varying equivalents of Cu²⁺-NTA illustrating the decrease in modulation size with increasing amounts of Cu²⁺-NTA. The raw time ESEEM signal is shown in inset along with an example of how modulation depth, k, is measured. B) Plot of k as a function equivalents of Cu²⁺-NTA added for the α -helical dHis site. A sharp decrease is observed after 1.5 equivalents of Cu²⁺-NTA are added. C) Plot showing linear decrease in k upon addition of Cu²⁺-NTA within β -sheet dHis site. Inset shows 6H8H-GB1. D) k versus equivalents of Cu²⁺-NTA added for the edge-strand β -sheet dHis site 15H/17H-GB1 (shown in inset). 102

Figure 4-11 Raw time domain DEER data for 6H/8H/28H/32H-GB1 (top) and 15H/17H/28H/32H-GB1 (bottom). Inset shows the clear modulations for each time domain data. The background exponential decay fit is shown as the dashed line..... 103

Figure 4-12 A) Representation of 6H/8H/28H/32H-GB1 and 15H/17H/28H/32H-GB1 distances (PDB: 4WH4)⁴⁷ B) Background subtracted time domain DEER data for both 6H/8H/28H/32H-GB1 (top) and resultant distance distribution (bottom). C) Background subtracted time domain DEER data of 15H/17H/28H/32H-GB1(top) with respective distance distribution (bottom). 104

Figure 4-13 A) Crystal structure of human glutathione S-transferase A1-1 (PDB: 1K3L)²⁶³. dHis site was added in with selected rotamers from previous crystal structure (PDB:

4WH4)⁴⁷. B) Raw time domain DEER data for 211H/215H huGSTA1-1. The background exponential decay fit is shown as the dashed line. C) Background subtracted time domain DEER data showing the increased modulation depth of the dHis-Cu²⁺-NTA motif (2750 scans) compared to the dHis-Cu²⁺-IDA motif (4300 scans). 106

Figure 5-1 Dimer of S100A12 (PDB: 1ODB)..... 110

Figure 5-2 A) CW-ESR titrations performed at different Cu²⁺ equivalents B) Simulations show that the CW-ESR spectrum at 1 equivalent of Cu²⁺ is a single component with g_{||} and A_{||} values of 2.335 and 118 G, respectively C) At 3.5 equivalents of Cu²⁺, a distinct second component is observed..... 114

Figure 5-3 Plot showing the bound Cu²⁺ concentration/protein monomer concentration versus total equivalents of Cu²⁺. 116

Figure 5-4 A) Echo-detected field swept spectrum of S100A12 with 3.5 Cu²⁺ equivalents (black solid). The simulated spectrum corresponding to the first component (grey dotted) and second component (grey solid) are shown. B) ESEEM experiments performed at two fields (inset, red lines). 118

Figure 5-5 A) ESEEM spectrum at 2807 G showing no visible peak at 2.8 MHz corresponding to backbone coordination. Inset shows HYSORE data at 2807 G which shows no visible cross peaks either. B) ESEEM spectrum at 3430 G shows a sharp peak at 2.8 MHz which could be due to possible backbone coordination. HYSORE data (inset) shows cross peaks (red squares) at (4.3, 2.8 MHz) confirming the presence of backbone coordination..... 119

Figure 5-6 A) X-ray crystal of S100A12 bound to both Ca²⁺ and Cu²⁺ (PDB: 1ODB)²⁸⁴. The green spheres represent the Ca²⁺ ions while the orange spheres represent the Cu²⁺ ions. The blue dashed lines show the inter-dimer distances while the red dotted line shows the

intra-dimer distance. B) Background subtracted time-domain DEER data showing the presence of two frequencies. The experimental data is represented by the gray solid line and the black dashed line represents the best fit obtained from Tikhonov regularization. The modulation depth is observed to be ~12.6%. C) Distance distributions obtained using Tikhonov regularization. The distribution also shows validation techniques including upper (red) and lower (blue) acceptable limits. 120

PREFACE

I am not able to believe that I have already spent five years in this group. Time does fly! I still remember packing three big bags, getting my passport stamped and arriving in Pittsburgh for the first time all by myself – roughly 10000 miles from my hometown in India. I was a scared chicken with no clue how I would survive in this foreign land in the following years. Five years later, I am proud of the person I have become, both personally and professionally, and this journey could not have been accomplished without my advisor, Dr. Sunil Saxena. He has been absolutely instrumental to everything that I have been able to achieve in grad school. Not only did he help me grow as a researcher, but also as a confident person. He has always had immense faith and trust in me, even at times when I doubted myself. I did not know what I was capable of until he made me believe in myself. His mentorship has been so rewarding, and I would not have made it to the end without his guidance. I would like to sincerely thank him for all that he has done for me, but I know I can never thank him enough.

I would like to thank Dr. Sean Garrett-Roe, Dr. Haitao Liu and Dr. Junmei Wang for serving on my committee and for taking the time to guide me through this important step in my graduate career. I was fortunate to have Dr. Garrett-Roe and Dr. Liu as course instructors and absolutely enjoyed the learning experience in their classes. I am grateful to Dr. Wang for all those long meetings where he patiently taught DFT right from the very basics. They have all been extremely supportive throughout and have always been a big source of advice and help.

I have had the pleasure of working in the group with some amazing colleagues. Austin Gamble Jarvi has been such an immense support throughout these five years. All those group meetings before comprehensive were a lot less stressful because the questions were directed to

both of us and not just me alone! Also, EPR was a lot less challenging because I could work with him to figure things out. Kevin Singewald has helped me understand EPR in simple ways that I could not have imagined. Xiaowei Du, Zikri Hasanbasri, Josh Casto and Alysia Mandato have been amazing to work with. I could always count on them for anything. Special thanks to Josh for sticking with me till 11 pm without complains and listening to me rant, pretty much about everything!

I would also like to thank Dr. Jessica Sarver, Dr. Ishara Silva, Dr. Zhongyu Yang and Dr. Matthew Lawless, former lab members of the Saxena group for guiding me whenever I needed help. I have lost count of the innumerable times I have asked Matt for help, even after he graduated. I think that he probably forced himself to graduate to avoid seeing me walk up to his office every now and then. I also had the opportunity to work with some talented undergrads – Amit Shimshi, Merline Paul, Hanna Brubaker, Yujie Luo and Charu Arora.

I would like to thank my wonderful collaborators, Dr. Gunnar Jeschke, Dr. Sharon Ruthstein and Dr. Steven Damo, who guided me in every way they could. Special thanks to Dr. Ruthstein who always took the time and effort to explain things to me and help me figure out research issues.

A big part of my grad school survival is due to friends who were going through the same journey as I was. Arai and Xing Yee, two amazing and smart scientists, who started their journey at Pitt with me, have been two big pillars of support. We did everything to lift each other up, both personally and professionally. I am lucky to find friends like them in grad school and will cherish our friendship forever. Saborni, Zeeshan and Shivang, my friends outside of work, who were there to pick me up from my mess, every time, and I will do the same as they near their graduation too.

I would like to immensely thank Betsy Callomon, the lovely dogs at the Dog Therapy Sessions and the television show “F.R.I.E.N.D.S” for providing me with the mental support to stay calm and positive during the five years of graduate school.

Finally, I would like to thank my family. I am the first person in my family to pursue a PhD (and in a foreign country!) and this just shows how much support my parents have provided to me. Every day for five years I have been talking to Ma and Baba, first thing in the morning, so that no matter how my day went, I know that my family has got my back. I would also like to thank my brother, sister-in-law and my wonderful niece who have cheered for me at every corner of life. And lastly, I would like to thank my husband, Prem. A big reason I completed grad school is because of him. Over the years, he has been a friend, counselor, mentor, and any other role that he needed to play to support me. He is the best thing to have happened to me in grad school and I thank God every day for making our paths cross.

1.0 INTRODUCTION

1.1 ELECTRON SPIN RESONANCE SPECTROSCOPY

[Parts of the Introduction have appeared in J. Chem. Ed., 2019, volume 96, pages 1752-1769]

Electron spin resonance (ESR) is a spectroscopic technique that has gained importance due to its diverse applications in chemistry, biology, physics, geology, medicine, materials and food science. The most wide-scale use of ESR, however, has been in the biophysical field. Importantly, ESR is not limited by the size of the macromolecule or by the optical properties of the sample and therefore, can measure the structure and dynamics of complexes, that are often not readily measurable by other biophysical approaches. ESR also offers high sensitivity (as low as nanomolar concentrations in cases) and can be used in a variety of medium such as lipid vesicles, membranes, lipid bilayers, micelles, in-cell, and in-vivo, that often closely represents the true cellular environment.

To implement ESR on biomolecules, site-directed spin labeling (SDSL) strategy was introduced that incorporated unpaired electron spins in the form of spin labels at specific sites¹⁻³. Traditionally SDSL involves the removal of all native cysteine residues by substituting with another amino acid, such as alanine or serine. A unique cysteine residue is then incorporated at a specific site via site-directed mutagenesis. The cysteine then further undergoes a sulfhydryl-specific reaction with a nitroxide containing reagent to generate a spin-labeled protein. The development of SDSL broadly expanded the applicability of ESR, allowing macromolecules that do not have native unpaired electron species to be readily investigated by ESR.

SDSL-ESR can provide a plethora of structural and dynamic information on proteins^{2,4}. When the spin label is introduced in a protein, its behavior is influenced by the local environment of the label as well as the motion of the protein, which is further reflected in the ESR spectra. The spectroscopic data can then be analyzed to obtain structural information such as protein secondary structure⁵⁻⁷. In addition the technique can elucidate the location, tilt and orientation of membrane embedded helical segments in a membrane protein^{8,9}. Information on dynamics, such as conformational changes, side chain motions and backbone fluctuations can also be obtained from SDSL-ESR^{10,11}.

The success of the SDSL-ESR inspired further development of ESR methodologies. Pulsed ESR techniques in combination with SDSL-ESR became a powerful biophysical tool for elucidation of protein structure and dynamics. One key implementation of the pulsed ESR techniques has been the ability to exploit the dipolar interaction between unpaired electrons to obtain distance constraints¹²⁻¹⁶. A network of distance constraints at engineered sites of the protein can be used to measure structural information such as quaternary structure, relative location of two proteins, or a protein-nucleic acid complex and/or a substrate and a protein¹⁷⁻²⁰. Importantly, distance measurements can often be measured in metastable states of the protein. These distance constraints can then be used to elucidate the structure of transient but functionally important states of the protein.

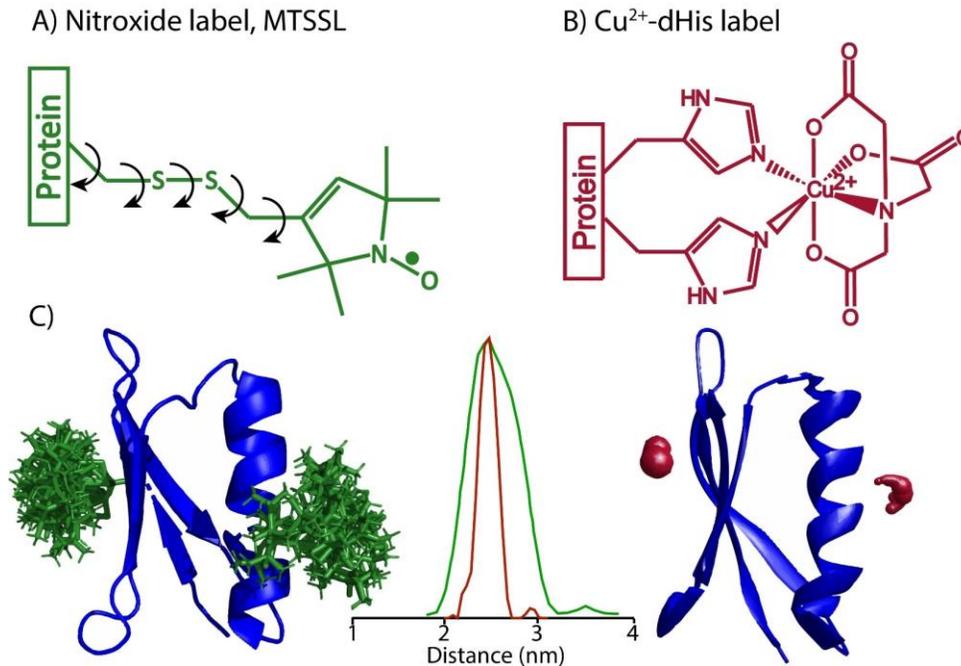


Figure 1-1 Structures of A) MTSSL and B) the Cu^{2+} -based label. C) The rotamer of the two labels when incorporated in the same protein and the corresponding distance distributions show that MTSSL is much more flexible than the Cu^{2+} label.

Distance measurements using SDSL-ESR primarily use nitroxides as the spin label for both proteins and nucleic acids. Figure 1-1A shows the most commonly used spin label in proteins, called the 1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl methanesulfonate spin label (MTSSL)¹¹. Despite the wide application of MTSSL, there exists some critical limitation of this labeling strategy. First, the use of cysteine mutagenesis for spin labeling poses limitations in proteins with functional cysteines such as kinases and transcription regulators. Second, the unpaired electron in MTSSL is delocalized in the N-O bond, separated from the protein backbone C_α by five bonds. These bonds are highly flexible, allowing the MTSSL to possess multiple rotameric states. On performing distance measurements using two or more such labels, the distance distributions are mostly dominated by the flexibility of the linker rather than the inherent fluctuations of the protein²¹⁻²³. As a result, extracting information on protein structure and

flexibility becomes ambiguous. Rigid nitroxide labels, such as the bifunctional label RX, have also been developed²⁴. However, the rigidity comes at the price of requiring two neighboring cysteine residues to bind and more complex labeling strategy.

For distance measurements in nucleic acids, the commonly used labeling strategies involve chemical modification of a natural base of the DNA, mostly cytosine²⁵⁻²⁷ or guanine^{28,29}, to incorporate a nitroxide moiety. Such labels are often highly rigid and corresponding distance measurements allow to extract information on label orientation³⁰⁻³². Despite the promising results, the dependence on a nucleotide limits the use of the nitroxide labels. Furthermore, the labels have a long linker³³⁻³⁵, causing the corresponding most probable distance to be longer than the backbone distance by a few nanometers, and requiring additional modeling to report on DNA backbone conformations.

Such limitations with the nitroxide labeling strategy led to the development of novel spin labeling techniques³⁶. Radicals like trityl³⁷ and paramagnetic metal ions like Gd³⁺, Fe³⁺, Mn²⁺ and Cu²⁺³⁸⁻⁴⁰ present a promising alternative labeling strategy and have also been extensively used for distance measurements. Thus, the potential and applicability of pulsed ESR distance measurement techniques were greatly enhanced.

Our group has pioneered Cu²⁺-based labeling strategies and ESR methodologies. Cu²⁺ is a paramagnetic species with a single unpaired electron, therefore the ESR spectrum is much simpler than most metal ions with multiple unpaired electrons⁴¹. Cu²⁺ is also one of the most abundant biological metals. Unlike other paramagnetic labels like Gd³⁺, Cu²⁺ binds to many metalloproteins, thereby expanding the use of Cu²⁺ as an ESR label. Besides, Cu²⁺ can be readily used on the common low field X-band (~9.5 GHz) spectrometers while metal ions like Gd³⁺ and Fe³⁺ often

require much higher fields such as the W-band (~95 GHz). Therefore, Cu²⁺ therefore presents a promising labeling strategy.

Accordingly, in our group, several experimental methodologies utilizing Cu²⁺-labels have been developed. The first double quantum coherence (DQC) based Cu²⁺-Cu²⁺ distance measurement was reported by our group⁴². Further improvement of the Cu²⁺-based DQC technique allowed detection of solely the dipolar interaction between the electron spins with minimum contribution from other interactions that affect the resolution of the signal^{43,44}. Additional work then elucidated key methodological principle with respect to the use of double electron electron resonance (DEER) to measure Cu²⁺- Cu²⁺ distances^{38,41,43,45,46}. Development of simulation programs and details of experimentation further expanded the applicability of Cu²⁺ to a wide range of biological systems⁴⁵. More recently, a number of Cu²⁺ based labels have been developed for site-specific spin labeling in peptides, proteins and DNA⁴⁷⁻⁵¹. A combination of these methodologies and labeling strategies have provided answers to key biological questions – conformational changes associated with a protein interacting with its specific, miscognate and non-specific DNA sequence^{21,52,53}, role of Cu²⁺ in the DNA cleavage activity of an endonuclease⁵⁴, recognizing a dynamic conformation of an unliganded protein that is structurally unresolved⁵⁵, and efficiently locating native metal binding sites in a protein^{54,56,57}, to name a few. To further expand the potential of these Cu²⁺ labels, computational techniques utilizing these labels have also been developed for proteins and PNA^{45,58-60} which provide detailed insight into the behavior of the labels and their influence on the macromolecules.

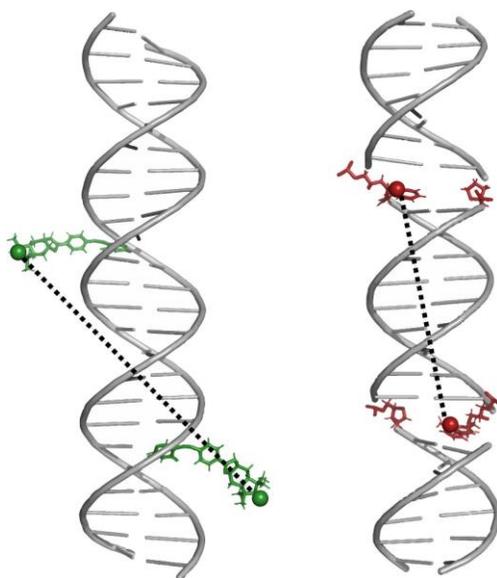


Figure 1-2 The Cu^{2+} -based label for DNA shown in red and the nitroxide label shown in green.

This dissertation builds on these methodological developments and focuses on the site-directed incorporation of Cu^{2+} spin labels in proteins and DNA. In chapter 2, we implement a Cu^{2+} -based labeling strategy for DNA that alleviates some of the key limitations of nitroxide labels. Figure 1-2 shows the Cu^{2+} -based label developed for DNA that involves a chelating ligand with a phosphoramidite linker. The label can be incorporated anywhere in the DNA duplex⁶¹. In this work, we first improve the labeling conditions to maximize Cu^{2+} binding to the ligand. Subsequently, we perform distance measurements on several DNA duplexes with varying base pair separations. We demonstrate that the Cu^{2+} -label can readily predict the DNA backbone distance constraints and can discern differences as small as one base pair, without requiring any additional modeling. Comparative measurements performed using nitroxides report a most probable distance that is at least a nanometer longer than the DNA backbone distance³³, due to the long linker as shown in Figure 1-2. Cu^{2+} -DPA, on the other hand, has a significantly smaller linker length, thereby having the ability to precisely report on the DNA backbone. The results in this

chapter display the potential of the structure and nucleotide-independent Cu^{2+} -label for elucidating DNA conformations.

In Chapter 3, we perform molecular dynamics (MD) simulation methodology on DNA duplexes using high quality force field parameters developed for the Cu^{2+} -label. Analysis of the MD simulations provides deeper insight into the ESR distance measurements obtained using this spin label. Moreover, the conjunction of ESR and MD provides a detailed atomic insight into the conformational fluctuations and orientations of the label in the DNA.

In Chapter 4, we develop a small, rigid Cu^{2+} -based spin label for proteins that does not require cysteine mutagenesis. The labeling strategy involves Cu^{2+} , complexed with a chelator, that binds to two strategically placed histidine residues in the protein⁴⁹. The simultaneous binding to two histidine residues restricts the motion of the label, thereby making it rigid. To obtain optimum conditions for labeling proteins with the Cu^{2+} -complex, we undertake a systematic analysis using ESR measurements. Subsequently, we perform ESR distance measurements on a doubly labeled protein under the optimized conditions. Figure 1-1 shows the distribution width from the Cu^{2+} label to be up to five times narrower when compared to analogous distribution obtained using MTSSL. The results also show a high affinity of the Cu^{2+} -label towards the engineered double histidine sites. The tight binding is further reflected in improved resolution of the ESR signal. This work, hence, reports a powerful labeling strategy using a Cu^{2+} -based label that is small and rigid, and most importantly, can readily relate to the protein backbone fluctuations.

In Chapter 5, we apply Cu^{2+} -based ESR measurements to investigate the effects of metal binding in a protein, calgranulin C⁶². Even though calgranulin C is a homodimer, we observe different coordination environment of the native Cu^{2+} sites. Furthermore, we provide insight into metal-induced protein oligomerization in solution using ESR distance measurements.

All the work presented in this dissertation involves the use of Cu²⁺-based ESR labels to provide information on conformational flexibility of biological macromolecules. In doing so, we use continuous wave (CW) ESR techniques as well as pulsed ESR techniques of electron spin echo envelope modulation (ESEEM) and double electron electron resonance (DEER). This chapter, therefore, focuses on providing basic underlying concepts of these techniques and how they have been utilized in this dissertation to obtain key information on the system of interest.

1.2 SPIN HAMILTONIAN

The spin Hamiltonian of an electron is a combination of the different energetic contributions experienced by the electron spin on interacting with other electron and nuclear spins. In ESR, the energetic interactions originate due to the presence of a magnetic field. Four interactions namely, the electron Zeeman (\hat{H}_{ez}) interaction, the nuclear Zeeman (\hat{H}_{nz}) interaction, the hyperfine interaction (\hat{H}_{hf}) and the nuclear quadrupole interaction (\hat{H}_{nq}) describe the spin Hamiltonian. The spin Hamiltonian is, thus, given as⁶³:

$$\hat{H} = \hat{H}_{ez} + \hat{H}_{nz} + \hat{H}_{hf} + \hat{H}_{nq} \quad (1-1)$$

$$\hat{H} = \beta_e \vec{B} \cdot \tilde{g} \cdot \hat{S} - \beta_n \sum_{i=1}^k (\vec{B} \cdot \tilde{g}_{i,n} \cdot \hat{I}_i) + \hbar \sum_{i=1}^k (\hat{S} \cdot \tilde{A}_i \cdot \hat{I}_i) + \hbar \sum_{i=1}^k (\hat{S} \cdot \tilde{Q}_i \cdot \hat{I}_i) \quad (1-2)$$

where \hat{S} and \hat{I}_i are the electron and i th nuclear spin angular momentum operators, β_e is the Bohr magneton, β_n is the nuclear magneton, \vec{B} is the applied magnetic field, \tilde{g} is the g-tensor of the electron spin, $\tilde{g}_{i,n}$ is the g-tensor of the i th nuclear spin, \hbar is the reduced Planck's constant, \tilde{A}_i and \tilde{Q}_i are the i th electron-nuclear hyperfine tensor and the quadrupole tensor of the i th nuclear spin, respectively. The nuclear quadrupole interaction is non-zero only for $I > 1/2$. Details on the different

tensor elements, the expansion of the Hamiltonian, and the solution of the Schrodinger equation to get the corresponding energy levels are available elsewhere⁶⁴.

1.3 CONTINUOUS WAVE (CW) ESR

1.3.1 CW-ESR spectrum for Cu^{2+}

One of the commonly used ESR experiment is the continuous wave (CW) ESR experiment. CW-ESR provides insight into several characteristics of the system, such as the coordination geometry, coordination environment of the ESR probe, presence of oligomers, binding affinities, and presence of multiple different species in the solution. In this section, we will focus on Cu^{2+} CW-ESR spectrum and how we can characterize a system using such spectrum.

In the absence of any magnetic field, the magnetic moment associated with the electron spin is randomly oriented and, for $S = 1/2$ as in the case of Cu^{2+} , the two energy levels are degenerate. When an external magnetic field B_0 is applied, the energy states lose their degeneracy and split into two, corresponding to $m_s = \pm 1/2$, where m_s is the quantum number for the z-component of electron spin angular momentum operator. In a CW-ESR experiment, the energy difference between the spin states is harnessed to produce the spectrum. A sample with unpaired electrons is placed in a constant-frequency microwave field and the magnetic field is swept. When the energy of the microwave radiation matches the energy separation between the spin states, absorption occurs, and spectrum is obtained. In CW-ESR, the dominant interactions are the electron Zeeman (~ 9.5 GHz at X-band frequency) and the hyperfine interaction, i.e., between Cu^{2+} electron and nuclear spin (~ 450 MHz). The remaining terms of nuclear Zeeman interaction (~ 4

MHz) and nuclear quadrupole interaction (<3 MHz) are negligible and hence, are not resolved by CW-ESR. Therefore, the transitions in a CW-ESR experiment can be given by the dominant terms in the electron spin Hamiltonian as:

$$\hat{H} = \beta_e \vec{B} \cdot \tilde{g} \cdot \hat{S} + \hbar \sum_{i=1}^k (\hat{S} \cdot \tilde{A}_i \cdot \hat{I}_i) \quad (1-3)$$

The first dominant term, i.e. the electron Zeeman interaction is the interaction between the electron spin and the external magnetic field. Besides, the interaction of the orbital angular momentum with the applied magnetic field and the spin-orbit coupling (SOC)⁶⁵ also induces additional magnetic field on the electron spin. This induced magnetic field depends on the shape of the orbital in which the unpaired electron is delocalized, and on the orientation of the molecule in the magnetic field. On the other hand, the dominant hyperfine interaction is the interaction between the unpaired electron spin and the nuclear spin of Cu²⁺. Cu²⁺ ion has an electron spin, S of ½ ($m_S = +1/2, -1/2$) and a nuclear spin, I of 3/2 ($m_I = +3/2, +1/2, -1/2, -3/2$); where m_S and m_I are the quantum numbers for the z-components of electron spin angular momentum and nuclear spin angular momentum operators, respectively. Based on the selection rules of $\Delta m_S = \pm 1$ and $\Delta m_I = 0$, four hyperfine transitions are observed. The energy level diagram for Cu²⁺ is shown in Figure 1-3.

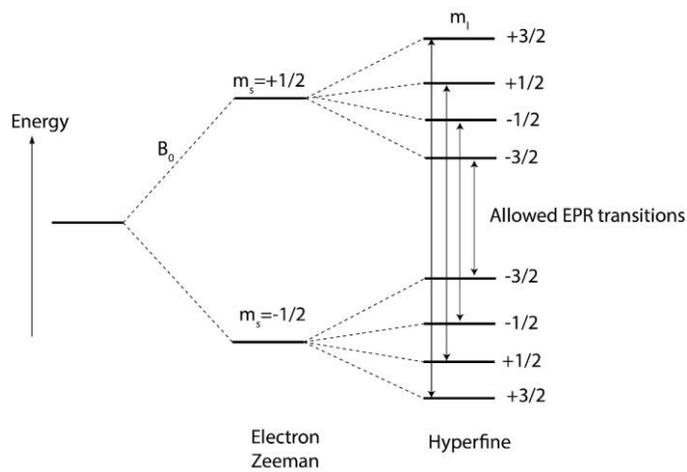


Figure 1-3 Energy level diagram of Cu²⁺.

CW-ESR experiments performed in this dissertation on Cu^{2+} -based systems are at low temperatures and in a frozen matrix. In such cases, the effective magnetic field experienced by the electron spin is orientation dependent. This orientation dependent variation of the Zeeman splitting is expressed by a g-tensor. There are three principal values of the g-tensor, g_{xx} , g_{yy} and g_{zz} . In the case of axial symmetry as seen in type-II Cu^{2+} centers, $g_{xx} = g_{yy} = g_{\perp}$ and $g_{zz} = g_{\parallel}$, as shown in Figure 1-4A. Thus, in a sample of randomly oriented molecules (Figure 1-4B), the resonance due to Zeeman interaction will occur at:

$$h\nu = g(\theta)\beta_e B_0 \quad (1-4)$$

where, ν is the microwave frequency and the angle θ is the angle between the applied magnetic field, B_0 and the molecular zz -axis. The $g(\theta)$ can be further expressed as:

$$g(\theta) = \sqrt{g_{\perp}^2 \sin^2 \theta + g_{\parallel}^2 \cos^2 \theta} \quad (1-5)$$

Since the sample is frozen, all θ angles will be captured, giving a range of resonant magnetic fields. However, statistically there are more molecules whose xx/yy axes will be aligned with the magnetic field than the zz axis (Figure 1-4C). Hence, the intensity of the spectrum increases sinusoidally from when the zz -axis is parallel to the magnetic field ($\theta=0^\circ$) to when the xx,yy -plane is parallel to the field ($\theta=90^\circ$), as shown in Figure 1-4D.

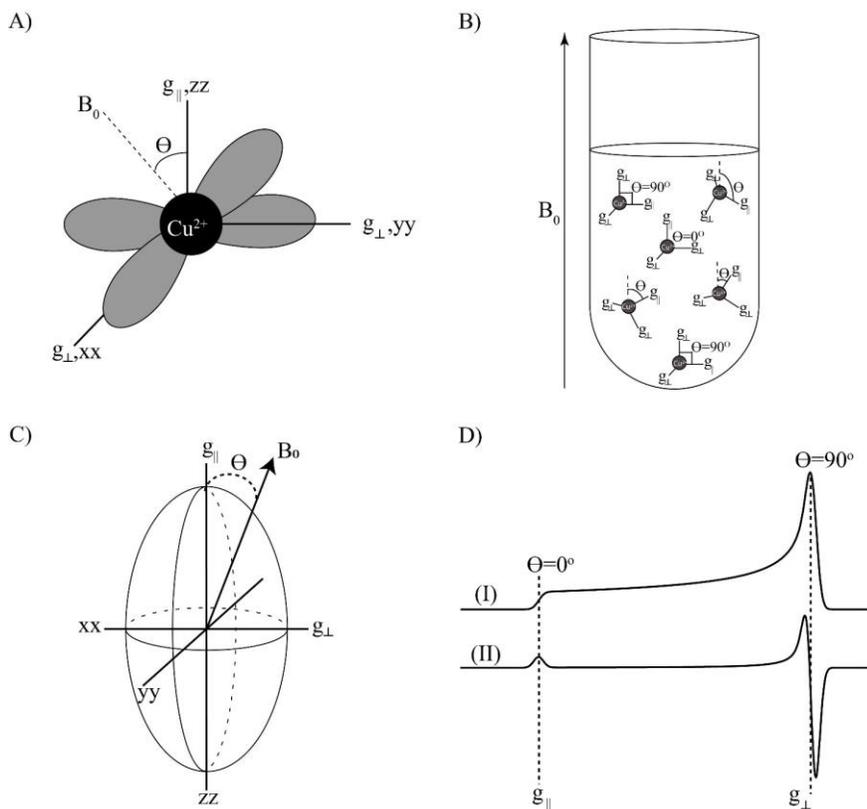


Figure 1-4 A) Unpaired electron in Cu^{2+} present in the $d_{x^2-y^2}$ orbital. Type (II) Cu^{2+} shows axial symmetry whereby, $g_{xx}=g_{yy}=g_{\perp}$ and $g_{zz}=g_{\parallel}$. The angle θ is the angle between the applied magnetic field, B_0 and the molecular zz -axis. B) A sample of randomly oriented molecules with random θ values. C) Schematic drawings of an axial g -tensor ellipsoid. For a sample with random orientations, there will be higher probability for the molecules to be at $\theta=90^\circ$ than $\theta=0^\circ$. D) The corresponding ESR spectrum. The intensity corresponding to $\theta=90^\circ$ will be the highest as shown in Trace (I). Trace (II) shows the first derivative of the absorption spectrum, as seen in CW-ESR.

The hyperfine interaction leads to further splitting of the lineshape as shown in Figure 1-3.

Thus, for a particular θ , the resonance condition is:

$$h\nu = g(\theta)\beta_e B_0 + A(\theta)m_I \quad (1-6)$$

where $A(\theta)$ is given as⁶⁶:

$$A(\theta) = \sqrt{\frac{A_{\perp}^2 g_{\perp}^4 \sin^2 \theta + A_{\parallel}^2 g_{\parallel}^4 \cos^2 \theta}{g_{\perp}^2 \sin^2 \theta + g_{\parallel}^2 \cos^2 \theta}} \quad (1-7)$$

The resultant Cu^{2+} CW-ESR spectrum is shown in Figure 1-5. The g_{\parallel} and A_{\parallel} are usually better resolved and hence, are used to determine the coordination environment of Cu^{2+} .

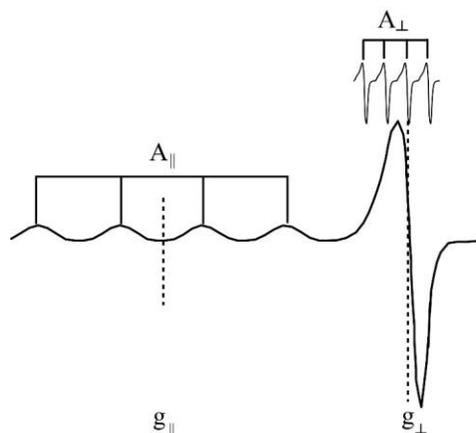


Figure 1-5 CW-ESR spectrum of type-II Cu^{2+} .

1.3.2 Characterization of Cu^{2+} coordination environment using CW-ESR spectrum

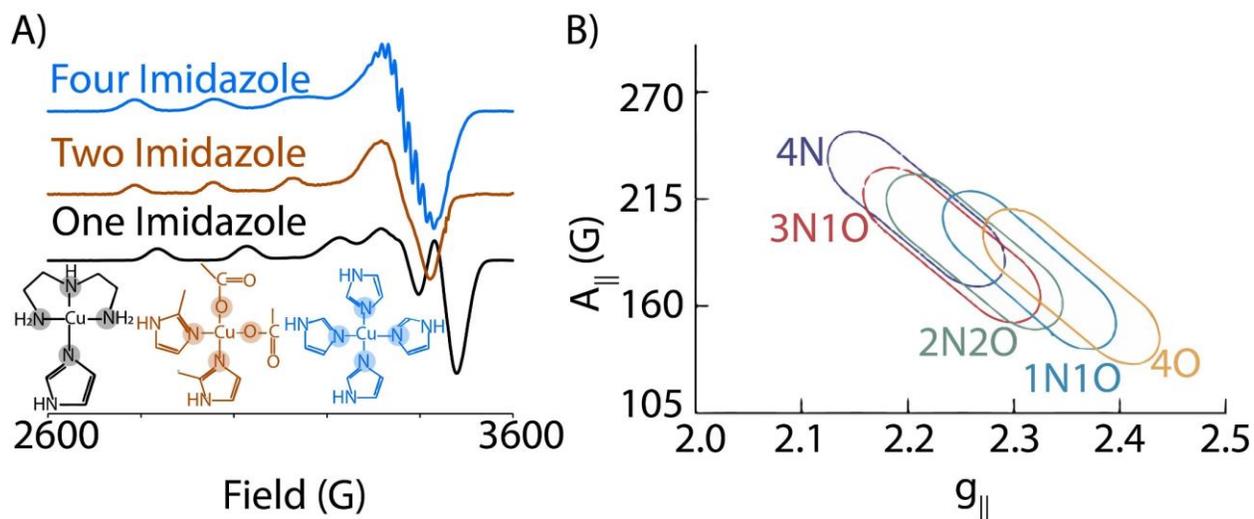


Figure 1-6 A) CW-ESR spectrum of differing Cu^{2+} coordination. The direct coordination environment of Cu^{2+} in each compound is shown by the shaded circles. The lineshape is different between the three spectra due to differing g_{\parallel} and A_{\parallel} values. These results are taken with permission from Ref. 69. B) The Peisach-Blumberg plot depicting the relationship between Cu^{2+} centers with different equatorially coordinated ligands and their respective g_{\parallel} and A_{\parallel} values.

Changes in the Cu^{2+} coordination environment are directly reflected in the g_{\parallel} and A_{\parallel} values⁶⁷ of the CW-ESR spectrum, as shown in Figure 1-6A. If the direct coordination of Cu^{2+} to oxygen is replaced by nitrogen, the g_{\parallel} value decreases and the A_{\parallel} increases. These trends are similar to that noted in literature⁶⁷⁻⁶⁹ and as shown in Figure 1-6B. Histidine is one of the strongest metal ion coordinators in proteins. Histidine contains an imidazole ligand with two nitrogen atoms to which Cu^{2+} can bind. Copper-containing proteins often have binding sites with irregular geometries containing one or more histidine ligands⁷⁰⁻⁷². Thus, for Cu^{2+} binding to histidine, the directly coordinated atom is the nitrogen atom of the imidazole moiety. This Cu^{2+} -N electron nuclear hyperfine interaction can be measured by CW-ESR.

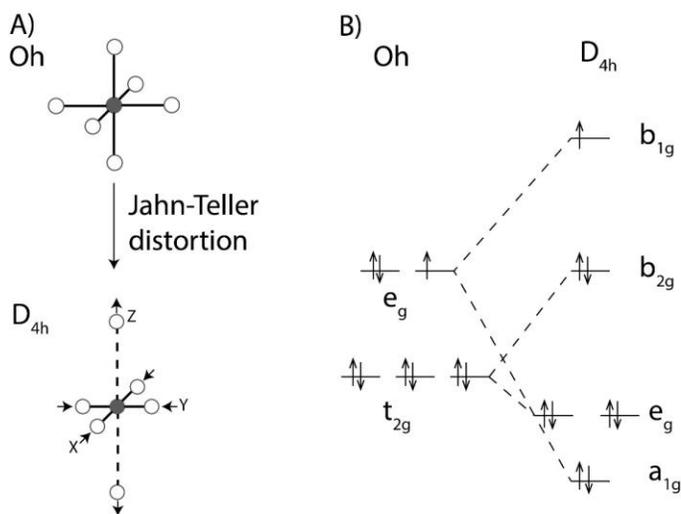


Figure 1-7 A) Jahn-Teller distortion of octahedral geometry leading to a tetragonal distortion B) Energy level diagram of octahedral to square planar geometry.

The dependence of direct ligand coordination to g values can be explained via the ligand field theory (LFT). For an octahedral complex, the d-orbitals are split into two sets⁷³: a 3-fold degenerate t_{2g} at a lower energy and a 2-fold degenerate e_g at a higher energy (Figure 1-7A). However, for Cu^{2+} , a d^9 system, there exists Jahn-Teller distortion, leading to a tetragonal distortion by axial elongation. As a result, the degenerate orbitals are further split, as shown in

Figure 1-7A. The splitting of the d-orbitals is based on the energy difference associated with the respective orbitals which in turn is dependent on the coordinated ligand with the metal. For Cu^{2+} , as shown in Figure 1-7B, the half-filled $d_{x^2-y^2}$ orbital has the highest energy. This is because the $d_{x^2-y^2}$ orbital is involved in an antibonding or repulsive interaction with the ligands in the equatorial plane (as its lobes are oriented along the ligand-metal bonds).

The SOC is also responsible for changes in g values with change in ligand coordination. For transition metal complexes, the SOC constants are large. For example, Cu^{2+} has a SOC constant of $\sim -830 \text{ cm}^{-1}$. As a result, there will be large mixing of the SOC with the excited states leading to large deviations of g values. When the d-orbitals of Cu^{2+} are coordinated to a ligand, they contain some ligand character. The electronic characteristics of the donor ligand atom influences the ligand character and thereby the covalency of the appropriate bonding. When Cu^{2+} is coordinated to nitrogen, the covalent character of the bond is much higher than when oxygen is present. Increase in covalency will increase the ligand character of the d-orbital. This in turn decreases the magnitude of the orbital angular momentum and thereby the SOC. Consequently, the g_{\parallel} decreases. For Cu^{2+} coordinated to oxygen, the SOC is more compared to nitrogen, as a result g_{\parallel} is more as well^{74,75}.

Similarly, for hyperfine interaction, there is dependence on the ground state of the metal ion as well as the covalency with the ligands. For Cu^{2+} ⁷⁶:

$$A_{\parallel} = A_{zz} = \beta^2 A_{fc} + \beta^2 A_{dd} + P \left[(g_{\parallel} - 2.00) + \frac{3}{7} (g_{\perp} - 2.00) \right] \quad (1-8)$$

$$A_{\perp} = A_{xx}, A_{yy} = \beta^2 A_{fc} - \frac{1}{2} \beta^2 A_{dd} + P \left[\frac{11}{14} (g_{\perp} - 2.00) \right] \quad (1-9)$$

where, β is the amount of metal character in the half-occupied $d_{x^2-y^2}$ orbital of Cu^{2+} , A_{fc} and A_{dd} are hyperfine constants due to Fermi contact and dipolar interaction, respectively and P is the free-ion dipole term (P for $\text{Cu}^{2+} \sim 0.036 \text{ cm}^{-1}$). The first term in the equation is the isotropic Fermi

contact interaction of the electron spin with the nuclear spin at the nucleus. The second term is the dipolar interaction between the electron and the nuclear spin. This term is anisotropic and is dependent on the orientation with respect to the magnetic field. The third term is the magnetic dipole interaction of the orbital angular momentum with the Cu^{2+} nuclear spin. Thus, if the covalency of the bonding increases, the electron density will be less delocalized. As a result, the interaction of the electron and the nuclear spin on Cu^{2+} will increase. Consequently, β or the metal character of the wavefunction will increase, increasing the value of A_{\parallel} . This explains the increasing trend in A_{\parallel} values with increased coordination with nitrogen.

Often there are distinct features in the A_{\perp} region which is due to superhyperfine interaction⁶⁸. Superhyperfine is the interaction of the electron spin of Cu^{2+} with the nuclear spin of the directly coordinated nitrogen. The features due to this interaction become more distinct as the number of directly coordinated imidazole rings increase, as seen in Figure 1-6A.

1.3.3 Binding affinity of Cu^{2+} to its binding site

CuCl_2 in *N*-ethylmorpholine (NEM) buffer at pH 7.4 is sparingly soluble in water and precipitates as ESR silent $[\text{Cu}(\text{OH})_2]_n$ ^{77,78}. Hence, any excess Cu^{2+} does not cause interference from the aquo species. Any Cu^{2+} signal that arises under such conditions is due to bound Cu^{2+} , as free Cu^{2+} will form precipitates. Using a standard ESR calibration curve, the bound Cu^{2+} concentration in solution can be determined from the integrated intensity of the spectrum. Now if the concentration of the binding site in question is known, then the fraction of loading of the binding site can be obtained. Performing titrations where the Cu^{2+} concentration is gradually incremented, will then provide the binding affinity of Cu^{2+} to the site in question. Once the site is completely loaded any addition of Cu^{2+} will not contribute to the ESR spectrum. Thus, using a

series of CW-ESR spectra, we can obtain the apparent dissociation constant, K_d of a particular binding site. In case of multiple binding sites, CW-ESR is advantageous. For two or more similar sites, the analysis remains the same as a one-component spectrum, as mentioned earlier. However, for two or more distinct Cu^{2+} binding environments, the spectral lineshape of each component will be different and thus, easy to distinguish. The components can then be easily simulated, and the binding affinity of each site can be determined.

1.4 ELECTRON SPIN ECHO ENVELOPE MODULATION (ESEEM)

While CW-ESR provides information about the immediate coordination environment, pulsed ESR techniques of electron spin echo envelope modulation (ESEEM) probes the interaction between the nuclear spin and the electron spin within a separation of 3-10 Å^{79,80}. The experiment is based on a three-pulse sequence^{79,81}. A useful summary of ESEEM on Cu^{2+} -histidine systems is also available elsewhere^{82,83}. ESEEM is useful for measuring weak hyperfine interactions, such as the interaction between Cu^{2+} electron spin and the remote ¹⁴N nuclear spin of the imidazole moiety. These remote nuclei are circled in Figure 1-8B. In case of Cu^{2+} -bound to a histidine residue, Cu^{2+} has a hyperfine interaction with both the nitrogen atoms in the imidazole ring of the histidine residues. For the directly coordinated nitrogen nuclei, the magnitude of the hyperfine interaction of the electron spin of Cu^{2+} with the nuclei is too large (~40 MHz) to be detected by the bandwidth of the typical pulse lengths used in ESEEM. However, for the remote nitrogen, the hyperfine interaction is much smaller and has a corresponding frequency that can be easily detected by ESEEM at X-band frequency. Figure 1-8A shows the characteristic background subtracted time domain ESEEM data for Cu^{2+} -bound to histidine. The spectrum in Figure 1-8B shows peaks

characteristic to the imidazole moiety. The remote nitrogen of the histidine results in three peaks below 2 MHz and a broad peak at ~ 4MHz^{68,81,84–86}, as shown in Figure 1-8B.

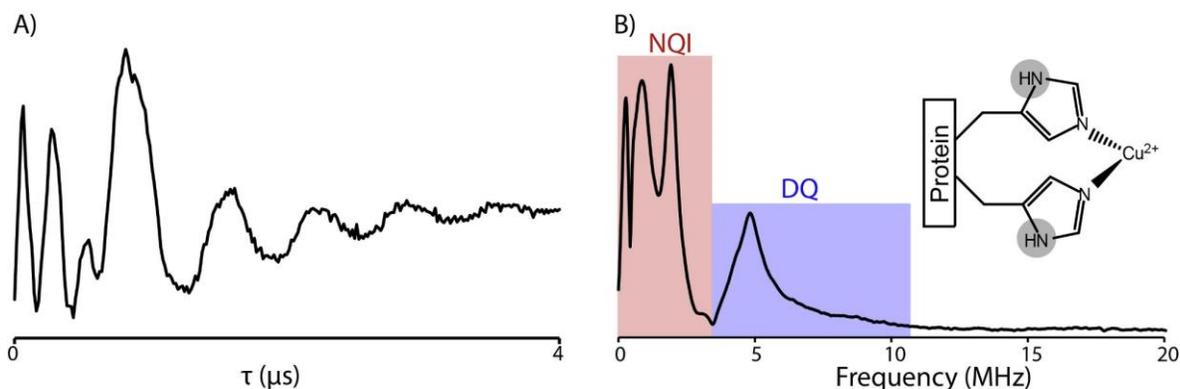


Figure 1-8 A) Background subtracted time domain ESEEM spectrum of Cu^{2+} coordinated to two histidine residues. B) Fourier transformed ESEEM spectrum shows the characteristic NQI peaks below 2 MHz and DQ peak at ~ 4 MHz. The inset shows Cu^{2+} bound to two histidine residues and the remote nitrogen of the imidazole ring that contributes to the ESEEM are circled (Ref. 51).

