# Site-Directed Cu$^{2+}$-Labeling of DNA and Proteins

by

**Matthew J. Lawless**

Bachelor of Arts, University of Delaware, 2012

Submitted to the Gradate Faculty of the Kenneth P. Dietrich

School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Matthew J. Lawless

It was defended on

January 3rd, 2018

and approved by

W. Seth Horne, Ph.D., Professor, Department of Chemistry

Sean Garrett-Roe, Ph.D., Professor, Department of Chemistry

Pei Tang, Ph.D., Professor, Department of Anesthesiology, Department of Pharmacology & Chemical Biology, Department of Computational and Systems Biology

Dissertation Director: Sunil Saxena, Ph.D., Professor, Department of Chemistry

# Site-Directed $Cu^{2+}$-Labeling of DNA and Proteins

Matthew J. Lawless, Ph.D.

University of Pittsburgh, 2018

Electron spin resonance (ESR) in combination with site-directed spin labeling can be used to determine structure, flexibility and conformational dynamics of proteins and DNA/RNA. This thesis focuses on using $Cu^{2+}$ as a spin label. First, we introduce a new method to incorporate $Cu^{2+}$ within DNA for the purpose of creating an accurate reporter of DNA structure. This method positions the paramagnetic label within the interior of the duplex as opposed to all current other labeling strategies. This methodology is also nucleotide and structure independent. Using this approach, the measured interspin distance is within 1 Å of the distance predicted by modeling and molecular dynamics simulations. This method is capable of reporting backbone-backbone distances without modeling.

$Cu^{2+}$-based strategies are also used to measure protein structure. We measure, by direct spectroscopic measurements, the optimum conditions to load $Cu^{2+}$ to engineered binding sites in both α-helices and β-sheets. The optimizing loading conditions lead to a two-fold increase in signal for the measurement of $Cu^{2+}$-$Cu^{2+}$ distances. The procedure is then used investigate the human glutathione S-transferase A1-1 protein. This enzyme promiscuously binds to and activates various forms of glutathione to reduce cellular radical species. We present a combination of nitroxide- and $Cu^{2+}$-based spin labeling to probe the conformation of the functionally important terminal helix. We find that the terminal helix exists in two distinct conformations, one of which

is largely dynamic. Such flexibility might be important for the high degree of substrate promiscuity. ESR measurements are then used as distance constraints to generate a model of the whole length protein.

We next developed strategies to perform structural measurements of proteins in their native in-cell environment. Measurements in-cell are challenging as the ESR signal quickly degrades within the cellular environment. To extend the applicability of ESR in-cell, first, we prolong the in-cell half-life of the typical ESR spin label with an oxidizing agent. Second, we use a combination of spectroscopic measurements to analyze the contribution of label degradation toward overall signal loss. Third, we collect time-based data to quantify the time required for molecular diffusion of a small globular protein within the cellular milieu.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# PREFACE

I came to Pittsburgh five and a half years ago as a boy wearing a backwards hat and usually opted not to wear sleeves. It was a difficult journey, but I write this preface as a man who is proud of whom he has become. Through this journey, I have met so many and owe my gratitude to each and every person who has helped me along the way. Let me start by saying thank you to my friend and advisor, Prof. Sunil Saxena. Trust me when I say, I am difficult to work with. Yet, Sunil never gave up on me, and pushed me to be the best that I could. No one has believed in me more, even in times when I didn't believe in myself. Certainly any advisor can generate a proficient Ph.D. scientist, but Sunil has done so much more for me. He has taught me not only about science, but what it truly means to be a professional and what it means to be an adult. Thank you for always putting up with me, Sunil, I will forever be grateful for all you have done.

I would next like to thank my committee members, Prof. W. Seth Horne, Prof. Sean Garrett-Roe and Prof. Pei Tang. I'd especially like to thank Prof. Horne and Prof. Tang for the years of collaborative work we put in together.

I'd also like to thank my many coworkers along the way. Ming Ji, Ishara Silva, and Tim Cunningham will always hold a special place in my heart. Without your help and guidance from day one, I would have never made it. I'd also like to thank Jessica Sarver, Monica Kinde and Nathan Tavenor for working with me on various collaborations, it was a privilege to work with

scientists so passionate about their studies. I'd like to thank my current colleagues still in the midst of their Ph.D. journey. Thank you to Austin Jarvi, Shreya Ghosh, Kevin Singewald and Xiaowei Du. Special thanks to Shreya for putting up with me working on all our crazy projects together.

Finally I'd like to thank my friends and family. It warms my heart to know that I always had a crowd cheering for me back home. To my sisters Lauren and Lindsey, thank you for always taking the time to talk to me. I always enjoyed picking up my phone and seeing a txt from you. And finally, thank you to my Mom and Dad. No one means more to me, and I could never have done this without your help. I have read and kept every letter that you ever sent. Thank you for all of your love and support over the years. We did it!

# 1.0     INTRODUCTION

## 1.1     ELECTRON SPIN RESONANCE

Electron spin resonance (ESR) has become a major tool used to determine the structure and conformational dynamics of biological macromolecules[3-9]. In the ESR approach, site-directed spin labeling (SDSL)[5] is used to position site-specific reporters within these molecules. ESR has no size limitation. As well, ESR does not require crystallization and can report on solution state structure and dynamics[10]. Furthermore, these spin labels are small and generally non-perturbative[11]. Thus, in combination with SDSL, ESR is an excellent technique to measure the structure and dynamics of large biological macromolecules such as protein-protein complexes, protein-DNA complexes and membrane bound proteins. Currently, there exist a wide variety of possible spin labels ranging from nitroxide radicals to paramagnetic metal ions. In this thesis, we develop new spin labeling methodologies using paramagnetic $Cu^{2+}$ for both proteins and DNA. We also implement these new techniques and show how these labels may be used to shed light on the function of an interesting cellular protein.

In Chapter 2 we introduce the use of $Cu^{2+}$ to determine structural constraints within a DNA duplex. Chapter 2 delves into the method of $Cu^{2+}$ labeling and its degree of DNA duplex perturbation, validation of site specific $Cu^{2+}$ labeling and finally its capability to generate distance constraints[12]. In Chapter 3, we then discuss the advancement of $Cu^{2+}$-based labeling of

proteins[13]. Specifically, we improve the overall signal of such measurements, and provide a comprehensive method of analysis to determine the extent of $Cu^{2+}$-labeling and optimum conditions for experimentation. These methods include a combination of both continuous wave (CW) ESR as well as pulsed ESR. Chapter 4 demonstrates these $Cu^{2+}$ based distance measurements in the human glutathione S-transferase enzyme. In particular, both nitroxide- and $Cu^{2+}$-based DEER constraints are used to elucidate the conformation of the terminal helix. The terminal helix was thought to exist between two distinct states, however the structure remained unresolved by NMR and X-ray crystallography for over a decade. Using precise distance constrains from nitroxide and $Cu^{2+}$-based ESR, we are able to confirm and model this hitherto unknown conformation.

This thesis will focus on all aspects of $Cu^{2+}$-based ESR and its use in elucidating structural information in biological macromolecules. In doing so, we will use several techniques such as CW-ESR, electron spin echo envelope modulation (ESEEM) and double electron-electron resonance (DEER). Therefore, it is essential to understand the basic premise behind such experiments. This chapter will serve to provide that information as well as a general overview of specific uses of such techniques with focus on $Cu^{2+}$-based ESR.

## 1.2    SPIN HAMILTONIAN

$Cu^{2+}$ is an ESR active species, meaning that it has an unpaired electron spin. This electron spin interacts with an applied magnetic field. The electron spin also interacts with any surrounding spins such as the $Cu^{2+}$ nucleus, which has a nuclear spin, $I = 3/2$. Other spins

typically present in biological systems are $^{14}$N and $^{1}$H. These interactions can be described by a spin Hamiltonian. Four terms comprise the spin Hamiltonian operator that describes any electron spin coupled to any $k$ nuclei (with non-zero spin): the electron Zeeman (ez) interaction, the nuclear Zeeman (nz) interaction, the electron-nuclear hyperfine (hf) interaction and the nuclear quadrupole (nq) interaction. The spin Hamiltonian operator is[14]:

$$\widehat{\mathcal{H}} = \beta_e B_0 \cdot \tilde{g} \cdot \hat{S} - \beta_n \sum_{i=1}^{k}(B_0 \cdot \tilde{g}_{i,n} \cdot \hat{I}_i) + \hbar \sum_{i=1}^{k}(\hat{S} \cdot \tilde{A}_i \cdot \hat{I}_i) + \hbar \sum_{i=1}^{k}(\hat{S} \cdot \tilde{Q}_i \cdot \hat{I}_i) \qquad (1\text{-}1)$$

where $\hat{S}$ and $\hat{I}_i$ are the electron and $i$th nuclear spin angular momentum operators, $\beta_e$ is the Bohr magneton, $\beta_n$ is the nuclear magneton, $B_0$ 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 > ½.

## 1.3    CONTINUOUS WAVE ESR

### 1.3.1 The CW Lineshape

Continuous Wave (CW) ESR spectroscopy on samples containing the paramagnetic metal ion $Cu^{2+}$ can provide information on four key aspects: the coordination geometry, directly coordinated atoms, the sample's $Cu^{2+}$ loading efficiency, and the percent contribution of multiple components in the case of different binding sites. In this section, we will take an in depth look as to how the CW spectrum supplies these key pieces of information.

First, we will look at what information the $Cu^{2+}$ CW-ESR lineshape can report on. The respective splitting and positioning of the features within the $Cu^{2+}$ CW spectrum are sensitive to the atoms that are liganded to the $Cu^{2+}$ atom[15]. For instance, a $Cu^{2+}$ with octahedral geometry has a directly coordinated environment of 6 atoms, 4 equatorially coordinated and 2 axially liganded. An example of this $Cu^{2+}$ coordination, is shown in Figure 1-1A with the atoms highlighted by dark shading. To understand the subtle nuances of the $Cu^{2+}$ CW-ESR spectrum, we must first look at the equations defining the system. The dominant interactions that affect the CW-ESR lineshape from Equation (1-1) are the electron Zeeman interaction (~9.6 GHz at X-band) and the hyperfine interaction between the electron spin and the nuclear spin of $Cu^{2+}$ (~450 MHz). The nuclear Zeeman at X-band frequencies has a much smaller magnitude of ~4 MHz. Furthermore, since we only detect transitions between electron spin states, this interaction does not alter the difference between energy levels of electron spin states and can be neglected. Furthermore the nuclear quadrupole interactions are small (< 3 MHz) and are rarely resolved in CW-ESR. Therefore, the nuclear quadrupole term can be neglected from Equation (2-1). Interactions with coordinated nitrogen nuclear spins, shown in Figure 1-1A, have an interaction of ~40 MHz and are sometimes not resolved. Neglecting both nuclear Zeeman and nuclear quadrupole terms gives the following[16]:

$$\hat{\mathcal{H}} = \beta_e B_0 \cdot \tilde{g} \cdot \hat{S} - \hbar \hat{S} \cdot \tilde{A} \cdot \hat{I} \tag{1-2}$$