Figure 1-9 shows these transitions in the energy level diagram. The three NQI peaks, as shown in Figure 1-8B, are a consequence of the energy due to nuclear Zeeman interaction and the hyperfine interaction cancelling each other out in the α -spin manifold ($m_s = +1/2$) (Figure 1-9). The three peaks result from three transitions, with frequencies given by⁸⁷:

$$\nu_- = \frac{e^2qQ(3-\eta)}{4h} ; \nu_0 = \frac{2\eta e^2qQ}{4h} ; \nu_+ = \frac{e^2qQ(3+\eta)}{4h} \quad (1-10)$$

where e is the electron charge, q is the electric gradient at the nuclear site, Q is the nuclear quadrupole moment, η is the asymmetry parameter and h is Planck's constant. The three NQI frequencies are related by the following equation:

$$\nu_+ = \nu_- + \nu_0 \quad (1-11)$$

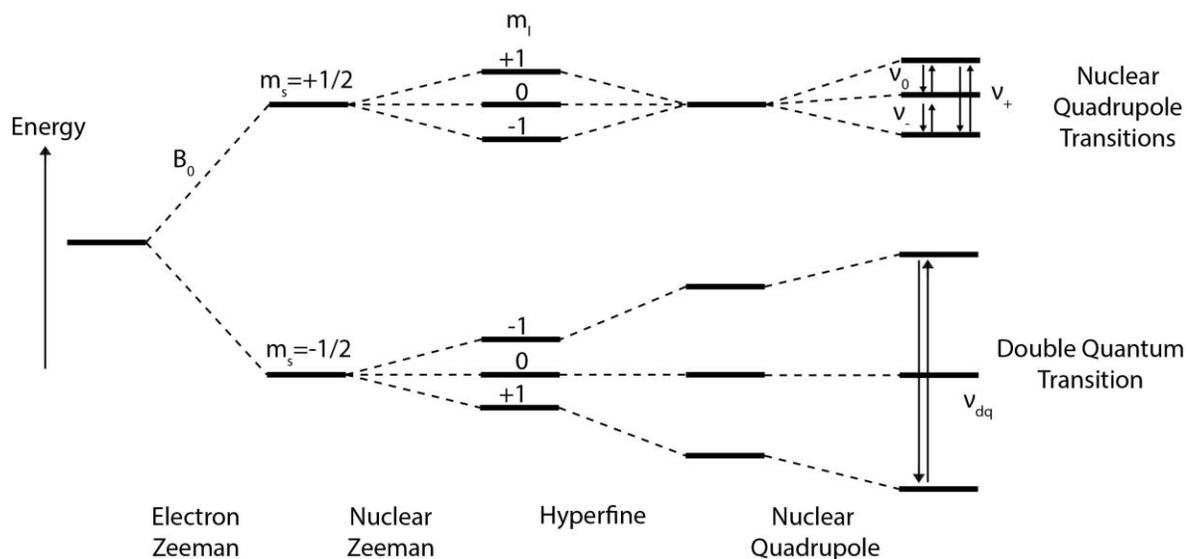


Figure 1-9 Interaction of the unpaired electron of Cu²⁺ with the nuclear spin of the remote nitrogen of the Cu²⁺-coordinated imidazole moiety

The DQ peak is observed in the β -spin manifold ($m_s = -1/2$) where the nuclear Zeeman and the hyperfine interactions are additive and hence, give rise to much broader resonances. Often, the only distinguishable feature is the double quantum (DQ) transition and the corresponding frequency is given as⁸⁸:

$$\nu_{dq} = 2\sqrt{\omega_{\beta}^2 + \left(\frac{e^2qQ}{4h}\right)^2 (3 + \eta^2)} \quad (1-12)$$

where ω_{β} is the nuclear transition frequency in the β electron spin manifold.

In this dissertation, ESEEM has been used for two main purposes – first, to identify histidine coordination to Cu²⁺ and second, to obtain information on optimum binding conditions of Cu²⁺ to the histidine residue. To elucidate histidine coordination, the Fourier-transformed spectrum of the ESEEM signal is analyzed and the presence of the NQI and DQ peaks is observed, as shown in Figure 1-8B. Besides, the spectrum also helps in quantifying the number of histidine residues coordinated to Cu²⁺. As the number of coordinating histidine residues increases, the DQ transition becomes more prominent and the DQ peak increases. Consequently, the integrated

intensity of the ^{14}N region in the spectrum increases. The ratio of the integrated intensities of the ^{14}N region in the spectrum (0-11 MHz) normalized to the ^1H region (between 13 and 16 MHz) can then help to quantify the number of histidine residues coordinated to the Cu^{2+} ^{68,89}. To probe the binding affinity of Cu^{2+} to the histidine site, a series of ESEEM titrations is performed where the histidine concentration, thereby, the protein concentration is kept constant and the Cu^{2+} concentration is gradually incremented. The ESEEM spectrum now is a sum of two components - Cu^{2+} -bound to histidine and unbound Cu^{2+} , given as:

$$V(t) = x.V_{\text{bound}}(t) + (1 - x).V_{\text{unbound}}(t) \quad (1-13)$$

where $V(t)$ is the total ESEEM signal, $V_{\text{bound}}(t)$ is the ESEEM signal where all the Cu^{2+} is completely bound to the histidine, $V_{\text{unbound}}(t)$ is the ESEEM signal for free Cu^{2+} in solution and x is the fraction of Cu^{2+} bound to the histidine. From the time-domain data, the depth of the ^{14}N modulations can be analyzed, given by the parameter, k ⁹⁰. The second modulation period of ^{14}N being most distinct is used for the modulation depth analysis. The modulation depth, k , is often defined as⁹⁰:

$$k = \frac{a}{(a+b)} \quad (1-14)$$

where, a is the echo intensity between the second and third crests of the signal and b is the echo intensity of the second trough of the signal, as shown in Figure 1-10A. As the histidine site gets loaded with Cu^{2+} , the population of Cu^{2+} bound to the imidazole moiety or $V_{\text{bound}}(t)$ increases. Additionally, there will also be contributions from the unbound Cu^{2+} -complex, $V_{\text{unbound}}(t)$, resulting in a featureless decay. Once the histidine site is fully loaded, further addition of Cu^{2+} leads to a considerable increase in unbound Cu^{2+} and consequently, in the $V_{\text{unbound}}(t)$. This abundance of free Cu^{2+} in solution significantly dampens the ^{14}N modulations, decreasing the

modulation depth. Therefore, a plot of k versus the Cu^{2+} concentration is useful to obtain the optimal binding conditions of Cu^{2+} to the histidine residues^{49,91}, as shown in Figure 1-10B.

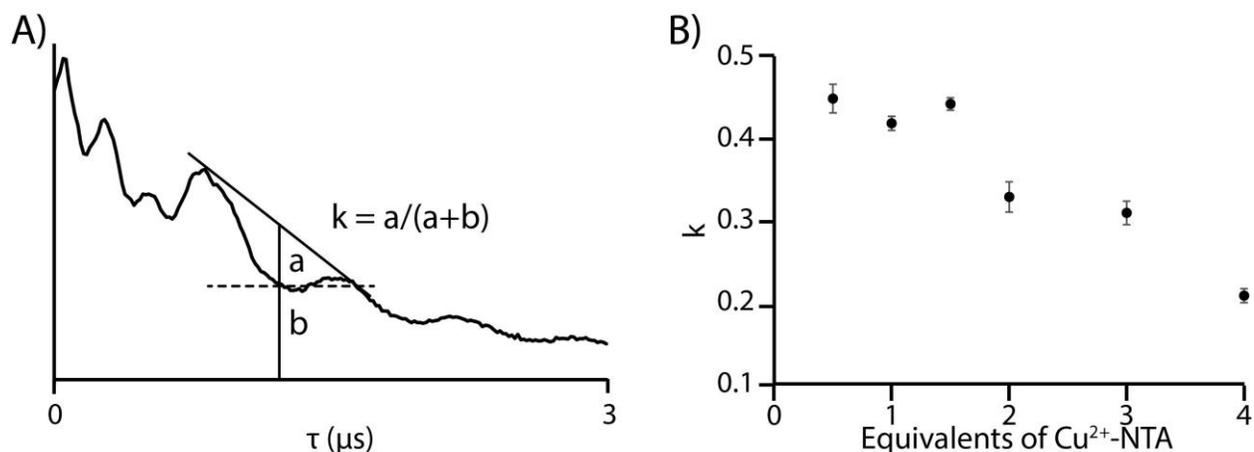


Figure 1-10 Raw time-domain ESEEM data showing the calculation of the parameters, a and b and the modulation depth, k . B) An example plot from Ref. 49 showing the plot of k versus the equivalents of a Cu^{2+} -complex that binds to histidine residues. The plot shows the contribution of unbound Cu^{2+} -complex towards the decrease in k .

In the one-dimensional three-pulse ESEEM, the first two pulses are separated by a fixed time τ , followed by a third pulse after time T which is incremented. Modulations in ESEEM arise as T increases due to nuclear transition frequencies of nuclei coupled with the electron spin. However, the intensity of each peak depends on the phase that is acquired during precession in the fixed period τ , and in the most adverse case a peak can be completely suppressed. This τ -dependent phenomenon is referred to as a blind spot in ESEEM. Therefore, ESEEM needs to be performed at several τ values. In such cases, hyperfine sublevel correlation (HYSCORE) is an efficient technique to overcome this blind spot effect. HYSCORE is essentially a two-dimensional ESEEM experiment where both the τ and T are incremented in separate dimension. The pulse sequence for HYSCORE is available in literature^{92,93}. A discussion of HYSCORE for histidine coordinated Cu^{2+} is available elsewhere⁸³.

In this dissertation, HYSCORE has been used to quantify the coordination environment for Cu^{2+} . Based on the position of diagonal and cross peaks in the HYSCORE spectrum, one can differentiate between axial or equatorial coordination with water^{94,95}. Furthermore, coordination of Cu^{2+} to the protein backbone can also be verified through this spectroscopic technique^{95,96}.

1.5 DOUBLE ELECTRON ELECTRON RESONANCE

Double electron electron resonance (DEER) spectroscopy is a widely used pulsed ESR experiment which measures the distance-dependent dipolar interaction between two or more unpaired electron spins. The commonly used DEER technique has a four-pulse sequence, as reported elsewhere¹⁶. A useful summary of Cu^{2+} based DEER can be found in literature^{82,97}. The double resonance technique uses two different frequencies whereby a set of spins is “observed” at one frequency and another set of spins is excited at the second frequency (called the pump pulse). The interaction is manifested as a modulation of the spin echo amplitude at different pulse separations between the “observer” pulse and the “pump” pulse.

We first consider a situation where a biomolecule has two labeled sites, as shown in Figure 1-11A. In solution, there are a combination of signals from both intramolecular and intermolecular interactions between the spins, as shown in Figure 1-11A. DEER measures all these interactions and the combined signal is given by:

$$V(t) = V_{intra}(t) * V_{inter}(t) \tag{1-15}$$

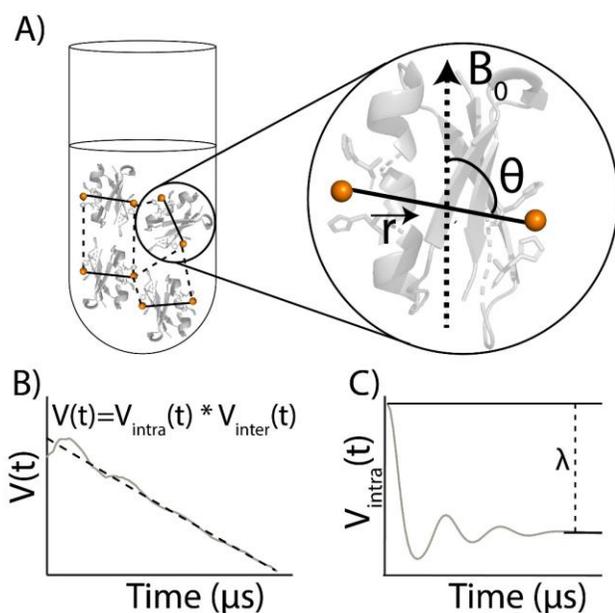


Figure 1-11 A) Intramolecular interaction (solid line) and intermolecular interaction (dashed line) between two spins in solution. The distance between two spins is represented by the interspin vector, \vec{r} . The angle between \vec{r} and the applied magnetic field B_0 is defined as θ . B) Raw time domain DEER data (solid) and the background decay (dashed). C) Background subtracted time domain DEER data showing intramolecular contributions only. The dashed line shows the modulation depth.

A representative raw time-domain DEER data, $V(t)$, is shown in Figure 1-11B. Since spin-labeled systems are randomly distributed in a sample, the large range of distances average out and make $V_{\text{inter}}(t)$ signal a featureless decay. For homogeneously distributed proteins in solution, the decay is exponential⁹⁸ (cf. Figure 1-11B, dashed line). Typically, the last 25% of the raw signal is fit to an exponential in order to estimate $V_{\text{inter}}(t)$. The signal of interest, $V_{\text{intra}}(t)$, can then be extracted by dividing the raw DEER signal with the estimated $V_{\text{inter}}(t)$ (cf. Figure 1-11C).

In general, the DEER data is acquired with two considerations in mind. An essential criterion is the ability to obtain the appropriate $V_{\text{inter}}(t)$ from $V(t)$. The $V_{\text{inter}}(t)$ dominates the end of the $V(t)$ signal while $V_{\text{intra}}(t)$ dominates $V(t)$ initially. Thus, if the total acquisition time of the experiment is not long enough, then the background fit will include the $V_{\text{intra}}(t)$ signal, giving erroneous results. To avoid this outcome, the total acquisition time of the experiment should be

long enough such that the last 25% of the raw data can be used to fit as the background, without including the $V_{\text{intra}}(t)$. In addition, in order to obtain an accurate distance distribution, the data should be acquired such there are at least two modulation periods corresponding to the longest distance^{99,100}. For Cu^{2+} this requirement is given by:

$$t = \left(\frac{r_{\text{max}}^3}{k} \right) \times 2 \quad (1-16)$$

where t is the total dipolar evolution time corresponding to two periods in μs , r_{max} is the longest distance in nm and k is a constant given by:

$$k = \frac{\mu_0}{4\pi\hbar} g_A g_B \beta_e^2 \quad (1-17)$$

where, g_A and g_B are the g values for the two different spins, A and B. For Cu^{2+} , k has a value of 62.84 is $\text{MHz}\cdot\text{nm}^{-3}$.

In general, there can exist species with different number of labeled spins. The total intramolecular signal from these species, $V_{\text{intra}}(t)$ is given by:

$$V_{\text{intra}}(t) = \sum_{n=1}^j f_n V_{\text{intra},n}(t) \quad (1-18)$$

where n is the number of spins in each species, f_n is the fraction of species with n spins, j is the total number of different labeled species present in solution and $V_{\text{intra},n}(t)$ is the signal corresponding to species with n spins. The signal $V_{\text{intra},n}(t)$ is further given by:

$$V_{\text{intra},n}(t) = \frac{1}{n} \sum_{A=1}^n \left\langle \prod_{\substack{B=1 \\ B \neq A}}^n \left(1 - p_b (1 - \cos(\omega_{AB}t)) \right) \right\rangle_{r,\theta} \quad (1-19)$$

where A and B are two different spins, θ is the angle between the interspin vector and the applied magnetic field, r is the distance between the spins A and B, p_b is the fraction of spin B excited by the pump pulse, t is the time by which the pump pulse is stepped out and ω_{AB} is the dipolar frequency of the coupled spins. The dipolar frequency, ω_{AB} is given as:

$$\omega_{AB} = \frac{\mu_0}{4\pi\hbar r^3} g_A g_B \beta_e^2 (1 - 3 \cos^2 \theta) \quad (1-20)$$

The background subtracted signal can be further simplified at conditions when $t \gg \omega_{AB}$ such that $t \rightarrow \infty$ and $\cos(\omega_{AB}t)$ term in equation 1-19 becomes 0. $V_{intra,n}(t)$ from equation 1-19 now simplifies to:

$$V_{intra,n}(t \rightarrow \infty) = (1 - p_b)^{n-1} \quad (1-21)$$

Substituting equation 1-21 in equation 1-18, we get:

$$V_{intra}(t \rightarrow \infty) = \sum_{n=1}^J f_n V_{intra,n}(t \rightarrow \infty) = \sum_{n=1}^J f_n (1 - p_b)^{n-1} \quad (1-22)$$

1.5.1 Determining the population of different labeled species and oligomeric states

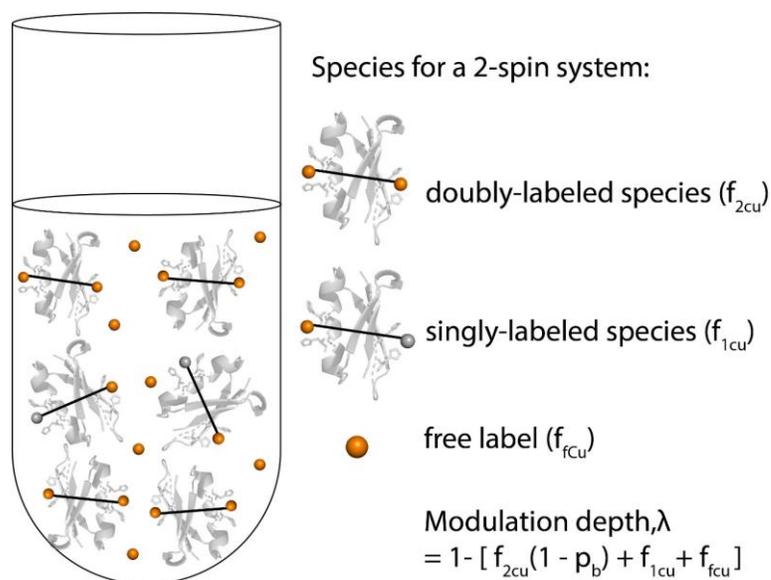


Figure 1-12 Dependence of modulation depth, λ , on the population of different labeled spin systems in solution.

The modulation depth, λ , of a DEER signal, as shown in Figure 1-11C, is an important parameter that is sensitive to the different labeled species in a system as well as the population of those species present in solution. The modulation depth is given as:

$$V_{intra}(t \rightarrow \infty) = 1 - \lambda \quad (1-23)$$

Combining equations 1-22 and 1-23, we get:

$$\lambda = 1 - \sum_{n=1}^j f_n (1 - p_b)^{n-1} \quad (1-24)$$

The above equation shows the dependence of modulation depth on the different species present in solution. This point is further illustrated in Figure 1-12, showing the different possible species for a protein with two sites for spin labeling. The protein can either be labeled at both sites (n=2) or only at one site (n=1). In addition, there can also be free spin labels in solution. In such a system, the modulation depth can be given as:

$$\lambda = 1 - ((f_{2cu}(1 - p_b)^1 + f_{1cu} + f_{fcu}) \quad (1-25)$$

where, f_{2cu} is the fraction of doubly labeled species, f_{1cu} is the fraction of singly labeled species and f_{fcu} is the fraction of free label in solution, such that:

$$f_{2cu} + f_{1cu} + f_{fcu} = 1 \quad (1-26)$$

Thus, from modulation depth in DEER, we can predict the fraction of different labeled species which, in turn, provides insight on the labeling efficiency^{49,91}. Moreover, equation 1-22 can be further expanded to account for different oligomeric assemblies in solution⁶².

1.5.2 Obtaining distance distribution

In solution, a macromolecule can exist in different conformations due to its inherent flexibility. As a result, the intramolecular interaction results in a distance distribution rather than a single distance. For a simple two-spin system (n=2), equation 1-19 can be represented as an integral, considering all possible distances with corresponding probabilities, P(r) as⁹⁸:

$$V_{intra}(t) = \int_{R_{min}}^{R_{max}} \int_0^1 [1 - p_b(1 - \cos(\omega_{AB}t))] \sin\theta d\theta P(r) dr \quad (1-27)$$

where $P(r)$ is the distance distribution function, and R_{\max} and R_{\min} are the maximum and minimum possible distances. Equation 1-27 can be further simplified as a Fredholm equation of the first kind¹⁰¹:

$$V_{intra}(t) = \int_{R_{min}}^{R_{max}} P(r) \kappa(r, t) dr \quad (1-28)$$

where $\kappa(r, t)$ is a kernel function given by:

$$\kappa(r, t) = \int_0^1 1 - p_b(1 - \cos(\omega_{AB}t)) \sin\theta d\theta \quad (1-29)$$

Once the background subtracted $V_{intra}(t)$ of the DEER signal is obtained, the data can then be converted into a distance distribution using several approaches such as Tikhonov regularization^{98,101–103}, Gaussian models^{104–106}, Bayesian inference¹⁰⁷, singular value decomposition¹⁰⁸, wavelet denoising¹⁰⁹, and neural networks¹¹⁰. Among these, Tikhonov regularization is the most widely employed and has also been primarily used in this dissertation. The Tikhonov regularization method solves for the distance distribution, $P(r)$ using the known kernel function, κ and with the criterion that all the elements of $P(r)$ should be non-negative and that the distributions should be smooth. Once the distance distribution is obtained, the width provides insight into the flexibility of the macromolecule as well as the label.

2.0 Cu²⁺-BASED DISTANCE MEASUREMENTS BY PULSED ESR PROVIDE DISTANCE CONSTRAINTS FOR DNA BACKBONE

This work, written in collaboration with Matthew J. Lawless, Hanna J. Brubaker, Kevin Singewald, Michael R. Kurpiewski, Linda Jen-Jacobson and Sunil Saxena, has been published in Nucleic Acids Research, 2020, volume 48, page e49. The dissertation author collected and analyzed a majority of the ESR data, performed modeling and simulations, and prepared the manuscript.

2.1 INTRODUCTION

DNA dynamics is an important factor that affects numerous cellular processes mediated by protein-DNA interactions^{111–115}. Often, upon interaction with a protein at specific sites, structural changes in the DNA such as bending, or twisting are induced within the DNA. The flexibility of the DNA duplex and its ability to adapt its shape are crucial for triggering countless cellular activities such as transcription¹¹⁶, replication¹¹⁷ and gene regulation¹¹⁸. Often due to the large size of protein-DNA complexes, low solubility and timescale of conformational changes, these processes are inaccessible to NMR and crystallographic techniques. On the other hand, electron spin resonance (ESR) techniques have become an invaluable method to probe conformational changes in such cases. Particularly, when two or more spins are present, pulsed ESR techniques can be employed to obtain point-to-point distances within a macromolecule. Such distance

constraints in conjunction with the available structures of the macromolecule can be used to model the conformations of the macromolecule in the different functional states^{52,119–126}.

To implement pulsed ESR techniques for distance measurements, one needs to incorporate two or more spin labels at specific sites in the DNA. To this end, a wide variety of spin labels have been developed for nucleic acids^{127,128}. These methods include modification of the nucleobase^{34,35,129–135}, backbone^{136–139} or terminal capping^{140,141}. Nitroxide based labels, particularly the cytidine analogue, C, ^{25–27} provide extremely rigid distance distributions as well as information on label orientation^{30,32,142}. These two pieces of information together have proven to be capable of reporting on inherent DNA motions in even small systems such as the cocaine aptamer¹⁴³. Radicals, such as the triarylmethyl (TAM) spin label, often attached at the oligonucleotide termini, have been used to demonstrate distance measurements in nucleic acids at physiological temperatures^{144,145}. Sterically shielded nitroxide labels, introduced post-synthetically,¹³⁹ and non-covalently bonded nitroxide labels, attached to an abasic site²⁸, that position the label closer to or within the helix have also been developed. Chelation of paramagnetic metal ions such as Gd³⁺, Mn²⁺ or Cu²⁺ ^{146–149} has been introduced as an alternative labeling methodology. Despite the success of such labeling strategies, there is a need for labeling schemes that are nucleotide independent, can be positioned anywhere within the DNA, and are small enough to reside within the helix.

Recently, we reported a Cu²⁺ based labeling method as a promising strategy to measure DNA backbone distances⁵⁰. The method involves the incorporation of a Cu²⁺-chelating ligand, a 2,2'-dipicolylamine (DPA) phosphoramidite, at two specific sites in the DNA duplex. This strategy introduces an abasic site (dSpacer) opposing the DPA in the complementary strand. While other methods may require specific secondary structures¹⁴⁶ or use labels with elongated tethers that place

the reporter on the exterior of the DNA³³, the DPA-DNA method is structure-independent and positions the probe in close proximity to the DNA backbone. Furthermore, the label is also nucleotide independent and can be positioned anywhere within the DNA molecule. In the initial work, a most probable distance of 2.7 nm was measured with the Cu²⁺-DPA motifs separated by eight base pairs. This distance was in good agreement with both the distance calculated using the known values of base pair separation for a B-DNA and with corresponding molecular dynamic simulations. This reported distance was resolved using a single measurement without the need of any post-experiment analysis.

In this work, we provide a comprehensive analysis of the Cu²⁺-DPA motif. First, we performed continuous wave (CW) ESR measurements to analyze the specificity of Cu²⁺ towards the DPA motif and to rule out any possibility of non-specific binding elsewhere on the DNA molecule. Second, we obtained systematic distance measurements using double electron-electron resonance (DEER) technique to probe the flexibility of the Cu²⁺-DPA motif. These measurements were performed at both Q-band and X-band frequencies over several magnetic fields to probe any orientation effects. Third, we performed molecular dynamics (MD) simulations on an unmodified DNA sequence to provide insight on how the distance measured using the Cu²⁺-DPA motif compares to the native DNA backbone. Finally, we used a physical model that depicts the DNA motion, to estimate the flexibility of the Cu²⁺-DPA label. Together, these efforts show that the Cu²⁺-DPA motif is a highly promising labeling technique that can be used to probe DNA conformations without the need of any extensive modeling.

2.2 MATERIALS AND METHODS

2.2.1 Oligonucleotides

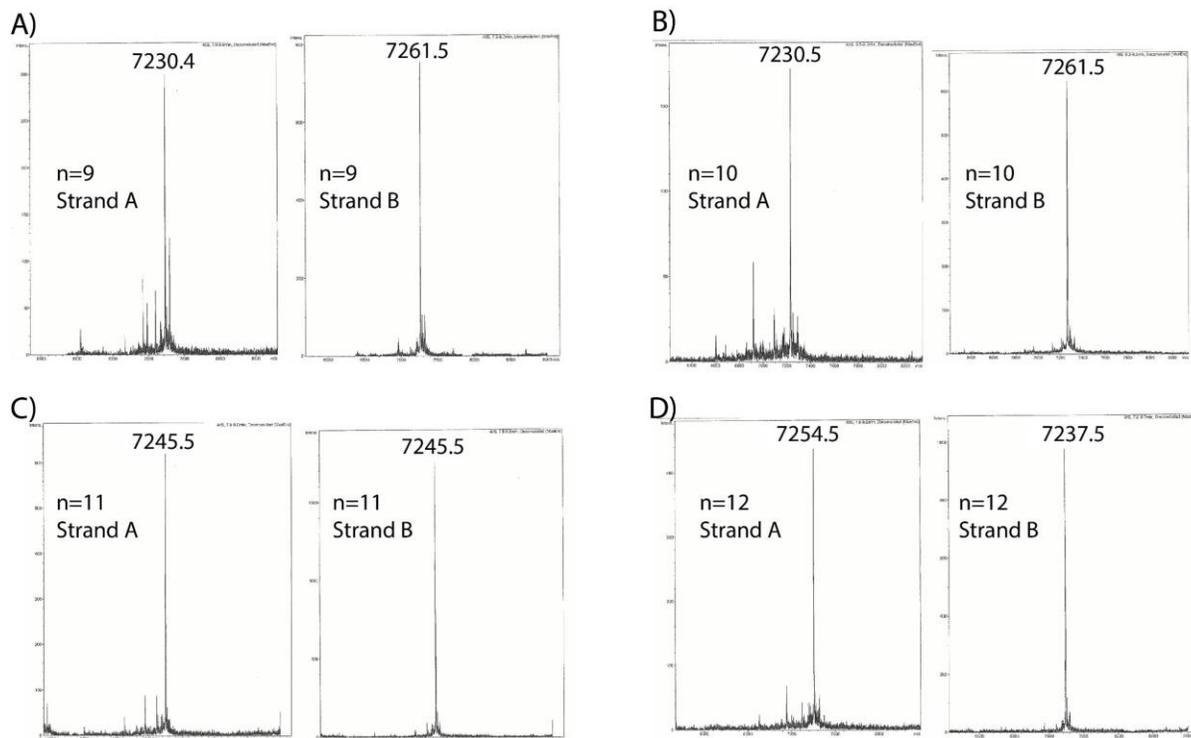


Figure 2-1 The mass spectrum of each strand was completed upon synthesis by ATDBio. The mass of each individual strand (in amu) of each duplex is shown above.

Single-strand oligonucleotides containing 2,2'-Dipicolylamine (DPA) were obtained from either Trilink Biotechnologies (duplex with eight base pair separation between DPA sites) or ATDBio Ltd. (duplexes with 9-12 base pair separations). The suppliers purified the strands using high performance liquid chromatography and characterized using mass spectrometry (Figure 2-1). The control unmodified DNA duplex was purchased from Integrated DNA Technologies (IDT).

Concentrations of purified single-stranded oligodeoxynucleotides were determined spectrophotometrically from molar extinction coefficients calculated by a nearest neighbor

method^{150–152}. Concentrations of duplexes formed from stoichiometric amounts of the complementary single strands were also determined spectrophotometrically using molar extinction coefficients calculated by the nearest neighbor method^{150–152}. The contribution of the DPA to the 24 base pair duplex was accounted for in three ways: a) Assuming an unpaired G in the position of the DPA b) Assuming an unpaired A in the position of the DPA c) Assuming a G-C base pair for the DPA and opposing dSpacer. Molar absorptivity coefficients for duplexes calculated according to a), b) and c) differed only by ~2%. The overall accuracy of these DNA concentrations is within 5%.

Equal amounts of complementary DNA strands were mixed and CuCl₂ were added such that there were slightly less than stoichiometric equivalents of Cu²⁺ per DPA binding site (~ 0.94 equivalents of Cu²⁺ per DPA). Duplexes were annealed in the presence of CuCl₂ to ensure proper formation and efficient Cu²⁺ chelation, using a GeneAmp PCR System 9700 with the following protocol: 90°C for 1 minute, 60°C for 3 minutes, 50°C for 4 minutes, 40°C for 4 minutes, 30°C for 5 minutes and then progressively cooled to 4°C. All duplexes were subjected to this annealing procedure.

2.2.2 ESR measurements

All ESR experiments were performed on either a Bruker ElexSys E580 or ElexSys E680 CW/FT spectrometer equipped with a Bruker ER4118X-MD5 or Bruker ER4118X-MD4 resonator for X-band frequencies, respectively and Bruker ER5106-QT2 resonator for Q-band frequencies. The E680 spectrometer is also equipped with a 300 W amplifier. Experiments were performed at either X-band (~9.68 GHz) or Q-band (~35 GHz) frequencies. Experiments were performed at either 20 K or 80 K.

X-band samples had a DNA duplex concentration ranging between 150-200 μM , with a total sample volume of 120 μL . Q-band sample had a DNA duplex concentration of 80 μM , with a total sample volume of 50 μL . All samples were prepared in 50 mM NEM buffer. The pH of the sample was 7.4 and 20% v/v glycerol was included as a cryoprotectant.

All continuous wave (CW) ESR data were acquired at 80 K. Experiments were performed at a center field of 3100 G with a sweep width of 2000 G for a total of 1024 data points. The data was collected using a modulation amplitude of 4 G, a modulation frequency of 100 kHz and a conversion time of 20.48 ms. The spectra were simulated using the EasySpin software¹⁵³.

Double electron-electron resonance (DEER) spectroscopy experiments were performed using the four pulse sequence $(\pi/2)v_1 - \tau_1 - (\pi)v_1 - \tau_1 + T - (\pi)v_2 - \tau_2 - T - (\pi)v_1 - \tau_2 - \text{echo}$ ¹⁵⁴. A 16-step phase cycling was used. For X-band DEER experiments, the observer pulse lengths, $(\pi/2)v_1$ and $(\pi)v_1$, were 16 ns and 32 ns respectively while the pump pulse length, $(\pi)v_2$, was 16 ns. The delay, T, was incremented using step sizes that varied from 8 ns up to 28 ns, depending on the DPA base pair separation. The pump frequency was placed at the maximum of the echo detected field swept spectrum, and the observer frequency was offset 100 MHz downfield. To probe orientational selectivity effects at X-band, we also performed DEER experiments with the pump frequency placed at the field corresponding to the g_{\parallel} region and the observer frequency was offset by 100 MHz upfield. For Q-band DEER experiments, the observer pulse lengths, $(\pi/2)v_1$ and $(\pi)v_1$ were 14 ns and 28 ns respectively and the pump pulse, $(\pi)v_2$ was 32 ns. The observer frequency was set 100 MHz higher in frequency than the pump. The delay, T, was incremented with a step size of 24 ns. DEER spectra were acquired from 11220 G to 11820 G at 8 different magnetic field

values. All DEER data were collected at 20 K. All DEER data were analyzed using DeerAnalysis2018¹⁰³.

2.2.3 MD simulations

Molecular dynamics (MD) simulations on the DNA duplex were performed using the AMBER parmbsc1 (bsc1)¹⁵⁵ force field. The Nucleic Acid Builder (NAB) module in the AMBER suite¹⁵⁶ was used to construct the B-DNA helix. The DNA duplex was then solvated in an explicit 12 Å TIP3P water box¹⁵⁷ and neutralized with Na⁺ and Cl⁻ ions. All simulations were performed using the pmemd program in the AMBER16 software package. The system was optimized, thermalized and pre-equilibrated for 2 ns before being set up for the unrestrained production MD run of 100 ns at 298.15 K. Periodic boundary conditions along with particle mesh Ewald (PME)¹⁵⁸ were applied to account for long-range electrostatic interactions under NPT (P = 1 atm) conditions. SHAKE¹⁵⁹ on bonds involving hydrogens was used in conjunction with an integration step of 2 fs. A non-bonded cut-off of 10 Å was applied. All visualizations for simulations were done on VMD¹⁶⁰. From the MD trajectories, distances were measured between the two C4' as well as the two C3' carbon atoms of the nucleotides that were replaced by DPA in the experiment.

2.3 RESULTS AND DISCUSSION

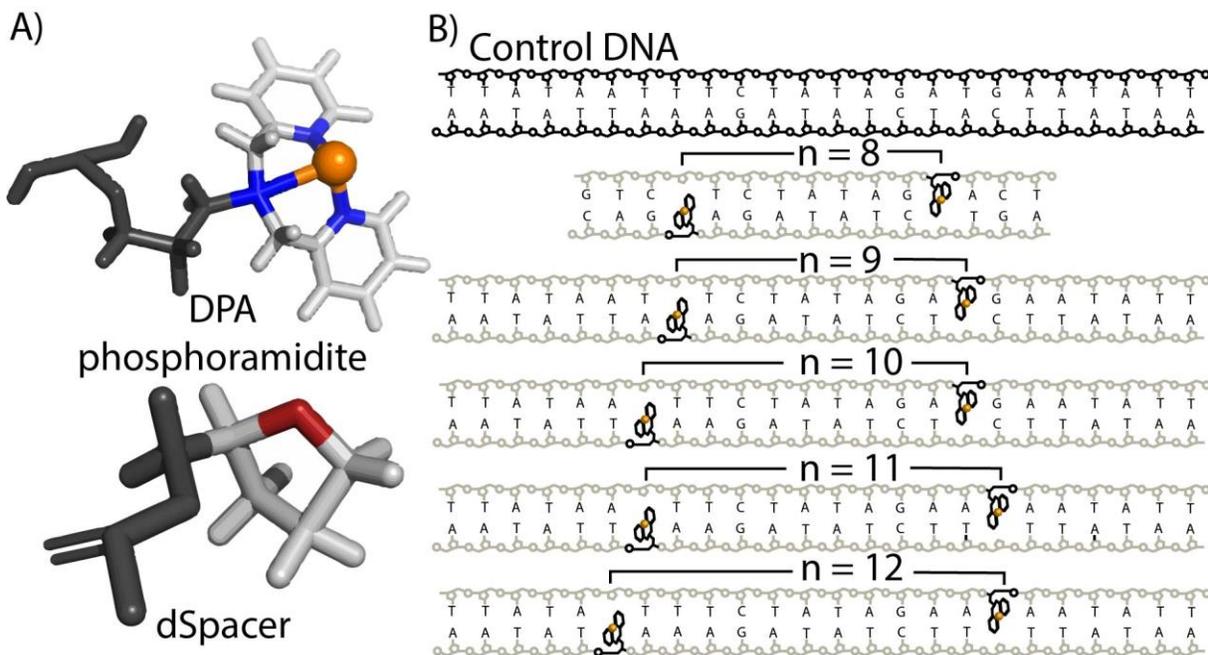


Figure 2-2 A) Structures of the Cu^{2+} -incorporated DPA phosphoramidite and the dSpacer. B) Sequences of the control DNA and the DPA-DNA duplexes with different base pair separation, n .

In this work, we examined five DPA-DNA duplexes. The structures of Cu^{2+} -DPA and dSpacer and the sequences of the DNA duplexes are shown in Figure 2-2. Each DNA duplex has two DPA phosphoramidites incorporated within the helix, one on each strand. A dSpacer (tetrahydrofuranyl residue mimicking the sugar phosphate backbone without a base) opposes the DPA on the complementary strand. The separation of the DPA motifs is monotonically increased in each DNA duplex by increasing the base pair separation (n) between them. All DNA constructs are heteroduplexes formed from single strands with non-palindromic sequences to prevent the chance of base pairing within the same DNA strand.

Results have been previously reported for a 15-nucleotide DNA duplex with a base pair separation, n of 8⁵⁰, between the DPA sites. In the $n=8$ duplex, as shown in Figure 2-2B, the Cu^{2+} -

binding site was flanked by only three bases. Because of the occurrence of slight unwinding at the ends of DNA duplexes, such unwinding proximal to the DPA-motif can increase the distribution of distances reported by the DPA-motif. In this work, the DPA motifs are flanked by at least six base pairs, thus reducing influence from terminal unwinding. Circular dichroism measurements of DPA-DNA and melting temperatures have previously shown that incorporation of the DPA motif inside the DNA duplex does not perturb the native helical structure of the DNA, nor does it influence the helical stability⁵⁰. Thus, the Cu²⁺-DPA motif is suitable for probing intra-DNA distances.

2.3.1 Characterization of Cu²⁺ bound to DPA-DNA

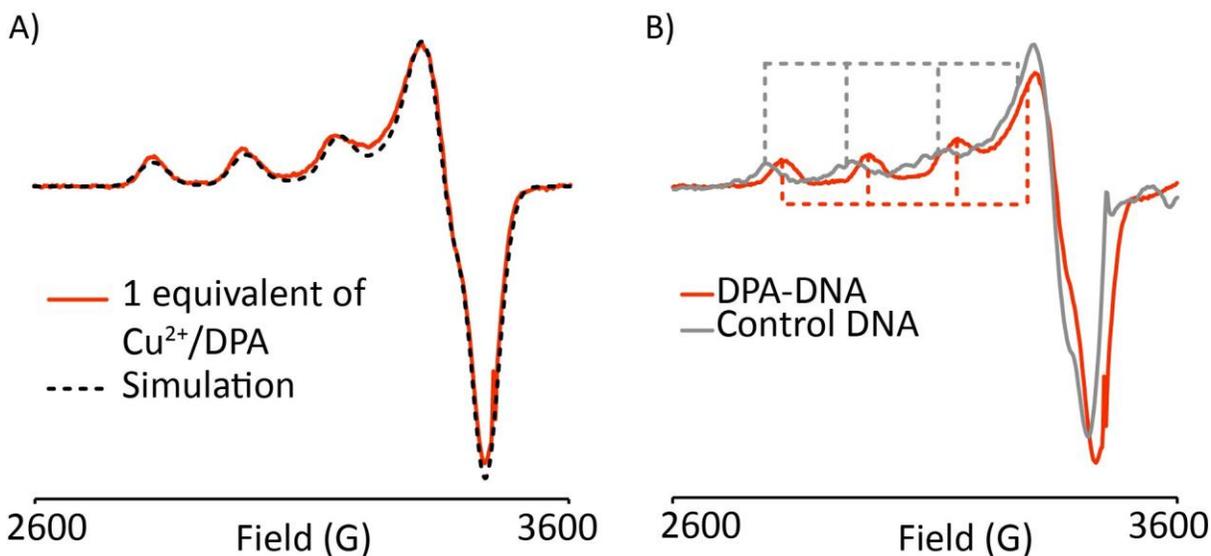


Figure 2-3 A) CW-ESR spectrum of Cu²⁺ bound DPA-DNA duplex (solid line). The single component fit, with g_{\parallel} and A_{\parallel} values of 2.240 and 170 G, respectively is overlaid (black dashed). B) Comparison of CW-ESR spectra between Cu²⁺ bound to control DNA (grey solid) and Cu²⁺ bound to DPA-DNA (red solid). The Cu²⁺-control DNA spectrum has g_{\parallel} and A_{\parallel} values of 2.280 and 163 G, respectively and is distinctly different from Cu²⁺-DPA-DNA spectrum. This confirms that there is no non-specific binding of Cu²⁺ in the DPA-DNA duplex.

To promote full solvent accessibility of the Cu^{2+} ion to the DPA chelating site, we annealed the individual strands of the DNA in the presence of Cu^{2+} . The temperature of the samples was raised to 95°C and incrementally brought down to 4°C as described in the Materials and Methods section. Cu^{2+} shows high binding affinity to a DPA ligand with an apparent association constant, K_a of $(5.0 \pm 2.0) * 10^6 \text{ M}^{-1}$ in NEM buffer at pH of 7.4¹⁶¹. On the other hand, Cu^{2+} also shows affinity towards native DNA with a K_a of $2.4 * 10^4 \text{ M}^{-1}$ ¹⁶² and particularly to the purine bases adenine and guanine, with a K_a of $1.7 * 10^3 \text{ M}^{-1}$ ^{50,163}. We, therefore, added a slightly sub-stoichiometric equivalent of Cu^{2+} (~ 0.94 equivalents of Cu^{2+} per DPA) to maximize loading of the DPA sites without causing non-specific binding elsewhere.

To verify that Cu^{2+} binds specifically to DPA, we first performed CW-ESR experiments on a Cu^{2+} -DPA-DNA duplex ($n=11$). Figure 2-3A (solid line) shows the CW-ESR spectrum of the Cu^{2+} -bound DPA-DNA duplex. We observed only a single component spectrum. Note that free unbound Cu^{2+} is ESR silent in NEM buffer^{78,164}. Further, simulation of the spectrum gave a single component fit, as shown in Figure 2-3A, with $g_{||}$ and $A_{||}$ values of 2.240 and 170 G, respectively. The $g_{||}$ and $A_{||}$ values are consistent with a 3 nitrogen/1 oxygen coordination⁶⁷ which corresponds to Cu^{2+} binding to the three nitrogen atoms of the DPA. Comparison of the spin concentration obtained from the CW-ESR data to the DNA concentration indicated a labeling efficiency of $\sim 80\%$, instead of the expected 93% from the K_a value of Cu^{2+} for free DPA. The K_a of Cu^{2+} to DPA may be slightly altered by the presence of adjacent bases in the duplex.

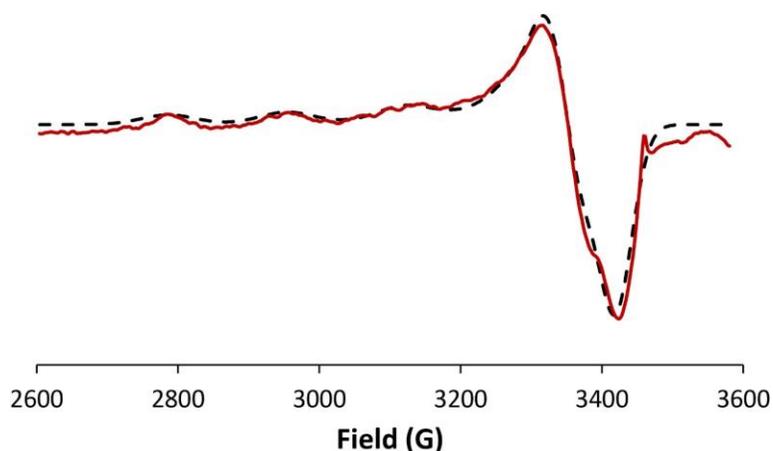


Figure 2-4 Simulation of the CW-ESR spectrum of the control DNA shows a single component fit with $g_{||}$ and $A_{||}$ values of 2.280 and 163 G, respectively.

To rule out the presence of non-specific binding in the presence of the Cu^{2+} -DPA motif, we also performed CW-ESR measurements on a control DNA, where the DPA and dSpacer positions were replaced by adenine and thymine, respectively (cf. Figure 2-2, top panel). The Cu^{2+} was added to the control DNA in a stoichiometric ratio. The spectrum of free DNA is distinctly different from that of DPA-DNA, as can be seen in Figure 2-3B. The CW-ESR spectrum of control DNA showed a single component fit with $g_{||}$ and $A_{||}$ of 2.280 and 163 G, respectively (Figure 2-4). Previous measurements on the $n=8$ duplex used a four-fold excess of Cu^{2+} to DPA which led to non-specific binding. In the current work, the data suggests that in the presence of DPA, the Cu^{2+} binds preferentially to the DPA moiety and that there is no interference from non-specific binding with stoichiometric or sub-stoichiometric loading (Figure 2-3B).

2.3.2 DEER measurements show no observable orientational selectivity at both X and Q-band frequencies

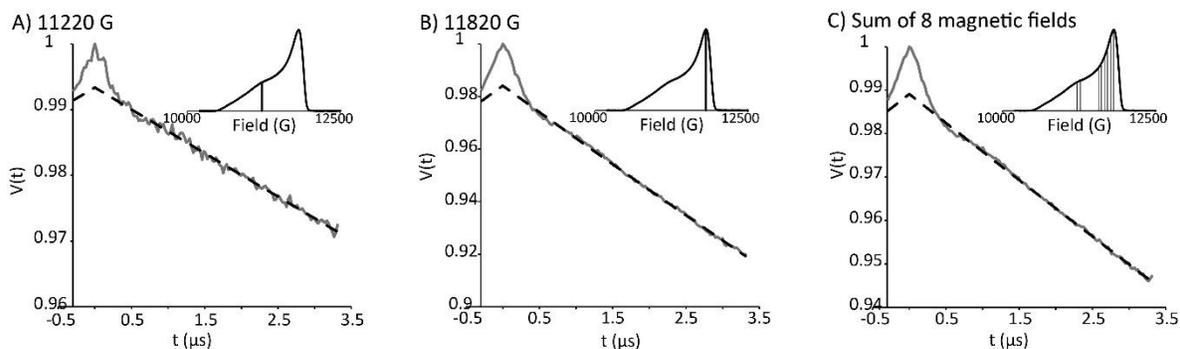


Figure 2-5 DEER signals of Cu²⁺-DPA-DNA (n=11) performed at Q-band frequency. DEER performed at A) g_{\parallel} region (11220 G) B) g_{\perp} region (11820 G) C) 8 different magnetic fields.

To probe orientational selectivity effects on the Cu²⁺-DPA motif, we first performed the DEER experiment on the DPA-DNA duplex with n=11. The data were acquired at eight magnetic field values over a range of 600 G (11220 G to 11820 G) between g_{\parallel} and g_{\perp} regions at Q-band frequency. All primary DEER time traces were background subtracted and analyzed via Tikhonov regularization using DeerAnalysis2018¹⁰³. The primary time domain data is shown in Figure 2-5. Figure 2-6A shows the field swept spectrum of Cu²⁺-DPA-DNA duplex and the magnetic field values at which DEER was performed. The corresponding background subtracted data is shown in Figure 2-6B.

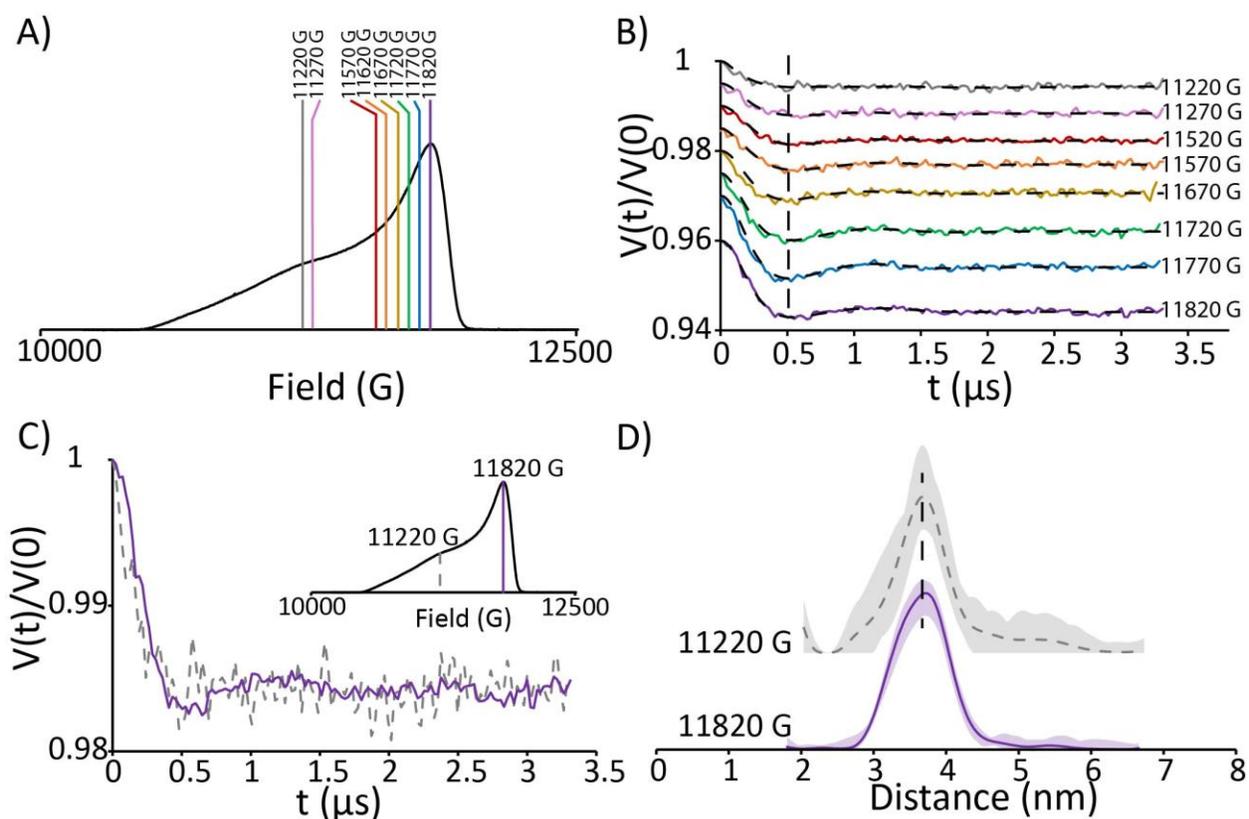


Figure 2-6 A) Field-swept electron spin echo spectrum of Cu^{2+} -DPA-DNA ($n=11$) at Q-band frequency. The lines show the different magnetic fields at which DEER was performed. B) Background subtracted time domain DEER signals at the different magnetic fields. The y-axis has been offset for ease of visualization. The data show minimal difference in dipolar frequency at the different fields. C) Background subtracted time domain data at g_{\parallel} (dashed, 11220 G) and g_{\perp} (solid, 11820 G) regions. These two regions also do not show any distinct difference in modulation frequency, suggesting Cu^{2+} -DPA shows minimal orientational effects. D) Distance distribution obtained via Tikhonov regularization, for both g_{\parallel} (dashed, 11220 G) and g_{\perp} (solid, 11820 G) regions. The similarity in the distributions further confirms that Cu^{2+} -DPA is not orientation selective for non-complexed DNA.

In DEER, selective excitation of the total spectrum can lead to excitation of only a small subset of all possible molecular orientations, leading to orientational selectivity effects^{46,148,165–170}. This effect can be mitigated by several factors and techniques. At X-band frequencies (9.5 GHz), most common systems such as nitroxides and Cu^{2+} labels have large hyperfine anisotropies relative to their g -anisotropy, which allows the mixing of orientations across the spectrum^{38,171}.

Additionally, flexible spin labels can reduce orientation selectivity effects as the inherent flexibility effectively randomizes the selected molecular orientations. A similar effect is also observed in more rigid Cu^{2+} binding motifs, in which the slight flexibility of the Cu^{2+} coordination environment produces a distribution in g-tensor orientations^{45,47,56}. Because orientational selectivity effects are due in part to the selective excitation, ultra-wide band pump pulses that effectively excite the entire spectrum have also shown dilution of orientational selectivity¹⁷². As can be seen from Figure 2-6B, the DEER time traces display dipolar modulations with similar dipolar frequency at all magnetic fields. This consistency suggests that orientational effects are negligible for this system at different magnetic fields. Furthermore, we compare the background subtracted time domain DEER, performed at g_{\parallel} with the g_{\perp} regions (Figure 2-6C), the modulation depth was adjusted for comparison. These regions should represent a maximal difference in the orientations excited. Again, we found minimal change in the modulation frequency of the DEER trace (Figure 2-6C). Consequently, the distance distribution obtained using Tikhonov regularization at the above two fields agree well within signal to noise, as shown in Figure 2-6D. Based on these data, we conclude that the inherent DNA flexibility along with that of the label suppresses any orientational selectivity effects in non-complexed DNA.

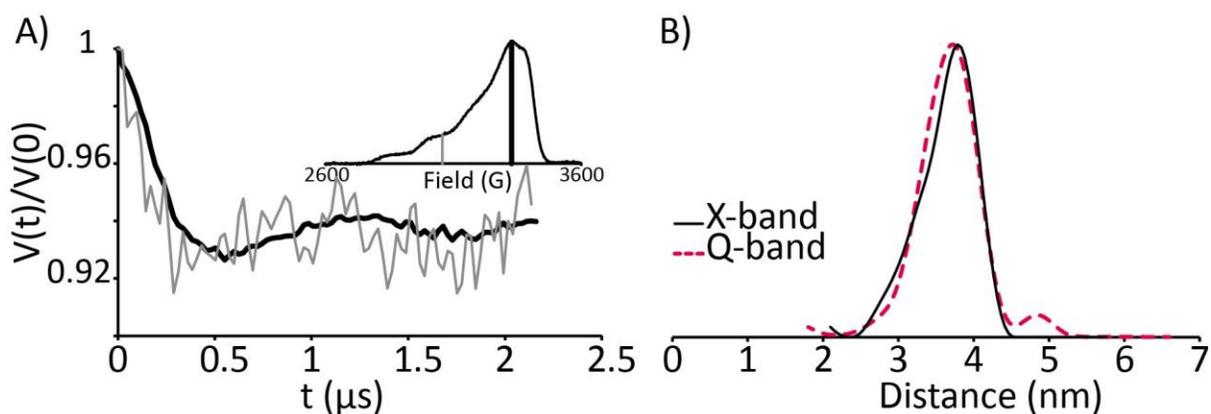


Figure 2-7 A) Background subtracted time domain DEER signal of Cu^{2+} -DPA-DNA ($n=11$) at X-band frequency at g_{\parallel} (grey) and g_{\perp} (black) regions. The field corresponding to the g_{\parallel} and g_{\perp} regions are shown in the field-swept electron spin echo spectrum in the inset. There is no distinguishable difference in the dipolar frequency between the two data. B) Distance distributions obtained as a sum of DEER time traces collected at different fields, ranging from g_{\parallel} and g_{\perp} regions at Q- band frequency (dashed red) and at g_{\perp} region at X-band frequency (solid black). Both distributions, analysed via Tikhonov regularization, show similar most probable distance and width.

Since DEER performed at Q-band frequency did not show any orientation effects for Cu^{2+} -DPA, we do not expect to see any such effects at X-band frequency. Nevertheless, we performed DEER on the same DPA-DNA duplex of $n=11$ at X-band frequency at both g_{\parallel} and g_{\perp} regions (cf. Figure 2-7A). The background subtracted time domain DEER of the two regions, shown in Figure 2-7A, confirms that orientational effects are minimal – again the modulation depths are adjusted for comparison. Interestingly, the distance distribution analyzed from the g_{\perp} region at X-band and that of the summed DEER signal over eight different magnetic fields at Q-band showed good agreement for the DPA-DNA duplex of $n=11$ (cf. Figure 2-7B). Nevertheless, we also probed for orientation selectivity effects at X-band frequency for other duplexes as well. In all cases, we observed minimal change in dipolar frequency in the time domain DEER data between g_{\parallel} and g_{\perp} regions (Figure 2-8).

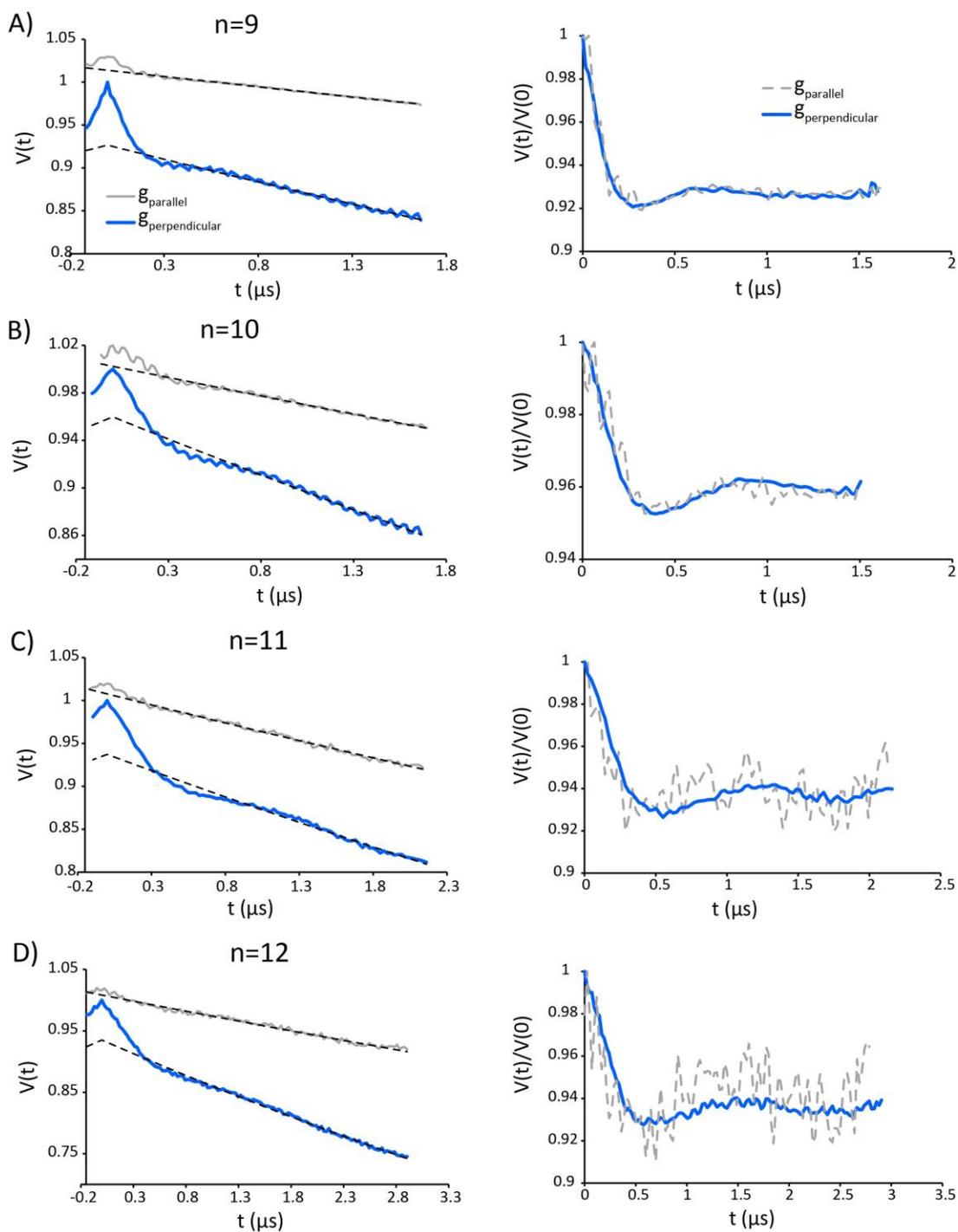


Figure 2-8 Raw time domain DEER signals for DNA duplexes with n ranging from 9 to 12, performed at fields corresponding to g_{\parallel} (grey) and g_{\perp} (blue) at X-band frequency and the corresponding background subtracted time domain DEER data. Minimal differences in dipolar frequency are observed, thereby showing no orientational selectivity effects at X-band frequency.

2.3.3 DEER measurements using Cu^{2+} -DPA motif can probe the DNA backbone

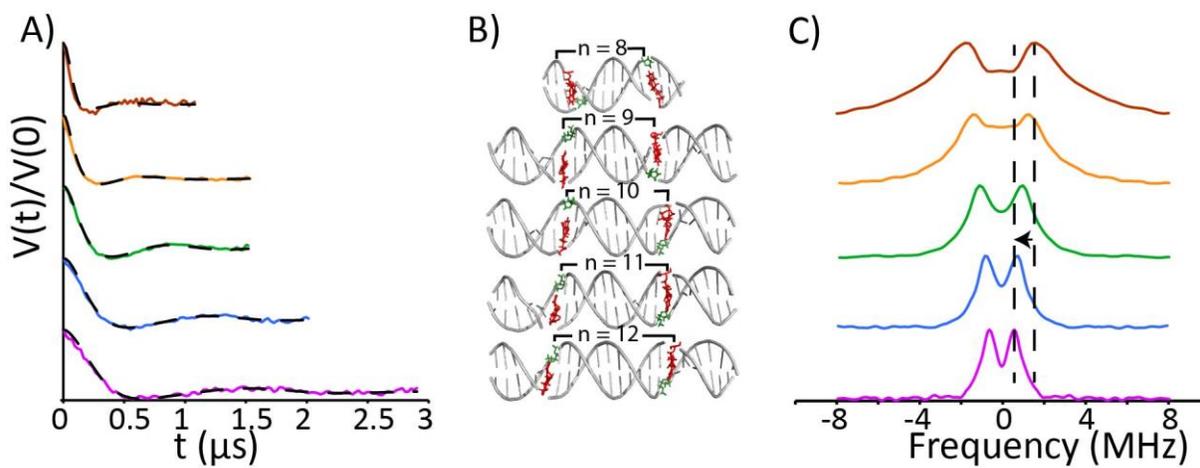


Figure 2-9 A) Background subtracted time domain DEER data of the DPA-DNA duplexes ranging from $n=8$ to 12 base pairs. The modulation frequency increases as the base pair separation increases. The y-axis has been offset for ease of visualization. B) Duplexes showing the Cu^{2+} -DPA and dSpacer positions for varying n . C) Pake pattern for duplexes $n=8$ to 12. The characteristic peak, corresponding to the perpendicular orientation, shows a gradual shift to a lower frequency as n increases.

Finally, to determine if the labeling strategy can accurately report on changes in distance upon incrementing base pair separation, we analyzed the DEER data acquired at g_{\perp} at X-band frequency (cf. Figure 2-8). Figure 2-9 shows the background subtracted DEER data for duplexes with n ranging from 9 to 12. Importantly, the absence of orientation effects was observed for all the five different spin-label positions. DEER measurement performed on duplex $n=8$ has been previously reported⁵⁰ and are also used in the following discussion. The corresponding Pake patterns, clearly show that as the distance increases with n , the dipolar frequency of the modulation gradually decreases as is expected (cf. Figure 2-9C).

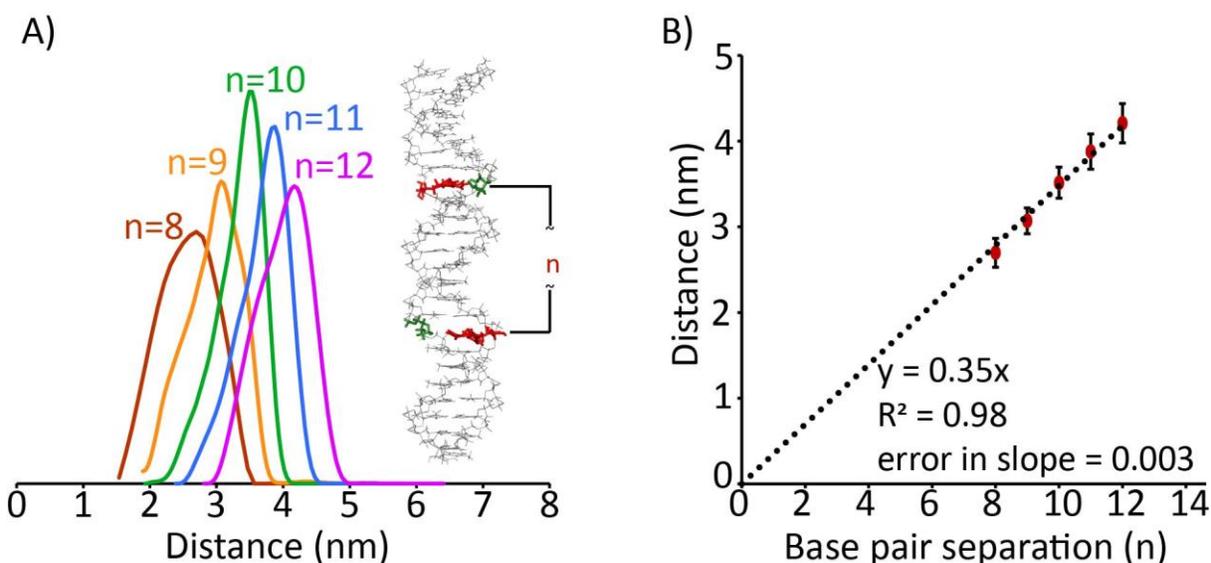


Figure 2-10 A) Area-normalized distance distributions obtained for duplexes $n=8$ to 12 via Tikhonov regularization. The distributions show an increase in the most probable distance with increasing n . B) Plot of most probable distance versus base pair separation, n . On fitting to a linear trend with a y -intercept of 0 , we obtained a slope of 0.35 nm with an error of 0.003 nm. The slope is in reasonable agreement with the reported base pair separation for a B-DNA.

We analyzed the time domain DEER data with Tikhonov regularization to obtain distance distributions (Figure 2-10). We observed an increase in the most probable distance as the base pair separation increases from 8 to 12 . Furthermore, we plotted the most probable distance against the base pair separation and observed a linear trend for a y -intercept of 0 (Figure 2-10B). The slope of the fit equals 0.35 nm which is in good agreement with the ~ 0.34 nm separation between adjacent bases for an idealized B-DNA¹⁷³. The data (Figure 2-10B) show that our experimental technique reports accurately on B-DNA double helix length over the range of 8 to 12 base pairs and discerns differences in length as small as one base pair.

2.3.4 MD simulations show good agreement with the most probable distance from DEER

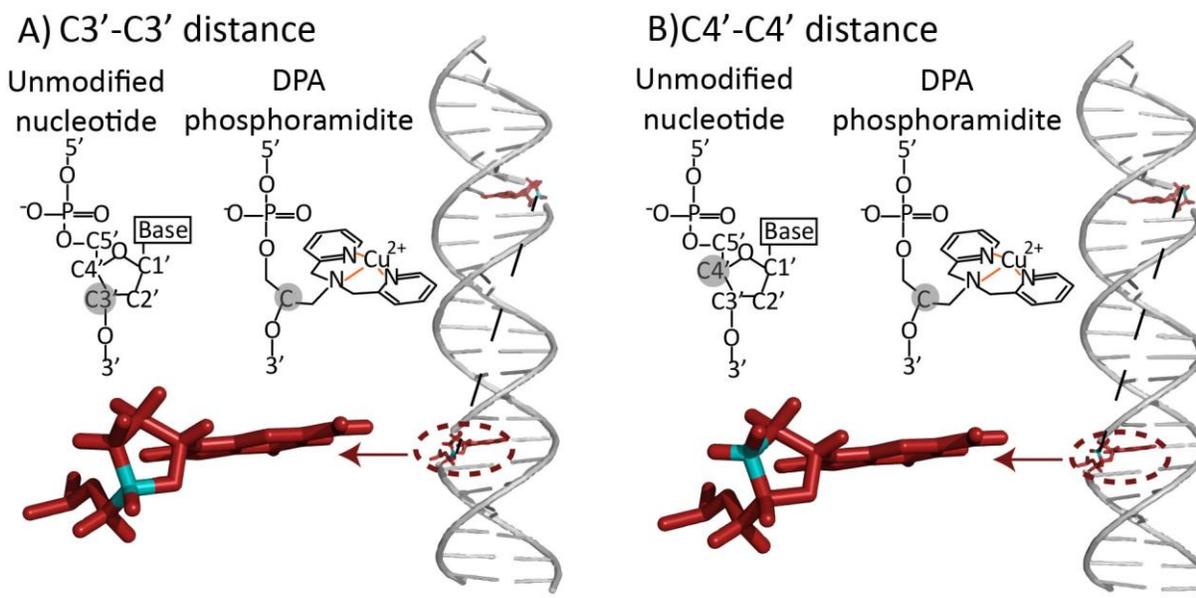


Figure 2-11 Backbone carbon atoms of an unmodified nucleotide compared to the DPA phosphoramidite. The backbone carbon atom in DPA (grey circle) can resemble either A) a C3' atom when viewed from the 3' end or B) a C4' atom when viewed from the 5' end. Accordingly, the distance distributions were analyzed between C3'-C3' and C4'-C4' backbone atoms of the residues that are replaced by DPA in the DPA-DNA duplexes.

Next, we performed MD simulations on the control unmodified DNA (cf. Figure 2-2) in order to compare the measured distances to native backbone conformations. We created a DNA duplex with the same sequence as the control DNA, where the DPA and dSpacer positions are replaced by adenine and thymine, respectively. For running MD simulations, we chose the newer AMBER parmbsc1 force field as it has previously shown good agreement to ESR experimental results obtained on doubly spin labeled DNA^{28,32}. We generated 100 ns of unrestrained MD trajectories of the DNA duplex. From the MD simulation we extracted backbone distance distributions to compare with our experimental results.

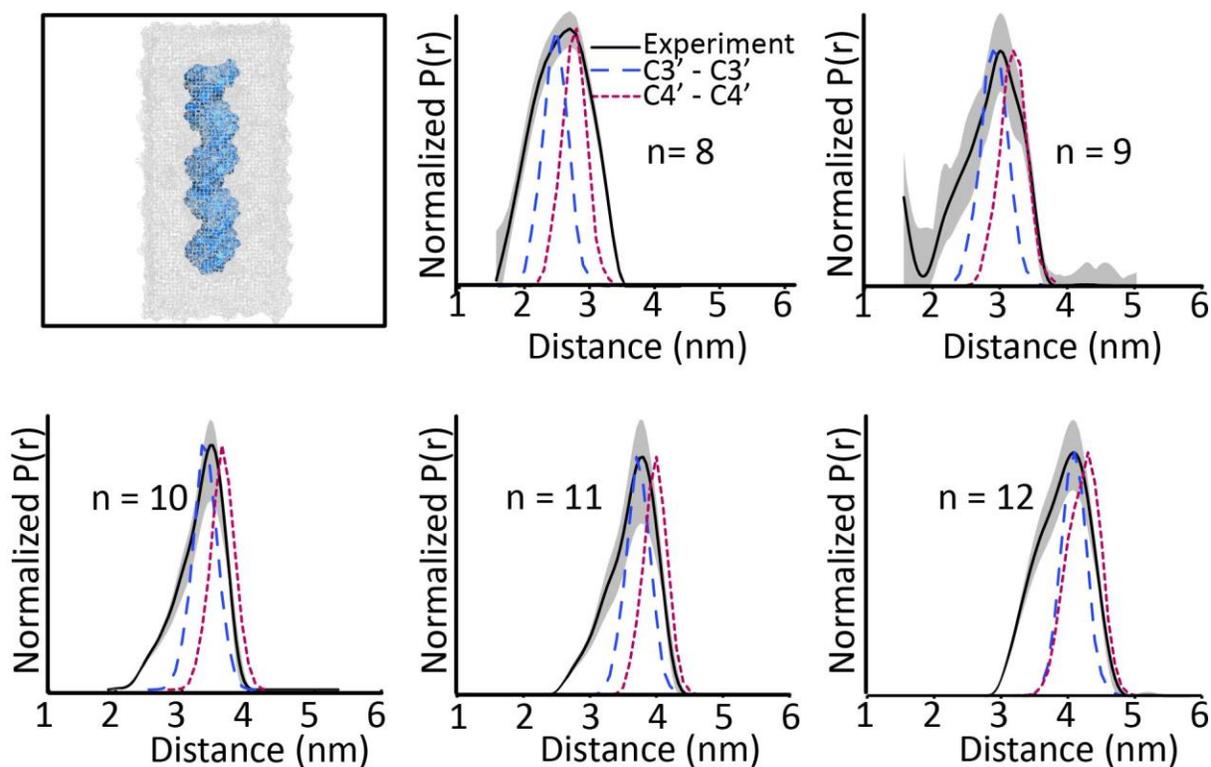


Figure 2-12 The first panel shows the DNA duplex inside a water box, used in our simulations. C3'-C3' (blue dashed) and C4'-C4' (red dotted) distance distributions obtained from MD simulations and experimental distance distributions (black solid). MD simulated distributions show a good agreement with the experimental most probable distance, within 1-2 Å.

First, we determined which backbone atoms of the unmodified DNA most nearly correspond to the distance reported by the Cu^{2+} -DPA label. We compared the structure of unmodified DNA to that of Cu^{2+} -DPA-DNA with an 11-bp separation (Figure 2-12A). The backbone carbon atom of the DPA motif is represented inside the grey circle. This particular atom forms the pivotal point in the phosphoramidite to which the DPA is attached. When comparing with a normal nucleotide, we considered two possibilities: a) when viewed from the 3' end, the DPA backbone carbon atom resembles a C3' atom in a normal nucleotide (Figure 2-11A) b) when viewed from the 5' end, the same backbone atom of DPA resembles a C4' atom of a normal

nucleotide (Figure 2-11B). Therefore, when evaluating the distance distributions from the MD trajectories, we considered both C3'-C3' and C4'-C4' distances between the suitable bases.

Figure 2-12 shows the MD and experimental distance distributions. The most probable distance between the experiment and the simulations showed a reasonable agreement within 1-2 Å, for all the duplexes. These results are further illustrated in Figure 2-13. The remarkable agreement in the most probable distances strongly supports the fact that the Cu²⁺-DPA motif is present within the DNA helix. The agreement also confirms that the incorporation of DPA into the DNA does not perturb the native DNA structure. It follows that the Cu²⁺-DPA motif is an accurate reporter of point-to-point distances in the DNA backbone without requiring any additional modeling.

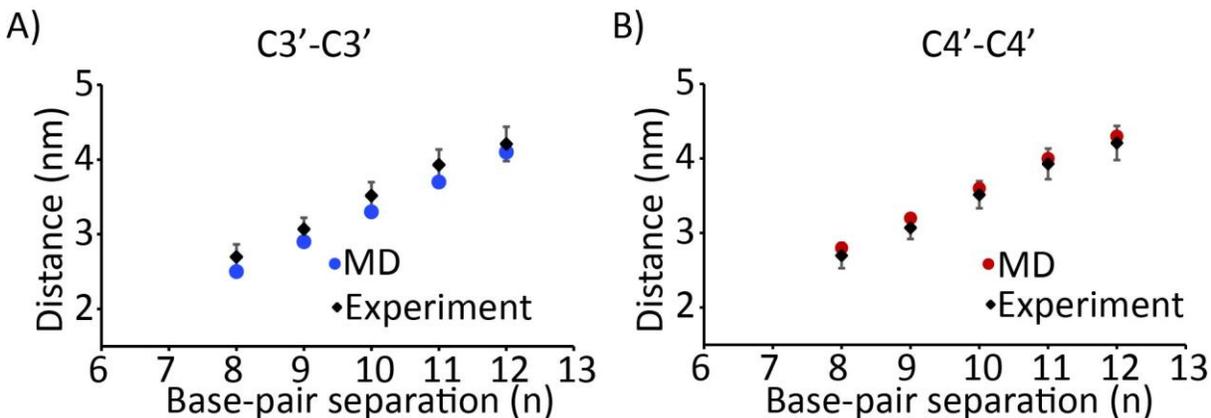


Figure 2-13 Comparison of the most probable distance from the experimental results with A) C3'-C3' distances and B) C4'-C4' distances from MD simulations. The data show good agreement for different base pair separations, n.

Despite the coincidence of most probable distance values, a discrepancy exists between the widths of the experimental and MD simulated distance distributions (cf. Figure 2-12). Clearly, the experimental distance distributions are broader. There are three possible explanations, which are not mutually exclusive. First, the DPA moiety itself has some inherent conformational flexibility, which adds to the distribution width. Second, the MD simulations used unmodified DNA. It is possible that the absence of interstrand hydrogen bonding at the DPA sites of the modified DNA

allow access to additional conformational fluctuations that are not accessible in the unmodified DNA. Third, the DNA in solution may experience conformational fluctuations on a time-scale much slower than the 100 ns of the MD simulation; these cannot be captured in short simulations. Accordingly, we analyzed our first hypothesis, i.e., the flexibility of the label.

2.3.5 DNA twist-stretch model enables estimation of label flexibility

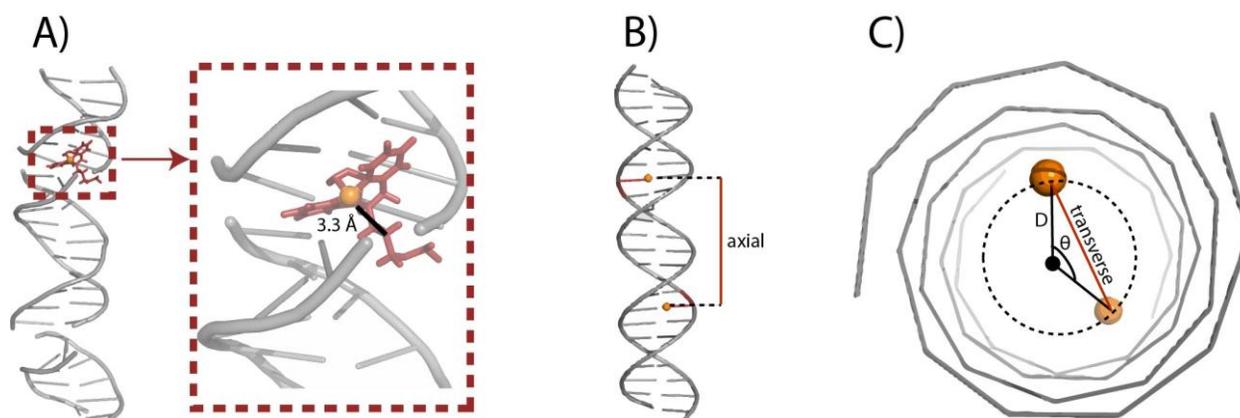


Figure 2-14 A) DPA structure, built on available crystal structure parameters, is incorporated in a DNA duplex and the approximate length of the label is calculated to be 3.3 Å. Geometric model of the B-DNA helix. The model comprises of two components B) axial and C) transverse. To calculate the angle θ , the DNA helix is viewed from the top. The Pythagorean sum of the two components give the final length between two labels placed inside a DNA.

In order to estimate the flexibility of the Cu^{2+} -DPA label, we modeled the twist-stretch motion of the DNA duplex, taking into account the label length and flexibility. Previous work has shown that the twist-stretch model is in agreement with DEER data on DNA obtained by using the rigid cytidine analogue nitroxide spin label¹⁴². Here, we adopted a similar approach but specific to our Cu^{2+} -DPA label. First, we needed to account for the length of the probe from the DNA backbone. To estimate the length of the label, we built a DPA-DNA duplex, where we incorporated the DPA at two specific sites in the DNA using PyMOL software¹⁷⁴. The DPA structure itself was

built on the bond length and bond angle information available from crystal structure¹⁷⁵. Figure 2-14A shows the estimated length of the DPA label, that is, the length of the Cu²⁺ from the DNA backbone is roughly 3.3 Å. Second, we constructed the physical model of the DNA twist-stretch motion. This DNA motion is built on two main parameters: an axial and a transverse distance between the two probes^{142,176}, as shown in Figure 2-14B and C. The axial distance is calculated as the following:

$$axial = axial_0 + rn \quad (2-1)$$

where $axial_0$ is the separation between the probes when separated by zero base pairs, r is the rise per base pair and n is the number of base pairs between the probes.

The transverse distance is calculated as:

$$transverse = 2D\sin\left(\frac{\theta}{2}\right) \quad (2-2)$$

where D is the radial displacement of the probes of off the helical axis and θ is the angle between the probes when viewed in projection along the axis of the helix. The angle θ can be further calculated as:

$$\theta = \theta_0 + 2\pi n/N \quad (2-3)$$

where θ_0 is the angle between the probes at zero base pair separation and N is the number of base pairs per turn (since N base pairs make one full turn or 2π).

The total length between the two probes can then be given as the Pythagorean sum of the axial and the transverse components as:

$$Length(n) = \sqrt{axial^2 + transverse^2} \quad (2-4)$$

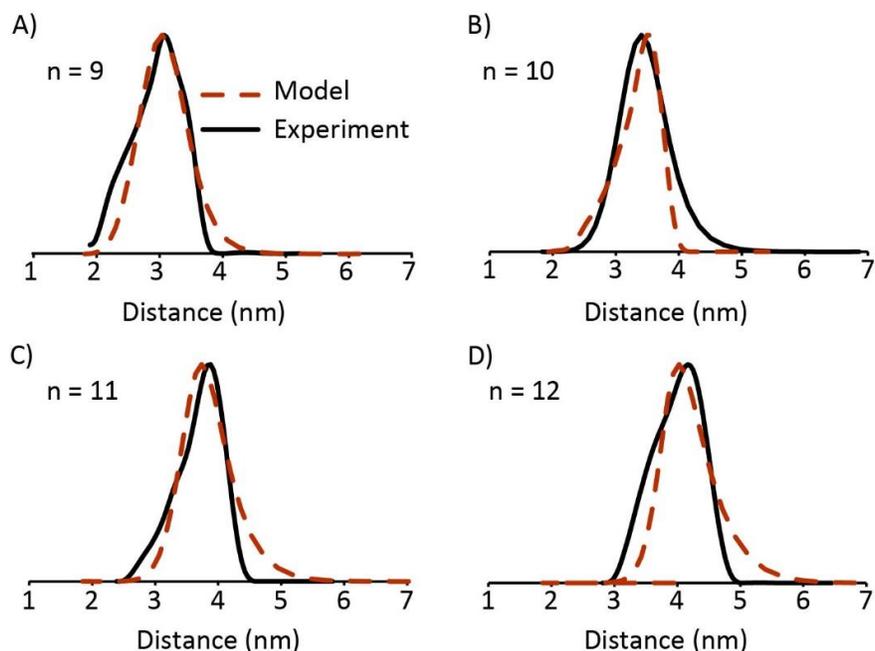


Figure 2-15 Distance distributions obtained from the twist-stretch model of the DNA compared to the experimental results for different base pair separations between the spin labels.

Finally, we obtained the distance distribution for different base pair separations from the model. We varied the radius of the DNA and calculated the corresponding change in radius from the equilibrium radius (Δr). Using a total flexibility of the DNA and the label (σ_r), we obtained a distance distribution, $P(r)$, for different base pair separations for the twist-stretch motion of the DNA, given by:

$$P(r) = \frac{1}{\sqrt{2\pi\sigma_r^2}} \exp\left(-\frac{\Delta r^2}{2\sigma_r^2}\right) \quad (2-5)$$

We then plotted the distribution $P(r)$ against the total length obtained from equation 2-4. Figure 2-15 shows the comparison of the experimental distance distributions with that from the model. The DNA duplex with $n=8$ was not included in the comparison since the absence of sufficient flanking sequences on either side of the label leads to additional flexibility of the label. One important aspect while obtaining these distributions, as mentioned previously, is the flexibility of the DNA as well as the label. Previous work on DNA breathing motion estimated the DNA

mobility to be $\sim 0.65 \text{ \AA}^{31}$. Here, along with the DNA mobility, we also added an additional label flexibility of $\sim 1.6 \text{ \AA}$. We observed a good agreement between the distributions from the twist-stretch model and the experiment, within the experimental error.

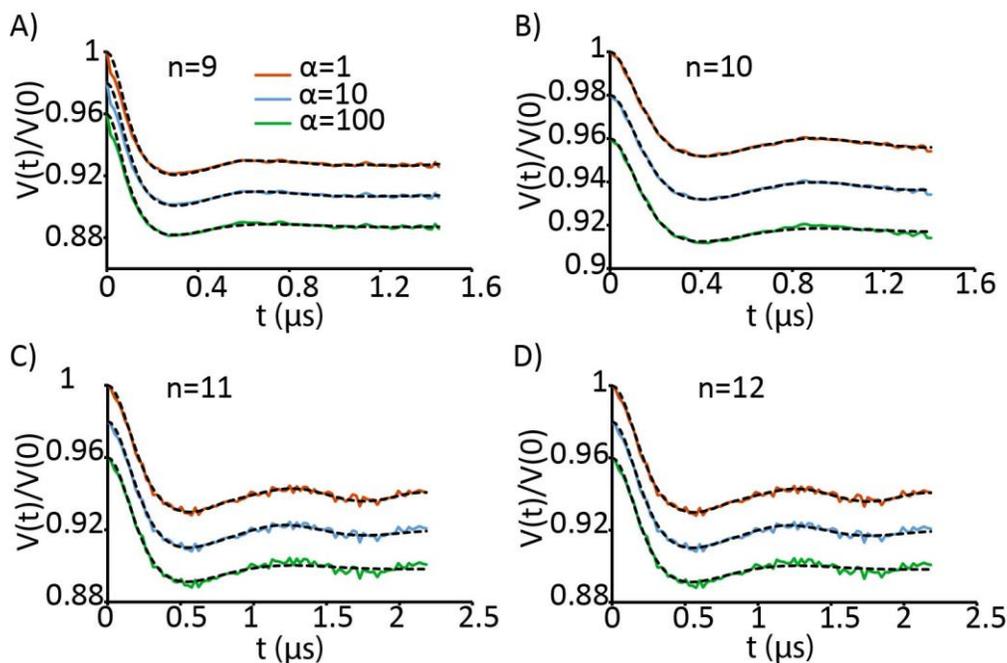


Figure 2-16 Background subtracted time-domain DEER data for the different DNA duplexes with variable α -values (1, 10 and 100). There is a good agreement of the experiment with the fit, over the different α -values used.

We then proceeded to compare the standard deviations between the model and the experimental distributions. When analyzing DEER distance distributions using Tikhonov regularization, the width of our distance distributions is strongly dependent on the regularization parameter, α . As a result, we varied the α -value between 1 to 100, making sure that we still obtain a good fit to our data. Figure 2-16 shows the variable α used to analyze each individual DEER data and the corresponding fits. We obtained the standard deviation for each of these fits and calculated the average (Figure 2-17A). Accordingly, we calculated the standard deviations of the above distributions and compared with that of the average. We see a reasonable agreement between the two as well, as shown in Figure 2-17B.

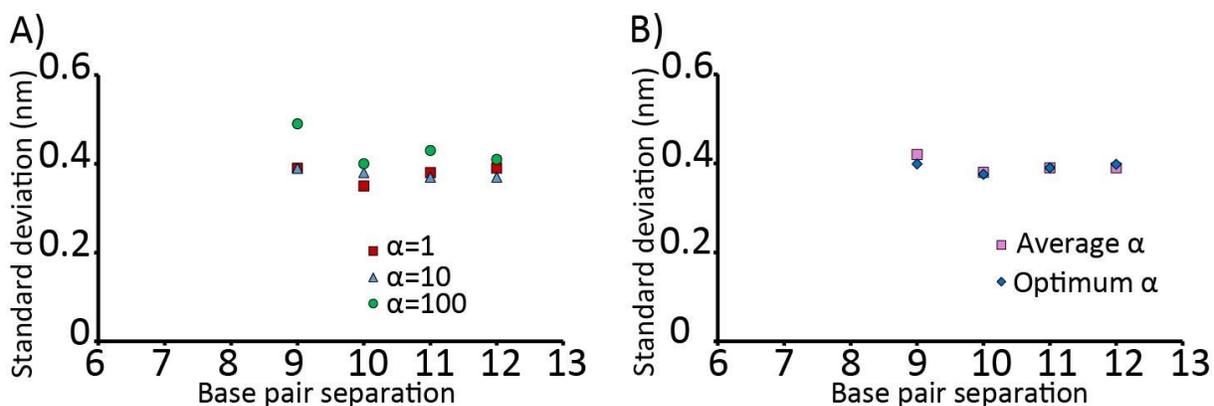


Figure 2-17 A) Comparison of standard deviations of distance distributions for varying α -values. B) Comparison of the average of the standard deviations for different α with that obtained from the optimum α -value.

Finally, Figure 2-18 shows the overall comparison of the most probable distance and the standard deviation of the modeled data with the experimental results for n ranging from 9 to 12. The experimental data agree with the model for an additional spin label flexibility of ~ 1.6 Å. Interestingly, the most rigid nitroxide labels, such as the cytidine analogue, has been reported to have a flexibility of 1 angstrom¹⁴². Note, that these comparisons are a rough estimation of the label flexibility since they consider only the twist-stretch motion¹⁴². However, there may be an additional influence of DNA bending³² which has not been taken into account. Nevertheless, these results indicate that the flexibility of the Cu^{2+} -DPA is only slightly higher than the most rigid nitroxide labels. Overall, this additional flexibility of the Cu^{2+} -DPA also explains the broader distance distributions obtained in our experiments versus MD (cf. Figure 2-10 and Figure 2-12).

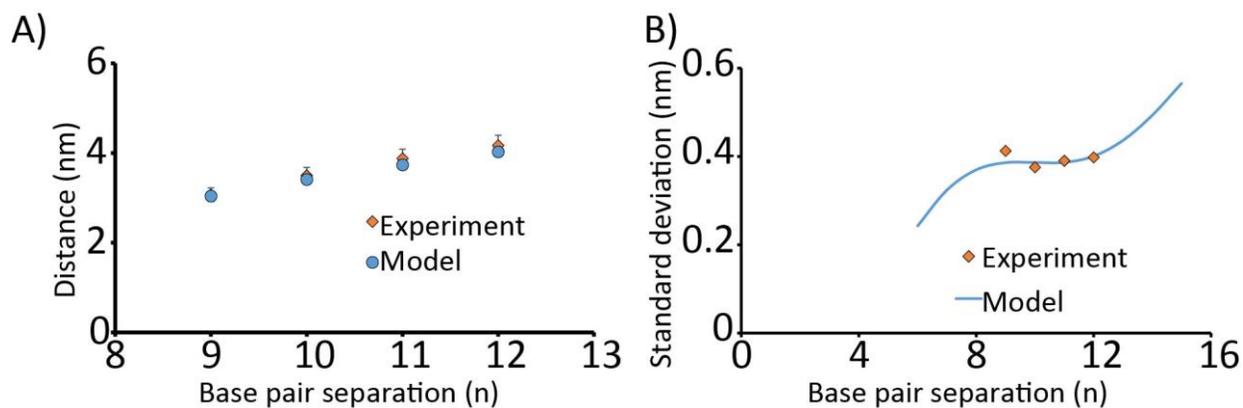


Figure 2-18 A) Plot of most probable distance versus base pair separation for the geometric model of DNA's twist-stretch motion (blue) and experiment (orange). The results agree within the error of the experiment. B) Plot of standard deviation versus base pair separation for model (blue) and experimental results (orange). The model considers an additional ~ 1.6 Å flexibility of the Cu^{2+} -DPA label and agrees with the experimental results.

More precise information on the label flexibility can also be obtained by elucidating the rotameric preferences of the label¹⁷⁷⁻¹⁸¹. Such analysis would require the knowledge of force field parameters of the Cu^{2+} -DPA label which is an important future direction.

2.3.6 Twist-stretch and in-silico models show that Cu^{2+} - Cu^{2+} distance can predict on DNA backbone despite having a radial offset

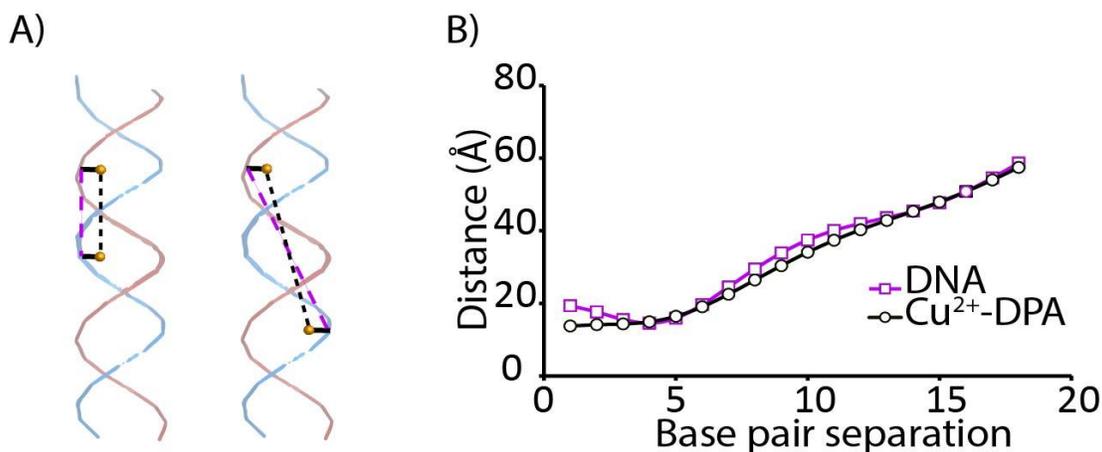


Figure 2-19 A) Two different models of the straight DNA where the Cu^{2+} -DPA motifs are a) on top of each other and b) opposing one another. The Cu^{2+} is modeled to be at 3.3 Å from the DNA backbone. In the first case, the Cu^{2+} - Cu^{2+} distance can be directly translated to DNA backbone distance as the 3.3 Å deviation does not play any role here. In the second case, the deviation, being significantly smaller than the axial distance, also does not contribute to the resultant distance. B) Plot of most probable distance of the DNA backbone and Cu^{2+} -DPA-DNA, obtained from the twist-stretch model, for different base pair separations. We observe that for a base pair separation of 4 and higher, the 3.3 Å between the Cu^{2+} and the DNA backbone has negligible contribution to the overall distance.