If the rotation of the molecule is faster than the ESR timescale, all anisotropy is averaged out. Under the high field approximation, Equation (1-2) becomes:

$$\hat{\mathcal{H}} = g_{iso}\beta_e B_0 \hat{S}_z - A_{iso}\hat{S}_z\hat{I}_z \tag{1-3a}$$

$$E_{m_S,m_I} = g_{iso}\beta_e B_0 m_s - A_{iso}m_S m_I \tag{1-3b}$$
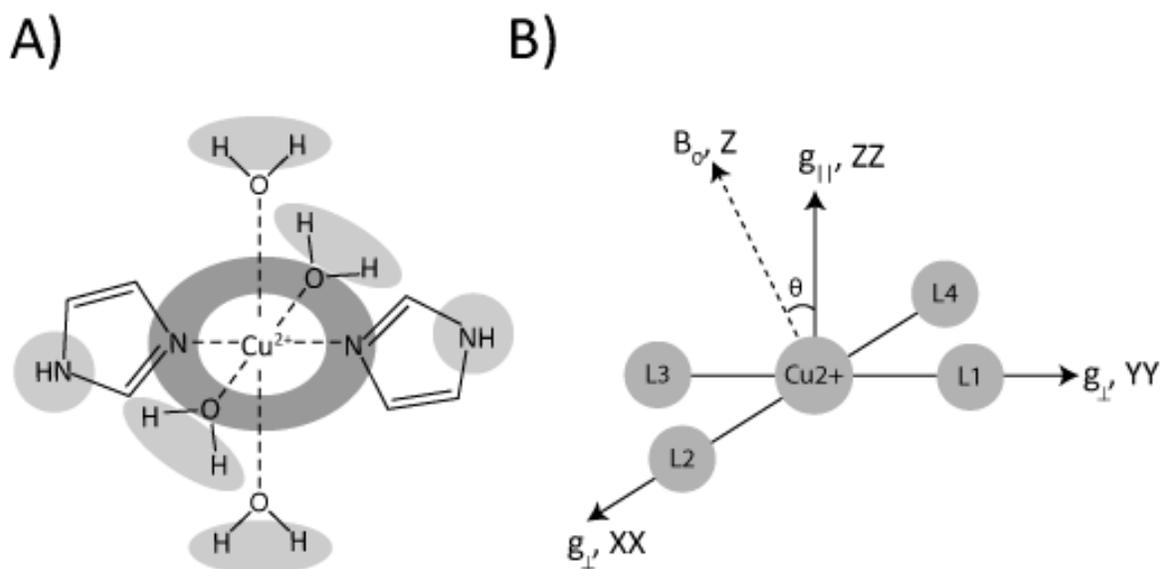
Figure 1-1. A) An example of a Type-II $Cu^{2+}$ center. Dark shaded region highlights directly coordinated equatorial ligands. Light shaded regions represent distally coordinated nuclei. B) Representation of a Type-II $Cu^{2+}$ center with four equatorial ligands (L1-4) within the lab frame denoted by the Z axis aligned with the magnetic field $B_0$. The molecular frame is defined by the g-tensor orientations (XX, YY and ZZ). $\theta$ represents the angle between the Z axis and ZZ axis.

Where $\hat{S}_z$ and $\hat{I}_z$ are the electron and nuclear spin angular moment operators along the z direction respectively, $E$ is the energy of the states, and $m_S$ and $m_I$ are the quantum numbers for the z component of electron spin and nuclear spin angular momentum operators, respectively. The z-axis is defined to be aligned with the externally applied magnetic field as shown in Figure 1-1B. For ESR experiments, only the electron spin is in resonance and therefore the selection rule is as follows:

$$\Delta m_S = \pm 1 \qquad \Delta m_I = 0 \tag{1-4}$$

Therefore, only the transitions between $m_S$ states are allowed. The difference in energy of these transitions for S = ½, and their associated frequencies, are:

$$\Delta E_{m_S,m_I} = h\nu = g_{iso}\beta_e B_0 - A_{iso}m_I \tag{1-5}$$

Since $I = 3/2$ for $Cu^{2+}$, four transitions are possible. These transitions, defined by Equations (1-3) and (1-5) are shown in Figure 1-2A. Equations (1-3) produce an isotropic CW-ESR spectrum with four features shown in Figure 1-2B. $g_{iso}$ and $A_{iso}$ are given by:

$$g_{iso} = \tfrac{1}{3}\left(g_{xx} + g_{yy} + g_{zz}\right) \tag{1-6a}$$

$$A_{iso} = \tfrac{1}{3}\left(A_{xx} + A_{yy} + A_{zz}\right) \tag{1-6b}$$

$g_{xx}$, $g_{yy}$, $g_{zz}$, $A_{xx}$, $A_{yy}$ and $A_{zz}$ refer to the principle components of the g- and A-tensor, respectively. The principle axis is shown in Figure 1-1B.

At low-temperature, or the rigid-limit, the $Cu^{2+}$ CW-ESR spectrum shows an interesting anisotropic lineshape that contains a plethora of information about the coordination environment of the metal ion[15]. At these temperatures, the $g$- and hyperfine tensors are not averaged out. The values of $g_{xx}$, $g_{yy}$, $g_{zz}$, $A_{xx}$, $A_{yy}$ and $A_{zz}$ depend on the coordination geometry. If the geometry of the $Cu^{2+}$ coordination is trigonal-planar or tetrahedral, often referred to as type-I $Cu^{2+}$, then $g_{xx}$
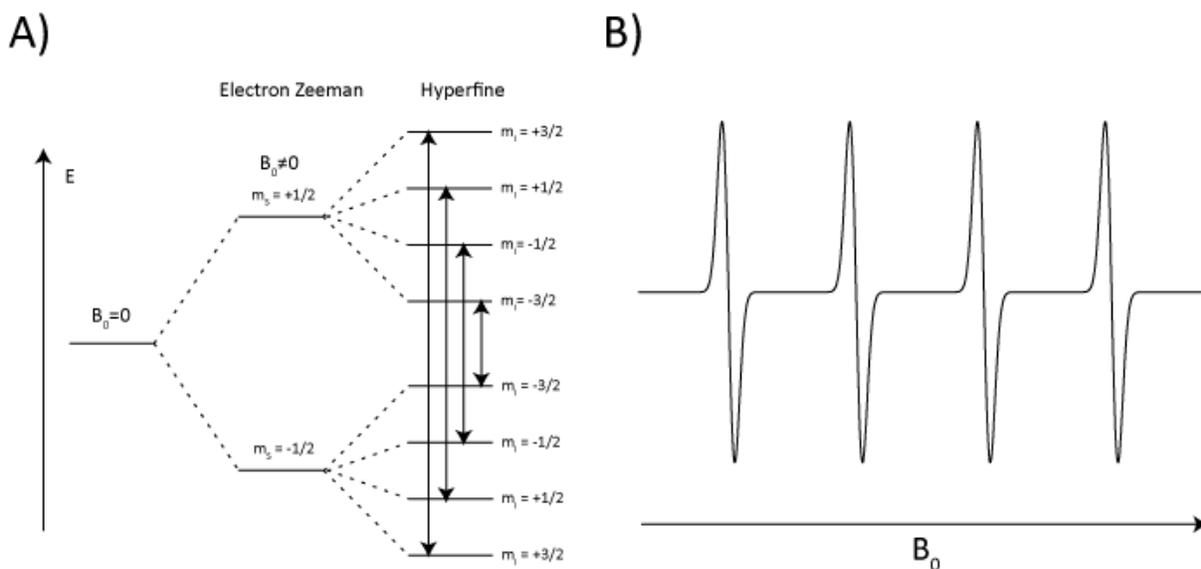
Figure 1-2. A) The $Cu^{2+}$ energy level diagram using isotropic values of g and A. Energy levels are degenerate in the absence of a magnetic field ($B_0$=0). Upon applying a magnetic field ($B_0 \neq 0$), there is a splitting of energy levels due to the electron Zeeman interaction. The nuclear spin of the $Cu^{2+}$ nucleus causes a further splitting due to the hyperfine interaction. The allowed transitions are illustrated by the black arrows. B) The four transitions of the isotropic $Cu^{2+}$ CW-ESR spectrum.

and $g_{yy}$ differ by up to 0.06 gauss and $A_{zz}$ is typically smaller than 100 G[17]. If the geometry is

octahedral, referred to as type-II Cu$^{2+}$, the $g_{xx} = g_{yy}$ and $A_{zz}$ is greater than 120 G[17]. This

thesis will largely consider type-II Cu$^{2+}$ systems. Here we denote $g_{xx} = g_{yy} = g_\perp$ and

that $g_{zz} = g_{||}$, and likewise for $A_\perp$ and $A_{||}$. This symmetry is usually denoted as axial symmetry.

To understand the powder CW-ESR lineshape for a type-II Cu$^{2+}$ system, we will, for

now, only consider an electron without any nuclear spin interaction. Without the hyperfine

interaction, and without averaging of $g_{iso}$, we rewrite the resonance condition defined by

Equation (1-5) as[18]:

$$hv = (g_e + \Delta g)\beta_e B_0 \qquad\qquad (1\text{-}7)$$

Where $g_e$ is the $g$-value of a free electron and $\Delta g$ represents a shift away from this value.

This shift occurs due to two factors[19]. First, is the interaction of the electron's orbital angular

momentum with the applied magnetic field. Second is a coupling between the electron's spin

angular momentum and orbital angular momentum. Both of these interactions produce a local

magnetic field experienced by the electron. The z-component of this local magnetic field, adds to

$B_0$. We account for this change in field using the factor $\Delta g$. This principle is similar to chemical

shift in NMR.

Next, let us consider an octahedral type-II Cu$^{2+}$ system within an applied magnetic field

as shown in Figure 1-3A. If the molecule is positioned within the magnetic field, such that the

xx,yy-plane is parallel to the applied magnetic field, $\Delta g$ have a different value than if the

molecule is positioned such that the zz axis is parallel to the field. This difference is because the

z-component of the local magnetic field will differ as the orientation of the orbital with respect to

the applied magnetic field is changed. Therefore, the resonant magnetic field will be different at

these two orientations. When the molecule is rotated between the xx- and zz-axis, the resonant
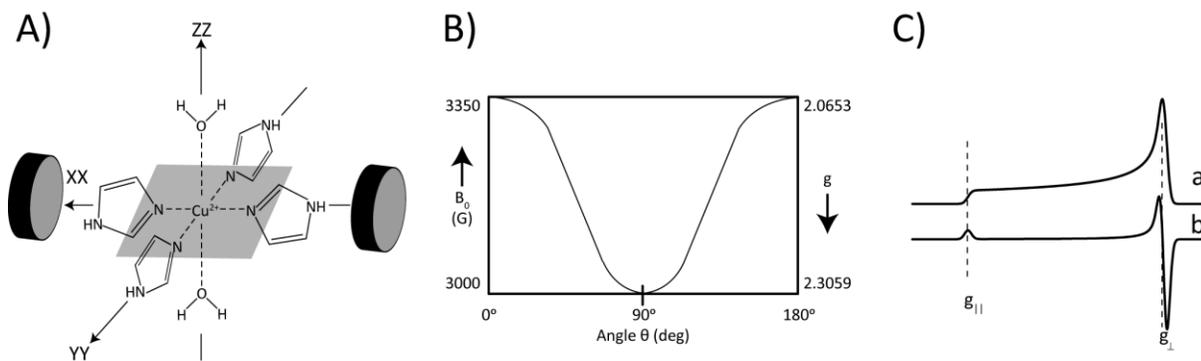
Figure 1-3. A) An axially symmetric type-II $Cu^{2+}$ system positioned within an applied magnetic field. B) A curve representative of how the resonant field from Equation (2-9) changes upon rotating the molecular axis with respect to the magnetic field. C) Trace (a) shows the absorption spectrum of the CW-ESR powder lineshape due to the electron Zeeman interaction for an axially symmetric system. Trace (b) shows the first derivative of the absorption spectrum which is the typical output of ESR spectrometers.

field changes accordingly, and vice-versa, the resonant field remains the same as the molecule is rotated between the xx- and yy-axis. Figure 1-3B shows how the resonant field changes as the angle, between the magnetic field and molecular z-axis, θ, changes from 0° to 90° to 180°. The angle θ is shown in Figure 1-1B.