An important outcome of the twist-stretch model and the DEER distance distributions is that despite the Cu^{2+} being 3.3 Å away from the DNA helical backbone, it can still predict precisely on the DNA backbone conformations. We explain this observation by considering the individual components of the model. The axial component in the model increases linearly, and the radial component is oscillatory as the base pair separation increases. Therefore, the axial becomes the more dominant component with increasing base pairs between the labels. Hence, the radial offset of 3.3 Å of Cu^{2+} from the backbone becomes negligible when obtaining the resultant distance, as shown in Figure 2-19A. Furthermore, we illustrated our point by plotting the most probable

distances from the twist-stretch model for DNA backbone and Cu^{2+} -DPA-DNA, shown in Figure 2-19B. Clearly, for a base pair separation of 4 or higher, the contribution of the radial component of our DPA label towards the resultant distance is small. As a result, the Cu^{2+} -DPA distance closely resembles the DNA backbone distance, specifically the C3' and the C4'. Conceptually, in our method, the spin label points towards each other into the helical axis rather than away from the helical axis. In addition, the size of the linker is much smaller than traditional labels which can be as much as a nanometer long. Together, these factors help reduce the effect of the offset.

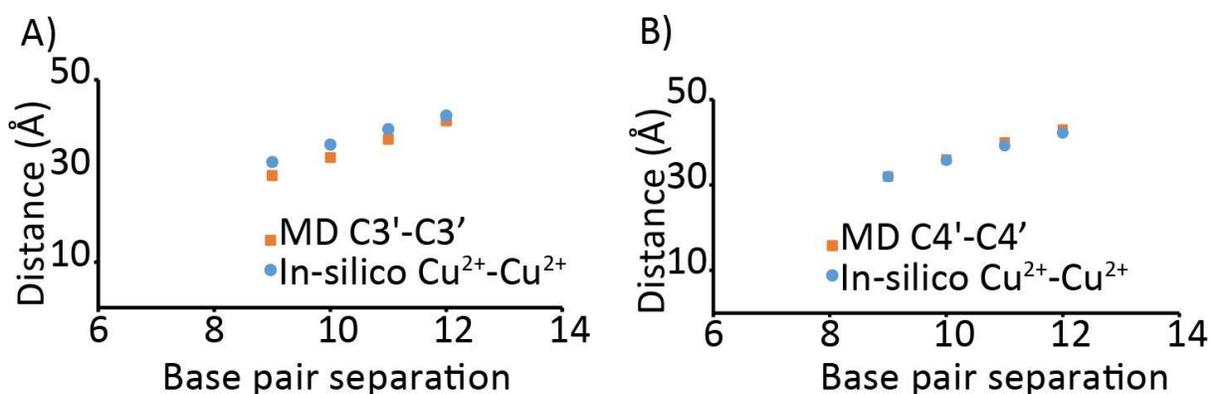


Figure 2-20 Comparison of distance from in-silico model of Cu^{2+} -DPA-DNA with most probable distance from MD simulations for A) C3'-C3' and B) C4'-C4'. We see a slightly better agreement of the Cu^{2+} - Cu^{2+} distance with the C4'-C4' backbone distance.

We also performed in-silico modeling of the Cu^{2+} -DPA on the linear DNA, with same sequence as used in our experiments. We built a DNA duplex, and incorporated the Cu^{2+} -DPA at two specific sites using PyMOL software¹⁷⁴ and performed the following analysis for base pair separations of 9-12. We measured the Cu^{2+} - Cu^{2+} distance from the in-silico model and compared with the most probable distances of the C3'-C3' and C4'-C4' distance distributions obtained from MD simulations. We observed a slightly better agreement of the Cu^{2+} - Cu^{2+} distances to the MD C4'-C4' distances than the C3'-C3', as shown in Figure 2-20. However, the C3'-C3' also agrees

within 2-3 Å. Importantly, this agreement also shows that the offset of Cu^{2+} from the DNA backbone does not significantly affect the overall distance.

2.3.7 In-silico model of non-linear DNA shows that Cu^{2+} -DPA can report on DNA backbone conformations of bent/kinked DNA

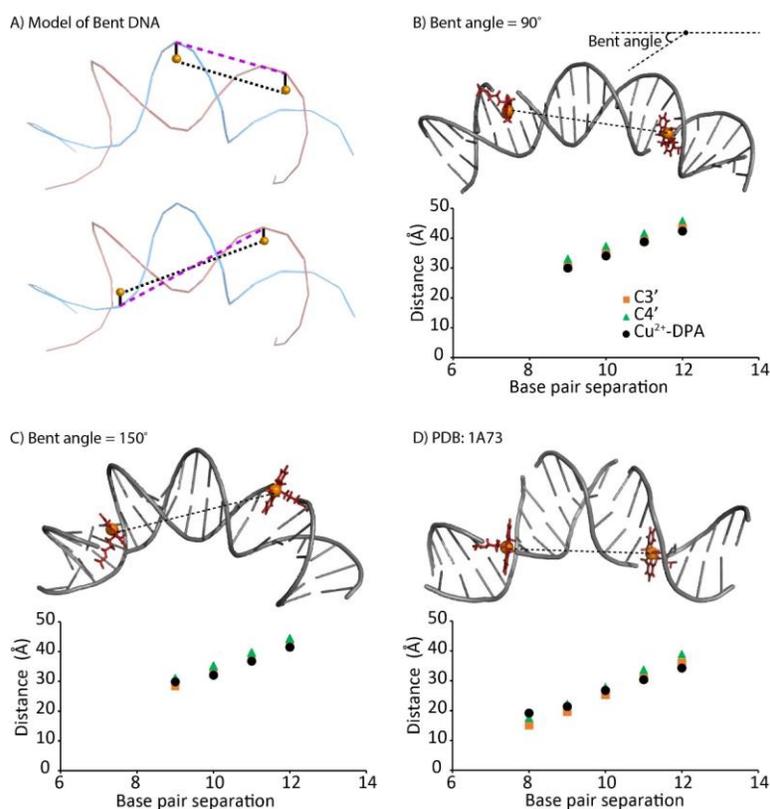


Figure 2-21 A) Model of a bent DNA showing the backbone distance and the Cu^{2+} - Cu^{2+} distance. In the first case, the position of the two labels being almost parallel causes the Cu^{2+} - Cu^{2+} distance to directly translate into the DNA backbone distance. In the second case, the labels are anti-parallel. However, the offset of Cu^{2+} from the backbone does not significantly contribute to the measured distance. B) and C) Modeled DNA structures with a bent angle of 90° and 150° , respectively. D) PDB of a bent DNA (PDB: 1A73)¹⁸³. C3' (orange squares) and C4' (green triangles) DNA backbone distances were measured. In-silico modeling was performed incorporating Cu^{2+} -DPA phosphoramidite into the DNA duplex for different base pair separations and the corresponding Cu^{2+} - Cu^{2+} distance was measured (black circles). All the three distances show reasonable agreement within the caveats of the model.

In order to examine how the Cu²⁺-DPA label would report on the backbone distances of a bent DNA, we took the following approach. We generated two PDB structures of the DNA using the same sequence as the control DNA (cf. Figure 2-2), with a bent angle of 90° and 150°, respectively, as shown in Figure 2-21. The bent structures were created using the 3D-DART software¹⁸². We calculated the C3'-C3' and C4'-C4' distances for base pair separations of 9 through 12. To compare how Cu²⁺-DPA reports on the bent DNA, we generated an in-silico model, where we incorporated the Cu²⁺-DPA at two specific sites of the bent DNA using PyMOL software¹⁷⁴. We measured the Cu²⁺-Cu²⁺ distance, repeated this step for different base pair separations and compared the DNA backbone distances with the Cu²⁺-DPA distance. Clearly, the Cu²⁺-DPA distances agree within 2-3 Å, within the caveats of the model. Finally, we used a bent DNA PDB structure, reported in literature (PDB 1A73)¹⁸³. This PDB structure comprises of an endonuclease bound DNA where the DNA is strongly bent for cleavage. We performed similar in-silico modeling on the PDB structure and measured the Cu²⁺-Cu²⁺ distance along with the backbone distances. Again, the Cu²⁺-DPA distances is within 2-3 Å of the backbone distances. Such agreement highlights the usefulness of the short linker and the orientation of the Cu²⁺-DPA label.

2.4 CONCLUSION

Herein, we provide a comprehensive analysis of the Cu²⁺-DPA motif as a promising spin label for DNA. We have shown that Cu²⁺ specifically binds to the DPA motif in the DNA duplex, eliminating any non-specific binding. Distance distributions of several DPA-DNA duplexes and MD simulations illustrate that the motif is capable of reporting on DNA backbone conformations

in solution, without the need for extensive modeling. Moreover, distance measurements on the Cu^{2+} -DPA motif in the non-complexed DNA does not show orientation effects for any base pair separation between the motifs. The methodology has the resolution to detect differences in DNA length as small as one base pair. These results also support the fact that the motif is small and is present inside the helix and does not perturb the native structure of DNA. Comparison of experimental distance distributions with those calculated from the geometric model of DNA shows that the Cu^{2+} -DPA is only slightly more flexible than the standard rigid nitroxide labels available for DNA. Notably, the spin labeling technique is nucleotide independent, which allows the incorporation of the label at any position. Such Cu^{2+} -based labeling strategy can be combined with nitroxide labeling of proteins in order to measure multiple structural constraints in a protein-DNA complex. The ability of the Cu^{2+} -DPA motif to accurately report on DNA backbone conformation will serve as an essential tool for probing backbone fluctuations and shape adaptations (e.g., bending or unwinding) as the DNA interacts with various proteins to execute biological functions.

2.5 ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation [NSF MCB-1613007] and US Israel Binational Science Foundation [BSF 2018029]. The ESR spectrometer was purchased through funds from the National Science Foundation [NSF MRI-1725678]. The author would like to thank the University of Pittsburgh for the Andrew Mellon Predoctoral Fellowship. We would also like to thank Dr. Jessica Sarver and Anthony Bogetti for the useful discussions on MD simulations. All simulations were carried out at the University of Pittsburgh's Center for Research Computing.

3.0 MOLECULAR DYNAMICS SIMULATIONS USING FORCE FIELD PARAMETERS OF Cu²⁺-BASED DNA LABELS PROVIDE ATOMIC INSIGHTS INTO PULSED ESR BASED DISTANCE CONSTRAINTS

This work has been performed in collaboration with Josh Casto, Xiaowei Bogetti, Charu Arora, Junmei Wang and Sunil Saxena. Junmei Wang developed the force field parameters of the labels. Josh Casto performed the ESR experiments. Xiaowei Bogetti performed the DFT calculations using ORCA package. The dissertation author performed the MD simulations, analyzed the data and prepared the manuscript.

3.1 INTRODUCTION

The DNA helix has a wide range of sequence dependent conformational variability and internal dynamics, that are essential for its role in biological processes^{111–115}. The DNA helix undergoes breathing, bending and twisting motions as well as fluctuations in the helical groove that impart flexibility to the helix. Such dynamics is crucial in many important physiological processes. For example, sequence-dependent bending of DNA is an important criterion for protein-DNA recognition and the function of several DNA-binding proteins such as transcription regulators and restriction enzymes. In recent times, electron spin resonance (ESR) spectroscopy has become increasingly important for measuring such conformational dynamics in DNA.

In particular, distance measurements by pulsed ESR techniques are an important approach for elucidating macromolecular structure and flexibility^{12,99}. These ESR techniques exploit the

dipolar coupling between unpaired electron spins to provide distances in the nanometer range with angstrom-level resolution^{13,36,38,42–44,99,154,184–187}. Such distance constraints can be employed to probe conformational changes in large and complex systems such as membrane proteins, chaperones, protein-protein and protein-nucleic acid complexes, that are otherwise inaccessible to other biophysical techniques^{21,188–195}.

For such measurements site-specific labeling^{2,3,15} of the biomolecule is needed. One such label that has shown great promise is the Cu²⁺-based label, where Cu²⁺ is chelated to the ligand 2,2'-dipicolylamine (DPA) phosphoramidite⁵⁰. The DPA moiety is nucleotide independent and can be easily introduced at any desired positions in the duplex during the synthesis. In addition, the label is structure independent. The complementary site to the Cu²⁺-DPA in the opposing strand contains an abasic site (dSpacer). We recently reported distance measurements performed on several duplexes that were labeled at two sites with the Cu²⁺-DPA⁶¹. The base pair separations between the labels varied from 9 to 12. The experimental most probable distances agreed with the predicted distances using the known values of base pair separation for a B-DNA. Moreover, molecular dynamics (MD) simulations performed on an unlabeled DNA also suggested that the most probable distance measured from the experiment agreed with the backbone (e.g. C3'-C3' or C4'-C4') distances of the labeled sites. These results suggest that the most probable distances obtained using Cu²⁺-DPA can be used to report on the DNA backbone conformations in solution and can discern label separation for differences as small as one base pair⁶¹.

However, ESR distance measurements provide sparse constraints and therefore, benefit from a complementary technique, such as modeling, to better describe the dynamical behavior. MD simulations combined with ESR distance constraints have provided valuable information on conformational dynamics of nucleic acids, relative orientation of double-stranded helices and

information on flexibility of DNA for a specific sequence^{28,32,155,196}. MD simulations can also provide insight into the global (DNA) and local (spin label) motions. The development of accurate force fields for the spin label is, however, a pre-requisite for such MD simulations.

In this work, we have developed force field parameters for Cu²⁺-DPA and its complementary base, dSpacer. Density functional theory (DFT) calculations were utilized to generate the optimized geometry of Cu²⁺-DPA and the dSpacer. Consequently, parameters of bond lengths, bond angles, dihedral angles, partial charges and force constants were developed. We then perform MD simulations with these force fields to better understand the atomic details of the flexibility of DNA as well as the label, Cu²⁺-DPA. The simulations were performed on duplexes with base pair separations of 9 to 12 between the Cu²⁺-DPA centers. We ran the MD simulations for as long as 1 μ s to sample the label dynamics that occur on much longer timescales. These long simulations provide insight into the shape and width of the experimental distance distributions that were observed in DPA-DNA^{50,61}, the experimental methodology used to measure distances by ESR, and the conformations of the DPA-label in the context of the DNA. By combining MD and ESR techniques, we can precisely characterize the structure, conformations, and flexibility of DNA in solution.

3.2 MATERIALS AND METHODS

3.2.1 HYSCORE

Four-pulse hyperfine sublevel correlation (HYSCORE)¹⁹⁷ experiments were performed at 20 K and at X-band frequencies. The pulse sequence $\pi/2 - \tau - \pi/2 - t1 - \pi - t2 - \pi/2 - \tau - \text{echo}$

was used. HYSORE was performed at the field of highest intensity. The initial time delays for τ , t_1 and t_2 was set at 128, 200 and 200 ns, respectively. The delay for t_1 and t_2 had a step size of 16 ns for a total of 256 points. The pulse lengths used were 16 ns and 32 ns for $\pi/2$ and π pulses, respectively. A four-step phase cycling was used to eliminate the unwanted echos^{198,199}. The raw 2D time-domain signals were analyzed and processed using the Hyscorean open-source software package²⁰⁰ in MATLAB to produce the 2D frequency-domain contour plot.

3.2.2 DFT optimization

The Cu^{2+} is coordinated to three nitrogen atoms of the DPA in the equatorial plane. Therefore, the remaining sites in the Cu^{2+} -DPA complex can be occupied by one, two or three water molecules, depending on the geometry. To find out the most possible Cu^{2+} -DPA structure, we performed *ab initio* calculations using a density functional theory (DFT) with the solvent effect being taking into account with the Polarizable Continuum Model implemented in Gaussian 16 software package²⁰¹. We first conducted geometry optimizations at the wB97xd/6-311++G(2d,p) level for three Cu^{2+} -DPA complexes with 1, 2 and 3 coordinated water molecules. The complexation energy of introducing a water molecule to a Cu^{2+} -DPA complex was then calculated. After the most probable Cu^{2+} -DPA complex was identified, we performed vibrational frequency analysis to facilitate us to derive the bond stretching and bond angle bending force constants for the complex at B3LYP/6-31G* level after geometry optimization. Lastly, HF/6-31G* single point calculations were performed to generate electrostatic potentials (ESP) for the Restrained ESP charge fitting. All zero-point energy (ZPE) calculations were performed using the same DFT model as optimization.

3.2.3 MD simulations

First, the Nucleic Acid Builder (NAB) module in the AMBER software suite¹⁵⁶ was used to create the unlabeled B-DNA models, with DPA and dSpacer positions replaced by adenine and thymine, respectively. The adenine and thymine residues at the specific sites were then removed in the PyMOL software¹⁷⁴ and the DFT-optimized structures of Cu²⁺-DPA and dSpacer were added. The Cu²⁺-DPA and dSpacer incorporated DNA were then subjected to MD simulations. The AMBER parmbsc1 force field¹⁵⁵ was used to treat the nucleic acids. For the modified nucleotide residue, we first developed force field parameters for a model compound as shown in Figure 3-1, using a strategy detailed previously⁵⁸. The residue topologies and the key force field parameters of Cu²⁺-DPA and dSpacer are provided in Appendix A. The solvent water was treated with TIP3P water model¹⁵⁷. The DNA duplexes were solvated in a 12 Å truncated octahedral water box. For each DNA duplex, 100 Cl⁻ ion were added to the water box so that the final Cl⁻ concentration is about 0.15 M and the counter ion Na⁺ were then added to neutralize the whole simulation system. The energy minimization and molecular dynamics simulations were performed using the pmemd program in the AMBER16 software package. The solvated systems were first energy-minimized with a harmonic restraint force applied to the DNA residues except for the Cu²⁺-DPA and dSpacer. The restraint force was gradually reduced from 20, 10, 5, 1 and finally to 0 kcal/(mol Å²). The systems were then gradually heated from 50, 100, 150, 200, 250 to 298.15 K. The systems were equilibrated for 2 ns before starting the production MD runs. The time step for integration for heating, equilibration and the production run were set to 2 fs. Periodic boundary conditions along with particle mesh Ewald (PME)¹⁵⁸ were applied to account for long-range electrostatic interactions under NPT (P=1 atm) conditions. SHAKE¹⁵⁹ was used to restrain all

bonds involving hydrogen and a nonbonded cutoff of 10 Å was applied. All visualizations for simulations were done using VMD¹⁶⁰.

3.3 RESULTS AND DISCUSSION

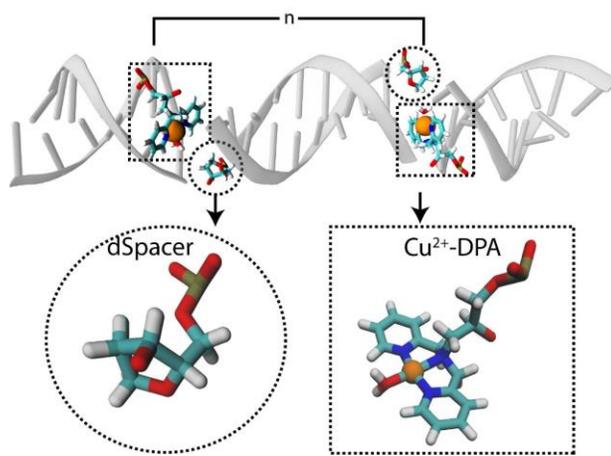


Figure 3-1 DPA-DNA duplex with the Cu²⁺-DPA motif (represented by rectangle) and the dSpacer (represented by circle). The base pair separation between the two Cu²⁺-DPA motifs is denoted by n.

In this work we systematically examined four DPA-DNA duplexes using MD simulations. Figure 3-1 shows a DPA-DNA duplex with two Cu²⁺-DPA motifs, one on each strand, and a dSpacer, complementary to the Cu²⁺-DPA. The base pair separations (n) between the Cu²⁺-DPA motifs were monotonically increased from 9 to 12. The distance distributions on these duplexes have been reported previously⁶¹.

3.3.1 DFT-optimized structure agrees well with crystal structure

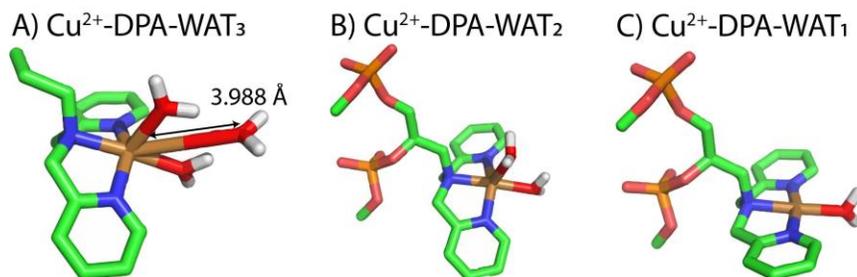
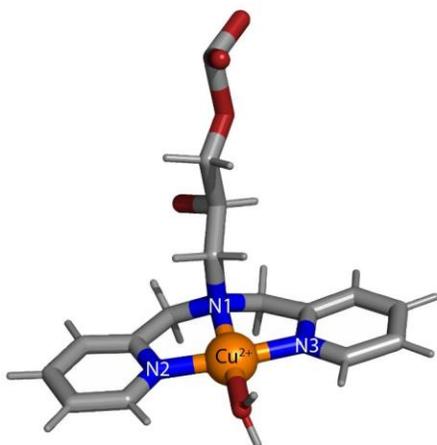


Figure 3-2 Cu^{2+} -DPA structures coordinated with A) three B) two and C) one water molecule.

We first identified the most probable structure for three Cu^{2+} -DPA complexes with 1, 2, and 3 water molecules, using high-level DFT-optimizations at the wB97xd/6-311++G(2d,p) level. The optimized geometries are shown in Figure 3-2. We observed that Cu^{2+} -DPA-WAT₃ is not stable as one Cu-OH₂ distance is 3.988 Å (Figure 3-2A), and this water molecule is likely a solvent water. The complexation energy of adding one water to Cu^{2+} -DPA to form Cu^{2+} -DPA-WAT₁ (Figure 3-2C) is -19.59 kcal/mol and -16.99 kcal/mol after ZPE correction. On the other hand, the energy after adding another water molecule in Cu^{2+} -DPA-WAT₂ (Figure 3-2B) is only -8.06 kcal/mol and -5.86 kcal/mol after ZPE correction. Considering the vaporization energy of water is -9.75 kcal/mol, Cu^{2+} -DPA-WAT₂ is unlikely to be formed in aqueous solvent. Thus, the most probable Cu^{2+} -DPA structure contains only one water molecule.

A)



B)

Bond (Å) or Angle (deg)	Crystal structure	QM
N1 - Cu ²⁺	2.075	2.034
N2 - Cu ²⁺	2.009	1.979
N3 - Cu ²⁺	2.002	1.968
N1 - Cu ²⁺ - N2	78.2	83.412
N1 - Cu ²⁺ - N3	98.2	83.791
N2 - Cu ²⁺ - N3	155.1	166.857

Figure 3-3 A) DFT optimized structure of Cu²⁺-DPA. B) Comparison of bond angles and bond lengths with crystal structure¹⁷⁵ of Cu²⁺-DPA show good agreement with the DFT-optimized structure.

Besides water, the most probable Cu²⁺-DPA complex, as shown in Figure 3-3A, also has the Cu²⁺ coordinated to 3N atoms of the DPA (N1, N2 and N3), similar to what the crystal structure showed¹⁷⁵. Previous continuous wave ESR data on Cu²⁺-DPA-DNA^{50,61} had reported g_{\parallel} and A_{\parallel} values that were also consistent with three nitrogen coordinating to Cu²⁺ at equatorial positions. A comparison of bond lengths and bond angles between the DFT-optimized and crystal structures show a reasonable agreement, as shown in Figure 3-3B. Furthermore, the optimized geometry did not accommodate any axial water coordination. This absence of axial water coordination indicates a square planar geometry of Cu²⁺-DPA that is in accordance with the crystal structure¹⁷⁵.

3.3.2 HYSORE indicates the presence of the fourth equatorially coordinating atom

In order to experimentally verify equatorial water coordination to DPA, we performed HYSORE experiments on Cu²⁺ bound to the DPA ligand. HYSORE provides an efficient way to measure the equatorial water coordination. Figure 3-4A shows the ¹H HYSORE spectrum of

Cu^{2+} -DPA which displays two proton features. First, the broad ridge spanning a width of 6-9 MHz, represented by rectangles in Figure 3-4A, is a characteristic feature of solvent coordination in the equatorial plane^{94,95,202,203}. Since three of the equatorial coordination sites of Cu^{2+} are occupied with DPA nitrogen atoms, it leaves space for only one solvent molecule to bind equatorially. Second, the short intense ridge around the proton Larmor frequency (~ 14 MHz), circled in Figure 3-4A, can be a consequence of weakly coupled protons of solvent molecules or axial coordination to water⁹⁴. Since crystal structure of Cu^{2+} -DPA shows a square planar geometry¹⁷⁵, the proton feature at ~ 14 MHz is likely due to the weakly coupled protons⁹⁵. We then performed HYSORE on one of the Cu^{2+} -DPA-DNA duplexes ($n=11$). The HYSORE spectrum of the Cu^{2+} -DPA-DNA duplex, shown in Figure 3-4B, closely resembles that of the Cu^{2+} -DPA, showing both the features of equatorial water coordination and weakly coupled solvent molecules.

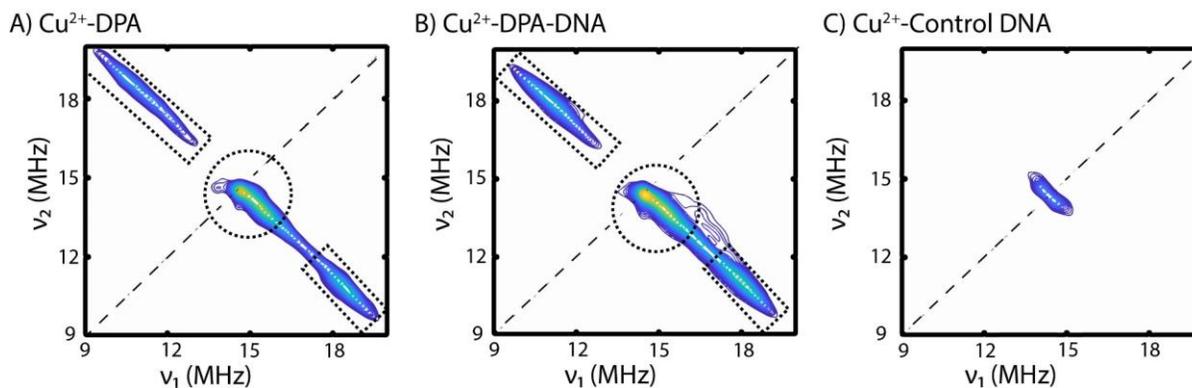


Figure 3-4 ^1H HYSORE spectra of A) Cu^{2+} -DPA B) Cu^{2+} -DPA-DNA and C) Cu^{2+} -control DNA, analyzed to the same contour levels. In Cu^{2+} -DPA and Cu^{2+} -DPA-DNA, the proton signal results from equatorially coordinated water molecule (rectangle) and weakly coupled solvent (circle). In control DNA, the proton signal results from only weakly coupled solvent molecules.

To compare how Cu^{2+} coordination to water differs in absence of DPA inside the DNA duplex, we performed HYSORE on a Cu^{2+} -bound control DNA, lacking any DPA phosphoramidite or dSpacer. The sequence of the control DNA is the same as the DPA-DNA

duplex, except the DPA and dSpacer positions were replaced by adenine and thymine, respectively. The proton signature in the HYSORE spectrum, shown in Figure 3-4C, clearly lacks the features of equatorial water coordination as seen in Cu^{2+} -DPA-DNA. Instead, the spectrum only shows the short intense ridge around ~ 14 MHz. Overall, from the HYSORE spectra we infer that the Cu^{2+} -DPA is coordinated to a solvent molecule in the equatorial plane, as also observed in the DFT-optimized Cu^{2+} -DPA structure.

3.3.3 MD simulations show that the Cu^{2+} - Cu^{2+} distance reasonably agrees with the backbone distance

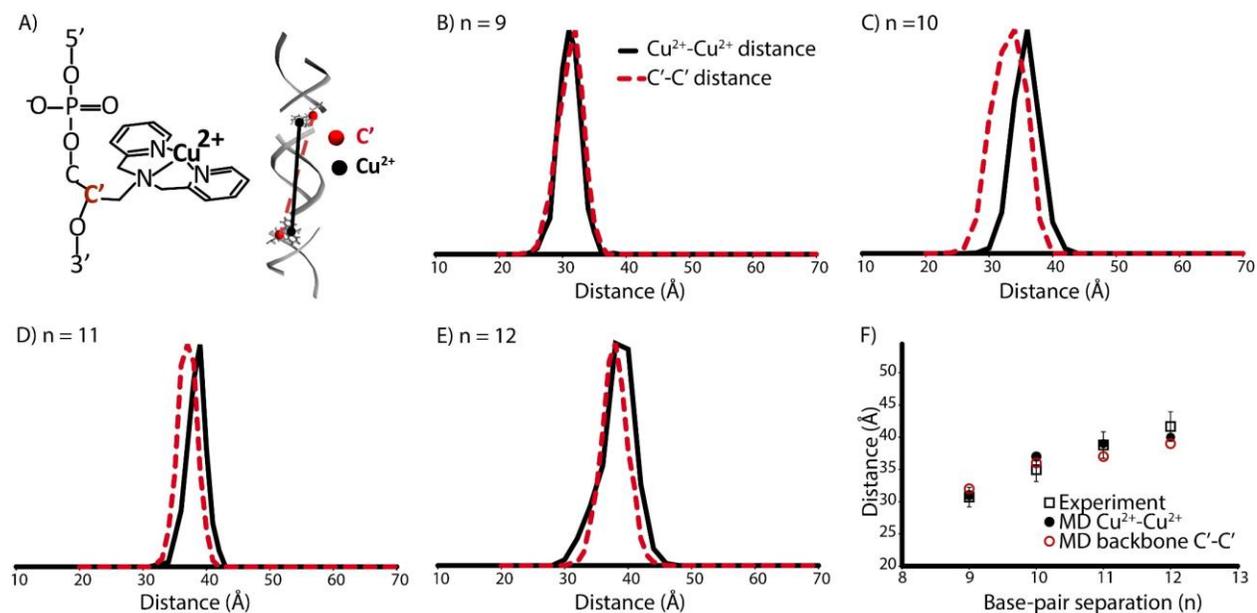


Figure 3-5 A) Cu^{2+} -DPA phosphoramidite with the backbone carbon atom, C' marked in red. The Cu^{2+} - Cu^{2+} distance is denoted by the black solid line and the backbone C'-C' distance is represented by the red dashed line in a DPA-DNA duplex. The Cu^{2+} - Cu^{2+} distance (black solid) is compared with the backbone distance (red dashed) for B) n=9 C) n=10 D) n=11 and E) n=12 base pair separations. F) Plot showing the experimental and MD Cu^{2+} - Cu^{2+} and MD C'-C' most probable distances. All the distance distributions agree reasonably well.

We first performed 1 μ s MD simulations on the Cu^{2+} -DPA-DNA duplexes, with n varying from 9 to 12. Figure 3-5A shows the Cu^{2+} -DPA phosphoramidite with the backbone carbon atom, C' marked in red. The corresponding Cu^{2+} - Cu^{2+} and C'-C' distances in a Cu^{2+} -DPA-DNA duplex are represented as solid black and dashed red lines, respectively. We chose C' as the backbone atom for the DPA phosphoramidite as it is the central point to which the DPA moiety is attached. Moreover, the C' atom has been previously shown to best represent the C3' or C4' backbone atoms of an unmodified regular nucleotide⁶¹.

Figure 3-5B-E shows the comparison between the Cu^{2+} - Cu^{2+} and C'-C' distance distributions obtained from 1 μ s trajectories for each DNA duplex. Figure 3-5F shows a plot of the most probable Cu^{2+} - Cu^{2+} distances from the experiment and MD as well as the most probable C'-C' distances from MD. As can be seen, all the distances agree well within ~ 1 Å. The plot, therefore, highlights that an important advantage of the Cu^{2+} -DPA labeling strategy is to directly report on the DNA backbone constraints, without any additional modeling.

Such an agreement is consistent with the linker length of the DPA moiety. From the MD simulations, the average length of the Cu^{2+} atom from the C' backbone atom, ranges from 3.5 to 4.5 Å. The linker length in Cu^{2+} -DPA is, therefore, considerably smaller than traditional labels, which can be a nanometer long^{33,204}. The agreement between the Cu^{2+} and backbone C' distance distributions is due to a combination of two factors. The size of the linker is much smaller than standard DNA labels. Second, the Cu^{2+} is arranged within the helix (cf. below). As a result, the offset due to the linker partially cancels. Finally, the Cu^{2+} - Cu^{2+} distance within the DNA helix can be considered as a sum of two components: an axial distance, which increases linearly with the base pair separation and a radial distance, which oscillates with base pair separation. As the two modified sites are separated by at least 9 base pairs, the axial vector (~ 31 Å for $n=9$) is significantly

longer than the radial counterpart, i.e., the Cu^{2+} - Cu^{2+} distance is dominated by the axial component. Indeed, previous work has shown that such considerations are valid for $n > 4$ ⁶¹. Consequently, the separation between the Cu^{2+} centers along the DNA axis is roughly the same as the separation between the DNA backbone atoms.

3.3.4 DPA-DNA based MD simulations capture the experimental most probable distance

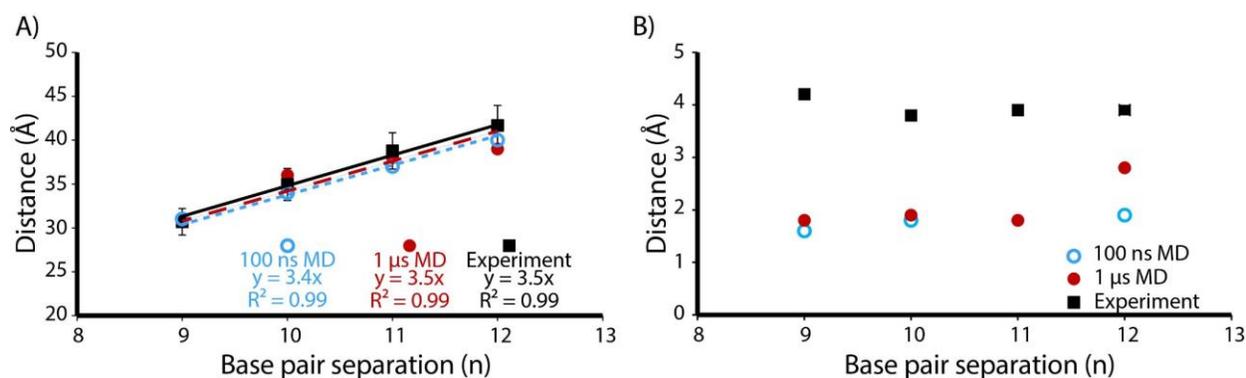


Figure 3-6 Plot of A) most probable distances and B) standard deviation of distance distributions from ESR (black), 100 ns MD (blue) and 1 μs MD (red) against base pair separation (n).

Next, we examined how the MD simulations on the Cu^{2+} -DPA-DNA duplexes compare to the experimental distance distributions. The experimental data and distance distributions for 9 to 12 base pair separations have been reported previously⁶¹. Figure 3-6A shows the comparison of the most probable Cu^{2+} - Cu^{2+} distances from 1 μs MD (red) and ESR (black) for the four Cu^{2+} -DPA-DNA duplexes. The distances agree well within 2 Å which is within the error of the experiments. Furthermore, to explore the shortest simulation time that can report on the experimental most probable conformation, we analyzed the first 100 ns of the MD simulations. The most probable distance from the 100 ns simulations, shown in blue in Figure 3-6A, is likely sufficient to predict the most probable distance from ESR measurements. A linear trend with a y-

intercept of 0, was obtained with a slope of 3.5 Å, 3.4 Å and 3.5 Å for the experiment, 100 ns and 1 μs MD simulations, respectively. These values agree well with each other as well as with the ~3.4 Å separation between adjacent bases, as observed in a B-DNA duplex¹⁷³.

Figure 3-6B shows the comparison of the standard deviation of the distance distributions obtained from the experimental data (black squares), 100 ns (blue circle) and 1 μs (red circles) MD simulations. Despite the agreement in most probable distance, the MD simulations do not appear to capture the full width of the experimental distribution. Extending the simulation time from 100 ns to 1 μs did not significantly change the distribution width for duplexes n=9 to 11. On the other hand, for the duplex n=12, there was an increase in the standard deviation from 1.9 Å for a 100 ns run to 2.8 Å for a 1 μs run. The experimentally measured standard deviation is 3.9 Å for n=12. This discrepancy in the distribution width between the experiment and MD simulations is discussed below.

3.3.5 Orientation of Cu^{2+} -DPA helps reduce the effect of linker offset

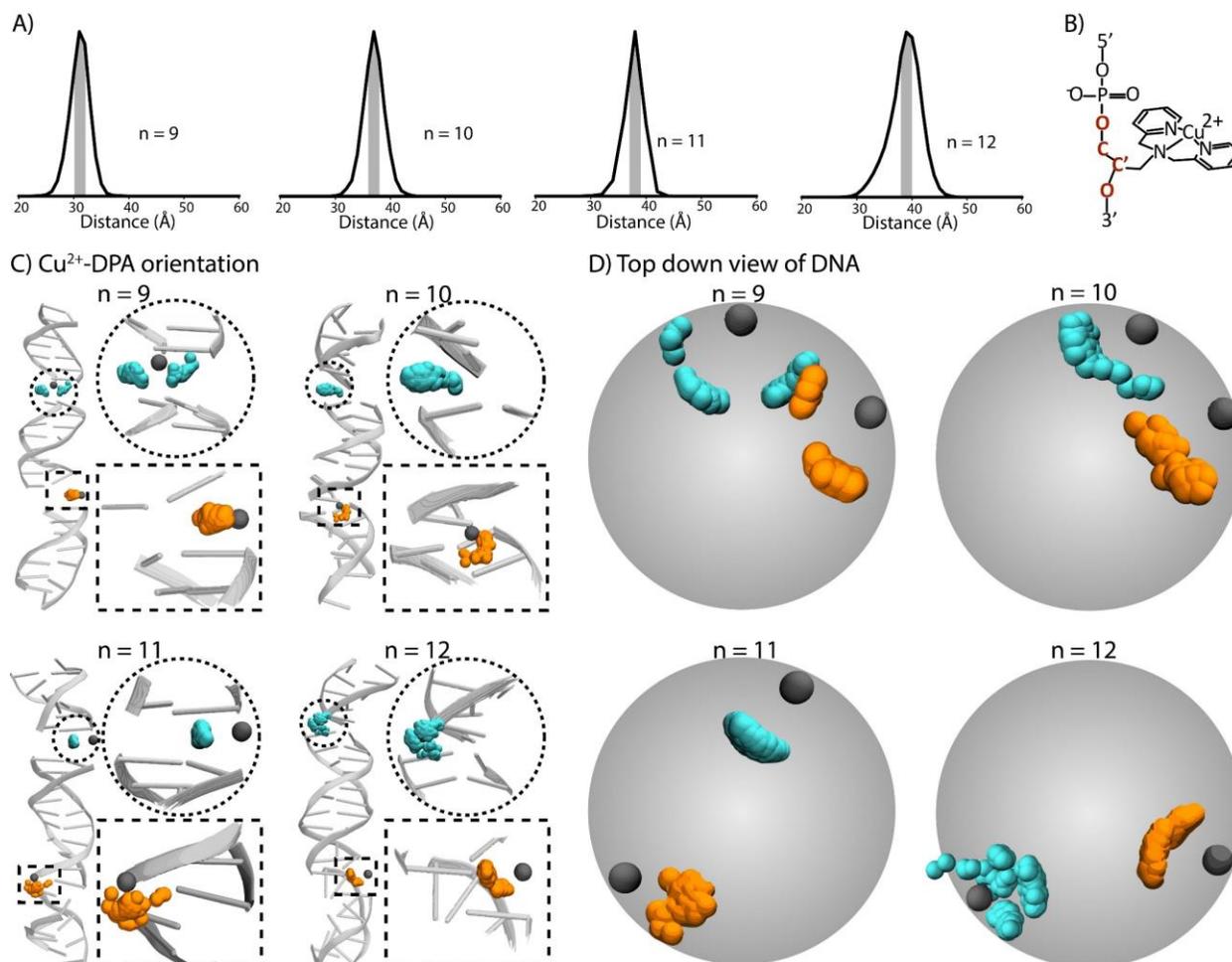


Figure 3-7 A) The frames from MD trajectories considered for obtaining the spatial distribution of Cu^{2+} are within $\pm 1 \text{ \AA}$ of the Cu^{2+} - Cu^{2+} most probable distance and shown by the shaded region. B) The frames are aligned with respect to the DPA backbone atoms marked in red. C) The distribution of the Cu^{2+} centers for the DPA site 1 (blue, circles) and site 2 (orange, rectangles) for duplexes $n=9-12$. D) The top-down view of the DNA where the blue and orange spheres represent the Cu^{2+} at DPA sites 1 and 2, respectively. The grey sphere represents the backbone carbon atom, C'.

In order to elucidate spin label conformations that yield the most probable distance, we extracted the frames from the MD trajectories for which the Cu^{2+} - Cu^{2+} distance is within $\pm 1 \text{ \AA}$ of the MD most probable distance. The analysis is shown in Figure 3-7A. Each frame was aligned

with respect to the O-C-C'-O backbone atoms, marked in red in Figure 3-7B. Figure 3-7C-D show the spatial distribution of Cu^{2+} for the DPA sites corresponding to the most probable distance. For all the duplexes, the Cu^{2+} -DPA is positioned within the DNA duplex.

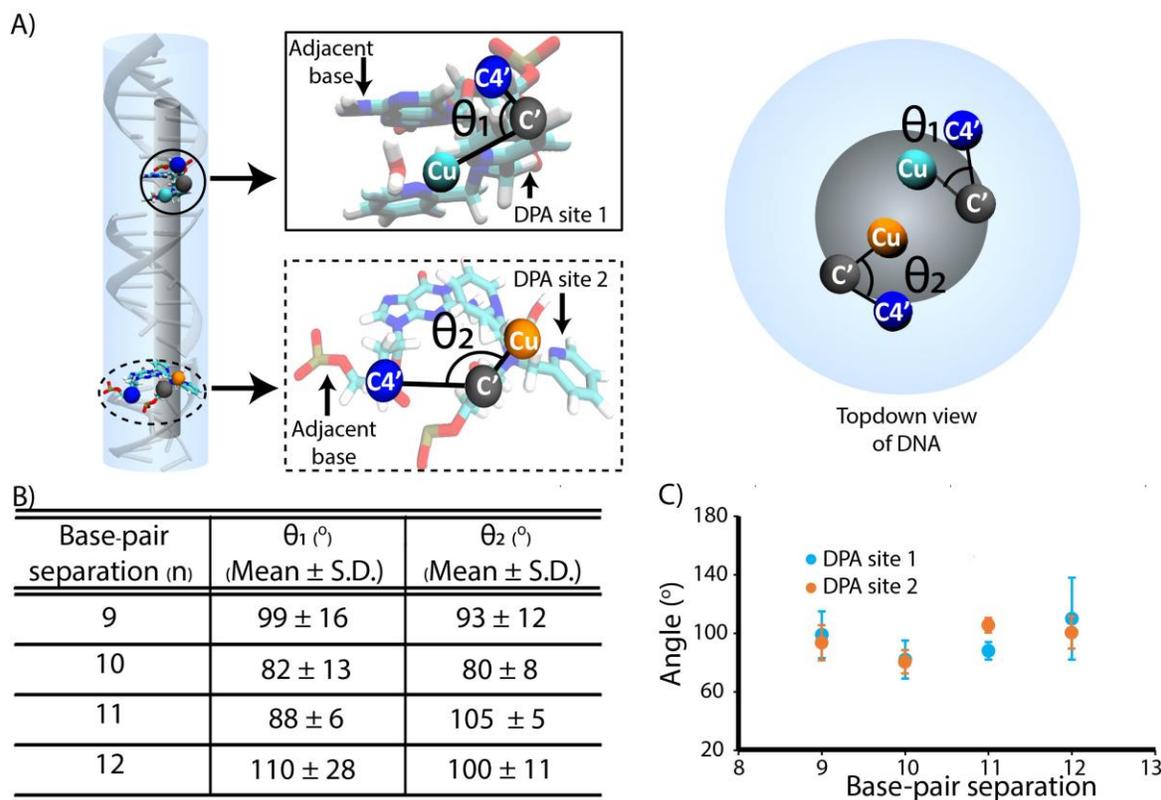


Figure 3-8 A) Angles between the C4' atom of the adjacent base to the DPA (blue sphere), the C' backbone atom of DPA (grey sphere) and the Cu^{2+} (cyan and orange for DPA site 1 and 2, respectively). B) Table showing the values for the two DPA sites C) Plot of the angles versus base pair separation for the two DPA sites.

We measured the angle between the DNA backbone and the Cu^{2+} for both sites in the duplex, as defined in Figure 3-8A. Figure 3-8B shows the value of these angles for the different duplexes. Within the standard deviation, the values agreed well. Figure 3-8C shows that the average Cu^{2+} to backbone angles for each duplex are all between 80° to 100° for both DPA sites. These results suggest that the most probable spin label conformation has the Cu^{2+} centers present almost perpendicular to the DNA backbone. Moreover, the pyridine rings of the DPA moiety are present planar with respect to the DNA backbone. This orientation would therefore lead to the least

steric interactions with the adjacent bases. In addition, the perpendicular orientation of the Cu^{2+} -DPA leads to the effects of linker offset to partially cancel out when the Cu^{2+} - Cu^{2+} is compared to the backbone distance.

3.3.6 MD simulations show that Cu^{2+} -DPA have varied mobility that depends on position

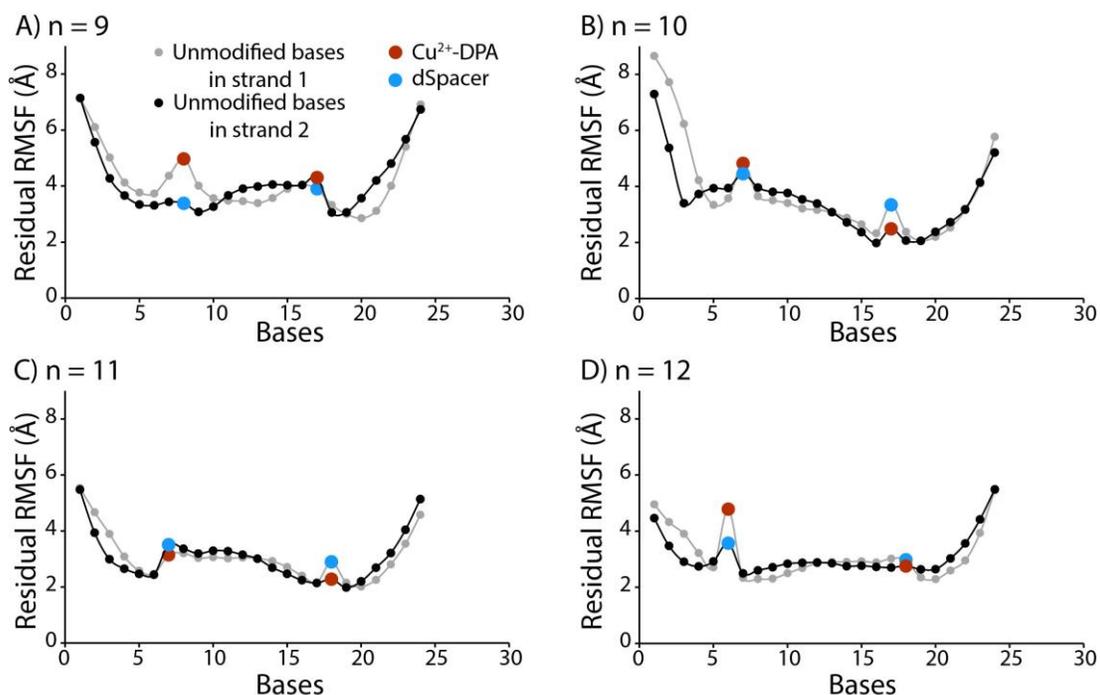


Figure 3-9 The root mean square fluctuations (RMSF) of all the bases in the DPA-DNA duplexes with A) n=9 B) n=10 C) n=11 and D) n=12. The grey and black denote the two strands in the DNA and the red and blue denote Cu^{2+} -DPA and dSpacer, respectively. As can be seen, the RMSF is high for the terminal bases, as expected. The RMSF for Cu^{2+} -DPA and the abasic dSpacer positions are generally high, indicating more flexibility than the natural bases.

In order to examine the motion of the DPA we first measured the root mean square fluctuation (RMSF) values of each base for all the DNA duplexes. The RMSF was calculated with reference to the average structure of each duplex and the results are shown in Figure 3-9. As expected, the bases at the 5' and 3' ends showed high RMSF and thereby, high flexibility.

Interestingly, the flexibility of the DPA and the dSpacer is also elevated than other bases and are comparable to the nitroxide derivative of guanine²⁸. These results are expected because DPA and the dSpacer lack intrastrand hydrogen bonding between them unlike regular base pairing. Furthermore, Figure 3-9D shows that the DPA site1 in duplex n=12 has a distinctly elevated RMSF compared to the adjacent bases. Indeed, this site also exhibited the largest fluctuations in θ_1 (cf. Figure 3-8B). Such higher mobility can help increase sampling of thermally accessible conformations of the label during the course of an MD run which is consistent with the increase in distance distribution width observed in the MD simulations from 100 ns to 1 μ s (cf. Figure 3-6B). Moreover, such differences in the mobility depending on the location of the DPA site is not unreasonable given that local steric interactions with the neighboring bases have a strong influence on the degree of motions of the label.

Finally, these results suggest that a 1 μ s simulation may be insufficient to fully capture the slow motion of the spin label^{129,205–207}. In addition, there may be contribution from other factors leading to the discrepancy between the distribution widths in MD vs the experiment. The experimental data was collected at 20 K which likely captures the conformations that exist in the glass state, while MD was performed at 298 K. Second, glycerol was used in the experiment in order to form a glass and as a cryoprotectant. It is possible that the presence of glycerol modified the solvation and dynamics of the DNA^{208,209}. Finally, the force fields for Cu²⁺-DPA may not be able to fully capture all tertiary interactions.

3.3.7 MD simulations provide insight into the use of DEER for such measurements

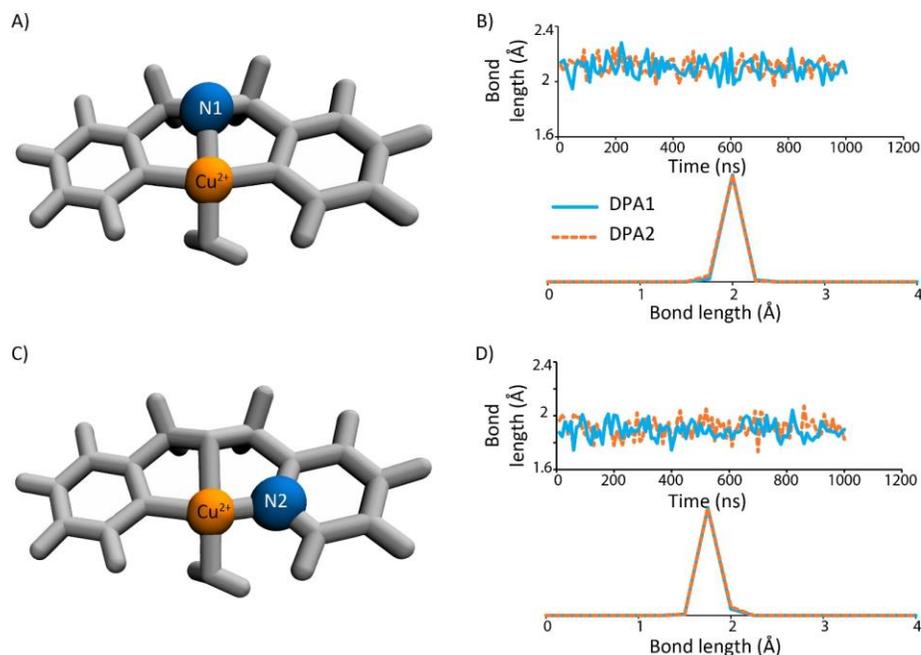


Figure 3-10 Fluctuations in bond length between Cu²⁺ and two coordinating nitrogen atoms – A) Bond between Cu²⁺ and the backbone N1 atom. B) Fluctuations in Cu²⁺-N1 bond length in the two DPA sites, sampled over 1 μ s of MD for the duplex n=11 (top chart) and the probability distribution (bottom chart). C) and D) Same analysis repeated for bond length between Cu²⁺ and the pyridine N2 atom. A change of \sim 25% in the bond length is observed for Cu²⁺-N1 and \sim 30% for Cu²⁺-N2.

The MD trajectories are also helpful to understand key principles of the use of double electron electron resonance (DEER) for measurements on Cu²⁺-DPA labeled species. The pulses used in DEER excite only a small fraction of the total ESR spectrum. Such selective excitation can lead to selection of only a small portion of all possible molecular orientations, leading to a dependence of the measured signal on the magnetic field. This is often referred to as orientational selectivity^{148,165,167–170,210–213}. Previous DEER results have shown the absence of orientational effects for Cu²⁺-DPA at both X-band and Q-band frequencies in the non-complexed state^{50,61}. This

is intriguing because Cu^{2+} -based measurements on proteins using the dHis motif¹⁶⁵ can be orientational selective at Q-band but not at X-band.

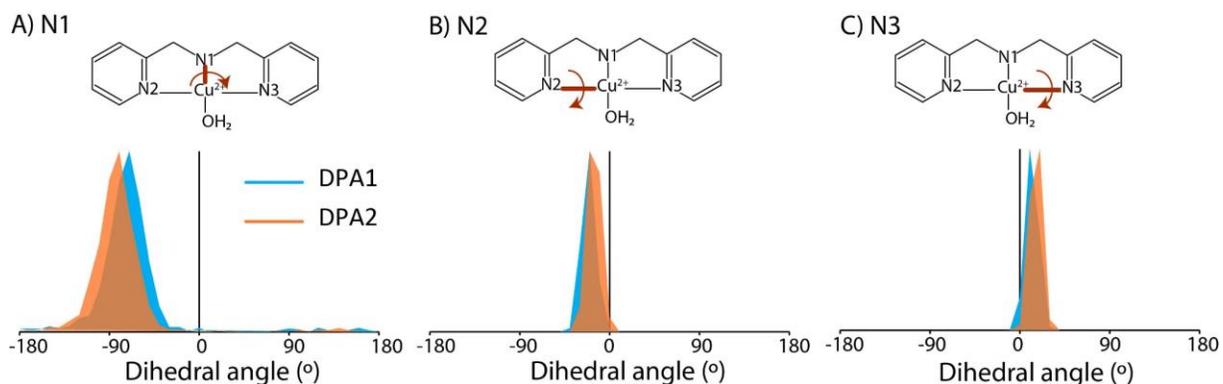


Figure 3-11 Dihedral angles measured between Cu^{2+} and its coordinating atoms for 1 μs of MD run on DPA-DNA duplex with $n=11$. The dihedral angle is denoted by the red arrow (top chart). Bottom chart shows the probability distribution of the dihedrals between the two DPA sites, DPA1 (blue) and DPA2 (orange).

We examined the MD data to gain insight into these observations. Figure 3-10 shows the fluctuations of the bond lengths for bonds involving Cu^{2+} . These bond lengths were extracted from the 1 μs MD trajectory on the $n=11$ duplex. The bond length between Cu^{2+} to N1 ranges from 1.9 to 2.3 \AA , with an average bond length of $2.1 \pm 0.1 \text{\AA}$ (average \pm s.d.). On the other hand, the bond length of Cu^{2+} to N2 was $1.9 \pm 0.1 \text{\AA}$. These mean values agree well with the crystal structure¹⁷⁵ (cf. Figure 3-3). Figure 3-11 shows the dihedral angles between Cu^{2+} and atoms in its immediate coordination environment. The dihedral angles, labeled as N1, N2 and N3, show values of $-75^\circ \pm 23^\circ$, $-20^\circ \pm 7^\circ$ and $19^\circ \pm 6^\circ$ (average \pm s.d.) respectively and are similar for the two DPA sites.

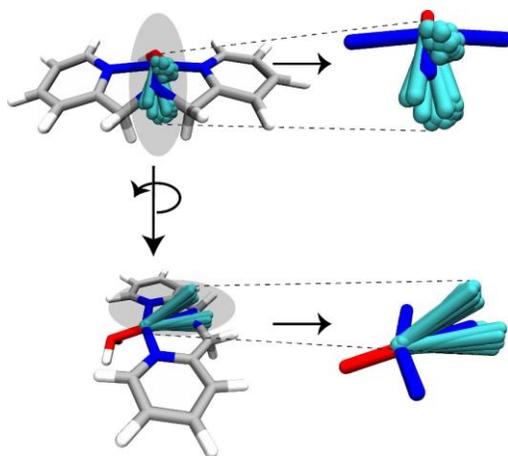


Figure 3-12 The distribution of $g_{||}$ directions (cyan) in the Cu^{2+} -DPA system, sampled every 10 ns of the 1 μs MD run. The blue and red bonds show the Cu^{2+} coordination with nitrogen and oxygen, respectively. The cyan represents the $g_{||}$. The g -tensor calculations were performed with ORCA^{214,215}.

Such variations in the coordination environment of Cu^{2+} have significant influence on the g -tensor of the label. Accordingly, we calculated the g -tensor values and orientations for 100 snapshots from our MD by using ORCA^{214,215}. For these calculations, we sampled every 10 ns of the 1 μs trajectory of duplex $n=11$. The distribution in $g_{||}$ orientations shows dramatic variations with as much as a 25° change, as shown in cyan in Figure 3-12. Due to its 180° symmetry^{165,212}, the $g_{||}$ orientations show both ‘up’ and ‘down’ orientations with respect to the plane of the DPA moiety. This is not unexpected and has been observed before^{165,212}. In addition to this distribution of $g_{||}$ in the two Cu^{2+} centers, there is also additional contribution due to the fluctuations of the dihedral angles of the DPA linker.

Together these two effects have significant implications on the relative orientations of the g -tensors of the two Cu^{2+} spins in the DNA duplex. The relative angles of the two g -tensors are defined by three angles that are shown in Figure 3-13. χ is the angle between the $g_{||}$ axis and the interspin vector, r . γ is the angle between the $g_{||}$ axis of spin A and its projection on spin B. η is the angle between the g_{\perp} axis of spin A and its projection on spin B.

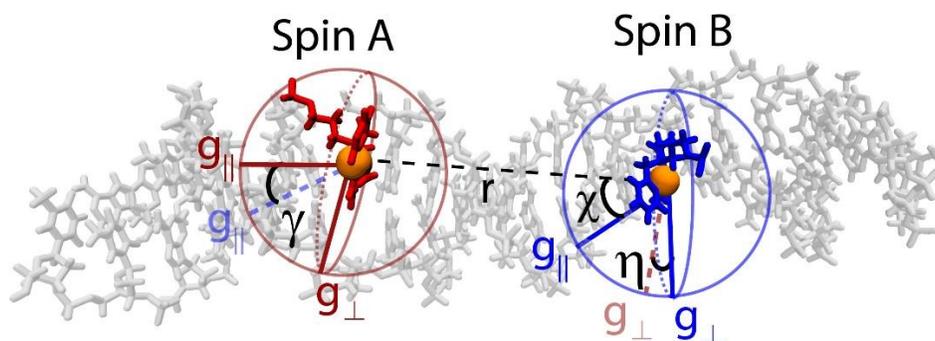


Figure 3-13 The relative orientations between the two spins, A and B, are indicated by three angles, χ , γ and η .

The three angles, namely χ , γ , and η , were calculated via the MDTraj software²¹⁶ for 100 snapshots of the DPA-DNA duplex with $n=11$. From our analysis, we obtained an average of $103^\circ \pm 69^\circ$ (average \pm s.d.), $87^\circ \pm 36^\circ$ and $93^\circ \pm 44^\circ$ for χ , η and γ , respectively. Similar analysis of angles for the dHis-Cu²⁺ protein label have shown a standard deviation of $\sim 10^\circ$ for each set of angles^{58,165}. Thus, the Cu²⁺-DPA labels have an orientational distribution that is significantly larger compared to the dHis-Cu²⁺ label used for proteins. The key distinguishing feature appears to be the lack of contribution from side-chain fluctuations as in the dHis motif^{58,165}.

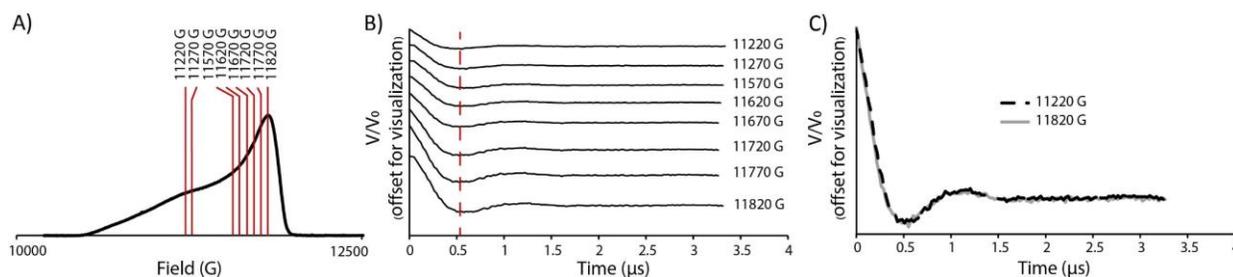


Figure 3-14 A) DEER performed on DPA-DNA duplex with $n=11$ at Q-band frequency. The red lines show the various fields at which DEER was performed. B) The simulated DEER time traces at each field at the Q-band frequency. The y-axis is offset for ease of visualization. The red dashed line represents the first period of the modulations for all fields. C) Background subtracted time domain data at $g_{||}$ (11220 G, dashed) and g_{\perp} (11820 G, solid) regions. The figure shows the lack of orientational selectivity effects at Q-band frequency for Cu²⁺-DPA.

We utilized the values of χ , γ , and η angles and performed simulations to obtain individual time-domain DEER signals at various fields using the methodology reported previously⁴⁵. The

parameters used for the simulations include the g and hyperfine tensor values for Cu^{2+} -DPA ($g_{\parallel} = 2.247$, $g_{\perp} = 2.054$, $A_{\parallel} = 170$ G and $A_{\perp} = 17$ G⁶¹). We also used the experimental most probable distance of 3.78 nm and a standard deviation of 0.38 nm for the distance distribution for duplex $n=11$. Figure 3-14A shows the eight different magnetic fields where we have previously reported DEER for $n=11$ ⁶¹ and where the simulations are carried out in this work. As can be seen from Figure 3-14B, the simulated DEER time traces have the same period at all fields. Furthermore, Figure 3-14C shows that there is no difference in the dipolar frequency at the g_{\perp} (11820 G) and g_{\parallel} (11220 G) regions, indicating the lack of any orientational selectivity effects at Q-band frequency.

3.3.8 Many orientations are excited even at a single magnetic field

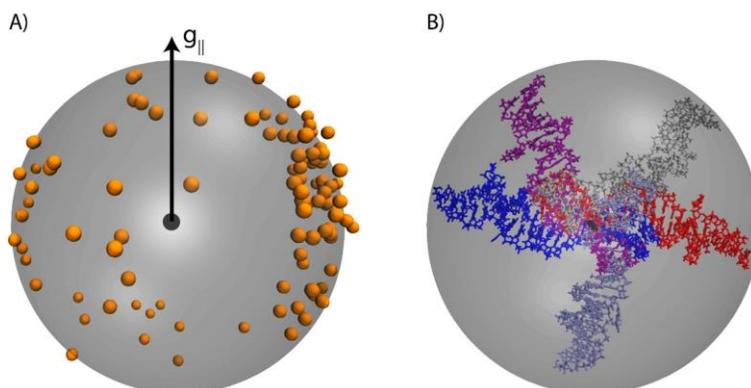


Figure 3-15 MD frames of DPA-DNA duplex ($n=11$) sampled every 10 ns for a total 1 μs . The black sphere in the center is the reference Cu^{2+} . All frames are aligned to the g_{\parallel} axis of the reference Cu^{2+} . The second Cu^{2+} position is represented as orange spheres. The g_{\parallel} direction, marked in the figure, is calculated from ORCA^{214,215}. B) The entire DNA duplex is shown for 5 snapshots for simple viewing. The figure suggests that there is a wide range of molecular orientations at a single g_{\parallel} orientation.

In order to visualize how the effects of orientational selectivity is mitigated, we chose one Cu^{2+} center as the reference and aligned its g_{\parallel} axis to overlay 100 structures of DPA-DNA, evenly sampled from the 1 μs MD simulations. The DPA-DNA duplex with a base pair separation of 11

is used for this analysis. Figure 3-15A shows the locations of the second Cu^{2+} (orange) and the reference Cu^{2+} center is shown as a black sphere. As is evident from the figure, even at a single magnetic field corresponding to g_{\parallel} , there is a large number of molecular orientations that can exist (Figure 3-15B). In general, the finite bandwidth of pump pulse in DEER leads to an excitation of an even wider range of molecular orientations at a particular magnetic field. Such occurrence is a combined consequence of fluctuations in the dihedral angles of the linker and in the coordination environment of Cu^{2+} leading to a large distribution in the g-tensor orientations.

3.4 CONCLUSION

In conclusion, we have developed force field parameters for the Cu^{2+} -DPA spin label and its complementary base, dSpacer in DNA. The DFT-optimized structure of the label agreed well with the crystal structure. Additionally, the most energy favorable DFT-structure had one equatorial water coordination and was consistent with HYSCORE results on the Cu^{2+} -DPA-DNA duplex. We have shown that the distance between the Cu^{2+} centers from the MD simulations can directly report on the DNA backbone distances without the need of additional modeling. In addition, the most probable Cu^{2+} - Cu^{2+} distance from the experiment and the MD agree within ~ 2 Å and for a simulation time as short as 100 ns. Further analyses showed that the Cu^{2+} -DPA moiety is inside the DNA helix and oriented roughly perpendicular to the DNA backbone. The motions of the label are slow such that simulation longer than 1 μs are needed to adequately capture the experimental distance distribution. Finally, the MD results illustrated that the fluctuations of the Cu^{2+} coordination environment, together with the linker flexibility, lead to a wide distribution of the relative orientation of the two Cu^{2+} g-tensor orientations. This distribution is large enough to

dilute any orientational selectivity effects, even at Q-band frequency. In conclusion, the development of force field parameters of the spin labels for MD simulations along with ESR can help elucidate DNA backbone structure and conformational dynamics.

3.5 ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation [NSF MCB-1613007, NSF 1955260] and the National Institutes of Health [NIH: P30 DA035778]. The ESR spectrometer was purchased through funds from the National Science Foundation [NSF MRI-1725678]. The author would like to thank the University of Pittsburgh for the Andrew Mellon Predoctoral Fellowship. We would also like to thank Dr. Lillian Chong and Anthony Bogetti for the useful discussions on MD simulations. All simulations were carried out at the University of Pittsburgh's Center for Research Computing.

4.0 THE Cu^{2+} -NITRILOTRIACETIC ACID COMPLEX IMPROVES LOADING OF α -HELICAL DOUBLE HISTIDINE SITE FOR PRECISE DISTANCE MEASUREMENTS BY PULSED ESR

This work, written in collaboration with Matthew J. Lawless, Gordon S. Rule and Sunil Saxena, has been published in Journal of Magnetic Resonance, 2018, volume 286, pages 163-171. The dissertation author collected and analyzed the ESR data and prepared the manuscript. The human glutathione S-transferase protein was provided by Gordon S. Rule.

4.1 INTRODCUTION

Electron spin resonance (ESR) in combination with site-directed spin labeling (SDSL) has developed into a widely used technique to provide a multifaceted view of protein structure and dynamics³. Recently, methods of distance measurements using paramagnetic metal ions such as Cu^{2+} ^{41,42,46,210,217,218}, Gd^{3+} ^{219–225}, and Mn^{2+} ^{226–229} have been established. These paramagnetic centers can be incorporated into proteins, and other biological macromolecules, using intrinsic metal binding centers^{41,54,230–233} or by metal-based spin labeling methods^{47,48,91,224,227,234}. Pulsed dipolar spectroscopic methods such as double quantum coherence (DQC)^{13,42,235,236} and double electron-electron resonance (DEER)^{99,154,217} are used to extract distance distributions to generate structural constraints. Since these methods report a distribution of distances, this pulsed ESR experiment can provide insight into protein structure and flexibility.

The paramagnetic Cu^{2+} metal ion has become of particular interest for DEER based structural determination^{38,41,50,212}. Cu^{2+} is used by numerous proteins in nature as an enzymatic cofactor, and thus, many proteins contain rigid, natural Cu^{2+} -binding sites^{41,54,230–233}. Furthermore, new spin-labeling methodologies have been developed to site specifically incorporate Cu^{2+} using covalently attached metal-chelating tags⁴⁸. Of special interest is the double histidine (dHis) motif which uses two histidine residues in proximity to chelate Cu^{2+} with surprisingly high rigidity^{47,91}. This motif uses an $i, i+4$ arrangement of histidine residues to attach Cu^{2+} to an α -helical site²³⁷. Likewise, β -strand sites use an $i, i+2$ arrangement which places the histidine residues on the same face of the sheet⁴⁷. However, this motif's major limitation is poor Cu^{2+} selectivity which leads to unspecific binding elsewhere in the protein.

The introduction of Cu^{2+} complexed to the chelating group iminodiacetic acid (IDA) drastically increases the binding specificity of Cu^{2+} towards the vacant dHis site and effectively reduces non-specific binding of Cu^{2+} ⁴⁷. The total dHis- Cu^{2+} -IDA scheme can be seen in Figure 4-1. Of particular note is the extreme rigidity seen by this motif. Distance distributions obtained from the dHis motif are narrow exhibiting ~ 0.5 angstrom standard deviations⁴⁷. Furthermore, the availability of crystal structure of Cu^{2+} -IDA coordinated to imidazole²³⁸ makes it relatively easy to extract direct protein backbone-backbone distances, as compared to the commonly used nitroxide side chain, R1³. This labeling strategy only requires the post-expression addition of Cu^{2+} -IDA without the need of additional covalent modification and purification procedures. In addition, data acquisition has been reported in as little as 1 hour at X-band for a ~ 300 μM of the B1 immunoglobulin-binding domain of protein G (GB1)⁹¹.

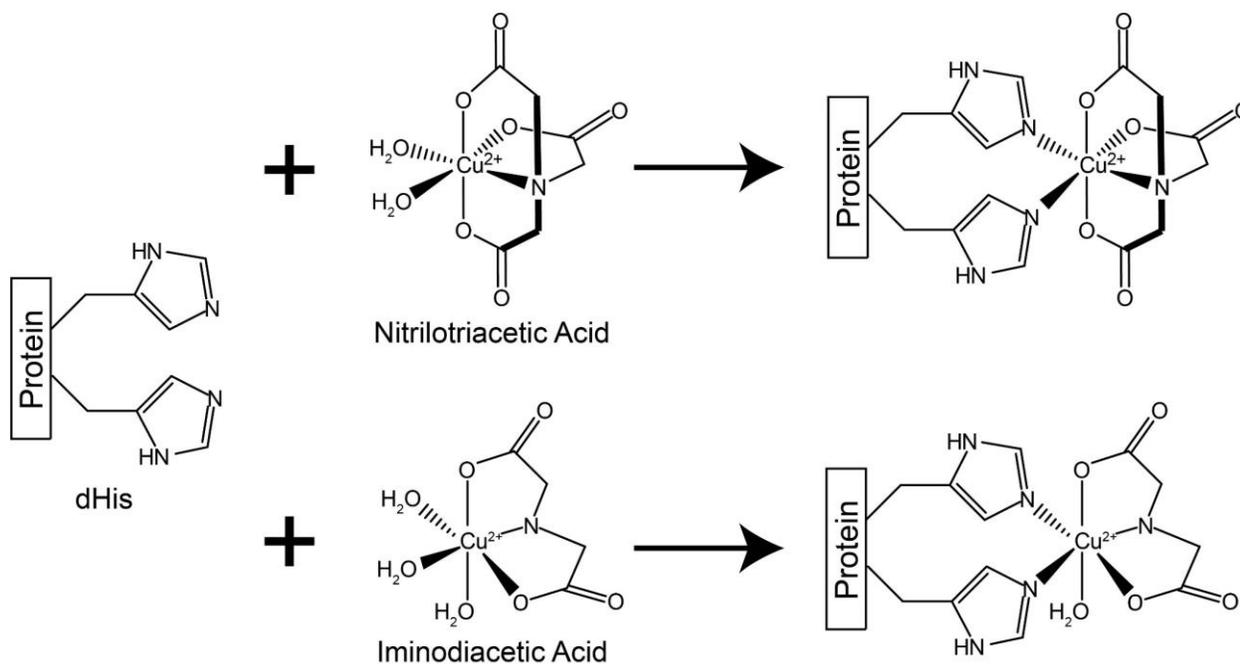


Figure 4-1 The dHis motif and the two Cu²⁺-complexes, NTA and IDA (top and bottom respectively) that create the dHis-Cu²⁺-NTA spin label (top) and the dHis-Cu²⁺-IDA spin label (bottom). NTA is a tetradentate chelating agent which includes an additional coordinating ligand compared to the tridentate IDA.

The first dHis-IDA measurement suffered from shallow dipolar modulations in the DEER time domain signal⁴⁷. The four-pulse DEER experiment probes the interspin distance between two paramagnetic centers by measuring the resultant dipolar frequency between them. This measurement generates an oscillating signal and the deeper these modulations, the easier it is to separate the dipolar frequency apart from noise and artifacts^{15,46,186,239}. Recently, the amplitude of these dipolar modulations were increased by a factor of two, effectively doubling the signal to noise ratio⁹¹. This increase was achieved by a thorough analysis of Cu²⁺-IDA binding to α -helical and β -strand dHis sites using various forms of spectroscopy. Through this analysis, the optimum number of equivalents needed to load a dHis site were obtained, which increased the total dHis-Cu²⁺-IDA present while minimizing free Cu²⁺-IDA in solution. It was determined that there were still two major hindrances of this Cu²⁺-IDA binding motif: 1) poor α -helical selectivity (two orders

of magnitude lower than that of a β -sheet dHis site) and 2) incomplete Cu^{2+} -IDA complexation in the required buffer conditions⁹¹. Thus, there is still room for improvement of this promising spin-labeling technique, especially for proteins containing α -helices. The degree of complexation and dHis site selectivity can depend on the choice of chelating agent for Cu^{2+} . Nitrilotriacetic acid (NTA) is an aminopolycarboxylic acid which has been extensively used for biological, medical and environment purposes^{240,241}. Metal-NTA complexes, particularly Ni^{2+} -NTA, have been widely used to bind to proteins with natural histidine residues, as well as attached His tags, for immobilization and purification purposes²⁴². NTA is a tetradentate chelating group which occupies four of the six sites in the Cu^{2+} coordination sphere, hence strongly binding the metal ion. Figure 4-1 shows the Cu^{2+} -NTA coordination. IDA, being a tridentate ligand, does not bind to the metal ion as effectively²⁴³. NTA binds to Cu^{2+} , occupying two axial and two equatorial coordination sites of the Cu^{2+} ²⁴⁴. This leaves the Cu^{2+} with exactly two equatorial sites available for cis-coordination with the dHis motif. The formation constant of the Cu^{2+} -IDA is $3.5 * 10^{10} \text{ M}^{-1}$ while that of Cu^{2+} -NTA is $1.3 * 10^{13} \text{ M}^{-1}$, under similar conditions of pH, temperature and ionic strength²⁴⁵. These formation constants indicate a higher complexation between Cu^{2+} and NTA as compared to IDA.

Herein, we explore NTA as an alternative chelating ligand for binding with Cu^{2+} . We see, by ESR, that we can effectively achieve 100% Cu^{2+} -NTA coordination. The complete formation of the Cu^{2+} -NTA complex allows for ratios of Cu^{2+} -NTA to protein that optimize loading of dHis sites and reduces unbound Cu^{2+} . Together, complete complexation improves the strength of the dipolar modulations in the DEER signal. We, also, determine the binding affinity of the Cu^{2+} -NTA complex with dHis sites within an α -helix, mid- β -sheet and edge- β -sheet. We report that the Cu^{2+} -NTA complex shows significantly higher selectivity towards α -helical dHis sites compared to Cu^{2+} -IDA, further improving the strength of DEER dipolar modulations. To validate this higher

binding affinity, we perform DEER on a protein with two α -helical dHis sites which showed a significant improvement in the modulation depth.

4.2 MATERIALS AND METHODS

4.2.1 Protein expression, purification and labeling

The proteins used are the B1 immunoglobulin-binding domain of protein G (GB1) and human glutathione S-transferase A1-1 (huGSTA1-1). The latter protein is homodimeric, thus one dHis site provides two Cu^{2+} ions. The mutation, expression and purification of the 28H/32H, 6H/8H, 6H/8H/28H/32H and 15H/17H/28H/32H constructs of GB1 and 211H/215H construct of huGSTA1-1 were performed as previously described^{47,246,247}. Cu^{2+} -NTA was prepared in a 1:1 ratio as per published protocols²⁴⁸. The pH of the final Cu^{2+} -NTA stock solution was adjusted to 7.4. All samples were prepared in 50 mM N-ethylmorpholine (NEM) buffer at a pH of 7.4 to eliminate any signal that can originate from free Cu^{2+} ⁷⁸. The protein concentration ranged from 270 – 400 μM .

4.2.2 ESR measurements

All sample volumes were 120 μL . The protein samples were prepared in 50 mM NEM buffer, containing 20% v/v glycerol. The pH was 7.4. The glycerol was added to the sample as a cryoprotectant. All ESR experiments were performed on either a Bruker ElexSys E680 X-band FT/CW spectrometer equipped with a Bruker EN4118X-MD4 resonator or a Bruker ElexSys 580

X-band FT/CW spectrometer equipped with a Bruker ER4118X-MD5 resonator. The temperature for all experiments was controlled using an Oxford ITC503 temperature controller with an Oxford ER 4118CF gas flow cryostat.

Continuous wave (CW) ESR experiments were carried out at ~9.68 GHz (X-band) at 80 K. Data were collected for 1024 points over a sweep width of 2000 G. All CW-ESR data used a modulation amplitude of 4 G, a modulation frequency of 100 kHz, a time constant of 10.24 ms, a conversion time of 20.48 ms and an incident power of 0.1992 mw. All CW-ESR spectra were simulated and fit using the Bruker Simfonia software.

Three-pulse electron-spin echo envelope modulation (ESEEM) experiments were performed at X-band frequencies at 80 K. A $\pi/2 - \tau - \pi/2 - T - \pi/2$ - echo pulse sequence with a $\pi/2$ pulse length of 16 ns was used. The first time delay, τ , was set to 144 ns and the second time delay, T , was set to 288 ns and was incremented by a step size of 16 ns. All experiments were performed at the magnetic field corresponding to the greatest intensity on the echo-detected field swept spectrum. A four-step phase cycling eliminated all unwanted echoes^{199,249}. Data acquisition took ~12 hours. The resultant signal was phase corrected, baseline subtracted and Fourier-transformed using the Bruker WinEPR software.

Double electron electron resonance (DEER) experiments were performed at 20 K at X-band frequencies. The four pulse sequence used was: $(\pi/2)v_1 - \tau_1 - (\pi)v_1 - (\tau_1 + t) - (\pi)v_2 - (\tau_1 + \tau_2 - t) - (\pi)v_1 - \tau_2 - \text{echo}$ ¹⁵⁴. For all the GB1 samples, the observer pulses, $(\pi/2)v_1$ and $(\pi)v_1$, were 6 ns and 12 ns, respectively and the pump pulse, $(\pi)v_2$, was 12 ns. The delay, t , was incremented using a step size of 10 ns for a total of 128 points. For the huGSTA1-1 sample, observer pulses, $(\pi/2)v_1$ and $(\pi)v_1$ were 16 ns and 32 ns, respectively and pump pulse was 16 ns. The delay, t , was incremented using a step size of 24 ns for a total of 128 points. The pump frequency, ν_2 , was placed

at the maximum of the echo detected field swept spectrum with the observer frequency, ν_1 , offset 150 MHz downfield. The raw time domain DEER data were analyzed by DeerAnalysis2015b¹⁰³ via Tikhonov regularization. Distance distribution was corrected using proper g-values²¹². Data acquisition took ~ 1 hour for the GB1 samples and 48 hours for the huGSTA1-1 sample.

4.2.3 Analysis of the modulation depth of DEER

DEER distance measurements were performed on the tetramutants 6H/8H/28H/32H and 15H/17H/28H/32H of GB1, which position a dHis site within both α -helix and β -sheet of GB1, and on the 211H/215H mutant of dimeric huGSTA1-1. Circular Dichroism (CD), DEER and crystallography have previously shown that these mutations do not alter the native structure of the protein⁴⁷ and that the dHis mutant of huGSTA1-1 is fully active.

For a system with N identical spins per molecule, the intramolecular DEER echo signal, $V_{\text{intra}}(t)$, also known as $V_p(t)$, is given as²⁵⁰:

$$V_p(t) = \frac{1}{N} \sum_{A=1}^N \left\langle \prod_{\substack{B=1 \\ B \neq A}}^N \left(1 - p_b (1 - \cos(\omega_{AB}t)) \right) \right\rangle_{r,\theta} \quad (4-1)$$

where t is the time the pump pulse is applied, A and B are two different spin systems, ω_{AB} is the dipolar interaction between the spins, p_b is the fraction of spin B excited by the pump pulse and N is the number of spins per molecule.

When $t \gg \omega_{AB}$, that is, $t \rightarrow \infty$, the dipolar term $\cos(\omega_{AB}t) = 0$. $V_p(t \rightarrow \infty)$ is thus where the oscillation due to the dipolar frequency dampens out. Equation 4-1 then simplifies further to²⁵¹⁻²⁵³:

$$V_p(t \rightarrow \infty) = (1 - p_b)^{N-1} = 1 - \lambda \quad (4-2)$$

where λ is the modulation depth parameter. For a mixture where molecules with different number of spins are present, the DEER signal will be a contribution from each of those species. Hence, equation 4-2 becomes²⁵²:

$$V_p(t \rightarrow \infty) = \sum_{i=1}^j f_i V_{pi}(t \rightarrow \infty) \quad (4-3)$$

where f_i is the fraction of each species present in solution and j is the number of different species present in solution.

Substituting equation 4-2 in equation 4-3, we get:

$$V_p(t \rightarrow \infty) = \sum_{i=1}^j f_i (1 - p_b)^{N_i-1} \quad (4-4)$$

For both tetramutant GB1 samples, there can be three different components in solution. These are: 1) a system with both α -helical and β -sheet dHis sites occupied, 2) a system with either α -helical or β -sheet dHis site loaded, 3) free Cu^{2+} -NTA as a mononuclear species. Only mononuclear species of Cu^{2+} -NTA is formed²⁴⁸, eliminating out the possibility of formation of any bis-complex. Furthermore, the NEM buffer eliminates free Cu^{2+} in the solution⁷⁸.

Thus, for our system with three different species, equation 4-4 becomes:

$$V_p(t \rightarrow \infty) = [f_{dHis2}(1 - p_b)^1 + f_{dHis1} + f_{dHis0}] \quad (4-5)$$

where f_{dHis2} is the fraction of Cu^{2+} -NTA bound to both the α -helix and the β -sheet dHis sites of the tetramutant protein making the protein doubly loaded, f_{dHis1} is the fraction of Cu^{2+} -NTA bound to either the α -helix or the β -sheet dHis site of the tetramutant protein making the protein singly loaded and f_{dHis0} is the fraction of Cu^{2+} -NTA present in solution as a mononuclear complex⁹¹.

Equating equations 4-2 and 4-5, we get:

$$\lambda = 1 - [f_{dHis2}(1 - p_b)^1 + f_{dHis1} + f_{dHis0}] \quad (4-6)$$

$$f_{dHis2} + f_{dHis1} + f_{dHis0} = 1 \quad (4-7)$$

To analyze the determined modulation depths from the DEER experiment, we followed the protocol of Lawless et al⁹¹. From CW-ESR spectra, we are able to calculate the probabilities of the occurrence of each of these fractions:

$$f_{dHis2} = \left[\frac{P_{\alpha} * P_{\beta}}{(P_{\alpha} * P_{\beta} + (P_{\alpha} P_{N\beta} + P_{\beta} P_{N\alpha}))} \right] * P_{sig} \quad (4-8)$$

$$f_{dHis0} = 1 - (f_{dHis1} + f_{dHis2}) \quad (4-9)$$

$$P_{sig} = 1 - f_{dHis0} \quad (4-10)$$

where P_{α} and P_{β} are the probabilities of the Cu^{2+} -NTA to bind to the dHis site in the α -helix and β -sheet, respectively. Both P_{α} and P_{β} can be determined from the CW-ESR spectrum. The total concentration of bound Cu^{2+} -NTA can be determined by double integration of the CW-ESR spectrum which is divided by the known dHis site concentration to give P_{α} or P_{β} . Similarly, $P_{N\alpha}$ and $P_{N\beta}$ are the probabilities of the Cu^{2+} -NTA of not binding to the α -helix and β -sheet, respectively. P_{sig} is the fraction of bound Cu^{2+} -NTA, contributing to the DEER signal. Also,

$$P_{\alpha} + P_{N\alpha} = 1, P_{\beta} + P_{N\beta} = 1 \quad (4-11)$$

4.3 RESULTS AND DISCUSSION

To characterize the optimum loading conditions of the Cu^{2+} -NTA complex to dHis sites within both α -helices and β -sheets, three mutant constructs of the immunoglobulin binding domain of protein G (GB1) were prepared. GB1 is useful due to its structural characterization by both NMR and X-ray crystallography²⁵⁴⁻²⁵⁸ and, in general, due to our vast experience with the system^{47,48,91,246,247}. The construct 28H/32H-GB1 contains a dHis site within an α -helix. The 6H/8H construct is a mid-strand β -sheet site while 15H/17H is a non-hydrogen bonded edge-strand

β -sheet site. Further, to expand this methodology to other proteins, we used a double mutant, 211H/215H human glutathione S-transferase A1-1 (huGSTA1-1) protein. huGSTA1-1 is a homodimer, so the double mutant provides two sites for Cu^{2+} -labeling, one in each subunit. The dHis motif in the huGSTA1-1 is present on an α -helix.

4.3.1 CW-ESR spectroscopy to show complexation between Cu^{2+} and NTA

We first performed CW-ESR to show that Cu^{2+} in the presence of NTA remains entirely in the complex state. Figure 4-2 shows that Cu^{2+} , when complexed with NTA, show a significantly different spectrum than free Cu^{2+} in water due to change in the direct coordination environment of Cu^{2+} . The g_{\parallel} and A_{\parallel} values changed from 2.415 and 121 G respectively for CuCl_2 to 2.327 and 152 G for Cu^{2+} -NTA. Further, no free Cu^{2+} component is visible in the Cu^{2+} -NTA spectrum indicating that all Cu^{2+} has complexed with the NTA. Previous work done on IDA showed incomplete complexation between Cu^{2+} and IDA which consequently lead to poor modulation depth⁹¹. This result is anticipated because the formation constant of Cu^{2+} -NTA is 3 orders of magnitude greater than Cu^{2+} -IDA²⁴³.

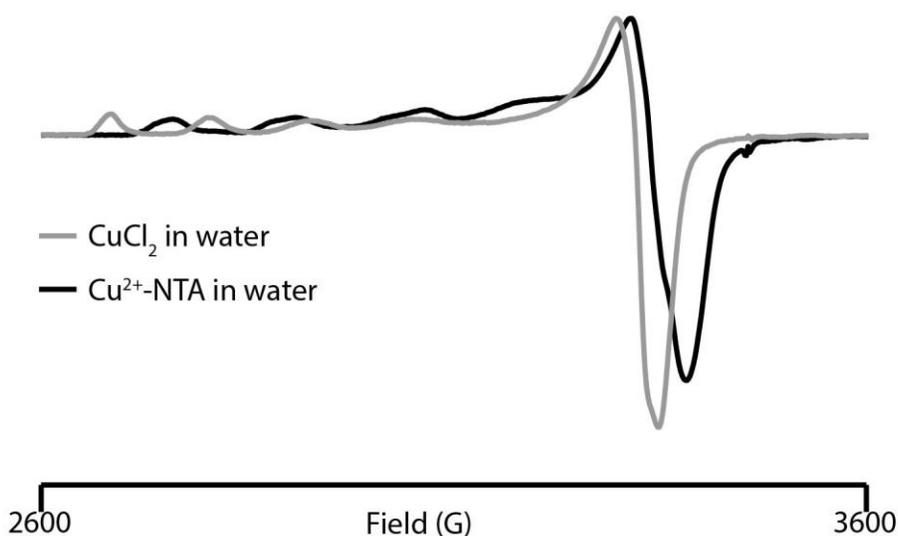


Figure 4-2 The spectrum of CuCl_2 in water (gray solid line) is markedly different from that of Cu^{2+} -NTA in water. Cu^{2+} -NTA in water shows no free CuCl_2 component, thereby indicating complete complexation between Cu^{2+} and NTA. Experiments were performed at X-band.

4.3.2 Binding affinity of Cu^{2+} -NTA to dHis site using CW-ESR spectroscopy

We next performed CW-ESR to examine the binding affinities of Cu^{2+} -NTA to each of the dHis sites within both α -helix and β -sheet. Figure 4-3A shows the CW-ESR spectrum of Cu^{2+} -NTA bound to an α -helical dHis site, 28H/32H with a dHis: Cu^{2+} -NTA = 1:0.25. Simulations of the spectrum showed a single component fit, implying all Cu^{2+} -NTA is bound to the protein. Figure 4-3B shows the CW-ESR spectrum with an increased dHis: Cu^{2+} -NTA = 1:2. The spectrum is comprised of two components. The g_{\parallel} and A_{\parallel} values for Component 1 are consistent with that obtained from the CW-ESR spectrum of Cu^{2+} -NTA in NEM buffer. Hence, Component 1 can be attributed to free Cu^{2+} -NTA in NEM buffer. Component 2 is Cu^{2+} -NTA bound to the α -helical dHis site, consistent with the CW-ESR spectrum of the Cu^{2+} -NTA:dHis at 1:0.25 in NEM buffer. We varied the ratio of the two components and added them together to obtain the best fit for all

data sets. The inset in Figure 4-3B shows the CW-ESR spectra of the two components that were used to simulate one Cu²⁺-NTA:dHis (1:2) spectrum. By altering the ratio of the two components, we obtained the best fit for each CW spectra in the titration at different Cu²⁺-NTA equivalents. From the ratio, we could determine the percentage of Cu²⁺-NTA bound to the protein.

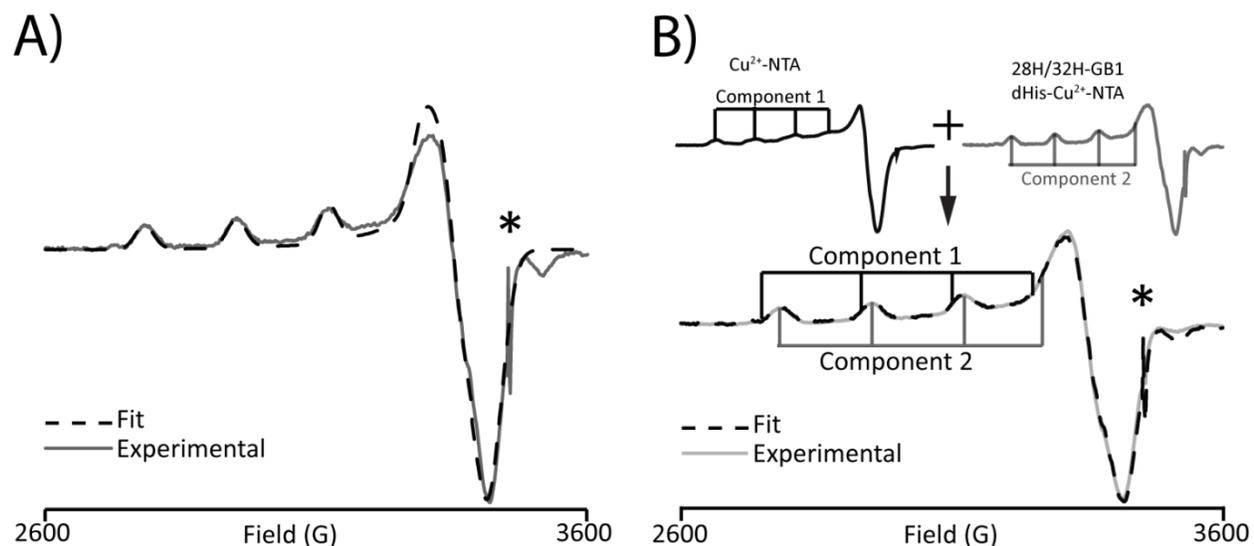


Figure 4-3 A) CW-ESR spectrum of Cu²⁺-NTA bound to the α -helical dHis site in GB1, where dHis:Cu²⁺-NTA = 1:0.25. The spectrum shows a single component fit indicating that Cu²⁺-NTA is completely bound to the protein. B) CW-ESR spectrum of Cu²⁺-NTA bound to the α -helical dHis site in GB1, where dHis:Cu²⁺-NTA = 1:2. This spectrum has contributions from both Cu²⁺-NTA bound to the dHis as well as free Cu²⁺-NTA in solution. Inset shows the two components used to simulate the spectrum. The first component is the free Cu²⁺-NTA, consistent with the CW-ESR spectrum of Cu²⁺-NTA in buffer. The second component is the protein bound Cu²⁺-NTA, consistent with spectrum of dHis:Cu²⁺-NTA = 1:0.25. Together, the two components were varied at different ratios to simulate the CW-ESR spectrum of Cu²⁺-NTA bound α -helical dHis site at 2 equivalents. Signal from quartz capillary tube denoted by *. All experiments were performed at X-band.