In a physical sample, which contains a minimum of $10^{13}$ spins at X-band, the sample is frozen and all θ angles are captured. Therefore, any resonant field between the two maximums of θ = 0° and 90° is possible. A sample with randomly oriented molecules will have significantly more molecules whose xx,yy-plane is aligned with the magnetic field, and statistically less molecules whose zz-axis is aligned with the magnetic field. Therefore the intensity of the perpendicular orientations will be much larger. The intensity increases sinusoidally from when the zz-axis is parallel to the magnetic field is rotated until the xx,yy-plane is parallel to the field. This physical description aptly describes the lineshape brought about by axially symmetric g-anisotropy shown in Figure 1-3C. Trace (a) in Figure 1-3C shows the actual absorption spectrum due to $g$-anisotropy illustrating any value between the two maximum is possible. Note that the lineshape in trace (b) of Figure 1-3C is the first derivative of the absorption spectrum which is the typical read out of ESR instrumentation.

We next quantitatively define the resonant field for a molecule of any orientation with respect to the laboratory frame. The laboratory frame z-axis is taken to be the direction of the applied magnetic field as shown in Figure 1-1B. First, we solve Equation (1-7) for the resonant field at a constant frequency:

$$B = \frac{h\nu}{g\beta_e} \tag{1-8}$$

Where $g$ is taken as $g_e + \Delta g$. For an axially symmetric system, such as a type-II $Cu^{2+}$ center, $g$ can be defined as[18]:

$$g(\theta) = \sqrt{g_\perp^2 sin^2(\theta) + g_\parallel^2 cos^2(\theta)} \tag{1-9}$$

Equation (1-9) also defines the shape of the curve shown in Figure 1-3B.

The anisotropy of the electron Zeeman is further split by the hyperfine interaction between the electron spin and the $Cu^{2+}$ nuclear spin. Similar to the notation for $g$ anisotropy, for an axially symmetric system, $A_{xx} = A_{yy} = A_\parallel$ and $A_{zz} = A_\perp$. Therefore, we expand the resonance condition given by Equation (1-5) to include this hyperfine interaction:

$$hv = g\beta_e B + Am_I \tag{1-10}$$

Where $A$ represents the hyperfine value which depends on the orientation of the molecular frame with respect to the applied magnetic field. Again, as the molecular axis is rotated with respect to the magnetic field, the value of $A$ will change and thus, the resonant field will also change. Like $g$ in Equation (1-9), for an axially symmetric system, $A$ can be defined in terms of $\theta$ by[18]:

$$A(\theta) = \sqrt{\frac{A^2 g_\perp^4 sin^2(\theta) + A^2 g_\parallel^4 cos^2(\theta)}{g_\perp^2 sin^2(\theta) + g_\parallel^2 cos^2(\theta)}} \tag{1-11}$$

Taking the selection rules into account from Equation (1-4), $I = 3/2$ for $Cu^{2+}$ and therefore four transitions are possible. The lineshape due to $g$ anisotropy in a system with axial symmetry, shown in Figure 1-3C, will be further split into a quartet by these four possible transitions. In the case of type-II $Cu^{2+}$, $A_\parallel$ will produce a splitting larger than 120 G, while $A_\perp$ will produce a splitting with values under 20 G. Example hyperfine splittings for $\theta = 0°$ and $\theta = 90°$ are shown Figure 1-4A. Figure 1-4B shows the actual powder lineshape of the $Cu^{2+}$ CW-ESR spectrum which accounts for combination of both $g$ and $A$ anisotropy at all $\theta$ angles. Typically, for X-band CW-ESR, the $A_\perp$ hyperfine splittings are not resolved due to broadening as shown in Figure 1-4B.
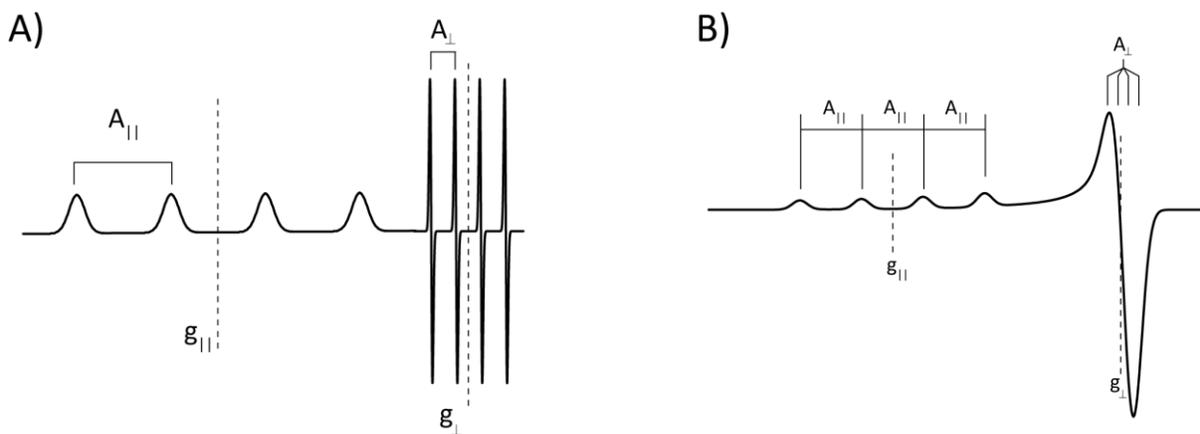
Figure 1-4. A) Overlay of both *g*- and *A*- splittings observed for systems with axial symmetry. B) A typical type-II Cu$^{2+}$ CW-ESR spectrum. The splittings due to $A_{\parallel}$ and $A_{\perp}$ are shown as well as the position of $g_{\parallel}$ and $g_{\perp}$.

The $g_\parallel$ and $A_\parallel$ values of the central $Cu^{2+}$ nucleus are highly dependent on equatorial coordination[15]. Figure 1-5A shows an example comparison of two different type-II $Cu^{2+}$ centers with different numbers of imidazole nitrogen atoms equatorially coordinated. We note the differences between spectra by the relative positions of the peaks with respect to the magnetic field, and by the splitting between two adjacent peaks.

Peisach and Blumberg tabulated $g_\parallel$ and $A_\parallel$ values of type-II $Cu^{2+}$ centers with different numbers of directly coordinated nitrogen and oxygen atoms. Replacing a liganded nitrogen atom with a more electronegative atom such as oxygen affects both $g$ and $A$ values. This increased electronegativity draws the electron density away from the central nucleus, decreasing the local magnetic field. The z-component of this local field, opposes the applied magnetic field. Therefore, the larger local magnetic field is taken as an increase in $\Delta g$, thereby changing the resonance condition determined by Equation (1-7). Likewise, oxygen's greater electronegativity reduces the electron density of the central $Cu^{2+}$ ion, reducing hyperfine interaction[15]. The trend follows that increased nitrogen coordination leads to a smaller $g_\parallel$ and a larger $A_\parallel$[15]. Figure 1-5B shows the dependence of both $g_\parallel$ and $A_\parallel$ values with respect to varying numbers of equatorially coordinated nitrogen and oxygen atoms. This relationship is widely used to determine coordinating ligands. Also, if two distinct binding sites are present in a sample, the CW spectrum will appear as a superposition of the two.

### 1.3.2   CW Spin Counting

The intricate spectral lineshape of $Cu^{2+}$ is not the only utility of the CW spectrum. The doubly integrated intensity of the CW spectrum is directly proportional to the number of spins
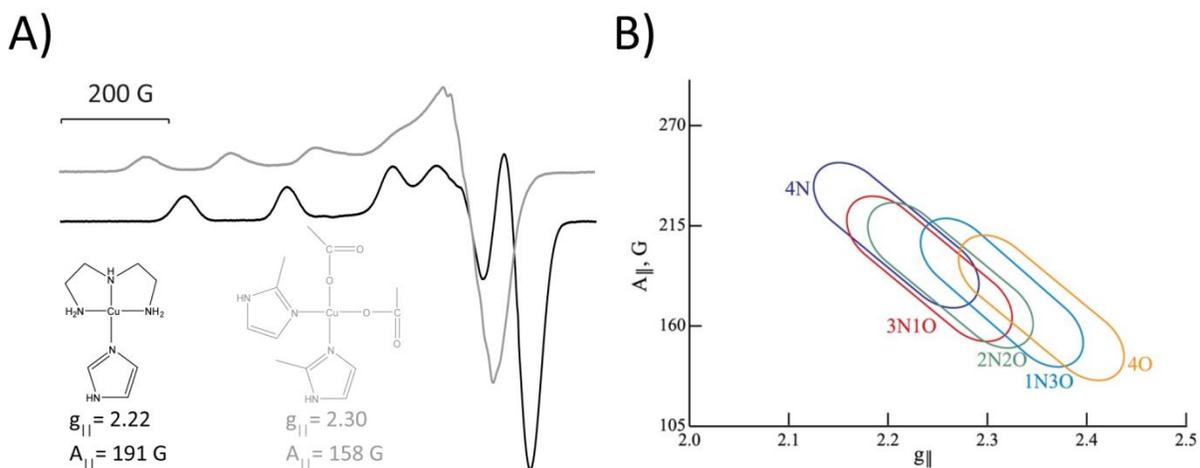
Figure 1-5. A) CW-ESR spectrum of differing $Cu^{2+}$ coordination. The lineshape is different between the two spectrum due to differing $\mathbf{g}_{||}$ and $A_{||}$ values. This data was taken with permission from ref. 37. B) The Peisach-Blumberg plot depicting the relationship between $Cu^{2+}$ centers with different equatorially coordinated ligands and their respective $\mathbf{g}_{||}$ and $A_{||}$ values.

present in a sample[20]. Therefore, the overall $Cu^{2+}$ concentration of the sample can easily be determined by comparison to a standard calibration curve of known $Cu^{2+}$ concentrations. The ability to determine concentrations is twofold: determination of $Cu^{2+}$ sample loading and the quantification of individual spectral components.

Most $Cu^{2+}$ ESR experiments in this thesis are performed in the n-ethylmorpholine (NEM) buffer. At pH 7.4-7.8, this buffer complexes any free $Cu^{2+}$ in solution into ESR silent $Cu(OH)_2$ removing its presence from the overall signal[21]. Therefore, in the NEM buffer, any signal present in the CW spectrum must be from bound $Cu^{2+}$, and the bound $Cu^{2+}$ concentration can be quantified. If, in general, the concentration of the binding site in question is known, then direct comparison to the $Cu^{2+}$ signal will provide evidence of binding site loading. Furthermore, one can now directly compare $Cu^{2+}$ binding site loading to the amount of $Cu^{2+}$ added and can extrapolate the apparent dissociation constant, $K_d$[12,13,21-24].

Should there be two distinct $Cu^{2+}$ binding environments, or if there is free $Cu^{2+}$ in solution due to the inability to use NEM on a given system, the spectral lineshape will appear as two components superimposed together. The two components are easily simulated. Here, the integrated intensities of the individual components can be determined and compared to analyze sample composition[13,25,26].

## 1.4    ELECTRON SPIN ECHO ENVELOPE MODULATION SPECTROSCOPY

While CW-ESR offers an experiment sensitive to the directly coordinated environment, the pulsed ESR technique electron spin-echo envelope modulation (ESEEM) is dependent on nearby nuclear spins. Figure 1-1A illustrates a type-II $Cu^{2+}$ coordination and highlights such

atoms with nuclear spins targeted by the ESEEM experiment (light shading). This multi-pulse experiment probes such a hyperfine interaction. For biological systems, $Cu^{2+}$-based ESEEM spectroscopy may provide two useful pieces of data. First, the frequency of the interaction with these distal nuclear spins are a blueprint indicative of specific amino acid coordination. In particular, this thesis largely focuses on the frequencies related to imidazole coordination. Second, the depth of these modulations corresponds to the amount or degree of coordination. This section will discuss how we obtain these two values, and how we analyze them.

### 1.4.1 ESEEM for the simplest case of I=1/2

To understand how the ESEEM technique can provide both pieces of information mentioned above, we must first understand two important concepts. First, the frequency of the hyperfine interaction between the electron spin and nearby nuclear spins can help identify the ligands bound to $Cu^{2+}$. We start with the simplest case of a $S$=1/2 and $I$=1/2 system.

The ESEEM experiment uses multiple pulses to probe the hyperfine interaction between the electron spin and nearby nuclear spins. This hyperfine interaction decreases as the separation between electron spin and nuclear spin increases. Past 10 Å separation between the electron spin and nuclear spin, the hyperfine interaction is too small to measure by typical pulsed X-band[14,27]. On the other hand, in a type-II $Cu^{2+}$ system, the directly coordinated atoms, described in section 3, have a hyperfine interaction of 40 MHz. This frequency is too large for the limited bandwidth of the ESEEM experiment using typical pulse lengths. The interaction with the $Cu^{2+}$ nuclear spin is of even greater magnitude and also undetected. Therefore, the ESEEM experiment is capable of measuring the hyperfine interaction between the electron spin and nuclear spins in the range

of 3-10 Å away from of the $Cu^{2+}$ center[14,27]. We call such nuclear spins the distally coordinated environment which is shown by the light shading in Figure 1-1A. The hyperfine interaction between the electron spin and such nuclear spins has a corresponding frequency < 20 MHz at X-band. The signal from these distally coordinated atoms can provide crucial insight into the binding environment of the $Cu^{2+}$ metal center.

While the CW experiment was only sensitive to transitions between electron spin states (Figure 6A, black arrows), the ESEEM experiment is also sensitive to nuclear transitions (Figure 1-6A, grey arrows). In the three pulse ESEEM experiment, with pulse sequence shown in Figure 1-6B, these nuclear transitions cause modulations apparent in the time domain signal (see Figure 1-7A). We first look at the equation for the signal in the time domain which contains these modulations. For the simplest case where $S = ½$ and $I = ½$, the echo intensity in the time domain signal can be written as[14]:

$$V_{3p}(\tau, t) = 1 - \frac{k}{4}\left\{[1 - \cos(\omega_\alpha\tau)]\left[1 - \cos\left(\omega_\beta(t + \tau)\right) + [1 - \cos(\omega_\beta\tau)][1 - \cos(\omega_\alpha(t + \tau))]\right]\right\}$$

(1-12)

Where $\tau$ and $t$ are the separation among pulses shown in Figure 1-6A, and k is the modulation depth parameter. Here, $\omega_\alpha$ and $\omega_\beta$ are the frequencies of the nuclear transitions shown in Figure 1-6B. An example of nuclear modulations in the time domain signal is shown in Figure 1-7A. Fourier transformation of the time domain signal in Equation (1-12) gives the nuclear transition frequencies $\omega_{\alpha,\beta}$ shown in Figure 1-7B. These frequencies are given by the following[14,28,29]:
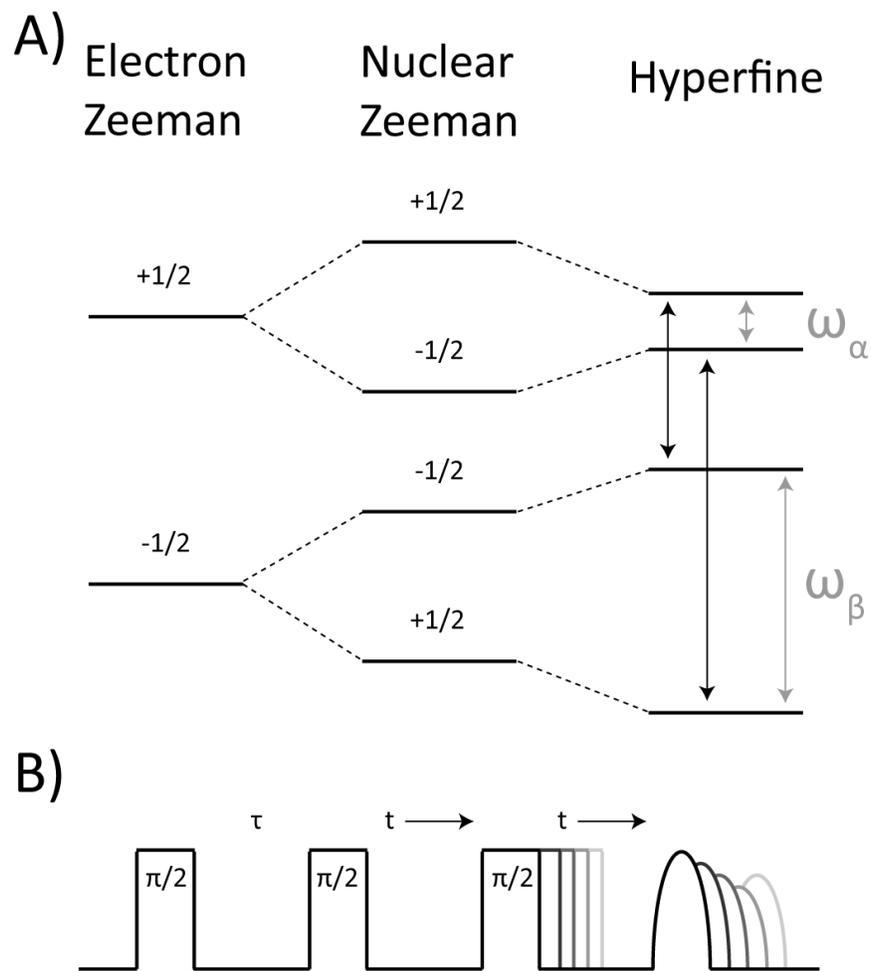
Figure 1-6. A) Energy level diagram for a system with S= ½ and I= ½. Black arrows indicate allowed transitions while grey arrows indicate forbidden transitions $\omega_\alpha$ and $\omega_\beta$. B) Pulse sequence of the three pulse electron spin echo envelope modulation (ESEEM) spectroscopy technique comprised of three $\pi/2$ pulses. The first two pulses are separated by a fixed value, $\tau$, where the separation of the second and third pulses, t, is incremented over time.
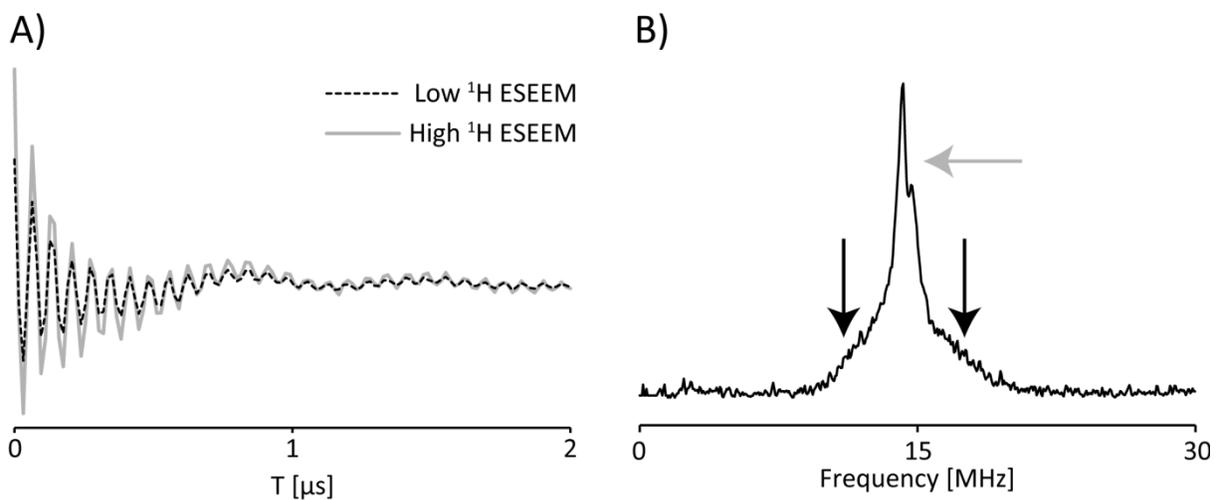
Figure 1-7. A) The background corrected ESEEM time domain showing modulations due to
$^1$H. Two different experimental signals (low $^1$H ESEEM in black dashed and high $^1$H ESEEM
in grey solid) are shown illustrating differences in modulation depth. B) Fourier transformed
spectrum showing $^1$H peaks centered at 14 MHz. This doublet is indicative of directly
coordinated water, shown by grey arrow. The signal is broadened (black arrows), typical for
biological samples with numerous and varied $^1$H coordination such as distal imidazole $^1$H and
solvent $^1$H.

$$\omega_{\alpha,\beta} = 2\pi\upsilon_{\alpha,\beta} = \sqrt{\left(\upsilon_I \pm \frac{A}{2}\right)^2 + \left(\frac{B}{2}\right)^2}$$

(1-13)

Where,

$$A = T(3\cos^2\theta - 1) + a_{iso} \quad B = 3T\sin\theta\cos\theta \tag{1-14}$$

Where, $\upsilon_I$ is the nuclear Larmor frequency, T is the dipolar coupling, $a_{iso}$ is the isotropic hyperfine coupling and $\theta$ is the angle between the electron-nuclear distance vector and the applied magnetic field vector, $B_0$. The dependence of $\omega_{\alpha,\beta}$ on the nuclear Larmor frequency makes ESEEM very diagnostic of nuclear identity. For example, at X-band, the two frequencies of $^1H$ nuclei, $\omega_\alpha$ and $\omega_\beta$ described by Equation (1-13), result in a doublet centered at ~14 MHz as shown by grey arrows in Figure 1-7B. An example of a typical Fourier transformed $^1H$ ESEEM signal is shown in Figure 7A. In biological samples, there also exist a large amount of nearby $^1H$ atoms at different distances. Frequencies due to differing $^1H$ coordination, such as solvent water or imidazole nitrogen, cause a superposition of doublets leading to broadened peak (black arrows, Figure 1-7B).

Furthermore, the depth of these modulations can be described by the modulation depth parameter mentioned in Equation (1-12). For a system with I = ½, the modulation depth parameter can be expressed as[30]:

$$k(\theta) = \left(\frac{B\upsilon_I}{\upsilon_\alpha\upsilon_\beta}\right)^2$$

(1-15)

An example of two ESEEM signals with differing modulation depths due to $^1H$ is shown in Figure 1-7A. We will see in the next section how analysis of the modulation depth can provide information on the extent of coordination in $Cu^{2+}$-binding sites.

## 1.4.2 ESEEM for a Cu$^{2+}$ system with I=1

Typically, in biological macromolecules, one of the most common interactions observed are of modulations resulting from $^{14}$N. In this chapter, we will largely focus on histidine imidazole coordination.

In the case of Cu$^{2+}$ interacting with distal nitrogen from histidine imidazole rings, I = 1, the nuclear quadrupole interaction provides characteristic transitions. Imidazole coordination follows the exact cancellation condition where nuclear Zeeman ≈ the hyperfine interaction. The energy level diagram and transitions are shown in Figure 1-8. When the exact cancellation condition is met, Cu$^{2+}$ systems with distal nitrogen coordination (~3-8 Å for X-band) contain three characteristic peaks between 0 and 2 MHz from the nuclear quadrupole interaction (NQI) transitions[25,31-33]. The frequencies of these nuclear transitions are given by[34]:

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

$$(1\text{-}16)$$

Where $e$ is the charge of an electron, $q$ is $z$-component of the electron charge gradient across the nucleus, $h$ is Planck's constant and $\eta$ is the asymmetry parameter. The three frequencies produce three peaks in the Fourier transformed spectrum which may be very well defined as shown in Figure 1-9B. The relationship between these three frequencies satisfies $\nu_+ = \nu_- + \nu_0$ which can be used to reaffirm the identity of the NQI peaks.