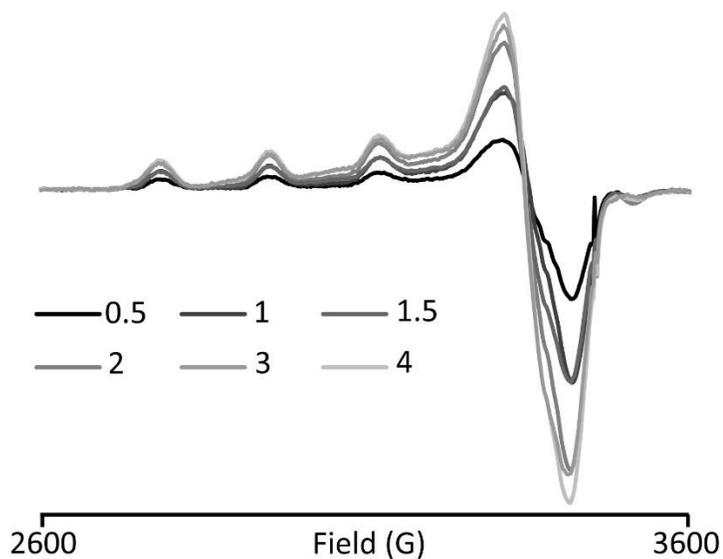


Figure 4-4 CW-ESR titrations for 28H/32H GB1 at different equivalents of Cu²⁺-NTA. The protein concentration is 400 μM.

Figure 4-4 shows the ESR data and simulations for different equivalents of Cu²⁺-NTA to protein. From each CW-ESR spectrum in the titration, we calculated the total concentration of Cu²⁺-NTA present in solution. The doubly integrated intensity of the CW-ESR spectrum is directly proportional to the number of spins present²⁵⁹. After calculating the doubly integrated intensity of each spectrum of the titration, comparison to a standard calibration curve yielded the total Cu²⁺-NTA concentration. Since we obtained the percentage of bound versus unbound Cu²⁺-NTA from the simulations, we could, determine the concentration of Cu²⁺-NTA bound to the respective dHis site. Knowing the protein concentration, we generated a plot of bound Cu²⁺-NTA per protein versus the Cu²⁺-NTA equivalents, shown in Figure 4-5A. We then used these data to calculate the apparent dissociation constant, K_d , of Cu²⁺-NTA to 28H/32H. The value of K_d was obtained using equation⁵⁷:

$$[B] = \frac{1}{2} \left[\left(((n + 1) * [P]) + K_d \right) - \sqrt{\left(((n + 1) * [P]) + K_d \right)^2 - (4 * n * [P]^2)} \right] \quad (4-12)$$

where, $[B]$ is the bound Cu^{2+} -NTA concentration, $[P]$ is the total concentration of the Cu^{2+} -NTA binding sites (dHis), n is the number of Cu^{2+} -NTA equivalents and K_d is the apparent dissociation constant. As shown in Figure 4-5A, K_d of the Cu^{2+} -NTA to the α -helical dHis site, 28H/32H is 285 μM . Finally, Figure 4-5B shows the amount of bound Cu^{2+} -NTA per unbound Cu^{2+} -NTA at different ratios of dHis: Cu^{2+} -NTA.

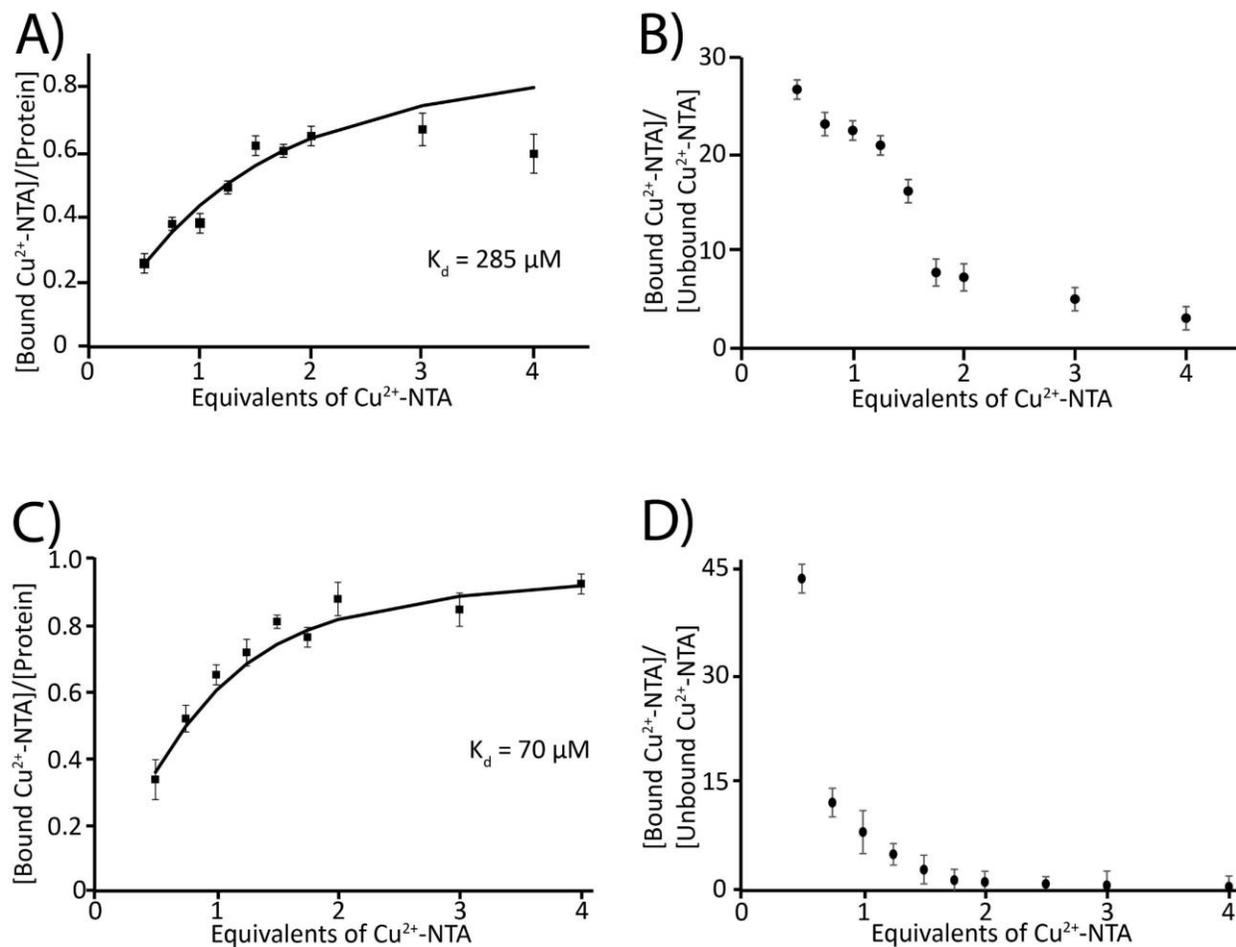


Figure 4-5 A) Plot of bound Cu^{2+} -NTA per protein versus Cu^{2+} -NTA equivalents gives the apparent binding affinity of Cu^{2+} -NTA to the dHis site in the α -helix, 28H/32H ($K_d = 285 \mu\text{M}$) in GB1. B) Plot showing the amount of bound Cu^{2+} -NTA per unbound Cu^{2+} -NTA at different Cu^{2+} -NTA equivalents for 28H/32H-GB1. C) Plot showing the apparent binding affinity of Cu^{2+} -NTA to the β -sheet, 6H/8H ($K_d = 70 \mu\text{M}$). D) Plot showing the amount of bound Cu^{2+} -NTA per unbound Cu^{2+} -NTA at different Cu^{2+} -NTA equivalents for 6H/8H.

Figure 4-5A shows that at 1.5 equivalents of Cu^{2+} -NTA the protein is ~60% loaded and the loading efficiency starts reaching its maximum hereafter. Higher equivalents than 1.5 increases the percentage of protein loading only slightly. At the same time, beyond 1.5 equivalents, there is a sharp increase in the unbound Cu^{2+} -NTA, as shown in Figure 4-5B. Equivalents of Cu^{2+} -NTA lower than 1.5 have higher ratios of Cu^{2+} -NTA bound to dHis but the percentage of protein loaded is less than 50%. Thus, considering both the factors of percentage of protein loaded and ratio of bound and unbound Cu^{2+} -NTA, 1.5 equivalents is the optimum ratio for the α -helix dHis.

The α -helical dHis site 28H/32H shows much promise with the Cu^{2+} -NTA with the K_d being almost four times lower than that of Cu^{2+} -IDA. The K_d decreased from 1210 μM for Cu^{2+} -IDA⁹¹ to 285 μM for Cu^{2+} -NTA. This increased affinity of Cu^{2+} -NTA for the α -helical dHis site is highly significant as this labeling methodology can be applied to proteins which consist of only α -helices as their primary structure. The apparent dissociation constant can vary based on a number of factors such as temperature, pH, and buffer. The apparent dissociation constants of Cu^{2+} to an α -helical dHis site ranges from 2 to 200 μM ²⁶⁰. Also, one can saturate the binding sites by overloading the protein and then filtering out the excess Cu^{2+} -NTA complex.

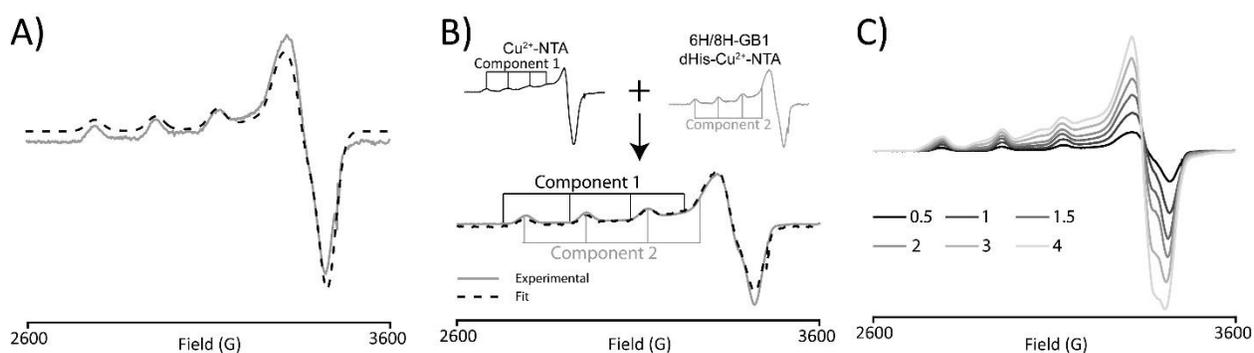


Figure 4-6 A) The CW-ESR spectrum of 6H/8H-GB1 (grey line) containing a 1:0.25 ratio of dHis: Cu^{2+} -NTA. Simulation shows a single component fit, indicating that Cu^{2+} -NTA is entirely bound to the protein. B) CW-ESR spectrum of 6H/8H-GB1 at a dHis: Cu^{2+} -NTA ratio of 1:2 (grey line) with its respective simulation (black dashed). The simulation was generated by the addition of two individual spectra: 1) free Cu^{2+} NTA (inset, black) and 2) Cu^{2+} -

NTA bound to its respective α -helical or β -sheet dHis site (inset, grey). We call the free Cu^{2+} -NTA Component 1 and the bound Cu^{2+} -NTA Component 2 as seen in the inset. C) CW-ESR titration of 6H/8H-GB1 at different equivalents of Cu^{2+} -NTA. The protein concentration is 400 μM .

For the 6H/8H β -sheet site, similar experiments were performed to determine binding affinity as well as optimum loading ratios for distance measurements. The CW-ESR data and analysis for the 6H/8H is shown in Figure 4-6. The apparent K_d of Cu^{2+} -NTA for a β -sheet dHis site, 6H/8H is 70 μM , as can be seen from Figure 4-5C. Figure 4-5D shows the plot of $[\text{Bound } \text{Cu}^{2+}\text{-NTA}]/[\text{Unbound } \text{Cu}^{2+}\text{-NTA}]$ at different equivalents of Cu^{2+} -NTA. For the β -sheet, 1.5 equivalent is the optimum ratio used for loading a dHis site. At this equivalent, around ~80% of the protein is loaded.

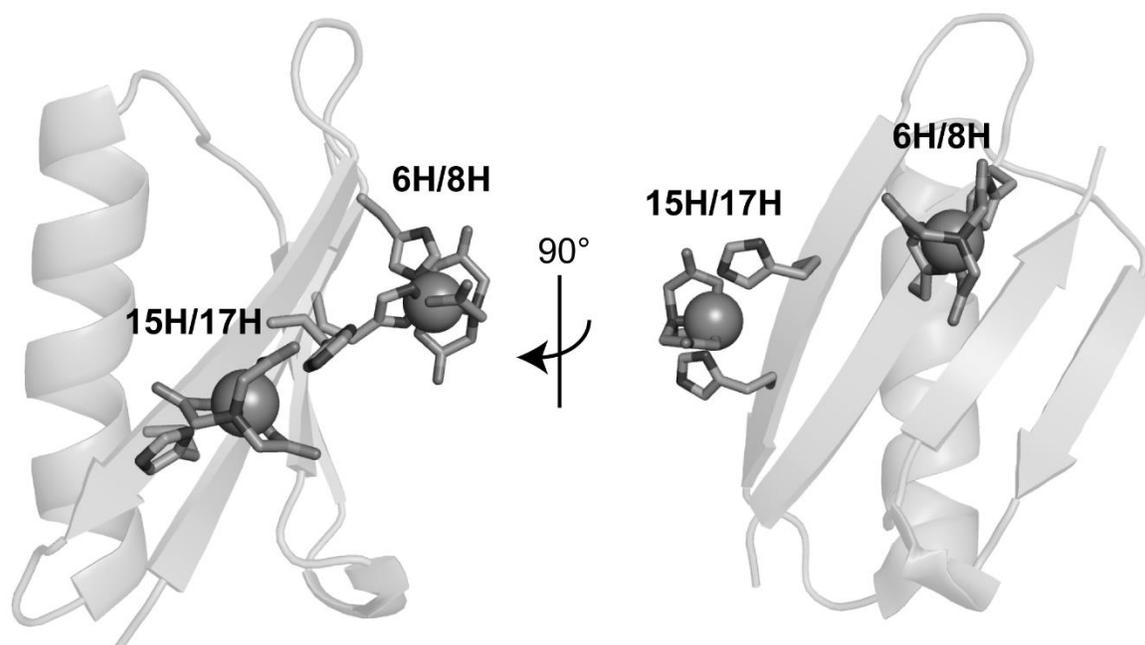


Figure 4-7 Both β -sheet dHis sites used; 6H/8H which positions the dHis binding site within a typical β -sheet location and 15H/17H which positions the dHis site within an edge-strand β -sheet site. The edge-strand has a non-hydrogen binding face leading to an orientation toward the sheet. This possibly increases steric interactions, lowering the binding affinity.

Previous work done on dHis sites using Cu²⁺-IDA has shown the binding affinity of Cu²⁺-IDA to be relatively high for a β -sheet 6H/8H, with a K_d value of 30 μ M⁹¹. In comparison, Cu²⁺-NTA shows a slightly lower binding affinity to the β -sheet with a K_d of 70 μ M. Being a tetradentate ligand as compared to the tridentate IDA, the NTA is bulkier. The dHis site in a β -sheet has the histidine residues at a separation of *i* and *i*+2. The NTA may experience some steric interaction with the histidine residues, as they are in close vicinity. This possible steric hindrance could cause the dHis site of the β -sheet to show less affinity to the Cu²⁺-NTA than Cu²⁺-IDA, thereby increasing the K_d.

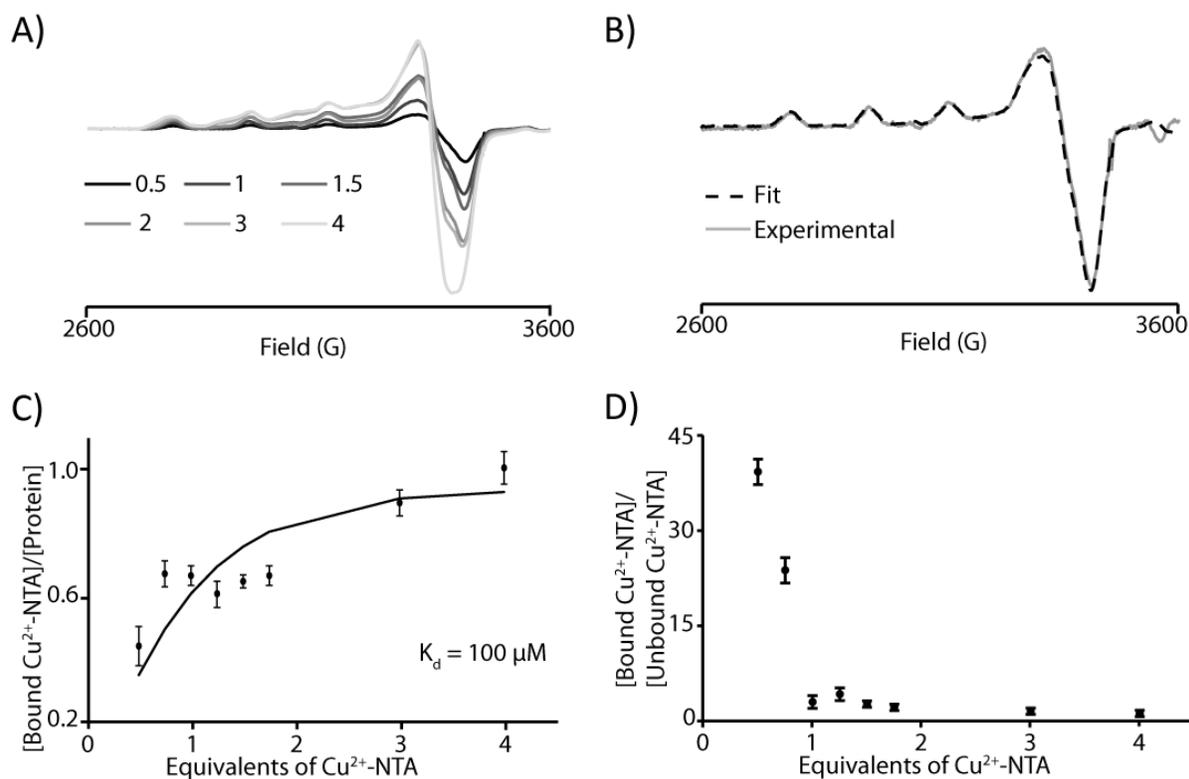


Figure 4-8 CW-ESR data for the 15H/17H-GB1 construct which places the dHis Cu²⁺-binding site with an edge-strand β -sheet. A) CW-ESR spectra of 15H/17H-GB1 with varying equivalents of Cu²⁺-NTA incrementally added. B) Example of two component simulation of 15H/17H-GB1 CW-ESR spectrum. C) Plot used to determine the apparent dissociation constant, K_d = 100 μ M. D) Plot illustrating the ratio of bound Cu²⁺-NTA per unbound Cu²⁺-NTA in

solution as different equivalents of Cu²⁺-NTA are added. There is a large increase in unbound Cu²⁺-NTA at 1 equivalent. The protein concentration is 270 μM.

To further examine the applicability of the motif, a Cu²⁺-binding dHis site was incorporated into a non-hydrogen bonded edge-strand β-sheet site using the 15H/17H-GB1 construct. Similar CW-ESR titrations performed using various equivalents of Cu²⁺-NTA showed a K_d of 100 μM (Figure 4-8). Non-hydrogen bonded residues on edge-strand sites have a tendency to orient toward the sheet, increasing possibility of steric interactions influencing rotameric behavior of the His sidechains²⁴⁷. This behavior possibly inhibits dHis-Cu²⁺-NTA formation.

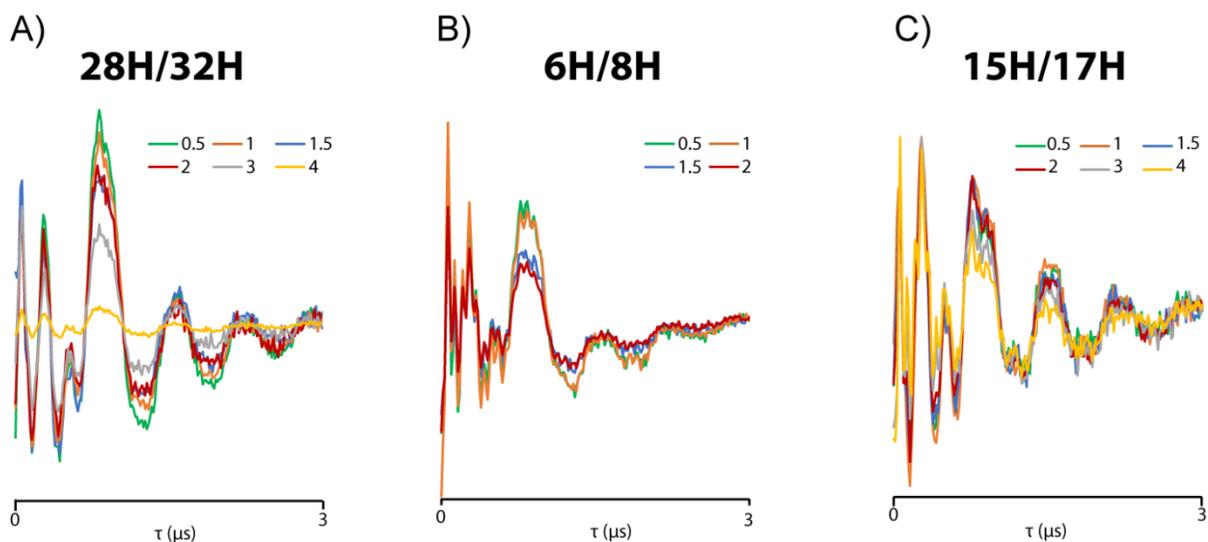


Figure 4-9 Background corrected ESEEM signal of 28H/32H-GB1 (A), 6H/8H-GB1 (B) and 15H/17H-GB1 (C) with varying equivalents of Cu²⁺-NTA. As the amount of unbound Cu²⁺-NTA in solution increases, the ESEEM modulations associated with dHis-Cu²⁺-NTA coordination become diminished.

To further validate the CW-ESR titration results, we used the pulsed ESR Electron spin-echo envelope modulation (ESEEM) experiment (Figure 4-9). Cu²⁺ coordinates to each outer imidazole nitrogen of the two histidine residues which causes the Cu²⁺ to show the characteristic ESEEM signal for the distal nitrogen of the imidazole^{68,81,89,261}. The relative modulation depth parameter, k²⁶², was calculated for each of the ESEEM data using the second modulation period of the nitrogen modulations⁹⁰ for all α-helical and β-strand dHis sites. As the fraction of bound

Cu²⁺-NTA increases, the k value also increases due to increased coordination with distal imidazole nitrogen^{68,89} (Figure 4-10A). Once the fraction of unbound Cu²⁺-NTA starts increasing, modulations due to imidazole coordination become washed out and hence, the k value decreases. Thus, the k value is a direct reflection of the increase in amount of unbound Cu²⁺-NTA as a function of equivalents of Cu²⁺-NTA. ESEEM titrations support that 1.5 equivalents of Cu²⁺-NTA per dHis site will enable the strongest dipolar modulations from both α -helical and β -strand dHis sites (Figure 4-10).

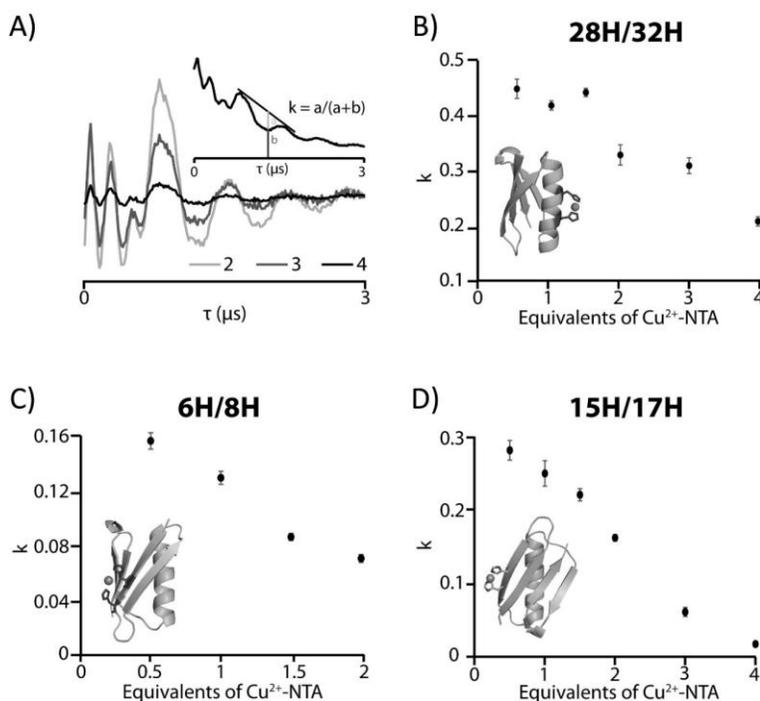


Figure 4-10 A) Background corrected ESEEM signal of 28H/32H-GB1 with varying equivalents of Cu²⁺-NTA illustrating the decrease in modulation size with increasing amounts of Cu²⁺-NTA. The raw time ESEEM signal is shown in inset along with an example of how modulation depth, k, is measured. B) Plot of k as a function equivalents of Cu²⁺-NTA added for the α -helical dHis site. A sharp decrease is observed after 1.5 equivalents of Cu²⁺-NTA are added. C) Plot showing linear decrease in k upon addition of Cu²⁺-NTA within β -sheet dHis site. Inset shows 6H8H-GB1. D) k versus equivalents of Cu²⁺-NTA added for the edge-strand β -sheet dHis site 15H/17H-GB1 (shown in inset).

4.3.3 Cu²⁺-NTA further improves dHis-based DEER signal

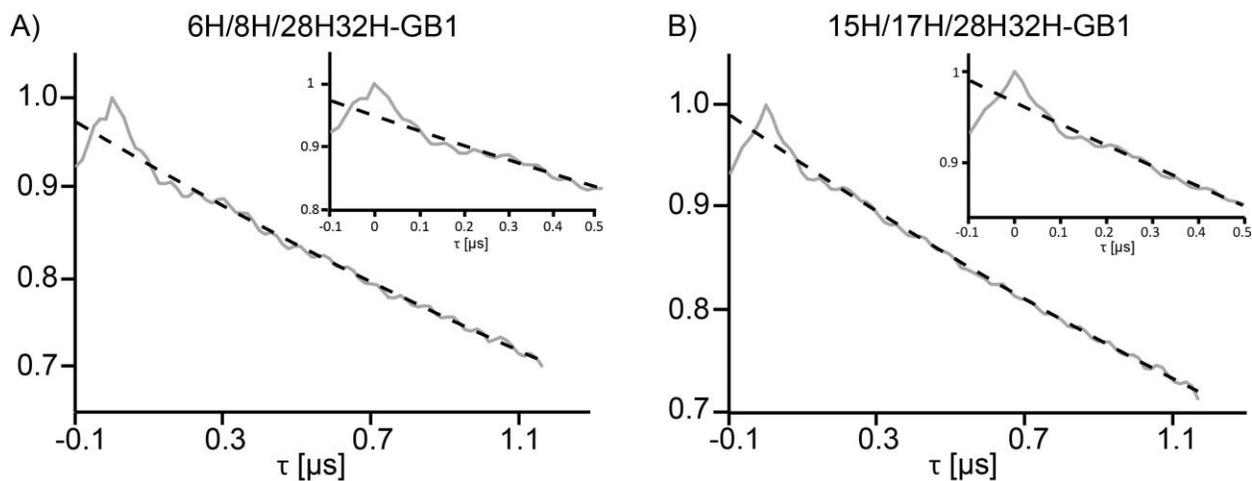


Figure 4-11 Raw time domain DEER data for 6H/8H/28H/32H-GB1 (top) and 15H/17H/28H/32H-GB1 (bottom). Inset shows the clear modulations for each time domain data. The background exponential decay fit is shown as the dashed line.

The GB1 tetramutants 6H/8H/28H/32H and 15H/17H/28H/32H, which each has a dHis site in both an α -helix and a β -sheet, was used for DEER measurements and subsequent modulation depth analysis (Figure 4-12A). As shown previously using CD, DEER and crystallography, the tetramutant retains the native structure⁴⁷.

The DEER experiment on the GB1 tetramutant construct 6H/8H/28H/32H with the Cu²⁺-NTA complex shows improved modulation depth compared to the Cu²⁺-IDA complex. Figure 4-11A shows the raw DEER data. The background subtracted DEER signal is shown in Figure 4-12B. The modulation depth was obtained to be $4.8 \pm 0.5\%$.

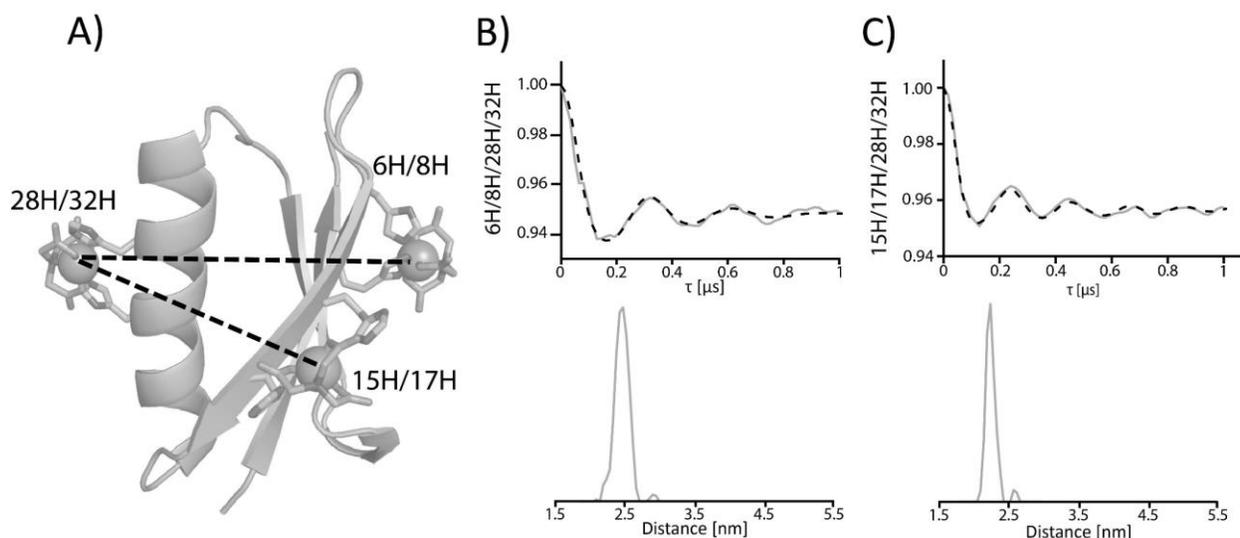


Figure 4-12 A) Representation of 6H/8H/28H/32H-GB1 and 15H/17H/28H/32H-GB1 distances (PDB: 4WH4)⁴⁷ B) Background subtracted time domain DEER data for both 6H/8H/28H/32H-GB1 (top) and resultant distance distribution (bottom). C) Background subtracted time domain DEER data of 15H/17H/28H/32H-GB1(top) with respective distance distribution (bottom). Dashed lines represent fit via Tikhonov regularization. Dashed lines in distance distribution represent predicted Cu²⁺-Cu²⁺ distances.

Using equations 4-4 to 4-11, we can account for this experimentally obtained modulation depth⁹¹. P_α and P_β are the fraction of Cu²⁺-NTA bound to the α -helical and β -sheet dHis site, respectively at 1.5 equivalents. From Figure 4-5A and Figure 4-5C, P_α and P_β are found to be 0.62 and 0.80, respectively. Accordingly, using equation 4-11, $P_{N\alpha}$ and $P_{N\beta}$ are 0.38 and 0.20, respectively. Using equations 4-8 to 4-11, we can calculate f_{dHis2} to be 0.44, f_{dHis1} is 0.39 and f_{dHis0} is 0.17. Furthermore, we can analyze the percentage of GB1 doubly labeled by the simple probability calculation = $P_\alpha * P_\beta$. Therefore, under these conditions, 50% of the GB1 was doubly labeled. Furthermore, the p_b value needed to solve for theoretical modulation depth can be calculated from the echo-detected field swept spectrum as previously described⁴⁶. The p_b value is calculated to be 0.12 for 6H/8H/28H/32H-GB1 using Cu²⁺-NTA with a pump pulse of 12 ns. Using the fraction values and p_b calculated above, we use equation 4-6 to determine the theoretical

modulation depth of $5.2 \pm 0.7\%$. This result is in good agreement with our experimentally obtained modulation depth of $4.8 \pm 0.5\%$.

The increase in modulation depth is significant as the sensitivity of the DEER signal depends on the modulation depth as^{15,186,220,239}:

$$S(DEER) \propto \lambda * V(0) * \exp\left(\frac{-2t_{max}}{T_M}\right) * \sqrt{T_1^{-1}} \quad (4-13)$$

where $S(DEER)$ is the sensitivity of the DEER experiment, λ is the modulation depth, $t_{max} = (\tau_1 + \tau_2)$, $V(0)$ is the echo intensity when $t_{max}=0$, T_M is the phase memory relaxation time and T_1 is the spin lattice relaxation time. Two factors were shown to hinder dHis based distance measurements by pulsed ESR: poor selectivity toward α -helical sites and incomplete complexation. As seen from Figure 4-2, the latter of those two has been solved by complete complexation of Cu^{2+} -NTA. Our results from Figure 4-5 indicate that the Cu^{2+} -NTA complex also has increased selectivity towards the α -helical dHis site. Taken together, we show in Figure 4-12, these two improvements have further increased the capabilities of the dHis motif.

To show that the method is feasible for another dHis site within the protein, DEER measurements were performed on the tetramutant 15H/17H/28H/32H. Figure 4-11B shows the raw time domain and Figure 4-12C shows the background subtracted time domain DEER data of 15H/17H/28H/32H-GB1. At 1.5 equivalents of Cu^{2+} -NTA per dHis site, the modulation depth was obtained to be $4.2 \pm 0.5\%$. In accordance with previous published dHis distance distributions, the breadth of the distribution is extremely narrow, and the most probable distance obtained is 2.2 nm, which agrees with the crystal structure²⁴⁷.

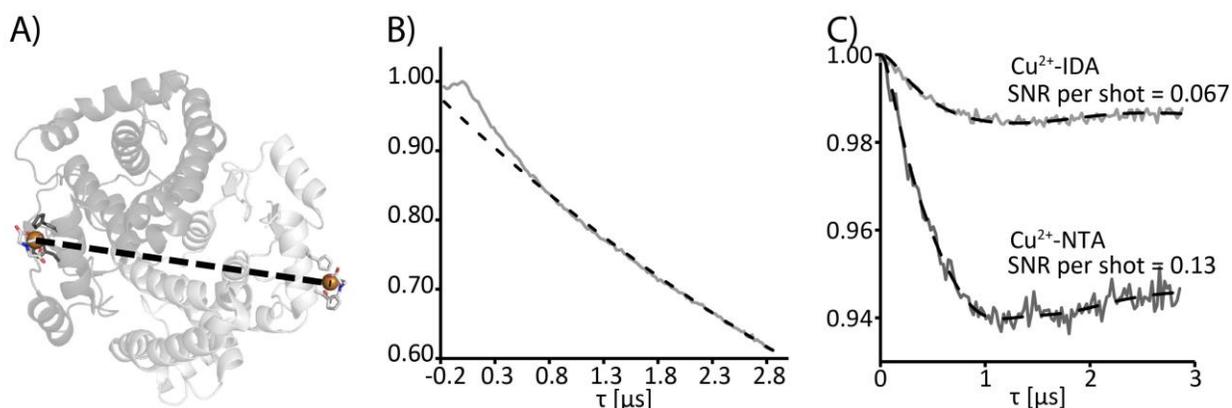


Figure 4-13 A) Crystal structure of human glutathione S-transferase A1-1 (PDB: 1K3L)²⁶³. dHis site was added in with selected rotamers from previous crystal structure (PDB: 4WH4)⁴⁷. B) Raw time domain DEER data for 211H/215H huGSTA1-1. The background exponential decay fit is shown as the dashed line. C) Background subtracted time domain DEER data showing the increased modulation depth of the dHis-Cu²⁺-NTA motif (2750 scans) compared to the dHis-Cu²⁺-IDA motif (4300 scans).

The above data suggests that modulation depth between two α -helical sites using Cu²⁺-NTA could be higher than 5%. To further probe the higher binding affinity of dHis-Cu²⁺-NTA in between two α -helical sites, we incorporated dHis into the α -helix rich huGSTA1-1. As shown in Figure 4-13A (PDB: 1K3L), we incorporated the dHis site into the C-terminal helix of each subunit within the dimeric enzyme²⁶³. On adding 1.5 equivalents of Cu²⁺-NTA per dHis site, the modulation depth was found to be almost three-fold higher when compared to previously obtained DEER data on the same protein but in presence of Cu²⁺-IDA (Figure 4-13B). At the same time, the signal to noise (S/N) ratio per shot increased from 0.06 within 4300 scans for Cu²⁺-IDA to 0.13 for 2750 scans of Cu²⁺-NTA (Figure 4-13C). The DEER data obtained from Cu²⁺-NTA did not require as strong S/N since the larger modulation depth enhanced the visibility of the dipolar modulations compared to the noise. While we saw a three-fold increase in λ , only two-fold increase in S/N per shot was observed. This discrepancy can be attributed to differing values of T_M and T_1 . T_M and T_1 changed from 8.4 μ s to 6.2 μ s and 340 μ s to 280 μ s, respectively, between Cu²⁺-IDA

and Cu²⁺-NTA samples. Using equation 4-13, this predicts only a two-fold increase in the S/N ratio per shot. This increase exemplifies the improvements achieved for inter- α -helical distance measurements using the Cu²⁺-NTA complex as opposed to Cu²⁺-IDA.

Taken together, these new findings for applications of the dHis motif with increased modulation depth present a promising future for accurate Cu²⁺-Cu²⁺ distance measurements in biological macromolecules. Typically, at X-band frequencies (~9.5 GHz) with a 16 ns pump pulse, Cu²⁺-based DEER measurements result in a 10-12% maximum modulation depth (5). However, using high-bandwidth loop-gap resonators and arbitrary wave generators, Cu²⁺-based DEER modulations depths have been reported as up to 40%²⁶⁴. The ability of the dHis motif to attach Cu²⁺ rigidly and site-specifically in both α -helices and β -sheets makes it an excellent labeling strategy for future use, especially with the improved modulation depths.

4.4 CONCLUSION

We have shown that the Cu²⁺-NTA complex has a higher selectivity toward α -helical dHis sites as compared to the Cu²⁺-IDA complex. The α -helix- α -helix dHis measurement in the human glutathione S-transferase using the NTA-complex resulted in a three-fold increase in modulation depth when compared to the IDA-complex. Even with the apparent lower selectivity toward β -sheet dHis binding sites, the DEER data showed an increased modulation depth for a system containing both α -helical and β -sheet dHis sites. This increase in DEER sensitivity allowed for acquisition of acceptable S/N ratios in as little as 1 hour for GB1. We also show that the dHis site is compatible in various β -sheet environments using the $i, i+2$ motif. Cu²⁺-NTA also exhibits total complexation thereby completely reducing free Cu²⁺ in solution and enabling precise additions of

complex equivalents per dHis sites. Furthermore, this complex exists only as a mononuclear species as per published protocols, eliminating possibilities of any bis-complex formation in solution. The dHis motif, in combination with these new findings and others show tremendous promise for precise Cu^{2+} - Cu^{2+} distance measurements.

4.5 ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation (NSF MCB-1613007).

5.0 Cu²⁺ ESR REVEALS TWO DISTINCT BINDING SITES AND OLIGOMERIZATION OF INNATE IMMUNE PROTEIN CALGRANULIN C

This work, written in collaboration with Velia Garcia, Kevin Singewald, Steven M. Damo and Sunil Saxena, has been published in Applied Magnetic Resonance, 2018, volume 49, pages 1299-1311. The protein Calgranulin C was expressed and purified by Velia Garcia and Steven C. Damo. The dissertation author collected and analyzed the ESR data and prepared the manuscript.

5.1 INTRODUCTION

Calgranulin C or S100A12 is a member of the S100 class of EF-hand calcium (Ca²⁺) binding proteins. There are ~20 S100 proteins which are expressed in a tissue and cell-specific type manner and act intracellularly to regulate a diverse array of biological functions including Ca²⁺ homeostasis, cell proliferation, and energy metabolism²⁶⁵. Additionally, a number of S100 proteins are exported from cells where they activate signaling mechanisms and act as antimicrobial agents^{266,267}. The biological relevance of S100 proteins is underscored by the misregulation of their expression in a number of diseases such as Alzheimer's disease, diabetes, cancer and inflammatory disorders²⁶⁸⁻²⁷⁰.

In addition to binding calcium, several S100 proteins bind divalent transition metals, most notably Zn²⁺. S100A7 and S100A15 bind zinc, and the S100A8/S100A9 heterodimer calprotectin binds zinc and manganese as part of the host innate immune response to starve pathogens, a mechanism termed nutritional immunity²⁷¹⁻²⁷⁵. S100A12 zinc sequestration has been implicated

in the control of *H. pylori* and *Campylobacter jejuni* infections^{276,277}. Additionally, Zn²⁺ binding to S100A12 and S100A7 has been implicated in activation of the receptor for advanced glycation end products, an inflammation receptor^{278,279}.

Compared to Zn²⁺, less is known about the role of Cu²⁺ in the function of S100 proteins. Copper is an important metal in the innate immune response. Copper is accumulated at sites of infection in the phagosomes of macrophages and can produce reactive oxygen species^{280,281}. The general mechanism of Cu²⁺ is based on its toxicity to bacterial pathogens²⁸². However, recently it was shown that calprotectin starves Cu²⁺ and Zn²⁺ as part of the innate immune response toward *Candida albicans*²⁸³. Several S100 proteins bind copper, including S100B, S100A13, S100A5, calprotectin, and S100A12²⁸³⁻²⁸⁸. Whether copper sequestration by S100 proteins is a general innate immune mechanism remains to be determined. Given the similarity between Zn²⁺ and Cu²⁺, the possibility that Cu²⁺ potentiates activation of inflammatory receptors is also feasible.

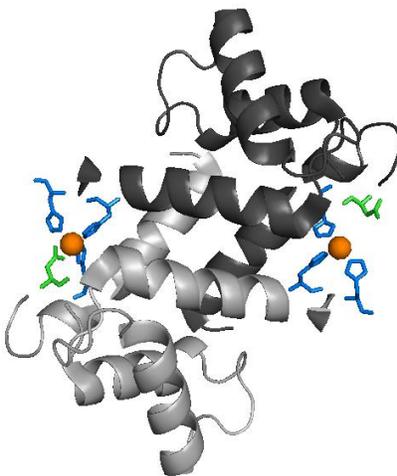


Figure 5-1 Dimer of S100A12 (PDB: 1ODB). The two subunits are shown in dark and light gray. The blue sticks denote the histidine residues while the green denotes aspartic acid. The orange spheres represent the Cu²⁺ ions. So far crystal structure (PDB: 1ODB)²⁸⁴ predicts that in a S100A12 dimer, each metal binding site is comprised of His15 and Asp25 of one subunit and His85 and His89 from the other. However, in solution, we see that the binding environment of the two Cu²⁺ binding sites are different.

S100A12 adopts the classic S100 protein fold, as shown in Figure 5-1. Two subunits of 92 amino acids, each comprised of two helix-loop-helix EF-hand motifs separated by a hinge region, are arranged in an antiparallel homodimer. Similar to the related proteins S100B, S100A7, S100A8 and S100A9, there are two transition metal binding sites located at opposite ends of the dimer interface. Each metal binding site is believed to comprise of two histidines from one subunit and a histidine and aspartic (or glutamic) acid in the other. These canonical transition metal binding sites are highly conserved in S100 proteins. In S100A12, each metal binding site is comprised of His15 and Asp25 of one subunit and His85 and His89 from the other, as shown in Figure 5-1. The X-ray crystal structure of the Ca^{2+} - Cu^{2+} -S100A12 has been determined to 2.19 Å resolution²⁸⁴. In comparison to the Ca^{2+} -S100A12 structure, there is little change to the overall tertiary structure upon copper binding. The most significant conformational change is the elongation of the C-terminal helix of Cu^{2+} - Ca^{2+} -S100A12 by four residues which includes the Cu^{2+} -chelating residue His89. In stark contrast, the quaternary structure of Cu^{2+} - Ca^{2+} -S100A12 is drastically different from apo-S100A12. Similar to Zn^{2+} and Ca^{2+} complexes of S100A12 which form dimers, tetramers, and hexamers,²⁸⁹⁻²⁹¹, Cu^{2+} induces oligomerization of Ca^{2+} -S100A12 into a trimer of dimers. It has been proposed that oligomeric forms of S100A12 are required for its activation of the RAGE inflammation receptor^{291,292}. However, to date, no link between Cu^{2+} and S100A12-RAGE signaling has been established.

Despite the importance of Cu^{2+} , relatively few studies have focused on the interaction of Cu^{2+} with S100A12 in solution. Here, we describe electron spin resonance (ESR) studies to characterize the Cu^{2+} binding site of Ca^{2+} -S100A12, and the effect of Cu^{2+} on S100A12 oligomerization. First, we use continuous wave (CW)-ESR to show that there are two distinct binding sites. This interpretation is supported by pulsed electron spin echo envelope modulation

(ESEEM) experiments that show differences in the spectra between the two sites, with backbone coordination to Cu^{2+} shown at only one of the sites. Hyperfine sublevel correlation (HYSCORE) experiments also validate the presence of backbone coordination for only one of the Cu^{2+} binding site. Finally, we explore the structure and oligomerization using double electron electron resonance (DEER) measurements.