In addition to these NQI transition, there is a broad peak at 3.8 MHz due to the double quantum (DQ) transition[33]. The DQ transition is shown in Figure 1-8. The frequency associated with this DQ transition is given by[35]:
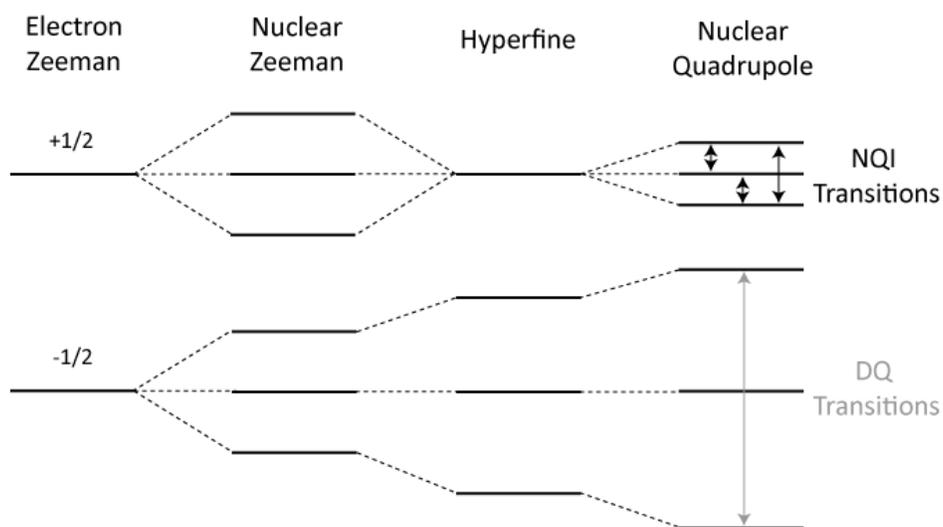
Figure 1-8. Energy level diagram for a system with S= ½ and I= 1 consistent with imidazole coordination. This type of coordination follows the exact cancellation condition for $Cu^{2+}$ where the nuclear Zeeman ≈ the hyperfine interaction. Nuclear quadrupole interactions (NQI) are denoted as black arrows and the double quantum transition is denoted as the grey arrow.
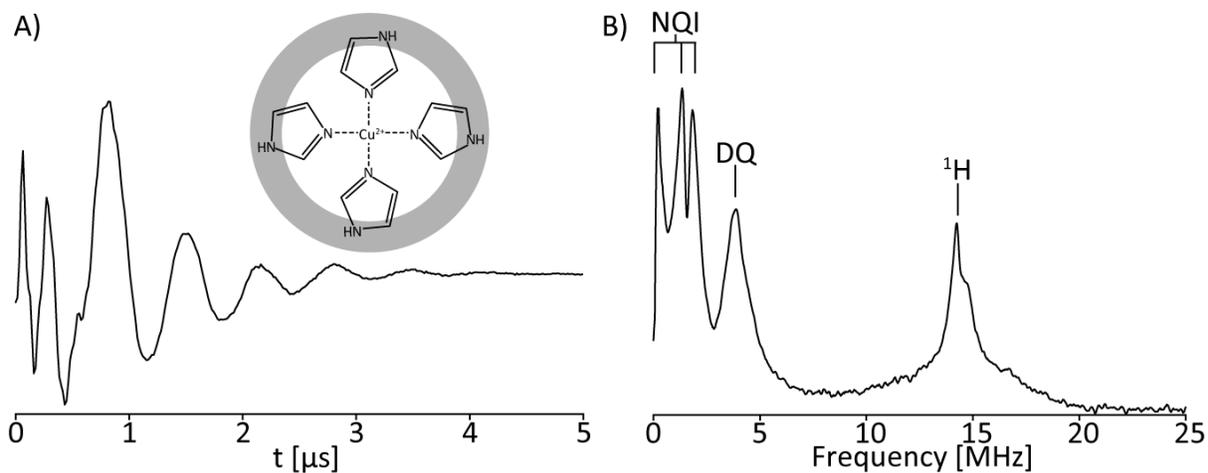
Figure 1-9. A) The background subtracted time domain ESEEM signal of $Cu^{2+}$ coordinated to imidazole shown in inset (highlighted region shows detectable distally coordinated environment). B) The Fourier transformed ESEEM spectrum illustrating the three NQI peaks (below 2 MHz), the DQ (~3.8 MHz) peak as well as the peak associated with the ^{1}H nuclei (14 MHz). The ^{1}H peak can easily be suppressed if necessary by choosing appropriate $\tau$ value to satisfy the blind spot condition.

$$v_{DQ} = 2\sqrt{\omega_\beta^2 + \left(\frac{e^2qQ}{4h}\right)(3 + \eta^2)}$$ (1-17)

where $\omega_\beta$ is defined in Equation (1-13). Since the nuclear Zeeman and hyperfine interactions are additive, it gives rise to much broader resonances[35,36]. The broad peak associated with the DQ transition around 3.8 MHz is shown in Figure 1-9B. Taken together, these distinct NQI and DQ features, shown in Figure 1-9B, of $Cu^{2+}$ coordination to imidazole, or its derivatives, are a common diagnostic tool in $Cu^{2+}$-based ESEEM spectroscopy[25,27,32,33,35,37-40].

### 1.4.3   Modulation depth on the extent of coordination

ESEEM also has the power to resolve the number of imidazole rings coordinated to a $Cu^{2+}$ ion. The first evidence of multiple imidazole coordination is the presence of combination peaks in the frequency domain spectrum[30,41]. These combination peaks rise from relatively low-probability, simultaneous spin-flip transitions between multiple remote nitrogen nuclei. In this case, prominent peaks in the frequency domain should be evident at $(v_0 + v_+)$, $(v_- + v_+)$ and $(2v_+)$ which all typically occur below ~4 MHz[42]. Additionally, a weak feature from between ~7.0-9.0 MHz should arise which corresponds to the combination of the $\Delta m_I = \pm2$ known as the DQ combination feature at $(2v_{DQ})$[38,39,43].

Furthermore, there exists a method to quantify the number of $^{14}N$ nuclei coordinated to a $Cu^{2+}$ center based on the integrated intensities of the ESEEM spectrum. Spectral simulations are typically used for such matter[43,44], however, Saxena and co-workers have presented an alternative method that compares the normalized integrated intensities of the frequency domain ESEEM spectra. With increasing non-directly-coordinated distal nitrogen, the $^{14}N$ ESEEM

intensity increases. This method is done by using the integrated intensity of the frequency domain ESEEM spectrum due to ^{14}N which is the region between 0-11 MHz[25,40]. This value is then divided by the integrated intensity of the ^{1}H ESEEM which is the region between 13-16 MHz, to normalize the integrated intensity due to [14]N. The regions of integration of the ESEEM spectrum are shown in Figure 1-10. It has been shown that as the number of imidazole coordinated to $Cu^{2+}$ increases, this normalized integrated intensity also increases[40,45].

### 1.4.4 Binding site affinity through modulation depth analysis

Not only are the frequencies of the nuclear modulations useful, but the relative amplitudes of these modulations can also provide key information on $Cu^{2+}$ coordination as well. The amplitude of these modulations are quantified by a relative modulation parameter, $k$. An example of two samples with differing values of $k$ is shown in Figure 1-11A. Through analysis of the nitrogen modulations in the time domain ESEEM signal, $k$ can be quantified[46]. Here, $k$ increases with distal nitrogen coordination[13,25,40]. The time domain ESEEM signal is a superposition of all modulations due to the NQI and DQ transitions which can make the determination of a unique modulation depth rather difficult. To this end, the second modulation period of the ^{14}N modulations is generally used for analysis of $k$[46]. Here, the effective $k$ is taken as[46]:

$$k = \frac{a}{a+b}$$

(1-18)

Where $a$ is the interpolated echo intensity between the second and third maxima of the time domain and $b$ is the echo intensity at the second minimum. Figure 1-11A shows how $k$ is extracted from the raw time domain ESEEM signal in terms of $a$ and $b$.
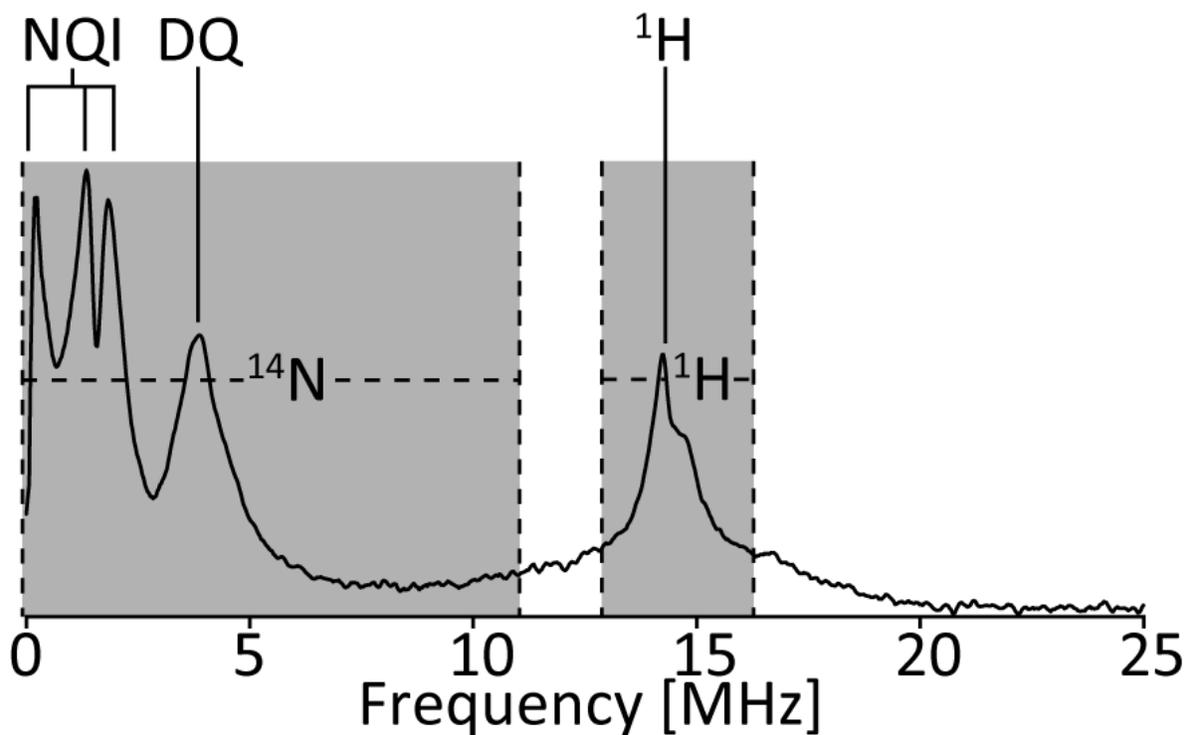
Figure 1-10. Experimentally obtained FFT ESEEM spectrum of $Cu^{2+}$ coordinated to histidine imidazole nitrogen. The area under the first shaded region (0-11 MHz) is indicative of the number of coordinated imidazole rings[20,34]. The second shaded region (13-16 MHz) can be used to normalize multiple spectra[20,34].
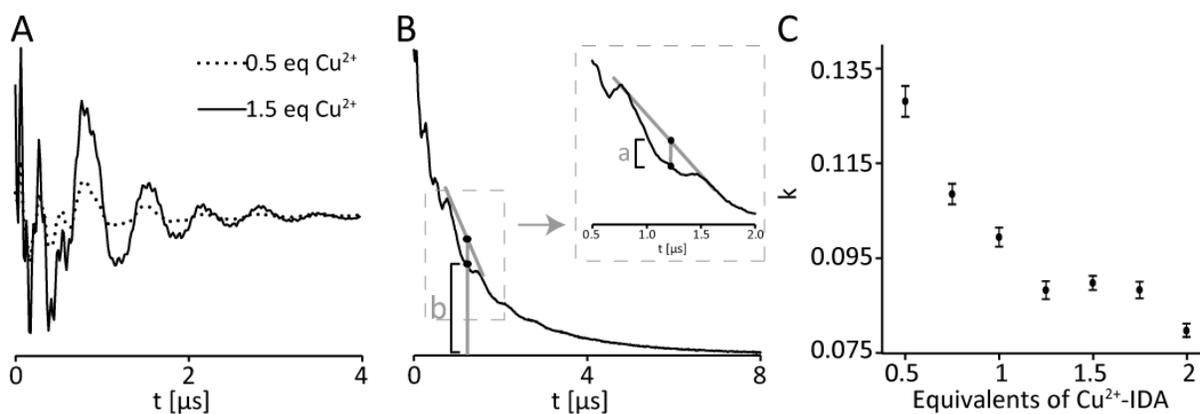
Figure 1-11. A) Background corrected time domain of two different samples with 0.5 equivalents of $Cu^{2+}$ (dotted) and 1.5 equivalents of $Cu^{2+}$ (solid) added. Note the increase in amplitude of modulations with no change in modulation frequency. B) Raw time domain signal of the ESEEM experiment of a $Cu^{2+}$-imidazole complex. Determination of both a and b for modulation depth, k, calculation is shown. B) An example plot from ref. 7 illustrating the decrease in k upon addition of background contributing $Cu^{2+}$-IDA.