5.2 MATERIALS AND METHODS

5.2.1 Sample preparation

The expression and purification of S100A12 was performed as previously described²⁹³. CaCl_2 solution, prepared in MQ H_2O , was added to the protein in a 4:1 ratio. CuCl_2 solution, prepared in water, was added to the calcium-bound protein, as necessary.

5.2.2 ESR measurements

The samples were prepared in 50 mM N-ethylmorpholine (NEM) buffer, pH 7.4 to eliminate any free Cu^{2+} signal⁷⁸. Glycerol (20% v/v) was added as a cryoprotectant. All sample volumes were 120 μL . ESR experiments were performed on either a Bruker ElexSys E680 X-band FT/CW spectrometer equipped with a Bruker EN4118X-MD4 resonator or a Bruker ElexSys 580 X-band FT/CW spectrometer equipped with a Bruker ER4118X-MD5 resonator. The temperature for all experiments was controlled using an Oxford ITC503 temperature controller with an Oxford ER 4118CF gas flow cryostat.

Continuous wave (CW) ESR experiments were carried out at X-band frequencies at 80 K. Data were collected for 1024 points over a sweep width of 2000 G, from 2100 G to 4100 G, with a modulation amplitude of 4 G. The CW-ESR data were acquired with a modulation frequency of 100 kHz, a time constant of 10.24 ms, a conversion time of 20.48 ms and an incident power of 0.1992 mw. All CW-ESR spectra were simulated and fit using the EasySpin software¹⁵³.

Three-pulse electron-spin envelope echo modulation (ESEEM) experiments were performed at X-band frequencies at 80 K. A $\pi/2 - \tau - \pi/2 - T - \pi/2$ - echo pulse sequence with a $\pi/2$ pulse length of 16 ns was used. The first time delay, τ , was set to 144 ns for ESEEM performed at 2807 G and 152 ns for 3430 G. The second time delay, T , was varied, starting from an initial value of 288 ns with a step size of 16 ns. A four-step phase cycling was employed to eliminate all unwanted echoes^{199,249}. Data acquisition took ~12 hours. The resultant signal was phase corrected, baseline subtracted and Fourier-transformed using the Bruker WinEPR software.

Four-pulse hyperfine sublevel correlation (HYSCORE) experiments were performed at 20 K at X-band frequencies. The pulse sequence $\pi/2 - \tau - \pi/2 - t_1 - \pi - t_2 - \pi/2 - \tau$ - echo was used. HYSCORE was performed at two fields. At 2807 G, the first pulse separation, τ was set at 144 ns while at 3430 G, the τ was set at 152 ns. The pulse separations, t_1 and t_2 were varied from 200 ns with a step size of 16 ns for a total of 256 points. The pulse lengths used were 16 ns and 32 ns for $\pi/2$ and π pulses, respectively. A four-step phase cycling eliminated unwanted echo. The real parts of the collected two-dimensional data were phase corrected, baseline subtracted and zero filled to 512 points in both dimensions using the Bruker WinEPR software. The data was then fast Fourier transformed and reported as a contour plot.

Double electron electron resonance (DEER) experiments were performed at 20 K at X-band frequencies. The four pulse sequence used was: $(\pi/2)v_1 - \tau_1 - (\pi)v_1 - (\tau_1 + t) - (\pi)v_2 - (\tau_1 + \tau_2 -$

t) - $(\pi)v_1 - \tau_2 - \text{echo}^{154}$. The observer pulses, $(\pi/2)v_1$ and $(\pi)v_1$ were 16 ns and 32 ns, respectively and the pump pulse $(\pi)v_2$ was 16 ns. The delay, t, was incremented using a step size of 14 ns for a total of 128 points. The pump frequency, ν_2 , was placed at the maximum of the echo detected field swept spectrum with the observer frequency, ν_1 , offset 150 MHz downfield. The raw time domain DEER data were analyzed by DEERAnalysis2015b¹⁰³ via Tikhonov regularization. Distance distribution was corrected using proper g-values¹⁶⁶. Data acquisition took ~24 hours.

5.3 RESULTS AND DISCUSSION

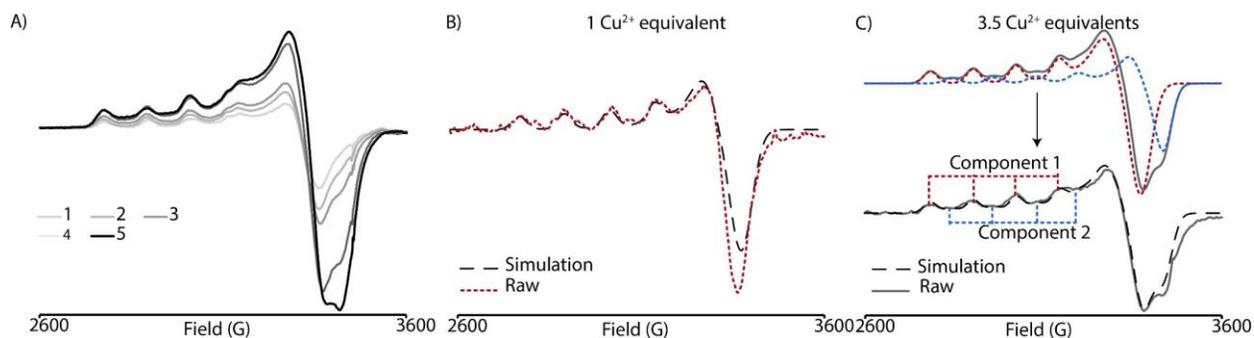


Figure 5-2 A) CW-ESR titrations performed at different Cu^{2+} equivalents B) Simulations show that the CW-ESR spectrum at 1 equivalent of Cu^{2+} is a single component with g_{\parallel} and A_{\parallel} values of 2.335 and 118 G, respectively C) At 3.5 equivalents of Cu^{2+} , a distinct second component is observed. Hence, we use another component with g_{\parallel} and A_{\parallel} of 2.292 and 115 G (blue dotted lines, upper panel). By varying the ratios of the two components we get the best fit for the spectrum, as shown in the lower panel.

We first performed CW-ESR titrations to obtain the maximum loading efficiency of the Cu^{2+} to the Ca^{2+} bound S100A12 protein and consequently, the binding affinity of Cu^{2+} to the native Cu^{2+} -binding sites in the protein. Figure 5-2A shows the CW-ESR spectra of the protein at different Cu^{2+} equivalents. Figure 5-2B shows the CW spectrum of 1 Cu^{2+} equivalent along with

the best fit. The simulation indicates a single component fit with g_{\parallel} and A_{\parallel} values of 2.335 and 118 G, respectively. We refer to this binding environment as component 1. Around 3 Cu^{2+} equivalents, the CW spectrum showed a distinct second component (cf. Figure 5-2C), which is easily distinguishable in the g_{\perp} region. Simulations of the spectrum at 3.5 Cu^{2+} equivalents confirmed the presence of two components. The second component corresponded to g_{\parallel} and A_{\parallel} values of 2.292 and 115 G, respectively. The difference in g_{\parallel} (2.335 and 2.292) between the first and the second component, indicates differences in the coordination environment of copper. The A_{\parallel} values of the two components indicate that the copper in S100A12 is a type II copper center which forms tetrahedral geometry, which is consistent with the Ca^{2+} - Cu^{2+} crystal structure²⁸⁴ and previously reported data from ESR studies on the Cu^{2+} -S100A12 (without Ca^{2+})²⁹⁴.

The CW-ESR results are interesting when compared to the crystal structure of Ca^{2+} -S100A12 crystallized in the presence of excess Cu^{2+} (PDB 1ODB)²⁸⁴. Three homodimers of S100A12 arranged as a hexamer are present in the asymmetric unit which results in six total Cu^{2+} ions in the structure. Five of the Cu^{2+} binding sites are canonical S100 transition metal binding sites comprised of a histidine and aspartic acid (His15 and Asp25) from one subunit of the dimer and two histidine residues from the other subunit (His85 and His89). However, in one of the dimers, the canonical Cu^{2+} binding site is altered. Asp25 is replaced by Glu55 from a symmetry related molecule. This results in a larger Cu-O distance as compared to the Asp- Cu^{2+} coordination in the other sites (2.4 vs 2.05)²⁷⁸. This noncanonical coordination may be an artifact from crystal packing. However, a study of Zn^{2+} binding to Ca^{2+} -S100B revealed that different ligands coordinate Zn^{2+} at different pH, which suggest the residues in the transition metal binding sites of S100 proteins may be dynamic²⁹⁵. Other studies have shown the canonical transition metal binding site may be modified by external ligands. For example, S100A15 has the required aspartic acid

residue replaced with a glycine. A chloride ion serves as the fourth zinc ligand in the Zn^{2+} -S100A15 crystal structure²⁹⁶. Additionally, the Zn^{2+} - Ca^{2+} -S100A8 homodimer structure has His27 replaced with a chloride ion in two out of the eight molecules in the asymmetric unit²⁹⁷. In total, these data suggest that the transition metal binding sites in S100 proteins may be more flexible than previously thought. Hence it is possible to observe slight differences in the predicted equivalent Cu^{2+} binding sites of S100A12. From the A_{\parallel} and g_{\parallel} values that we report, Cu^{2+} shows difference in its binding environment. This suggests that in solution the two symmetrical Cu^{2+} binding sites in S100A12 behave differently. Indeed, literature reports that even though proteins of S100 family are known to be typically homodimers, they can adopt overall asymmetric conformations upon metal binding²⁷¹.

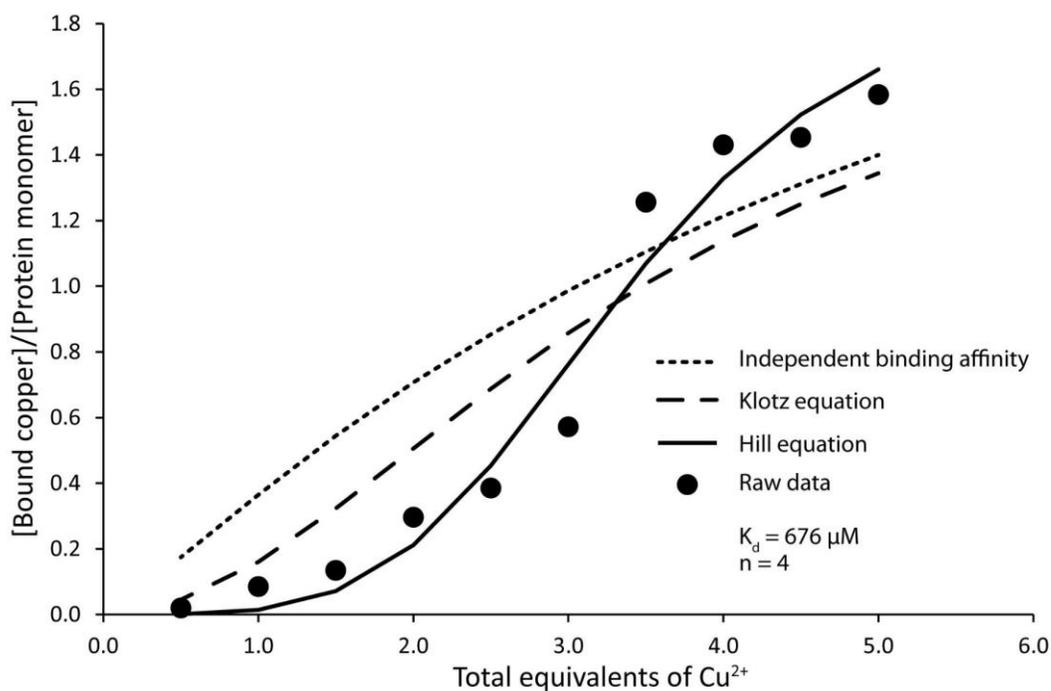


Figure 5-3 Plot showing the bound Cu^{2+} concentration/protein monomer concentration versus total equivalents of Cu^{2+} . Protein monomer concentration is 200 μM . The raw data was fitted to three different models. The model corresponding to the Hill equation best fits the data. The apparent dissociation constant using the Hill equation was found to be 676 μM with a positive cooperativity of 4.

Next, we established the mode of binding of Cu^{2+} to Ca^{2+} -bound S100A12 in solution. The double integrated intensity of the CW-ESR spectrum is related to the concentration of the bound Cu^{2+} . Note that the data was acquired in NEM buffer where free Cu^{2+} is ESR silent^{78,298}. By comparing to a standard, the concentration of bound Cu^{2+} can be determined^{49,57,91,259}. Figure 5-3 shows the ratio of bound Cu^{2+} : protein as a function of total equivalents of added Cu^{2+} . We fit this data to different binding models. The first model we considered was independent binding to the two sites⁵⁷. The best fit is shown by a dotted line in Figure 5-3. The second model we tried was the Klotz equation^{299,300} (Figure 5-3, dashed line). The Klotz model considers cooperative binding between the two binding sites with different binding affinities for each site. Lastly, we applied the Hill's equation³⁰¹⁻³⁰³ to fit our data (Figure 5-3, solid line). As can be seen, our data best fits the binding affinity model described by Hill's equation. The apparent dissociation constant, K_d can be then calculated from the Hill's equation as:

$$\theta = \frac{[L]^n}{K_d + [L]^n} \quad (5-1)$$

where, θ is the fraction of protein bound to ligand, $[L]$ is the concentration of the free (unbound) ligand, K_d is the apparent dissociation constant and n is the Hill's coefficient. A value of n greater than 1 is associated with positive cooperative binding. For our data, the Hill's coefficient of 4 indicates that the binding of a ligand molecule at one site facilitates the binding of the same ligand at a different site. Thus, affinity of the second site towards the ligand increases and results in positive cooperative binding. The apparent K_d value of Cu^{2+} bound to Ca^{2+} -S100A12 is 676 μM . This apparent K_d likely reflects a complex process involving binding of Cu^{2+} to Ca^{2+} -S100A12 and oligomerization of Cu^{2+} - Ca^{2+} -S100A12 (see below). Factors such as temperature, buffer, and pH largely affect the apparent K_d . Most importantly, our measured value is at 80 K and the affinity at physiological conditions is expected to be much higher⁹¹.

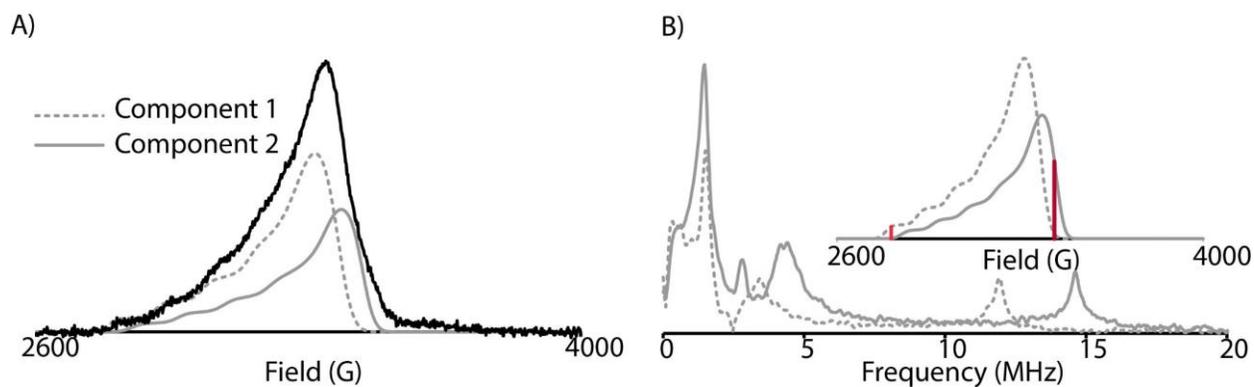


Figure 5-4 A) Echo-detected field swept spectrum of S100A12 with 3.5 Cu^{2+} equivalents (black solid). The simulated spectrum corresponding to the first component (grey dotted) and second component (grey solid) are shown. B) ESEEM experiments performed at two fields (inset, red lines). At 2807 G, the first component (grey dotted) is predominant while at 3430 G, the second component (grey solid) is the main constituent. The spectra have been normalized to the hydrogen peak intensity.

In order to confirm our findings, we performed ESEEM experiments at two different fields (Figure 5-4B, inset, red lines) on the S100A12 with 3.5 equivalents of Cu^{2+} (Figure 5-4). Analysis of the CW-ESR simulations indicated that at a field of 2807 G, the spectrum consists of 98% of the first component while at 3430 G, the second component constitutes ~92% of the spectrum (Figure 5-4A and B, inset). Therefore, ESEEM experiments were performed at these two fields to probe Cu^{2+} coordination at each site. These positions are shown in red in the inset of Figure 5-4B. ESEEM is mainly sensitive to weak hyperfine interactions, which normally are not resolved in the broad CW-ESR spectra. For Cu^{2+} bound to histidines a characteristic ESEEM spectrum is observed^{68,89}. Nuclear quadrupole interactions (NQI) of ^{14}N give rise to three sharp peaks below 2 MHz in the ESEEM spectra. A broad feature around 4 MHz is due to the double quantum (DQ) transition of the remote nitrogen in an imidazole ring. The ESEEM spectra also show significant changes with changes in the number of coordinated histidine residues. In particular, as the number of distal nitrogen coordination increases with increase in histidine coordination, the overall

intensity of the ^{14}N peak in the ESEEM spectrum will increase^{68,84,304}. As can be seen in Figure 5-4B, the ESEEM spectrum obtained at the 2807 G differ significantly from that obtained at 3430 G. Furthermore, the ESEEM spectrum at 3430 G shows a peak at 2.8 MHz. When Cu^{2+} ion is directly coordinated to the carbonyl of an amino acid residue, the amide nitrogen results in a peak at 2.8 MHz in the ESEEM spectrum⁹⁵. Thus, the Cu^{2+} binding site corresponding to the second component possibly shows backbone coordination with the carbonyl group of the amino acid residue.

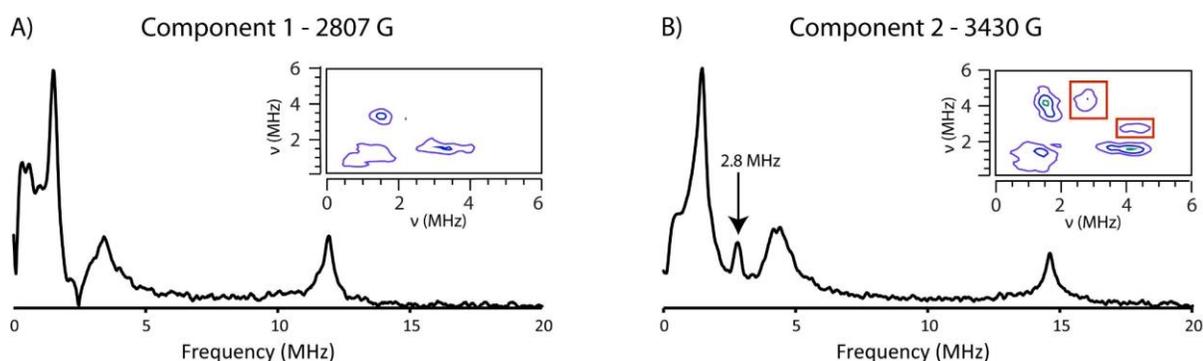


Figure 5-5 A) ESEEM spectrum at 2807 G showing no visible peak at 2.8 MHz corresponding to backbone coordination. Inset shows HYSORE data at 2807 G which shows no visible cross peaks either. B) ESEEM spectrum at 3430 G shows a sharp peak at 2.8 MHz which could be due to possible backbone coordination. HYSORE data (inset) shows cross peaks (red squares) at (4.3, 2.8 MHz) confirming the presence of backbone coordination.

To ascertain the difference in coordination at the two binding sites and to confirm the presence of backbone coordination, we performed HYSORE at the two fields. At 2807 G, comprising mostly of the first component, there is no observable peak at 2.8 MHz corresponding to the backbone coordination as shown in Figure 5-5A. Furthermore, HYSORE also shows no trace of backbone coordination. However, at 3430 G, ESEEM spectrum shows a sharp peak at 2.8 MHz (Figure 5-5B). On analyzing HYSORE data at 3430 G, we observed cross peaks at (4.3, 2.8 MHz). The position of these cross peaks are indicative of an amide nitrogen atom of the peptide backbone when the adjacent carbonyl group is coordinated to the Cu^{2+} ion^{95,96}. Thus, only one of

the Cu^{2+} binding sites is shown to undergo backbone coordination with the amino acid residue, confirming the observation that the two Cu^{2+} binding sites show different coordination environment.

Since metals have been shown to induce oligomerization of S100 proteins^{284,305-309}, we further analyzed the distances between the Cu^{2+} -binding sites from the X-ray crystal structure of S100A12 (PDB: 1ODB)²⁸⁴ as shown in Figure 5-6A. The crystal structure shows the asymmetric unit is a hexamer assembled from a trimer of dimers. The intra-dimer Cu^{2+} - Cu^{2+} distance is ~ 3 nm. There are also inter-dimer distances of ~ 1.4 nm, 4 nm and 4.8 nm contained within the hexamer.

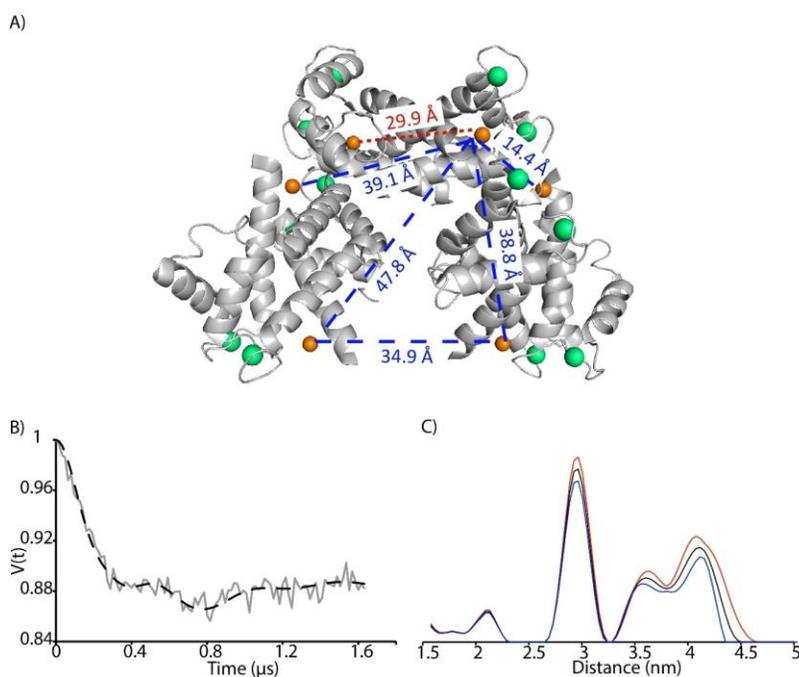


Figure 5-6 A) X-ray crystal of S100A12 bound to both Ca^{2+} and Cu^{2+} (PDB: 1ODB)²⁸⁴. The green spheres represent the Ca^{2+} ions while the orange spheres represent the Cu^{2+} ions. The blue dashed lines show the inter-dimer distances while the red dotted line shows the intra-dimer distance. B) Background subtracted time-domain DEER data showing the presence of two frequencies. The experimental data is represented by the gray solid line and the black dashed line represents the best fit obtained from Tikhonov regularization. The modulation depth is observed to be $\sim 12.6\%$. C)

Distance distributions obtained using Tikhonov regularization. The distribution also shows validation techniques including upper (red) and lower (blue) acceptable limits.

To compare these crystal structure data to that in solution, we performed Cu²⁺-DEER^{38,45,54} on the Ca²⁺ loaded S100A12 with 2 equivalents of Cu²⁺. The baseline corrected DEER signal and the distance distribution is shown in Figure 5-6B and C. First, we noticed that for a pump pulse of 16 ns, the modulation depth experimentally obtained is ~12.6%, as shown in Figure 5-6B. This is higher than the theoretical modulation depth for a system of two spins (~8.2%), calculated from the echo-detected field swept spectrum as previously described⁴⁶. The high modulation depth confirms the presence of oligomers in our sample. Second, the distance distribution, in Figure 5-6C, clearly shows a sharp peak at 3 nm and a broad peak around 4 nm. This is in good agreement with the intra-dimer distance of 3 nm and the several inter-dimer distances around 4 nm obtained from crystal structure (PDB 1ODB)²⁸⁴. We also performed a two-Gaussian fit and we find that a bimodal distance of 3 nm and 4 nm best fits our data. Note that the 1.4 nm distance is unlikely to be sampled by DEER. Using the modulation depth, a rough estimate can be made on the amount of dimers and trimers of dimers in solution.

Considering our sample to be a mix of two and six spin systems, the modulation depth is^{49,91}:

$$\lambda = x * \lambda_2 + (1 - x) * \lambda_6 \quad (5-2)$$

where λ is the modulation depth and x is the mole fraction of the dimer. Also λ_N is given by²⁵⁰:

$$\lambda_N = 1 - (1 - p_b)^{N-1} \quad (5-3)$$

where N is the number of spins in the system, and p_b is the fraction of spins excited by the pump pulse. For a pump pulse of 16 ns, we obtain a p_b of 0.082⁴⁶. As shown in the crystal structure (PDB 1ODB), S100A12 exists as a trimer of dimers²⁸⁴, which would make the sample a six-spin system.

Using these equations, we estimate that roughly 83% of the protein exists as dimers in solution, under the assumption that only dimers and trimers of dimers exist.

These results are interesting because they show that oligomers exist even in the solution state. To the best of our knowledge, no previous studies have investigated the role of Cu^{2+} in oligomerization of S100A12. Previous work on calgranulin C, in presence of Zn^{2+} , has suggested the presence of oligomers and that oligomeric forms of S100A12 are important in the activation of the inflammation receptor RAGE^{278,284,291,310}. Since Cu^{2+} occupies the same binding site as Zn^{2+} and shows similar coordination geometry, we expect that the addition of Cu^{2+} will also lead to oligomerization. The agreement between the DEER distance and those measured from the X-ray structure suggests that the Cu^{2+} - Ca^{2+} -S100A12 adopts a trimer of dimers hexameric structure. Additionally, these data suggest that the Cu^{2+} binds only to the transition metal binding sites at the dimer interface and not elsewhere. This is consistent with the observation that Cu^{2+} was only observed at the canonical transition metal binding sites of S100A12, despite the fact that the crystal was grown under a large excess of Cu^{2+} .

5.4 CONCLUSION

We have shown that the two binding sites of the S100A12 bind differently to Cu^{2+} . Even though S100A12 is a homodimer and the two canonical transition metal binding sites are related by symmetry, we have shown that, in solution, the two Cu^{2+} binding sites are distinct. The CW-ESR spectrum of Cu^{2+} bound to the protein clearly shows the presence of two components, corresponding to two different binding environments. ESEEM and HYSORE validate these findings and show that only one of the components undergoes backbone coordination, indicating

that the binding environment of the Cu^{2+} ion is different at the two sites. DEER data further confirmed that the difference in coordination is not due to Cu^{2+} binding to previously unidentified sites. The agreement of each of the most probable distance of the DEER distribution to that of the X-ray crystal structure confirms that the Cu^{2+} binds only to the canonical S100 metal binding sites located at the dimer interface. Moreover, the distances at 3 and 4 nm confirms the formation of oligomers in presence of Cu^{2+} and Ca^{2+} and highlights the utility of ESR spectroscopy to characterize metal-induced protein oligomerization in solution.

5.5 ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation (awards NSF MCB-1613007, NSF MRI-1725678, NSF HRD-1400969 and NSF HRD-1547757). Velia Garcia was supported by an NSF LSAMP Bridge to the Doctorate award HRD- 1500320. We thank Dr. Matthew Lawless for useful discussion.

APPENDIX A: RESIDUE TOPOLOGY AND KEY FORCE FIELD PARAMETERS

APPENDIX A.1 RESIDUE TOPOLOGY OF Cu²⁺-DPA-WAT₁

File Name: DPA.prepi

0 0 2

This is a remark line

dpa.res

DPA INT 0

CORRECT OMIT DU BEG

0.0000

1	DUMM	DU	M	0	-1	-2	0.000	.0	.0	.00000
2	DUMM	DU	M	1	0	-1	1.449	.0	.0	.00000
3	DUMM	DU	M	2	1	0	1.523	111.21	.0	.00000
4	P2	P	M	3	2	1	1.540	111.208	-180.000	1.369015
5	O6	O2	E	4	3	2	1.503	138.219	-132.370	-0.834137
6	O7	O2	E	4	3	2	1.504	69.118	111.415	-0.834137
7	O5	OS	M	4	3	2	1.666	41.750	-83.842	-0.434069
8	C13	CT	M	7	4	3	1.423	117.834	24.136	-0.069712
9	H20	H1	E	8	7	4	1.097	110.010	-66.592	0.048214
10	H21	H1	E	8	7	4	1.098	110.806	53.014	0.048214
11	C11	CT	M	8	7	4	1.529	108.550	172.648	0.523010
12	C16	CT	3	11	8	7	1.537	108.764	164.916	-0.072228
13	N1	N3	S	12	11	8	1.502	117.117	176.539	0.060271
14	Cu1	Cu	3	13	12	11	2.010	102.855	163.593	0.930057
15	N2	NX	S	14	13	12	1.966	84.525	-91.969	-0.533625
16	C2	CA	B	15	14	13	1.354	112.881	-12.764	0.418204
17	C1	CT	B	16	15	14	1.508	115.313	-6.306	-0.382868
18	H1	HP	E	17	16	15	1.096	108.214	-86.447	0.166229
19	H2	HP	E	17	16	15	1.096	112.326	155.202	0.166229

20	C3	CA	B	16	15	14	1.390	121.578	176.281	-0.248187
21	H3	HA	E	20	16	15	1.084	119.664	-179.204	0.194440
22	C4	CA	B	20	16	15	1.395	118.963	0.690	0.058679
23	H4	HA	E	22	20	16	1.085	120.170	179.937	0.169276
24	C5	CA	B	22	20	16	1.395	119.256	-0.439	-0.273038
25	H5	HA	E	24	22	20	1.084	121.597	-179.698	0.185970
26	C6	CA	S	24	22	20	1.389	118.706	-0.256	0.188894
27	H6	H4	E	26	24	22	1.082	121.359	-179.268	0.067462
28	N3	NZ	S	14	13	12	1.959	84.734	94.839	-0.533625
29	C8	CA	B	28	14	13	1.353	113.489	10.902	0.418204
30	C7	CT	B	29	28	14	1.509	115.858	5.706	-0.382868
31	H7	HP	E	30	29	28	1.090	112.312	-149.878	0.166229
32	H8	HP	E	30	29	28	1.095	108.521	91.444	0.166229
33	C9	CA	B	29	28	14	1.391	121.329	-176.852	-0.248187
34	H9	HA	E	33	29	28	1.085	119.725	179.384	0.194440
35	C10	CA	B	33	29	28	1.394	119.024	-0.668	0.058679
36	C14	CA	B	35	33	29	1.395	119.252	0.461	-0.273038
37	H10	HA	E	36	35	33	1.084	121.472	179.710	0.185970
38	C15	CA	S	36	35	33	1.388	118.813	0.200	0.188894
39	H11	H4	E	38	36	35	1.082	122.105	179.267	0.067462
40	H22	HA	E	35	33	29	1.085	120.171	-179.944	0.169276
41	O9	ow	B	14	13	12	2.017	169.111	13.639	-0.592041
42	H18	hw	E	41	14	13	0.973	121.699	-173.567	0.414588
43	H19	hw	E	41	14	13	0.972	120.565	-33.308	0.414588
44	H24	HP	E	12	11	8	1.094	107.367	-62.153	0.030515
45	H25	HP	E	12	11	8	1.094	110.203	54.612	0.030515
46	H23	H1	E	11	8	7	1.093	108.135	45.383	0.020529
47	O1	OS	M	11	8	7	1.427	109.983	-73.869	-0.408523

LOOP

C1 N1

C7 N1

C6 N2

```

C15  N3
IMPROPER
C2  C6  N2  Cu1
C3  C1  C2  N2
C2  C4  C3  H3
C3  C5  C4  H4
C4  C6  C5  H5
C5  H6  C6  N2
C8  C15  N3  Cu1
C9  C7  C8  N3
C8  C10  C9  H9
C9  C14  C10  H22
C10  C15  C14  H10
C14  H11  C15  N3
DONE
STOP

```

APPENDIX A.2 RESIDUE TOPOLOGY OF DSPACER

File Name: DSP.prepi

```
0 0 2
```

This is a remark line

DSP.res

```
DSP  INT  0
```

```
CORRECT  OMIT  DU  BEG
```

```
0.0000
```

1	DUMM	DU	M	0	-1	-2	0.000	.0	.0	.00000
2	DUMM	DU	M	1	0	-1	1.449	.0	.0	.00000
3	DUMM	DU	M	2	1	0	1.523	111.21	.0	.00000
4	P2	P	M	3	2	1	1.540	111.208	-180.000	1.295965

5	O4	O2	E	4	3	2	1.473	54.315	-77.788	-0.826887
6	O6	O2	E	4	3	2	1.478	93.607	49.757	-0.826887
7	O3	OS	M	4	3	2	1.619	155.989	-126.296	-0.522060
8	C5	CT	M	7	4	3	1.401	120.164	-146.267	0.096810
9	H4	H1	E	8	7	4	1.085	109.662	-60.978	0.033214
10	H5	H1	E	8	7	4	1.086	110.081	57.875	0.033214
11	C1	CT	M	8	7	4	1.521	109.828	178.845	0.389113
12	O1	OS	S	11	8	7	1.420	110.213	-63.586	-0.523884
13	C4	CT	3	12	11	8	1.403	110.936	-121.973	0.146514
14	C3	CT	B	13	12	11	1.521	105.399	23.279	-0.038815
15	H9	HC	E	14	13	12	1.087	110.254	79.352	0.013253
16	H10	HC	E	14	13	12	1.081	113.757	-157.319	0.013253
17	H2	H1	E	13	12	11	1.086	108.288	145.251	0.027202
18	H3	H1	E	13	12	11	1.086	110.563	-96.056	0.027202
19	H1	H1	E	11	8	7	1.082	109.213	55.803	0.019398
20	C2	CT	M	11	8	7	1.539	112.747	178.156	0.161221
21	H14	H1	E	20	11	8	1.081	112.923	-24.157	0.031635
22	O9	OS	M	20	11	8	1.408	108.473	-145.576	-0.549462

LOOP

C2 C3

IMPROPER

DONE

STOP

APPENDIX A.3 KEY FORCE FIELD PARAMETERS OF Cu²⁺-DPA-WAT₁

File Name: DPA.frcmod2

Remark line goes here

MASS

Cu 63.010 0.000

NX 14.010	0.530	same as nc
NZ 14.010	0.530	same as nc
hw 1.008	0.000	H in TIP3P water
ow 16.00	0.000	oxygen in TIP3P water

BOND

Cu-N3	47.11	2.010	calculated for Cu-N3
Cu-NX	52.79	1.966	calculated for Cu-NX
Cu-NZ	53.77	1.959	calculated for Cu-NZ
Cu-ow	39.51	2.017	calculated for Cu-OW
CA-NX	394.60	1.352	same as ca-nc, penalty score= 0.0
CA-NZ	394.60	1.352	same as ca-nc, penalty score= 0.0
ow-hw	553.0	0.9572	! TIP3P water
hw-hw	553.0	1.5136	TIP3P water

ANGLE

CT-N3-Cu	20.541	105.102
N3-Cu-NX	45.470	84.527
N3-Cu-NZ	45.491	84.734
N3-Cu-ow	31.748	169.109
NX-Cu-NZ	32.749	167.291
NX-Cu-ow	42.088	98.298
NZ-Cu-ow	43.223	93.513
CA-NX-Cu	19.811	120.227
CA-NZ-Cu	19.891	120.080
Cu-ow-hw	11.709	121.130
CA-CT-N3	81.700	113.800
CT-CA-NX	85.113	115.317
CA-CA-NX	87.600	119.720
H4-CA-NX	63.000	118.360
CA-NX-CA	72.000	109.950
CA-CT-HP	47.300	110.470
CT-CA-NZ	84.914	115.858
CA-CA-NZ	87.600	119.720

H4-CA-NZ 63.000 118.360

CA-NZ-CA 72.000 109.950

hw-ow-hw 100. 104.52

hw-hw-ow 0. 127.74

DIHEDRAL

NX-Cu-N3-CT 1 0.000 0.000 2.000

NZ-Cu-N3-CT 1 0.000 0.000 2.000

ow-Cu-N3-CT 1 0.000 0.000 2.000

N3-Cu-NX-CA 1 0.000 0.000 2.000

N3-Cu-NZ-CA 1 0.000 0.000 2.000

N3-Cu-ow-hw 1 0.000 0.000 2.000

CT-CA-NX-Cu 2 9.600 180.000 2.000

CA-CA-NX-Cu 2 9.600 180.000 2.000

H4-CA-NX-Cu 2 9.600 180.000 2.000

CT-CA-NZ-Cu 2 9.600 180.000 2.000

CA-CA-NZ-Cu 2 9.600 180.000 2.000

H4-CA-NZ-Cu 2 9.600 180.000 2.000

NX-Cu-NZ-CA 1 0.000 0.000 2.000

NX-Cu-ow-hw 1 0.000 0.000 2.000

CA-CA-NX-CA 2 9.600 180.000 2.000

H4-CA-NX-CA 2 9.600 180.000 2.000

CT-CA-NX-CA 2 9.600 180.000 2.000

NZ-Cu-NX-CA 1 0.000 0.000 2.000

NZ-Cu-ow-hw 1 0.000 0.000 2.000

CA-CA-NZ-CA 2 9.600 180.000 2.000

H4-CA-NZ-CA 2 9.600 180.000 2.000

CT-CA-NZ-CA 2 9.600 180.000 2.000

ow-Cu-NX-CA 1 0.000 0.000 2.000

ow-Cu-NZ-CA 1 0.000 0.000 2.000

hw-hw-ow-Cu 1 0.000 0.000 2.000

IMPROPER

CA-CA-NX-Cu 1.1 180.0 2.0

CA-CT-CA-NX	1.1	180.0	2.0
CA-CA-CA-HA	1.1	180.0	2.0
CA-H4-CA-NX	1.1	180.0	2.0
CA-CA-NZ-Cu	1.1	180.0	2.0
CA-CT-CA-NZ	1.1	180.0	2.0
CA-H4-CA-NZ	1.1	180.0	2.0

NONBON

Cu	2.2100	0.1729
NX	1.8993	0.0941
NZ	1.8993	0.0941
ow	1.8200	0.0930
hw	0.3019	0.0047

BIBLIOGRAPHY

- (1) Altenbach, C.; Marti, T.; Khorana, H. G.; Hubbell, W. L. *Science* **1990**, *248* (4959), 1088.
- (2) Hubbell, W. L.; Mchaourab, H. S.; Altenbach, C.; Lietzow, M. A. *Structure* **1996**, *4* (7), 779.
- (3) Hubbell, W. L.; Lopez, C. J.; Altenbach, C.; Yang, Z. *Cur. Opin. Struct. Biol.* **2013**, *23* (5), 725.
- (4) Hubbell, W. L.; Gross, A.; Langen, R.; Lietzow, M. A. *Curr. Opin. Struct. Biol.* **1998**, *8* (5), 649.
- (5) Hubbell, W. L.; Altenbach, C. *Curr. Opin. Struct. Biol.* **1994**, *4* (4), 566.
- (6) Voss, J.; He, M. M.; Hubbell, W. L.; Ronald Kaback, H. *Biochemistry* **1996**, *35* (39), 12915.
- (7) Lietzow, M. A.; Hubbell, W. L. *Biochemistry* **2004**, *43* (11), 3137.
- (8) Altenbach, C.; Greenhalgh, D. A.; Khorana, H. G.; Hubbell, W. L. *Proc. Natl. Acad. Sci. USA* **1994**, *91* (5), 1667.
- (9) Klug, C. S.; Su, W.; Feix, J. B. *Biochemistry* **1997**, *36* (42), 13027.
- (10) Columbus, L.; Hubbell, W. L. *Trends Biochem. Sci.* **2002**, *27* (6), 288.
- (11) Hubbell, W. L.; Cafiso, D. S.; Altenbach, C. *Nat. Struct. Biol.* **2000**, *7* (9), 735.
- (12) Borbat, P. P.; Freed, J. H. In *Biological Magnetic Resonance*; Timmel, C., Harmer, J., Eds.; Kluwer Academic/Plenum Publishers: New York, 2000; pp 1–82.
- (13) Saxena, S.; Freed, J. H. *Chem. Phys. Lett.* **1996**, *251* (1–2), 102.
- (14) Milikisyants, S.; Scarpelli, F.; Finiguerra, M. G.; Ubbink, M.; Huber, M. *J. Magn. Reson.* **2009**, *201* (1), 48.
- (15) Jeschke, G.; Polyhach, Y. *Phys. Chem. Chem. Phys.* **2007**, *9* (16), 1895.

- (16) Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W. *J. Magn. Reson.* **2000**, *142* (2), 331.
- (17) Chakrapani, S. *Methods Enzymol.* **2015**, *557*, 279.
- (18) Berliner, L. J.; Eaton, S. S.; Eaton, G. R. *Distance Measurements in Biological Systems by EPR*; Springer Science and Business Media, 2006.
- (19) Glaenger, J.; Peter, M. F.; Hagelueken, G. *Methods* **2018**, *147*, 163.
- (20) Schiemann, O.; Prisner, T. F. *Q. Rev. Biophys.* **2007**, *40* (01), 1.
- (21) Stone, K. M.; Townsend, J. E.; Sarver, J.; Sapienza, P. J.; Saxena, S.; Jen-Jacobson, L. *Angew. Chemie Int. Ed.* **2008**, *47*, 10192.
- (22) Sarver, J. L.; Townsend, J. E.; Rajapakse, G.; Jen-Jacobson, L.; Saxena, S. *J. Phys. Chem. B* **2012**, *116* (13), 4024.
- (23) Fajer, M. I.; Li, H.; Yang, W.; Piotr G. Fajer. *J. Am. Chem. Soc.* **2007**, *129* (45), 13840.
- (24) Fleissner, M. R.; Bridges, M. D.; Brooks, E. K.; Cascio, D.; Kalai, T.; Hideg, K.; Hubbell, W. L. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (39), 16241.
- (25) Barhate, N.; Cekan, P.; Massey, A. P.; Sigurdsson, S. T. *Angew. Chemie Int. Ed.* **2007**, *46*, 2655.
- (26) Cekan, P.; Sigurdsson, S. T. *Biochem. Biophys. Res. Commun.* **2012**, *420*, 656.
- (27) Cekan, P.; Smith, A. L.; Barhate, N.; Robinson, B. H.; Sigurdsson, S. T. *Nucleic Acids Res.* **2008**, *36*, 5946.
- (28) Heinz, M.; Erlenbach, N.; Stelzl, L. S.; Thierolf, G.; Kamble, N. R.; Sigurdsson, S. T.; Prisner, T. F.; Hummer, G. *Nucleic Acids Res.* **2020**, *48* (2), 924.
- (29) Kamble, N. R.; Gränz, M.; Prisner, T. F.; Sigurdsson, S. T. *Chem. Commun.* **2016**, *52* (100), 14442.

- (30) Edwards, T. E.; Cekan, P.; Reginsson, G. W.; Shelke, S. A.; Ferré-D'Amaré, A. R.; Schiemann, O.; Sigurdsson, S. T. *Nucleic Acids Res.* **2011**, *39* (10), 4419.
- (31) Marko, A.; Denysenkov, V.; Margraf, D.; Cekan, P.; Schiemann, O.; Sigurdsson, S. T.; Prisner, T. F. *J. Am. Chem. Soc.* **2011**, *133*, 13375.
- (32) Stelzl, L. S.; Erlenbach, N.; Heinz, M.; Prisner, T. F.; Hummer, G. *J. Am. Chem. Soc.* **2017**, *139*, 11674.
- (33) Gophane, D. B.; Endeward, B.; Prisner, T. F.; Sigurdsson, S. T. *Chem. Eur. J.* **2014**, *20* (48), 15913.
- (34) Erlenbach, N.; Endeward, B.; Schöps, P.; Gophane, D. B.; Sigurdsson, S. T.; Prisner, T. F. *Phys. Chem. Chem. Phys.* **2016**, *18* (24), 16196.
- (35) Kerzhner, M.; Abdullin, D.; Wiecek, J.; Matsuoka, H.; Hagelueken, G.; Schiemann, O.; Famulok, M. *Chem. Eur. J.* **2016**, *22* (34), 12113.
- (36) Schmidt, T.; Wälti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M. *Angew. Chemie Int. Ed.* **2016**, *55* (51), 15905.
- (37) Jassoy, J. J.; Berndhäuser, A.; Duthie, F.; Kühn, S. P.; Hagelueken, G.; Schiemann, O. *Angew. Chemie Int. Ed.* **2017**, *56* (1), 177.
- (38) Ji, M.; Ruthstein, S.; Saxena, S. *Acc. Chem. Res.* **2014**, *47* (2), 688.
- (39) Feintuch, A.; Otting, G.; Goldfarb, D. In *Methods in Enzymology*; Academic Press Inc., 2015; Vol. 563, pp 415–457.
- (40) Abdullin, D.; Duthie, F.; Meyer, A.; Müller, E. S.; Hagelueken, G.; Schiemann, O. *J. Phys. Chem. B* **2015**, *119* (43), 13539.
- (41) Yang, Z.; Ji, M.; Cunningham, T. F.; Saxena, S. *Methods Enzymol.* **2015**, *563*, 459.
- (42) Becker, J. S.; Saxena, S. *Chem. Phys. Lett.* **2005**, *414* (1–3), 248.