The analysis of $k$ can be used to ascertain binding selectivity of a potential $Cu^{2+}$ binding site if histidine or its derivatives are involved within the coordinating environment. A titration can be performed by adding in known aliquots of $Cu^{2+}$. In this case, only imidazole-bound $Cu^{2+}$ will contribute to the strength of the modulations in the overall ESEEM signal[21]. However, unbound $Cu^{2+}$ will also have an ESEEM signal, albeit a featureless decay. The total ESEEM signal is therefore a superposition of modulations produced by imidazole-bound $Cu^{2+}$ and featureless free $Cu^{2+}$. Larger amounts of free $Cu^{2+}$ will, therefore, dampen the modulations due to imidazole bound $Cu^{2+}$, effectively reducing $k$. Plotting the titration as $k$ versus equivalents of $Cu^{2+}$ added will produce a curve highly dependent on the $Cu^{2+}$-binding affinity of the $Cu^{2+}$-binding site. This curve, in combination with CW-ESR analysis, can provide crucial information on the extent of loading of $Cu^{2+}$-binding sites[13,23].

## 1.5     DOUBLE ELECTRON-ELECTRON RESONANCE SPECTROSCOPY

One of the most practical tools in ESR is the ability to measure the dipolar interaction between two coupled electron spins. The Double Electron-Electron Resonance (DEER) experiment probes this interaction. This technique provides multiple information pertaining to a $Cu^{2+}$ system: 1) local concentrations, 2) spin counting, 3) $Cu^{2+}$ loading efficiency, and 4) distance constraints. In this section, we will learn how this pulsed ESR experiment can elucidate such information.

### 1.5.1　The DEER signal

The most common pulsed ESR technique used in structural biology is the dead time free four pulse DEER[47,48]. The two-frequency four pulse sequence used in DEER is shown in Figure 1-12A. This experiment probes the weak magnetic dipole-dipole interaction between two electron spins. The DEER signal is a time domain curve that modulates with the frequency of this dipolar interaction between the two spins. An example of the raw DEER time domain signal is shown in Figure 1-12B. This raw time domain DEER signal, $V(t)$, is comprised of the product of two signals, an intramolecular signal, $V_{intra}(t)$, and an intermolecular background signal, $V_{inter}(t)$[7,9]. A visual representation of $V_{intra}(t)$ and $V_{inter}(t)$ is shown in Figure 1-13A. The intramolecular signal is the dipolar interaction between two electron spins within the same protein or complex (solid black line, Figure 1-13A). However, in an actual sample with $> 10^{13}$ spins, proteins or complexes will be randomly distributed at all orientations and distances. The intermolecular signal is the net of all dipolar interactions of electron spins between neighboring proteins or complexes (dotted lines, Figure 1-13A). The overall DEER signal, $V(t)$, is given by[7,9]:

$$V(t) = V_{intra}(t)V_{inter}(t) \tag{1-19}$$

Where $t$ is the variable position of the pump pulse. For a system with two spins that are rigidly separated, the intramolecular signal is given by[6]:

$$V_{intra}(t) = 1 - p_b \left( 1 - \int_0^1 \cos(\omega_D t) \, d \cos \theta \right)$$

(1-21)

$\theta$ is the angle between the interspin vector and the magnetic field vector, $p_b$ is the fraction of spins excited by the pump pulse and $\omega_D$ is the dipolar frequency given by[49]:
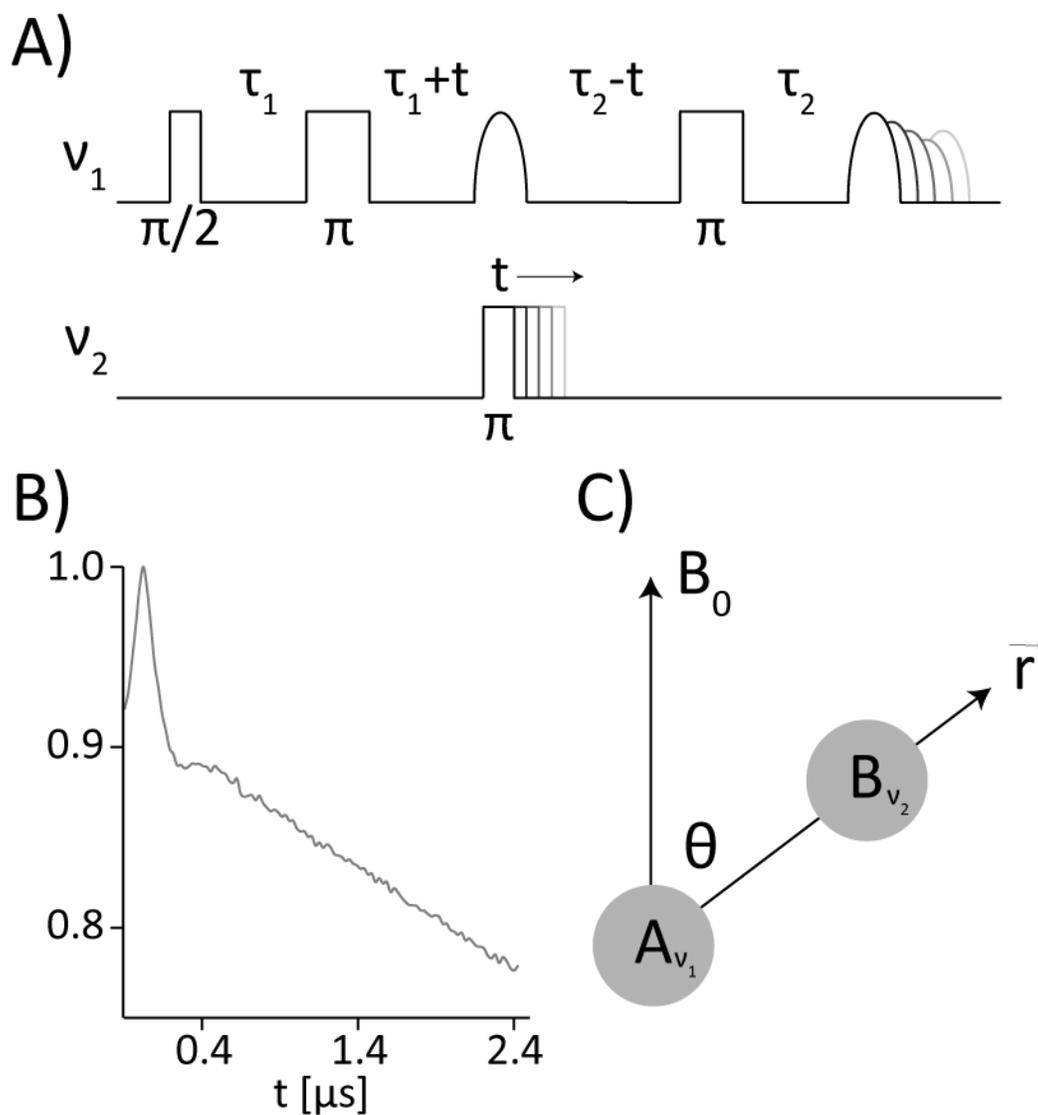
Figure 1-12. A) The two frequency four-pulse sequence of the DEER experiment. B) An example of a raw time domain signal from the DEER experiment of a $Cu^{2+}$-$Cu^{2+}$ distance. C) Representation of a two spin system, A and B, within an applied magnetic field.
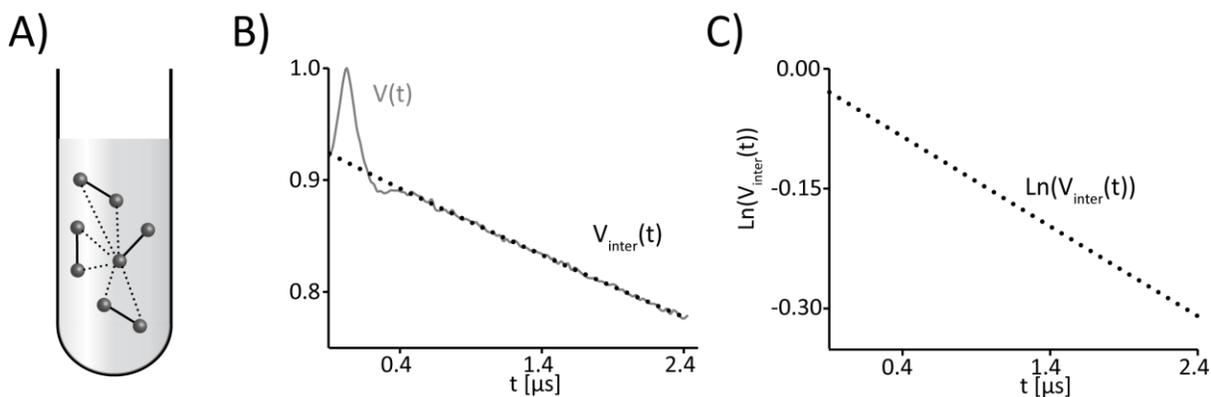
Figure 1-13. A) Visualization of multiple doubly labeled systems within a frozen sample. Intermolecular signal is shown as dotted line while intramolecular signal is shown as black line. B) Raw time domain signal, V(t) (grey), with the simulated background signal, $V_{inter}(t)$ (dotted). C) A logarithmic plot of $V_{inter}(t)$ producing a linear function whose slope is proportional to the local concentration.

$$\omega_D = \frac{C}{r^3}(1 - 3\cos^2\theta)$$

(1-22)

Where $C$ is proportional to the product of g values due to both spins and takes the value of either 52.04 or 62.84 MHz•nm$^{-3}$, for nitroxide or Cu$^{2+}$ systems, respectively, and $r$ is the interspin distance. A representation of a two spin system within a magnetic field is shown in Figure 1-12C. After subtraction of $V_{inter}(t)$, the signal is then Fourier transformed producing a Pake pattern from which the dipolar frequency can be extracted. Besides the obvious advantages of distance information between spins, the DEER experiment contains a plethora of information that will be presented in the following sections.

### 1.5.2 Local concentrations

The intermolecular signal can be analyzed to extract the local concentration of spins. An example of the raw time domain DEER data with $V_{inter}(t)$ illustrated is shown in Figure 1-13B. The intermolecular signal is an exponential decay due to the superposition of numerous intermolecular dipolar interactions. $V_{inter}(t)$ is given as follows[3,49,50]:

$$V_{inter}(t) = exp(-Kt)$$

(1-23)

with the decay rate, $K$, as[51]:

$$K^{-1} = B\frac{10^{-3}}{p_b c}$$
(1-24a)

$$B = \frac{8\pi^2}{9\sqrt{3}}\gamma^2\hbar = 1.0027$$

(1-24b)

Where $c$ is the local concentration of the sample, $\gamma$ is the gyromagnetic ratio and $\hbar$ is the reduced Planck constant. While CW spectroscopy is directly related to the total number of spins, the integrated intensity of the CW spectrum has little to no information on if the spins are clustered versus homogeneously dispersed.