- (43) Ruthstein, S.; Ji, M.; Mehta, P.; Jen-Jacobson, L.; Saxena, S. *J. Phys. Chem. B* **2013**, *117* (20), 6227.
- (44) Ruthstein, S.; Ji, M.; Shin, B. K.; Saxena, S. *J. Magn. Reson.* **2015**, *257*, 45.
- (45) Yang, Z.; Kise, D.; Saxena, S. *J. Phys. Chem. B* **2010**, *114* (18), 6165.
- (46) Yang, Z.; Ji, M.; Saxena, S. *Appl. Magn. Reson.* **2010**, *39* (4), 487.
- (47) Cunningham, T. F.; Putterman, M. R.; Desai, A.; Horne, W. S.; Saxena, S. *Angew. Chemie Int. Ed.* **2015**, *54*, 6330.
- (48) Cunningham, T. F.; Shannon, M. D.; Putterman, M. R.; Arachchige, R. J.; Sengupta, I.; Gao, M.; Jaroniec, C. P.; Saxena, S. *J. Phys. Chem. B* **2015**, *119* (7), 2839.
- (49) Ghosh, S.; Lawless, M. J.; Rule, G. S.; Saxena, S. *J. Magn. Reson.* **2018**, *286*, 163.
- (50) Lawless, M. J.; Sarver, J. L.; Saxena, S. *Angew. Chemie Int. Ed.* **2017**, *56*, 2115.
- (51) Jun, S.; Becker, J. S.; Yonkunas, M.; Coalson, R.; Saxena, S. *Biochemistry* **2006**, *45* (38), 11666.
- (52) Sameach, H.; Ghosh, S.; Gevorkyan-Airapetov, L.; Saxena, S.; Ruthstein, S. *Angew. Chemie Int. Ed.* **2019**, *58*, 3053.
- (53) Ji, M.; Tan, L.; Jen-Jacobson, L.; Saxena, S. *Mol. Phys.* **2014**, *112* (24), 3173.
- (54) Yang, Z.; Kurpiewski, M. R.; Ji, M.; Townsend, J. E.; Mehta, P.; Jen-Jacobson, L.; Saxena, S. *Proc. Natl. Acad. Sci. USA* **2012**, *109* (17), E993.
- (55) Lawless, M. J.; Pettersson, J. R.; Rule, G. S.; Lanni, F.; Saxena, S. *Biophys. J.* **2018**, *114* (3), 592.
- (56) Gamble Jarvi, A.; Cunningham, T. F.; Saxena, S. *Phys. Chem. Chem. Phys.* **2019**, *21* (20), 10238.
- (57) Ruthstein, S.; Stone, K. M.; Cunningham, T. F.; Ji, M.; Cascio, M.; Saxena, S. *Biophys. J.*

- 2010**, 99, 2497.
- (58) Bogetti, X.; Ghosh, S.; Gamble Jarvi, A.; Wang, J.; Saxena, S. *J. Phys. Chem. B* **2020**, 124 (14), 2788.
- (59) Gamble Jarvi, A.; Sargun, A.; Bogetti, X.; Wang, J.; Achim, C.; Saxena, S. *J. Phys. Chem. B* **2020**, (*in press*).
- (60) Ghosh, S.; Saxena, S.; Jeschke, G. *Appl. Magn. Reson.* **2018**, 49 (11), 1281.
- (61) Ghosh, S.; Lawless, M. J.; Brubaker, H. J.; Singewald, K.; Kurpiewski, M. R.; Jen-Jacobson, L.; Saxena, S. *Nucleic Acids Res.* **2020**, 48 (e49).
- (62) Ghosh, S.; Garcia, V.; Singewald, K.; Damo, S. M.; Saxena, S. *Appl. Magn. Reson.* **2018**, 49 (11), 1299.
- (63) Abragam, A.; Pryce, M. H. L. *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.* **1951**, 205 (1080), 135.
- (64) Weil, J. A.; Bolton, J. R. *Electron paramagnetic resonance: elementary theory and practical applications.*; Wiley-Interscience, 2007.
- (65) Slichter, C. P. *Principles of Magnetic Resonance*; Springer Series in Solid-State Sciences; Springer Berlin Heidelberg: Berlin, Heidelberg, 1990; Vol. 1.
- (66) Raymond Hagen, W. *Biomolecular EPR Spectroscopy*; CRC Press, Boca Raton, 2008.
- (67) Peisach, J.; Blumberg, W. E. *Arch. Biochem. Biophys.* **1974**, 165 (2), 691.
- (68) Silva, K. I.; Michael, B. C.; Geib, S. J.; Saxena, S. *J. Phys. Chem. B* **2014**, 118, 8935.
- (69) Wagner, E. P.; Gronborg, K. C.; Ghosh, S.; Saxena, S. *J. Chem. Educ.* **2019**, 96, 1752.
- (70) Solomon, E. I.; Hare, J. W.; Dooley, D. M.; Dawson, J. H.; Stephens, P. J.; Gray, H. B. *J. Am. Chem. Soc.* **1980**, 102 (1), 168.
- (71) Rienzo, F. De; Gabdoulline, R. R.; Menziani, M. C.; Wade, R. C. *Protein Sci.* **2000**, 9 (8),

1439.

- (72) Ryde, U.; Olsson, M. H. M.; Pierloot, K. *Theor. Comput. Chem.* **2001**, *9*, 1.
- (73) Cotton, F. A.; Wilkinson, G. *Advanced Inorganic Chemistry*, Fourth.; John Wiley & Sons, Inc., 1980.
- (74) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. *Chem. Rev.* **1992**, *92* (4), 521.
- (75) Solomon, E. I.; Heppner, D. E.; Johnston, E. M.; Ginsbach, J. W.; Cirera, J.; Qayyum, M.; Kieber-Emmons, M. T.; Kjaergaard, C. H.; Hadt, R. G.; Tian, L. *Chem. Rev.* **2014**, *114* (7), 3659.
- (76) Abragam, A.; Pryce, M. H. L. *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.* **1951**, *206* (1085), 173.
- (77) 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.; Millhauser, G. L. *Biochemistry* **2000**, *39*, 13760.
- (78) Syme, C. D.; Nadal, R. C.; Rigby, S. E. J.; Viles, J. H. *J. Biol. Chem.* **2004**, *279*, 18169.
- (79) Mims, W. B. *Phys. Rev. B* **1972**, *5* (7), 2409.
- (80) Mims, W. B.; Peisach, J. *J. Chem. Phys.* **1978**, *69* (11), 4921.
- (81) McCracken, J.; Peisach, J.; Dooley, D. M.; Dooley DM. *J. Am. Chem. Soc.* **1987**, *109* (13), 4064.
- (82) Lawless, M. J. Site-Directed Cu²⁺ Labeling of DNA and Proteins, University of Pittsburgh, 2018.
- (83) Shin, B. Electron Spin Resonance Analysis of Cu(II) Coordination in Alzheimer's Disease-Related Peptides, University of Pittsburgh, 2011.
- (84) McCracken, J.; Desai, P. R.; Papadopoulos, N. J.; Villafranca, J. J.; Peisach, J. *Biochemistry*

- 1988**, 27, 4133.
- (85) Hernandez-Guzman, J.; Sun, L.; Mehta, A. K.; Dong, J.; Lynn, D. G.; Warncke, K. *ChemBioChem* **2013**, 14 (14), 1762.
- (86) Gunderson, W. A.; Hernández-Guzmán, J.; Karr, J. W.; Sun, L.; Szalai, V. A.; Warncke, K. *J. Am. Chem. Soc.* **2012**, 134 (44), 18330.
- (87) Lee, H. I.; Doan, P. E.; Hoffman, B. M. *J. Magn. Reson.* **1999**, 140 (1), 91.
- (88) Flanagan, H. L.; Singel, D. J. *J. Chem. Phys.* **1987**, 87 (10), 5606.
- (89) Shin, B. K.; Saxena, S. *J. Phys. Chem. A* **2011**, 115, 9590.
- (90) Carmieli, R.; Papo, N.; Zimmermann, H.; Potapov, A.; Shai, Y.; Goldfarb, D. *Biophys. J.* **2006**, 90 (2), 492.
- (91) Lawless, M. J.; Ghosh, S.; Cunningham, T. F.; Shimshi, A.; Saxena, S. *Phys. Chem. Chem. Phys.* **2017**, 19 (31), 20959.
- (92) Höfer, P.; Grupp, A.; Nebenführ, H.; Mehring, M. *Chem. Phys. Lett.* **1986**, 132 (3), 279.
- (93) Höfer, P. *J. Magn. Reson. Ser. A* **1994**, 111 (1), 77.
- (94) Santangelo, M. G.; Medina-Molner, A.; Schweiger, A.; Mitrikas, G.; Spingler, B. *J. Biol. Inorg. Chem.* **2007**, 12 (6), 767.
- (95) Burns, C. S.; Aronoff-Spencer, E.; Dunham, C. M.; Lario, P.; Avdievich, N. I.; Antholine, W. E.; Olmstead, M. M.; Vrieland, A.; Gerfen, G. J.; Peisach, J.; Scott, W. G.; Millhauser, G. L. *Biochemistry* **2002**, 41 (12), 3991.
- (96) Dorlet, P.; Gambarelli, S.; Faller, P.; Hureau, C. *Angew. Chemie Int. Ed.* **2009**, 121 (49), 9437.
- (97) Yang, Z. Copper Ion-Based Electron Spin Resonance Spectroscopic Rulers, University of Pittsburgh, 2004.

- (98) Jeschke, G.; Panek, G.; Godt, A.; Bender, A.; Paulsen, H. *Appl. Magn. Reson* **2004**, *26*, 223.
- (99) Jeschke, G. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419.
- (100) Fanucci, G. E.; Cafiso, D. S. *Curr. Opin. Struct. Biol.* **2006**, *16* (5), 644.
- (101) Chiang, Y. W.; Borbat, P. P.; Freed, J. H. *J. Magn. Reson.* **2005**, *172* (2), 279.
- (102) Bowman, M. K.; Maryasov, A. G.; Kim, N.; DeRose, V. J. *Appl. Magn. Reson.* **2004**, *26* (1–2), 23.
- (103) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. *Appl. Magn. Reson.* **2006**, *30*, 473.
- (104) Sen, K. I.; Logan, T. M.; Fajer, P. G. *Biochemistry* **2007**, *46* (41), 11639.
- (105) Brandon, S.; Beth, A. H.; Hustedt, E. J. *J. Magn. Reson.* **2012**, *218*, 93.
- (106) Stein, R. A.; Beth, A. H.; Hustedt, E. J. *Methods Enzymol.* **2015**, *563*, 531.
- (107) Edwards, T. H.; Stoll, S. *J. Magn. Reson.* **2016**, *270*, 87.
- (108) Srivastava, M.; Freed, J. H. *J. Phys. Chem. A* **2019**, *123* (1), 359.
- (109) Srivastava, M.; Georgieva, E. R.; Freed, J. H. *J. Phys. Chem. A* **2017**, *121* (12), 2452.
- (110) Worswick, S. G.; Spencer, J. A.; Jeschke, G.; Kuprov, I. *Sci. Adv.* **2018**, *4* (8), eaat5218.
- (111) Parker, S. C.; Hansen, L.; Abaan, H. O.; Tullius, T. D.; Margulies, E. H. *Science* **2009**, *324*, 389.
- (112) Rohs, R.; West, S. M.; Sosinsky, A.; Liu, P.; Mann, R. S.; Honig, B. *Nature* **2009**, *461*, 1248.
- (113) Harteis, S.; Schneider, S. *Int. J. Mol. Sci.* **2014**, *15*, 12335.
- (114) Reginsson, G. W.; Shelke, S. A.; Rouillon, C.; White, M. F.; Sigurdsson, S. T.; Schiemann, O. *Nucleic Acids Res.* **2013**, *41* (1), e11.
- (115) Krumkacheva, O. A.; Shevelev, G. Y.; Lomzov, A. A.; Dyrkheeva, N. S.; Kuzhelev, A. A.;

- Koval, V. V.; Tormyshev, V. M.; Polienko, Y. F.; Fedin, M. V.; Pyshnyi, D. V.; Lavrik, O. I.; Bagryanskaya, E. G. *Nucleic Acids Res.* **2019**, *47* (15), 7767.
- (116) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67.
- (117) Hwang, D. S.; Kornberg, A. *J. Biol. Chem.* **1992**, *267*, 23083.
- (118) Kitayner, M.; Rozenberg, H.; Rohs, R.; Suad, O.; Rabinovich, D.; Honig, B.; Shakked, Z. *Nat. Struct. Mol. Biol.* **2010**, *17*, 423.
- (119) Chuo, S. W.; Wang, L. P.; Britt, R. D.; Goodin, D. B. *Biochemistry* **2019**, *58* (18), 2353.
- (120) Park, S. Y.; Borbat, P. P.; Gonzalez-Bonet, G.; Bhatnagar, J.; Pollard, A. M.; Freed, J. H.; Bilwes, A. M.; Crane, B. R. *Nat. Struct. Mol. Biol.* **2006**, *13* (5), 400.
- (121) Borbat, P. P.; Freed, J. H. *Methods Enzymol.* **2007**, *423*, 52.
- (122) Evans, E. G. B.; Pushie, M. J.; Markham, K. A.; Lee, H.-W.; Millhauser, G. L. *Structure* **2016**, *24*, 1057.
- (123) Ceccon, A.; Schmidt, T.; Tugarinov, V.; Kotler, S. A.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2018**, *140* (20), 6199.
- (124) Boura, E.; Rózycki, B.; Herrick, D. Z.; Chung, H. S.; Vecer, J.; Eaton, W. A.; Cafiso, D. S.; Hummer, G.; Hurley, J. H.; Rozycki, B.; Herrick, D. Z.; Chung, H. S.; Vecer, J.; Eaton, W. A.; Cafiso, D. S.; Hummer, G.; Hurley, J. H.; Rózycki, B.; Herrick, D. Z.; Chung, H. S.; Vecer, J.; Eaton, W. A.; Cafiso, D. S.; Hummer, G.; Hurley, J. H. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (23), 9437.
- (125) Duss, O.; Yulikov, M.; Jeschke, G.; Allain, F. H. T. *Nat. Commun.* **2014**, *5* (1), 3669.
- (126) Nicklisch, S. C. T.; Wunnicke, D.; Borovykh, I. V.; Morbach, S.; Klare, J. P.; Steinhoff, H. J.; Krämer, R. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818* (3), 359.
- (127) Shelke, S. A.; Sigurdsson, S. T. *Eur. J. Org. Chem.* **2012**, *2012* (12), 2291.

- (128) Shelke, S. A.; Sigurdsson, S. T. In *Modified Nucleic Acids*; Springer, 2016; Vol. 31, pp 159–187.
- (129) Schiemann, O.; Piton, N.; Mu, Y.; Stock, G.; Engels, J. W.; Prisner, T. F. *J. Am. Chem. Soc.* **2004**, *126* (18), 5722.
- (130) Sicoli, G.; Wachowius, F.; Bennati, M.; Hobartner, C. *Angew. Chemie Int. Ed.* **2010**, *49*, 6443.
- (131) Babaylova, E. S.; Ivanov, A. V; Malygin, A. A.; Vorobjeva, M. A.; Venyaminova, A. G.; Polienko, Y. F.; Kirilyuk, I. A.; Krumkacheva, O. A.; Fedin, M. V; Karpova, G. G.; Bagryanskaya, E. G. *Org. Biomol. Chem.* **2014**, *12*, 3129.
- (132) Saha, S.; Jagtap, A. P.; Sigurdsson, S. T. *Chem. Commun.* **2015**, *51* (66), 13142.
- (133) Halbmaier, K.; Seikowski, J.; Tkach, I.; Hobartner, C.; Sezer, D.; Bennati, M. *Chem. Sci.* **2016**, *7*, 3172.
- (134) Frolow, O.; Endeward, B.; Schiemann, O.; Prisner, T. F.; Engels, J. W. *Nucleic Acids Symp. Ser.* **2008**, *52* (1), 153.
- (135) Cai, Q.; Kusnetzow, A. K.; Hubbell, W. L.; Haworth, I. S.; Gacho, G. P. C.; Van Eps, N.; Hideg, K.; Chambers, E. J.; Qin, P. Z. *Nucleic Acids Res.* **2006**, *34* (17), 4722.
- (136) Nguyen, P. H.; Popova, A.; Hideg, K.; Qin, P. Z. *BMC Biophys.* **2016**, *8*, 6.
- (137) Popova, A. M.; Kálai, T.; Hideg, K.; Qin, P. Z. *Biochemistry* **2009**, *48*, 8540.
- (138) Qin, P. Z.; Haworth, I. S.; Cai, Q.; Kusnetzow, A. K.; Grant, G. P. G.; Price, E. A.; Sowa, G. Z.; Popova, A.; Herreros, B.; He, H. *Nat. Protoc.* **2007**, *2* (10), 2354.
- (139) Haugland, M. M.; El-Sagheer, A. H.; Porter, R. J.; Pena, J.; Brown, T.; Anderson, E. A.; Lovett, J. E. *J. Am. Chem. Soc.* **2016**, *138*, 9069.
- (140) Fedin, M. V.; Shevelev, G. Y.; Pyshnyi, D. V.; Tormyshev, V. M.; Jeschke, G.; Yulikov,

- M.; Bagryanskaya, E. G. *Phys. Chem. Chem. Phys.* **2016**, *18* (42), 29549.
- (141) Shevelev, G. Y.; Krumkacheva, O.; Lomzov, A. A.; Kuzhelev, A. A.; Trukhin, D. V.; Rogozhnikova, O. Y.; Tormyshev, V. M.; Pyshnyi, D. V.; Fedin, M. V.; Bagryanskaya, E. G. *J. Phys. Chem. B* **2015**, *119*, 13641.
- (142) Marko, A.; Denysenkov, V.; Margraf, D.; Cekan, P.; Schiemann, O.; Sigurdsson, S. T.; Prisner, T. F. *J. Am. Chem. Soc.* **2011**, *133*, 13375.
- (143) Grytz, C. M.; Marko, A.; Cekan, P.; Sigurdsson, S. T.; Prisner, T. F. *Phys. Chem. Chem. Phys.* **2016**, *18* (4), 2993.
- (144) Shevelev, G. Y.; Krumkacheva, O. A.; Lomzov, A. A.; Kuzhelev, A. A.; Rogozhnikova, O. Y.; Trukhin, D. V.; Troitskaya, T. I.; Tormyshev, V. M.; Fedin, M. V.; Pyshnyi, D. V.; Bagryanskaya, E. G. *J. Am. Chem. Soc.* **2014**, *136* (28), 9874.
- (145) Kuzhelev, A. A.; Krumkacheva, O. A.; Shevelev, G. Y.; Yulikov, M.; Fedin, M. V.; Bagryanskaya, E. G. *Phys. Chem. Chem. Phys.* **2018**, *20* (15), 10224.
- (146) Donohue, M. P.; Szalai, V. A. *Phys. Chem. Chem. Phys.* **2016**, *18* (22), 15447.
- (147) Wojciechowski, F.; Gross, A.; Holder, I. T.; Knorr, L.; Drescher, M.; Hartig, J. S.; Groß, A.; Holder, I. T.; Knörr, L.; Drescher, M.; Hartig, J. S. *Chem. Comm.* **2015**, *51* (72), 13850.
- (148) Engelhard, D. M.; Meyer, A.; Berndhäuser, A.; Schiemann, O.; Clever, G. H. *Chem. Commun.* **2018**, *54* (54), 7455.
- (149) Song, Y.; Meade, T. J.; Astashkin, A. V.; Klein, E. L.; Enemark, J. H.; Raitsimring, A. J. *Magn. Reson.* **2011**, *210* (1), 59.
- (150) Cantor, C. R.; Tinoco, I. *Biopolymers* **1967**, *5* (9), 821.
- (151) Richards, E. G. *Handb. Biochem. Mol. Biol. 3rd Edn.* **1975**, *1*, 596.
- (152) Senior, M.; Jones, R. A.; Breslauer, K. J. *Biochemistry* **1988**, *27* (10), 3879.

- (153) Stoll, S.; Schweiger, A. *J. Magn. Reson.* **2006**, *178* (1), 42.
- (154) Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W. *J. Magn. Reson.* **2000**, *142* (2), 331.
- (155) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; Portella, G.; Battistini, F.; Gelpí, J. L.; González, C.; Vendruscolo, M.; Laughton, C. A.; Harris, S. A.; Case, D. A.; Orozco, M. *Nat. Methods* **2016**, *13* (1), 55.
- (156) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
- (157) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79* (2), 926.
- (158) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98* (12), 10089.
- (159) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. . *J. Comput. Phys.* **1977**, *23* (3), 327.
- (160) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14* (1), 33.
- (161) Aldakheel, F. *Coll. Sci. Heal. Theses Diss.* **2017**, 248.
- (162) Zhou, W.; Saran, R.; Liu, J. *Chem. Rev.* **2017**, *117* (12), 8272.
- (163) Izatt, R. M.; Christensen, J. J.; Rytting, J. H. *Chem. Rev.* **1971**, *71* (5), 439.
- (164) Silva, K. I.; Saxena, S. *J. Phys. Chem. B* **2013**, *117* (32), 9386.
- (165) Gamble Jarvi, A.; Ranguelova, K.; Ghosh, S.; Weber, R. T.; Saxena, S. *J. Phys. Chem. B* **2018**, *122* (47), 10669.
- (166) Bowen, A. M.; Jones, M. W.; Lovett, J. E.; Gaule, T. G.; McPherson, M. J.; Dilworth, J. R.; Timmel, C. R.; Harmer, J. R. *Phys. Chem. Chem. Phys.* **2016**, *18*, 5981.
- (167) Abé, C.; Klose, D.; Dietrich, F.; Ziegler, W. H.; Polyhach, Y.; Jeschke, G.; Steinhoff, H. J. *J. Magn. Reson.* **2012**, *216*, 53.

- (168) Schiemann, O.; Cekan, P.; Margraf, D.; Prisner, T. F.; Sigurdsson, S. T. *Angew. Chemie Int. Ed.* **2009**, *48* (18), 3292.
- (169) Bode, B. E.; Plackmeyer, J.; Prisner, T. F.; Schiemann, O. *J. Phys. Chem. A* **2008**, *112* (23), 5064.
- (170) Denysenkov, V. P.; Prisner, T. F.; Stubbe, J.; Bennati, M. *Proc. Natl. Acad. Sci. USA* **2006**, *103* (36), 13386.
- (171) Sarver, J. L.; Silva, K. I.; Saxena, S. *Appl. Magn. Reson.* **2013**, *44* (5), 583.
- (172) Breitgoff, F. D.; Keller, K.; Qi, M.; Klose, D.; Yulikov, M.; Godt, A.; Jeschke, G. *J. Magn. Reson.* **2019**, *308*, 106560.
- (173) Olson, W. K.; Gorin, A. A.; Lu, X. J.; Hock, L. M.; Zhurkin, V. B. *Proc. Natl. Acad. Sci. USA* **1998**, *95* (19), 11163.
- (174) Delano, W. L. *PyMOL: An Open-Source Molecular Graphics Tool*; 2002.
- (175) Choi, K.-Y.; Ryu, H.; Sung, N.-D.; Suh, M. *J. Chem. Crystallogr.* **2003**, *33* (12), 947.
- (176) Mathew-Fenn, R. S.; Das, R.; Harbury, P. A. B. *Science* **2008**, *322*, 446.
- (177) Polyhach, Y.; Bordignon, E.; Jeschke, G. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356.
- (178) Jeschke, G. *Protein Sci.* **2018**, *27* (1), 76.
- (179) Hagelueken, G.; Ward, R.; Naimsith, J. H.; Schiemann, O.; Naismith, J. H.; Schiemann, O. *Appl. Magn. Reson.* **2012**, *42* (3), 377.
- (180) Hirst, S. J.; Alexander, N. S.; Mchaourab, H. S.; Meiler, J. *J. Struct. Biol.* **2011**, *173* (3), 506.
- (181) Reichel, K.; Stelzl, L. S.; Köfinger, J.; Hummer, G. *J. Phys. Chem. Lett.* **2018**, *9* (19), 5748.
- (182) van Dijk, M.; Bonvin, A. M. J. J. *Nucleic Acids Res.* **2009**, *37* (Web Server), W235.
- (183) Flick, K. E.; Jurica, M. S.; Monnat, R. J.; Stoddard, B. L. *Nature* **1998**, *394* (6688), 96.

- (184) Milikisyants, S.; Scarpelli, F.; Finiguerra, M. G.; Ubbink, M.; Huber, M. *J. Magn. Reson.* **2009**, *201*, 48.
- (185) Jeschke, G.; Pannier, M.; Godt, A.; Spiess, H. W. *Chem. Phys. Lett.* **2000**, *331* (2–4), 243.
- (186) Milov, A. D.; Maryasov, A. G.; Tsvetkov, Y. D. *Appl. Magn. Reson.* **1998**, *15* (1), 107.
- (187) Kulik, L. V. V.; Dzuba, S. A. A.; Grigoryev, I. A. A.; Tsvetkov, Y. D. *Chem. Phys. Lett.* **2001**, *343* (3–4), 315.
- (188) Grote, M.; Polyhach, Y.; Jeschke, G.; Steinhoff, H.; Schneider, E.; Bordignon, E. *J. Biol. Chem.* **2009**, *284* (26), 17521.
- (189) Hagelueken, G.; Ingledew, W. J.; Huang, H.; Petrovic-Stojanovska, B.; Whitfield, C.; ELMkami, H.; Schiemann, O.; Naismith, J. H. *Angew. Chemie Int. Ed.* **2009**, *48* (16), 2904.
- (190) Drescher, M.; Huber, M.; Subramaniam, V. *ChemBioChem* **2012**, *13* (6), 761.
- (191) Joseph, B.; Sikora, A.; Bordignon, E.; Jeschke, G.; Cafiso, D. S.; Prisner, T. F. *Angew. Chemie Int. Ed.* **2015**, *54* (21), 6196.
- (192) Puljung, M. C.; DeBerg, H. A.; Zagotta, W. N.; Stoll, S. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9816.
- (193) Sahu, I. D.; Dixit, G.; Reynolds, W.; Harding, B.; Jaycox, C.; Faleel, F. D. M.; McCarrick, R. M.; Sanders, C. R.; Lorigan, G. A. *Biophys. J.* **2019**, *116* (3), 26a.
- (194) Schmidt, T.; Schwieters, C. D.; Clore, G. M. *Proc. Natl. Acad. Sci. USA* **2019**, *116* (36), 17809.
- (195) Giannoulis, A.; Feintuch, A.; Barak, Y.; Mazal, H.; Albeck, S.; Unger, T.; Yang, F.; Su, X. C.; Goldfarb, D. *Proc. Natl. Acad. Sci. USA* **2020**, *117* (1), 395.
- (196) Prior, C.; Danilane, L.; Oganessian, V. S. *Phys. Chem. Chem. Phys.* **2018**, *20* (19), 13461.
- (197) Van Doorslaer, S. In *eMagRes*; John Wiley & Sons, Ltd: Chichester, UK, 2017; Vol. 6, pp

51–70.

- (198) Fauth, J. M.; Schweiger, A.; Braunschweiler, L.; Forrer, J.; Ernst, R. R. *J. Magn. Reson.* **1986**, *66*, 74.
- (199) Gemperle, C.; Aebli, G.; Schweiger, A.; Ernst, R. R. *J. Magn. Reson.* **1990**, *88*, 241.
- (200) Fábregas Ibáñez, L.; Soetbeer, J.; Klose, D.; Tinzl, M.; Hilvert, D.; Jeschke, G. *J. Magn. Reson.* **2019**, *307*, 106576.
- (201) Frisch, M. J.; Trucks, G. W.; Sclegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; M. J. Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revis. C.01, Gaussian Inc, Wellingt.* **2016**.
- (202) Schosseler, P. M.; Wehrli, B.; Schweiger, A. *Inorg. Chem.* **1997**, *36* (20), 4490.
- (203) Pöpl, A.; Kevan, L. *J. Phys. Chem.* **1996**, *100* (9), 3387.
- (204) Gophane, D. B.; Sigurdsson, S. T. *Chem. Commun.* **2013**, *49* (10), 999.
- (205) Galindo-Murillo, R.; Roe, D. R.; Cheatham, T. E. *Biochim. Biophys. Acta* **2015**, *1850* (5),

1041.

- (206) von Hippel, P. H.; Johnson, N. P.; Marcus, A. H. *Biopolymers* **2013**, *99* (12), 923.
- (207) Frank-Kamenetskii, M. D.; Prakash, S. *Phys. Life Rev.* **2014**, *11* (2), 153.
- (208) Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H.; Grigoryants, M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H. *J. Magn. Reson.* **2012**, *216*, 69.
- (209) Sinibaldi, R.; Ortore, M. G.; Spinozzi, F.; Carsughi, F.; Frielinghaus, H.; Cinelli, S.; Onori, G.; Mariani, P. *J. Chem. Phys.* **2007**, *126* (23), 235101.
- (210) Yang, Z.; Kise, D.; Saxena, S. *J. Phys. Chem. B* **2010**, *114* (18), 6165.
- (211) Marko, A.; Margraf, D.; Cekan, P.; Sigurdsson, S. T.; Schiemann, O.; Prisner, T. F. *Phys. Rev. E* **2010**, *81* (2), 021911.
- (212) Bowen, A. M.; Jones, M. W.; Lovett, J. E.; Gaule, T. G.; McPherson, M. J.; Dilworth, J. R.; Timmel, C. R.; Harmer, J. R. *Phys. Chem. Chem. Phys.* **2016**, *18* (8), 5981.
- (213) Lovett, J. E.; Bowen, A. M.; Timmel, C. R.; Jones, M. W.; Dilworth, J. R.; Caprotti, D.; Bell, S. G.; Wong, L. L.; Harmer, J. *Phys. Chem. Chem. Phys.* **2009**, *11* (31), 6840.
- (214) Neese, F. *WIREs Comput. Mol. Sci.* **2012**, *2* (1), 73.
- (215) Neese, F. *WIREs Comput. Mol. Sci.* **2018**, *8* (1), e1327.
- (216) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. *Biophys. J.* **2015**, *109* (8), 1528.
- (217) van Amsterdam, I. M. C.; Ubbink, M.; Canters, G. W.; Huber, M. *Angew. Chemie Int. Ed.* **2003**, *42* (1), 62.
- (218) Yang, Z.; Becker, J.; Saxena, S. *J. Magn. Reson.* **2007**, *188* (2), 337.

- (219) Raitsimring, A. M.; Gunanathan, C.; Potapov, A.; Efremenko, I.; Martin, J. M.; Milstein, D.; Goldfarb, D. *J. Am. Chem. Soc.* **2007**, *129*, 14138.
- (220) Goldfarb, D. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 9685.
- (221) Abdelkader, E. H.; Lee, M. D.; Feintuch, A.; Cohen, M. R.; Swarbrick, J. D.; Otting, G.; Graham, B.; Goldfarb, D. *J. Phys. Chem. Lett.* **2015**, *6*, 5016.
- (222) Barthelmes, D.; Granz, M.; Barthelmes, K.; Allen, K. N.; Imperiali, B.; Prisner, T. F.; Schwalbe, H. *J. Biol. NMR* **2015**, *63*, 275.
- (223) Doll, A.; Qi, M.; Wili, N.; Pribitzer, S.; Godt, A.; Jeschke, G. *J. Magn. Reson.* **2015**, *259*, 153.
- (224) Mascali, F. C.; Ching, H. Y. V.; Rasia, R. M.; Un, S.; Tabares, L. *Angew. Chemie Int. Ed.* **2016**, *128*, 11207.
- (225) Yulikov, M.; Lueders, P.; Warsi, M. F.; Chechik, V.; Jeschke, G. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10732.
- (226) Banerjee, D.; Yagi, H.; Huber, T.; Otting, G.; Goldfarb, D. *J. Phys. Chem. Lett.* **2012**, *3*, 157.
- (227) Martorana, A.; Yang, Y.; Zhao, Y.; Q., L.; Goldfarb, D. *Dalt. Trans.* **2015**, *44*, 20812.
- (228) Akhmetzyanov, D.; Plackmeyer, J.; Endeward, B.; Denysenkov, V.; Prisner, T. F. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6760.
- (229) Keller, K.; Zalibera, M.; Qi, M.; Koch, V.; Wegner, J.; Hintz, H.; Godt, A.; Jeschke, G.; Savitsky, A.; Yuliko, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 25120.
- (230) Kay, C. W. M.; El Mkami, H.; Cammack, R.; Evans, R. W.; Christopher W. M. Kay, H. E. M.; Richard Cammack, A.; Evans, R. W. *J. Am. Chem. Soc.* **2007**, *129*, 4868.
- (231) van Wonderen, J. H.; Kostrz, D. N.; Dennison, C.; MacMillan, F. *Angew. Chemie Int. Ed.*

- 2013**, 52, 1990.
- (232) Merz, G. E.; Borbat, P. P.; Pratt, A. J.; Getzoff, E. D.; Freed, J. H.; Crane, B. R. *Biophys. J.* **2014**, 107 (7), 1669.
- (233) Abdullin, D.; Florin, N.; Hagelueken, G.; Schiemann, O. *Angew. Chemie Int. Ed.* **2015**, 54 (6), 1827.
- (234) Welegedara, A. P.; Yang, Y.; Lee, M. D.; Swarbrick, J. D.; Huber, T.; Graham, B.; Goldfarb, D.; Otting, G. *Chem. - A Eur. J.* **2017**, 23 (48), 11694.
- (235) Borbat, P. P.; Freed, J. H. *Chem. Phys. Lett.* **1999**, 313 (1–2), 145.
- (236) Borbat, P. P.; Mchaourab, H. S.; Freed, J. H. *J. Am. Chem. Soc.* **2002**, 124 (19), 5304.
- (237) Todd, R. J.; Van Dam, M. E.; Casimiro, D.; Haymore, B. L.; Arnold, F. H. *Proteins* **1991**, 10 (2), 156.
- (238) Brandi-Blanco, M. P.; Benavides-Giménez, M. M. de; González-Pérez, J. M.; Choquesillo-Lazarte, D. *Acta Crystallogr. Sect. E* **2007**, 63 (6), m1678.
- (239) Raitsimring, A.; Astashkin, A. V.; Enemark, J. H.; Kaminker, I.; Goldfarb, D.; Walter, E. D.; Song, Y.; Meade, T. J. *Appl. Magn. Reson.* **2013**, 44 (6), 649.
- (240) Bucheli-Witschel, M.; Egli, T. *FEMS Microbiol. Rev.* **2001**, 25 (1), 69.
- (241) Lauer, S. A.; Nolan, J. P. *Cytometry* **2002**, 48 (3), 136.
- (242) Bresolin, I. T. L.; Borsoi-Ribeiro, M.; Tamashiro, W. M. S. C.; Augusto, E. F. P.; Vijayalakshmi, M. A.; Bueno, S. M. A. *Appl. Biochem. Biotechnol.* **2010**, 160 (7), 2148.
- (243) Diekmann, S.; Siegmund, G.; Roecker, A.; Klemm, D. O. *Cellulose* **2003**, 10 (1), 53.
- (244) Stadlbauer, S.; Riechers, A.; Späth, A.; König, B. *Chem. - A Eur. J.* **2008**, 14 (8), 2536.
- (245) Buglyó, P.; Kiss, T.; Dyba, M.; Jezowska-Bojczuk, M.; Kozłowski, H.; Bouhsina, S. *Polyhedron* **1997**, 16 (19), 3447.

- (246) Cunningham, T. F.; McGoff, M. S.; Sengupta, I.; Jaroniec, C. P.; Horne, W. S.; Saxena, S. *Biochemistry* **2012**, *51* (32), 6350.
- (247) Cunningham, T. F.; Pornsuwan, S.; Horne, W. S.; Saxena, S. *Protein Sci.* **2016**, *25* (5), 1049.
- (248) Mehlenbacher, M. R.; Bou-Abdallah, F.; Liu, X. X.; Melman, A. *Inorganica Chim. Acta* **2015**, *437*, 152.
- (249) Fauth, J. M.; Schweiger, A.; Braunschweiler, L.; Forrer, J.; Ernst, R. R. *J. Magn. Reson.* **1986**, *66*, 74.
- (250) Milov, A. D.; Tsvetkov, Y. D.; Formaggio, F.; Crisma, M.; Toniolo, C.; Raap, J. *J. Am. Chem. Soc.* **2000**, *112*, 3843.
- (251) Milov, A. D. D.; Ponomarev, A. B. B.; Tsvetkov, Y. D. D. *Chem. Phys. Lett.* **1984**, *110* (1), 67.
- (252) Bode, B. E.; Margraf, D.; Plackmeyer, J.; Durner, G.; Prisner, T. F.; Schiemann, O. *J. Am. Chem. Soc.* **2007**, *129*, 6736.
- (253) Tsvetkov, Y. D.; Milov, A. D.; Maryasov, A. G. *Russ. Chem. Rev.* **2008**, *77*, 487.
- (254) Clore, G. M.; Gronenborn, A. M. *Science* **1991**, *252* (5011), 1390.
- (255) Lian, L. Y.; Yang, J. C.; Derrick, J. P.; Sutcliffe, M. J.; Roberts, G. C. K.; Murphy, J. P.; Goward, C. R.; Atkinson, T. *Biochemistry* **1991**, *30*, 5335.
- (256) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721.
- (257) W. Trent Franks; Benjamin J. Wylie; Sara A. Stellfox, A.; Rienstra, C. M. *J. Am. Chem. Soc.* **2006**, *128* (10), 3154.
- (258) Wylie, B. J.; Sperling, L. J.; Nieuwkoop, A. J.; Franks, W. T.; Oldfield, E.; Rienstra, C. M. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (41), 16974.

- (259) Aasa, R.; Vanngard, T. *J. Magn. Reson.* **1975**, *19*, 308.
- (260) Regan, L. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 257.
- (261) Jiang, F.; McCracken, J.; Peisach, J. *J. Am. Chem. Soc.* **1990**, *112*, 9035.
- (262) Kevan, L. In *Time Domain Electron Spin Resonance*; John Wiley & Sons: New York, 1979; pp 279–341.
- (263) Le Trong, I.; Stenkamp, R. E.; Ibarra, C.; Atkins, W. M.; Adman, E. T. *Proteins Struct. Funct. Bioinforma.* **2002**, *48* (4), 618.
- (264) Tschaggelar, R.; Breitgoff, F. D.; Oberhänsli, O.; Qi, M.; Godt, A.; Jeschke, G. *Appl. Magn. Reson.* **2017**, *48* (11–12), 1273.
- (265) Donato, R.; Cannon, B. R.; Sorci, G.; Riuzzi, F.; Hsu, K.; Weber, D. J.; Geczy, C. L. *Curr Mol Med* **2013**, *13* (1), 24.
- (266) Gilston, B. A.; Skaar, E. P.; Chazin, W. J. *Sci China Life Sci* **2016**, *59* (8), 792.
- (267) Zackular, J. P.; Chazin, W. J.; Skaar, E. P. *J Biol Chem* **2015**, *290* (31), 18991.
- (268) Bresnick, A. R.; Weber, D. J.; Zimmer, D. B. *Nat Rev Cancer* **2015**, *15* (2), 96.
- (269) Manigrasso, M. B.; Juranek, J.; Ramasamy, R.; Schmidt, A. M. *Trends Endocrinol Metab* **2014**, *25* (1), 15.
- (270) Venegas, C.; Heneka, M. T. *J Leukoc Biol* **2017**, *101* (1), 87.
- (271) Brodersen, D. E.; Nyborg, J.; Kjeldgaard, M. *Biochemistry* **1999**, *38* (6), 1695.
- (272) Hayden, J. A.; Brophy, M. B.; Cunden, L. S.; Nolan, E. M. *J. Am. Chem. Soc.* **2013**, *135* (2), 775.
- (273) Glaser, R.; Harder, J.; Lange, H.; Bartels, J.; Christophers, E.; Schroder, J. M. *Nat Immunol* **2005**, *6* (1), 57.
- (274) Kehl-Fie, T. E.; Chitayat, S.; Hood, M. I.; Damo, S.; Restrepo, N.; Garcia, C.; Munro, K.

- A.; Chazin, W. J.; Skaar, E. P. *Cell Host Microbe* **2011**, *10* (2), 158.
- (275) Damo, S. M.; Kehl-Fie, T. E.; Sugitani, N.; Holt, M. E.; Rathi, S.; Murphy, W. J.; Zhang, Y.; Betz, C.; Hench, L.; Fritz, G.; Skaar, E. P.; Chazin, W. J. *Proc. Natl. Acad. Sci. USA* **2013**, *110* (10), 3841.
- (276) Haley, K. P.; Delgado, A. G.; Piazuolo, M. B.; Mortensen, B. L.; Correa, P.; Damo, S. M.; Chazin, W. J.; Skaar, E. P.; Gaddy, J. A. *Infect. Immun.* **2015**, *83* (7), 2944.
- (277) Shank, J. M.; Kelley, B. R.; Jackson, J. W.; Tweedie, J. L.; Franklin, D.; Damo, S. M.; Gaddy, J. A.; Murphy, C. N.; Johnson, J. G. *Infect Immun* **2018**, *86* (6).
- (278) Moroz, O. V.; Blagova, E. V.; Wilkinson, A. J.; Wilson, K. S.; Bronstein, I. B. *J. Mol. Biol.* **2009**, *391* (3), 536.
- (279) Wolf, R.; Howard, O. M.; Dong, H. F.; Voscopoulos, C.; Boeshans, K.; Winston, J.; Divi, R.; Gunsior, M.; Goldsmith, P.; Ahvazi, B.; Chavakis, T.; Oppenheim, J. J.; Yuspa, S. H. *J Immunol* **2008**, *181* (2), 1499.
- (280) Solioz, M.; Abicht, H. K.; Mermoud, M.; Mancini, S. *J Biol Inorg Chem* **2010**, *15* (1), 3.
- (281) White, C.; Lee, J.; Kambe, T.; Fritsche, K.; Petris, M. J. *J Biol Chem* **2009**, *284* (49), 33949.
- (282) Becker, K. W.; Skaar, E. P. *FEMS Microbiol Rev* **2014**, *38* (6), 1235.
- (283) Besold, A. N.; Gilston, B. A.; Radin, J. N.; Ramsomair, C.; Culbertson, E. M.; Li, C. X.; Cormack, B. P.; Chazin, W. J.; Kehl-Fie, T. E.; Culotta, V. C. *Infect Immun* **2017**.
- (284) Moroz, O. V.; Antson, A. A.; Grist, S. J.; Maitland, N. J.; Dodson, G. G.; Wilson, K. S.; Lukanidin, E.; Bronstein, I. B. *Acta Crystallogr. D. Biol. Crystallogr.* **2003**, *59* (Pt 5), 859.
- (285) Nishikawa, T.; Lee, I. S.; Shiraishi, N.; Ishikawa, T.; Ohta, Y.; Nishikimi, M. *J Biol Chem* **1997**, *272* (37), 23037.
- (286) Schafer, B. W.; Fritschy, J. M.; Murmann, P.; Troxler, H.; Durussel, I.; Heizmann, C. W.;

- Cox, J. A. *J Biol Chem* **2000**, 275 (39), 30623.
- (287) Sivaraja, V.; Kumar, T. K.; Rajalingam, D.; Graziani, I.; Prudovsky, I.; Yu, C. *Biophys. J.* **2006**, 91 (5), 1832.
- (288) Wheeler, L. C.; Harms, M. J. *BMC Biophys* **2017**, 10, 8.
- (289) Moroz, O. V.; Antson, A. A.; Dodson, E. J.; Burrell, H. J.; Grist, S. J.; Lloyd, R. M.; Maitland, N. J.; Dodson, G. G.; Wilson, K. S.; Lukanidin, E.; Bronstein, I. B. *Acta Crystallogr D Biol Crystallogr* **2002**, 58 (Pt 3), 407.
- (290) Moroz, O. V.; Antson, A. A.; Murshudov, G. N.; Maitland, N. J.; Dodson, G. G.; Wilson, K. S.; Skibshoj, I.; Lukanidin, E. M.; Bronstein, I. B. *Acta Crystallogr D Biol Crystallogr* **2001**, 57 (Pt 1), 20.
- (291) Moroz, O. V.; Burkitt, W.; Wittkowski, H.; He, W.; Ianoul, A.; Novitskaya, V.; Xie, J.; Polyakova, O.; Lednev, I. K.; Shekhtman, A.; Derrick, P. J.; Bjoerk, P.; Foell, D.; Bronstein, I. B. *BMC Biochem* **2009**, 10 (1), 11.
- (292) Xie, J.; Burz, D. S.; He, W.; Bronstein, I. B.; Lednev, I.; Shekhtman, A. *J Biol Chem* **2007**, 282 (6), 4218.
- (293) Jackson, E.; Little, S.; Franklin, D. S.; Gaddy, J. A.; Damo, S. M. *J Vis Exp* **2017**, 123, 55557.
- (294) Fielding, A. J.; Fox, S.; Millhauser, G. L.; Chattopadhyay, M.; Kroneck, P. M. H.; Fritz, G.; Eaton, G. R.; Eaton, S. S. *J. Magn. Reson.* **2006**, 179 (1), 92.
- (295) Ostendorp, T.; Diez, J.; Heizmann, C. W.; Fritz, G. *Biochim Biophys Acta* **2011**, 1813 (5), 1083.
- (296) Murray, J. I.; Tonkin, M. L.; Whiting, A. L.; Peng, F.; Farnell, B.; Cullen, J. T.; Hof, F.; Boulanger, M. J. *BMC Struct Biol* **2012**, 12, 16.

- (297) Lin, H.; Andersen, G. R.; Yatime, L. *BMC Struct Biol* **2016**, *16* (1), 8.
- (298) Shin, B.; Saxena, S. *Biochemistry* **2008**, *47* (35), 9117.
- (299) Klotz, I. M. *J. Biol. Chem.* **2004**, *279* (1), 1.
- (300) Klotz, I. M. *Arch. Biochem.* **1946**, *9*, 109.
- (301) Gesztelyi, R.; Zsuga, J.; Kemeny-Beke, A.; Varga, B.; Juhasz, B.; Tosaki, A. *Arch. Hist. Exact Sci.* **2012**, *66* (4), 427.
- (302) Stefan, M. I.; Le Novère, N. *PLoS Comput. Biol.* **2013**, *9* (6), e1003106.
- (303) Nelson, D. L. (David L.; Cox, M. M.; Lehninger, A. L. *Lehninger principles of biochemistry*; W.H. Freeman and Company, 2013.
- (304) Goldfarb, D.; Fauth, J.-M.; Tor, Y.; Shanzer, A. *J. Am. Chem. Soc.* **1991**, *113* (98), 1941.
- (305) Brophy, M. B.; Hayden, J. A.; Nolan, E. M. *J. Am. Chem. Soc.* **2012**, *134* (43), 18089.
- (306) Cunden, L. S.; Gaillard, A.; Nolan, E. M. *Chem Sci* **2016**, *7* (2), 1338.
- (307) Korndorfer, I. P.; Brueckner, F.; Skerra, A. *J Mol Biol* **2007**, *370* (5), 887.
- (308) Leukert, N.; Vogl, T.; Strupat, K.; Reichelt, R.; Sorg, C.; Roth, J. *J Mol Biol* **2006**, *359* (4), 961.
- (309) Ostendorp, T.; Leclerc, E.; Galichet, A.; Koch, M.; Demling, N.; Weigle, B.; Heizmann, C. W.; Kroneck, P. M.; Fritz, G. *EMBO J.* **2007**, *26* (16), 3868.
- (310) Hung, K.-W.; Hsu, C.-C.; Yu, C. *J Biomol NMR* **2013**, *57*, 313.