To extract the local concentration from the raw time domain DEER signal, the signal can be fit using an exponential decay function. While it is possible to fit this exponential decay using Equations (1-23) and (1-24), exponential fits are inherently more difficult to accurately fit compared to a straight line. Thus, a logarithmic fit of $V_{inter}(t)$ allows for a linear fit to determine the local concentration[52] (c.f. Figure 1-13C).

The local concentration can provide key insight into sample conditions such as aggregation and rate of diffusion. If a sample forms soluble aggregates, then the total concentration can be determined from the CW spectrum. However, the local concentration should be much higher which can suggest that the $Cu^{2+}$ ions are clustered. Furthermore, the local concentration is expected to decrease if molecular diffusion disperses solute that is initially concentrated in a region. This practical application has been utilized in cell to measure local concentrations of microinjected proteins in-cells at different times and will be discussed in Chapter 5[52,53].

### 1.5.3 Modulation depth analysis

A key parameter in the DEER experiment is known as the modulation depth, $\lambda$[47,54]. The modulation depth is defined as the difference from the maximum of the normalized background-corrected DEER signal to where the signal becomes constant, as shown in Figure 1-14. This value, $\lambda$, can be a quantitative measure of the depth of the dipolar modulations and is directly proportional to the DEER signal sensitivity[3,7,47,55] and is defined by the following equations[54,56,57]:

$$\lambda = 1 - V_p \tag{1-25}$$

$$V_p = (1 - p_b)^{(N-1)} \tag{1-26}$$

Where $V_p$ is the difference between the constant DEER signal at $t = \infty$ and 0, $p_b$ is the faction of spins excited by the pump pulse and $N$ is the number of coupled spins. $\lambda$ depends on several factors such as the length of the pump pulse, the frequency offset between the pump and observer pulses and the spectral lineshape[1,7].

### 1.5.3.1 Counting the number of monomers

The modulation depth, $\lambda$ is sensitive to the number of coupled spins in a given system[47,56]. Therefore, since $\lambda$ is determined experimentally, the number of $N$ coupled spins can be determined from the DEER time traces. This feature can be exploited to effectively count the number of spin labeled monomers present. If a protein has multiple spin labels, $\lambda$ will depend on the total number of spins present[56]. A simulated example of how $\lambda$ depends on the amount of
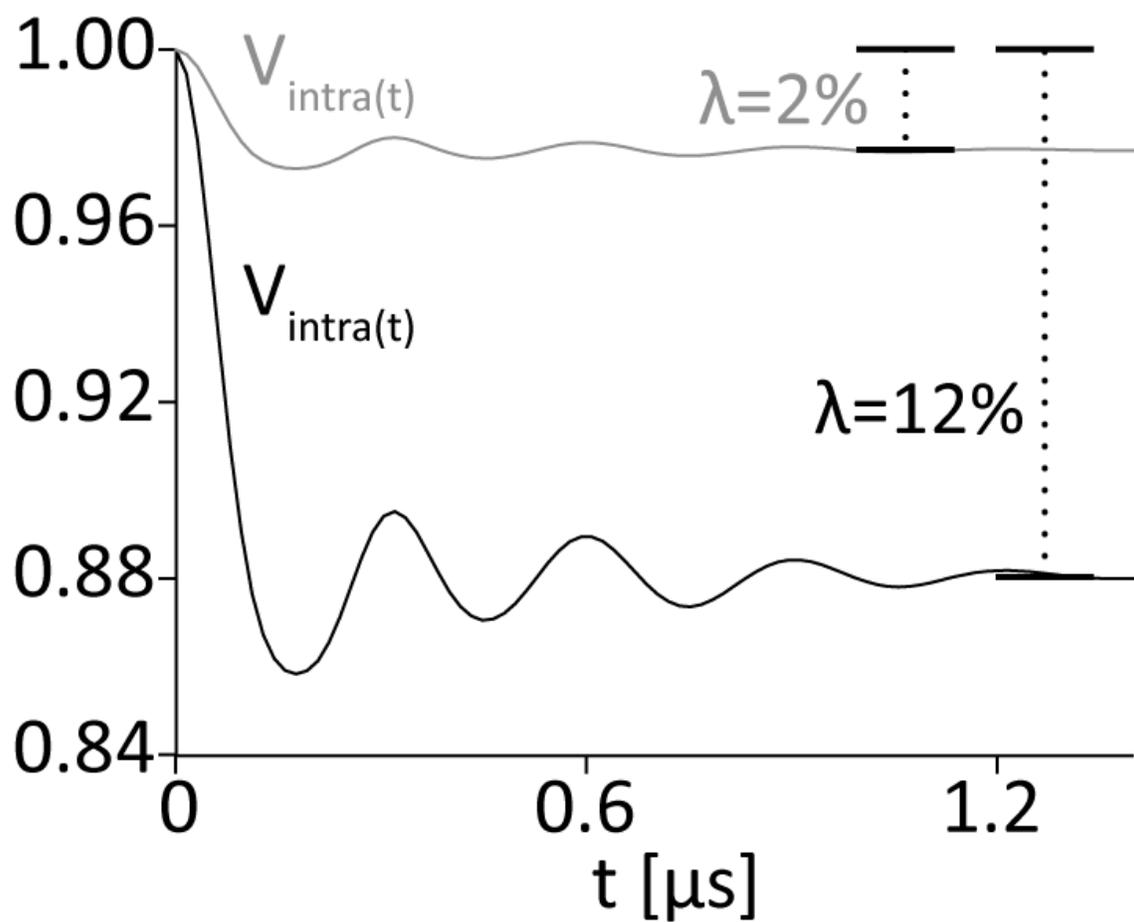
Figure 1-14. Example of the background subtracted time domain DEER signal at two different modulation depths, $\lambda$, of 2% (grey) and 12% (black). 12% is the expected modulation depth for a two $Cu^{2+}$ spin system using a 16 ns pump pulse at X-band frequencies.

coupled spins present is shown in Figure 1-15. This technique is not $Cu^{2+}$ specific, it can be used with any paramagnetic spin center, but it is an intrinsic method to determining the number of monomers that are conjugated together.

Given that $V_p$ and $p_b$ can both be determined experimentally, Equation (1-26) can be rewritten to solve for $N$ as[57]:

$$N = \frac{\ln V_p}{\ln(1-p_b)} + 1 \tag{1-27}$$

In the case of strong angular correlations, different $p_b$ values must be taken into account for each spin pair. The number of coupled electron spins can directly be determined from the modulation depth extracted from the DEER time trace with an overall error of only 5%[56]. An assumption for this direct method is that spin labeling/loading is complete. If labeling is incomplete, or a mixture of oligomeric states is possible, then direct analysis of $N$ yields a much larger error. It has been reported that the oligomeric state of biologically relevant complexes can be evaluated down to concentrations of ~50 µM at a volume of 80 µL for X-band DEER[56]. This technique works for homogenous samples containing one species. However, most biological samples are a heterogeneous mixture. We will discuss further how modulation depth can provide information on heterogeneous mixtures in the following section.

## 1.5.3.2 $Cu^{2+}$ Loading Efficiency

Since the modulation depth depends on the number of coupled spins in a given system, the modulation depth is thereby sensitive to a sample composed of species with different amounts of spins. For such a sample[56]:
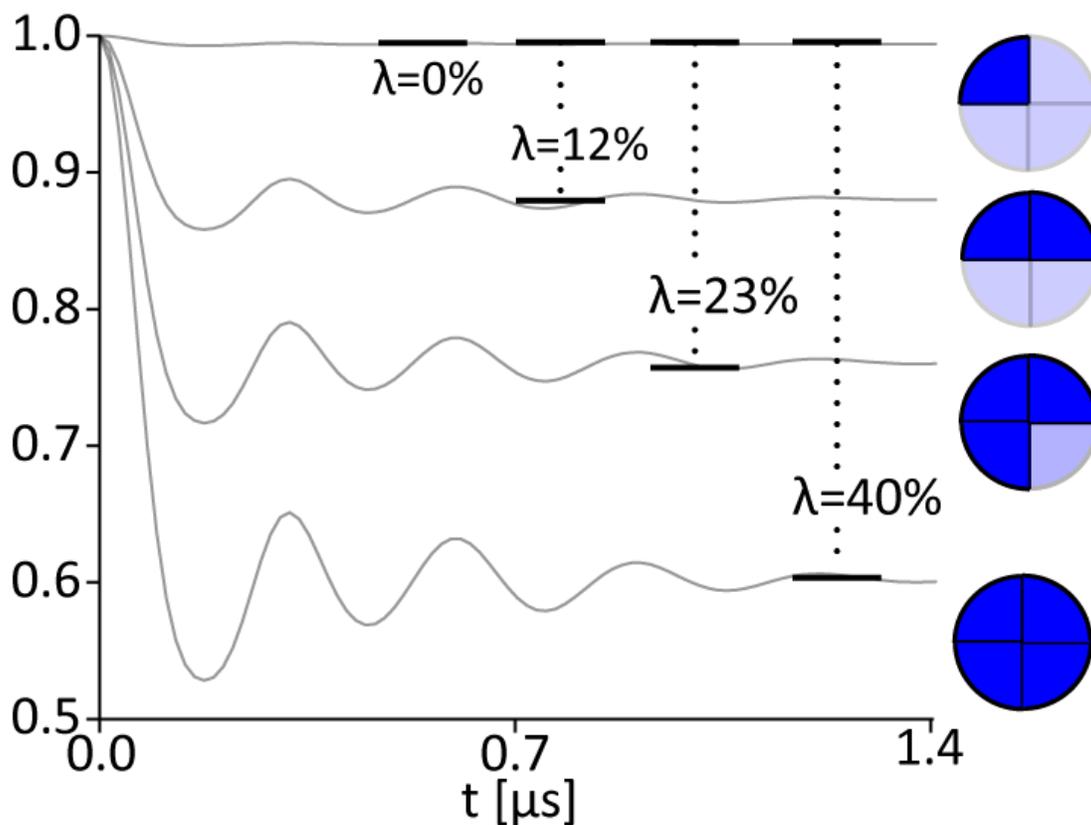
Figure 1-15. Example of how modulation depth, $\lambda$, depends on the number of $Cu^{2+}$ spins present per system for X-band DEER using a pump pulse of 16 ns. Single $Cu^{2+}$ containing monomers are represented by blue quarter spheres such that the number of spins, $N$, equals 1, 2, 3 and 4 respectively. When $N$ equals 1, there is no dipolar interaction which yields a $\lambda$ equal to 0%. For a $Cu^{2+}$ system with $N = 4$, the $\lambda$ increases to 40%. This example does not take into account the generation of secondary distances or multi-spin artifacts in the time domain.

$$V_p(t \to \infty) = \sum_{i=1}^{j} f_i (1 - p_b)^{Ni-1} \qquad (1\text{-}28)$$

Where $f_i$ is the fraction of each species in a given system and j is the total number of different species present.

This is particularly useful for determining the loading efficiency of $Cu^{2+}$ binding sites. The loading efficiency of $Cu^{2+}$ can be helpful in ascertaining the optimum $Cu^{2+}$ concentration to add to any given sample to yield the greatest DEER sensitivity. This calculation can be performed in two ways for $Cu^{2+}$ DEER: introduce $Cu^{2+}$ as an ion in the NEM buffer where any unbound $Cu^{2+}$ will be silenced[21], or introduce $Cu^{2+}$ as an ESR active complex[13,23,58]. For the first example, we will assume a sample with two unique $Cu^{2+}$ binding sites, A and B. For this example, if the $Cu^{2+}$ is introduced as an ion, there are three different consistent systems possible: a fully loaded two spin system (2Cu), a singly loaded two spin system (1Cu) and unbound Cu free in solution (fCu). In the NEM buffer, fCu is eliminated and contributes no signal. Therefore, Equation (1-28) can be taken as:

$$\lambda = 1 - (f_{2Cu}(1 - p_b)^1 + f_{1Cu}) \qquad (1\text{-}29)$$

$$f_{2Cu} + f_{1Cu} = 1 \qquad (1\text{-}30)$$

Where $f_{2Cu}$ is the fraction of $Cu^{2+}$ bound to both binding sites A and B, $f_{1Cu}$ is the fraction of $Cu^{2+}$ bound to only either A or B and their total sum must equal 1.

Using the experimentally determined $\lambda$, solving for $f_{2Cu}$ and $f_{1Cu}$ gives the total percentage of systems with fully loaded $Cu^{2+}$ binding sites. However, calculation of a theoretical $\lambda$ is useful as a maximum $\lambda$ can be determined ahead of sample preparation to utilize optimum conditions to increase DEER sensitivity. To calculate the theoretical $\lambda$, first $f_{2Cu}$ and $f_{1Cu}$ must be determined by CW spectroscopy and spectral simulations. Since the double integrated intensity of the CW spectrum is directly proportional to the number of spins present, comparison

to a calibration curve yields the total $Cu^{2+}$ concentration present. As shown by Lawless et al., comparison of the total $Cu^{2+}$ concentration to the known binding site concentration gives us[13,23]:

$$P_A = \frac{[Cu^{2+}]}{[binding\ site]} \tag{1-31}$$

$$1 = P_A + P_{NA} \tag{1-32}$$

Where $P_A$ is the probability that binding site A is occupied and $P_{NA}$ is the probability that binding site A is unoccupied, likewise for $P_B$ and $P_{NB}$ respectively. If binding sites A and B have the same $Cu^{2+}$ binding affinity, then $P_A = P_B$. Therefore, we can express $f_{2Cu}$ and $f_{1Cu}$ as:

$$f_{2Cu} = \frac{P_A P_B}{P_A P_B + P_A P_{NB} + P_{NA} P_B} \qquad\qquad f_{1Cu} = \frac{P_A P_{NB} + P_{NA} P_B}{P_A P_B + P_A P_{NB} + P_{NA} P_B} \tag{1-33}$$

In the case that the sample is not in the NEM buffer or if the $Cu^{2+}$ is introduced as a non-ESR silent complex, $f_{fCu}$ can no longer be taken as 0. The non-zero $f_{fCu}$ will cause Equation (1-33) to be rewritten as:

$$f_{2Cu} = \frac{P_A P_B}{P_A P_B + P_A P_{NB} + P_{NA} P_B} P_{sig} \qquad\qquad f_{1Cu} = \frac{P_A P_{NB} + P_{NA} P_B}{P_A P_B + P_A P_{NB} + P_{NA} P_B} P_{sig}$$

$$\tag{1-34}$$

$$P_{sig} = 1 - f_{fCu}$$

$$\tag{1-35}$$

Where $P_{sig}$ is a scaling factor to represent the total DEER signal originating from $Cu^{2+}$ bound to binding site A or B.

In the case of a more complex sample with more than two binding sites, an example could be an oligomeric assembly with dimer, trimer, etc. all possible in solution, Equation (1-29) simply expands to[13,23]:

$$\lambda = 1 - (f_{nCu}(1 - p_b)^{n-1} + \cdots + f_{2Cu}(1 - p_b)^1 + f_{1Cu}(1 - p_b)^0 + f_{fCu}(1 - p_b)^0)$$

$$\tag{4-17}$$

Where $n$ is the number of possible $Cu^{2+}$ binding sites in the system. This methodology can be employed for spin counting in samples that contain heterogeneous mixtures.

### 1.5.4   Measurement of distance distribution

In the late 90's, pulsed ESR distance measurements were first developed and performed on model complexes containing nitroxide spin labels[48,49,59-61]. These new techniques were quickly implemented into proteins systems[6,9,62], and for over a decade, the nitroxide sidechain R1 became the standard label[5,63]. During the mid to late 2000's, our group and others pioneered the use of the paramagnetic metal ion $Cu^{2+}$ for such pulsed ESR distance measurements.[1,8,64-69]. $Cu^{2+}$ is a cofactor for many enzymatic processes, and thus many biological systems have intrinsic $Cu^{2+}$-binding sites. Using such intrinsic binding sites, $Cu^{2+}$ based distance measurements using pulsed ESR quickly started appearing in scientific literature[70-76].

$Cu^{2+}$ based distance measurements, in combination with nitroxide labels, were used to locate a hitherto unknown intrinsic $Cu^{2+}$ binding site within the large protein-DNA complex[75]. Upon learning the location of the intrinsic $Cu^{2+}$ binding sites, a combination of pulsed ESR and molecular dynamic simulations were used to gain insight into the mechanism of catalytic inhibition the $Cu^{2+}$ ion induced within this EcoR1-DNA complex[77]. The approach of trilateration of $Cu^{2+}$ binding sites was refined shortly thereafter[70], and methods were developed to selective measure distances within systems containing multiple spins[74]. The ability to identify $Cu^{2+}$-binding sites has led to several works in subsequent years.  Such a combination of nitroxide labels and an intrinsic $Cu^{2+}$ binding site was used to report on the conformation of Cu-Zn superoxide dismutase[73]. ESR distance measurements were able to measure the unknown solution

state dynamics of Cu-Zn superoxide dismutase that were not easily examinable by other methods. This ability led to detection of an increased range of motion which could contribute to the disease progression of amyotrophic lateral sclerosis (ALS). Intrinsic metal-binding sites were also used to measure the molecular structure of the dicupric human serum transferrin[76]. The structure of this two-domain protein was previously resolved by X-ray crystallography, however the assembly of the two domains in solution was still in question. The rigid metal binding site in each domain was populated with $Cu^{2+}$ and distance constraints were measured confirming the X-ray structure. Also, the copper amine oxidase dimer from *Arthrobacter globiformis* was examined by $Cu^{2+}$-based ESR[71]. Distances made within this protein were compared to well defined porphyrin model systems which exemplified the utility of $Cu^{2+}$-based DEER measurements. Understanding of the $Cu^{2+}$-binding prion protein has also benefitted from $Cu^{2+}$-based distance measurements[72]. Using a combination of an intrinsic $Cu^{2+}$-binding site and nitroxide labels, NMR and ESR revealed how the globular domain of the prion protein regulates the N-terminal domain.

Not all biological macromolecules have intrinsic $Cu^{2+}$-binding sites. Thus, in order to generalize the $Cu^{2+}$ approach, our group, has recently developed new methodologies of site-directed spin labeling with $Cu^{2+}$. These methodologies include $Cu^{2+}$ chelating sidechains covalently attached to proteins[78] and a rigid motif in proteins that uses selectively placed histidine residues[13,23,58]. This section will provide a brief overview of the latter method which is used in the remainder of this thesis.

The double histidine (dHis) motif is an attractive $Cu^{2+}$-based spin-labelling methodology useful for ESR distance measurements in proteins. The dHis motif is shown in Figure 1-16. The two required histidine residues are incorporated via mutagenesis in either an α-helix, using an *i,*

*i*+4 mutation scheme, or a β-sheet, using an *i, i*+2 mutation scheme. $Cu^{2+}$ is then introduced complexed to the chelating agent iminodiacetic acid (IDA), shown in the top inset of Figure 16. The $Cu^{2+}$-IDA complex prevents any unspecific binding. This method requires no post-expression synthetic manipulations (save the addition of $Cu^{2+}$-IDA). The most probable distances reported by DEER between two dHis sites are more readily interpreted to extract the $C_\alpha$- $C_\alpha$ distance. Initial distance measurements using this strategy reported distance distributions that were up to 5-times narrower than comparable, typical nitroxide based labels[58]. A comparison of R1- and dHis-based distance distributions is shown in Figure 1-16. This data is more easily interpretable and readily related to protein backbone and flexibility. Furthermore, the dHis binding sites can be incorporated into systems in which cysteine incorporation and/or cysteine based labeling is problematic.

In Chapter 3, this thesis describes how this dHis method is further optimized to enhance signal to noise. In Chapter 4, this thesis shows an implementation of the dHis motif in the human glutathione S-transferase A1-1 enzyme. The terminal helix of this enzyme is unresolved by NMR and X-ray crystallography. The combination of nitroxide- and dHis-based measurements allows distance constraints to be made to resolve the 3-D structure of the helix for the first time. Furthermore, in an effort to expand these labeling strategies to other biological macromolecules, in Chapter 2 we present a method of incorporating $Cu^{2+}$ within DNA for precise distance measurements.
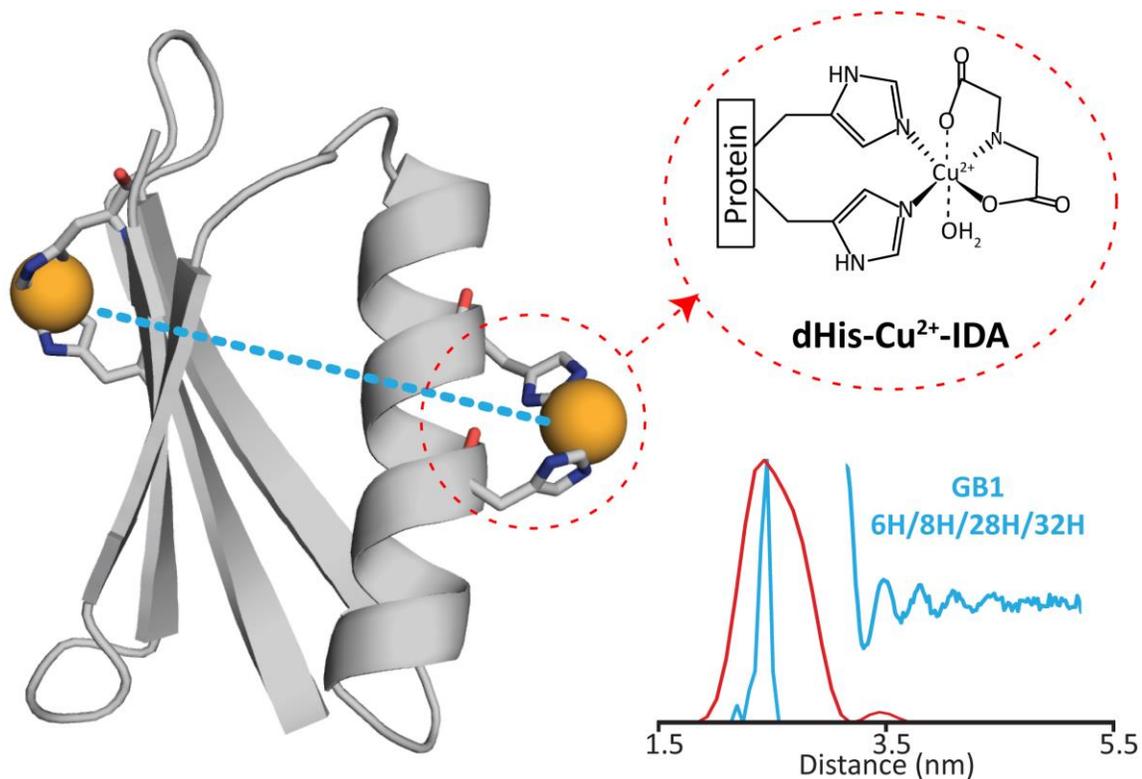
Figure 1-16. The double histidine motif within the protein GB1 (PDB code: 4WH4) loaded with $Cu^{2+}$. Inset shows the dHis-$Cu^{2+}$-IDA coordination scheme of the dHis $Cu^{2+}$ binding site. Bottom right shows the background subtracted DEER time trace with respective distance distribution (blue) compared to that of a comparable distance measured by nitroxide spin labels.

# 2.0 NUCLEOTIDE INDEPENDENT CU$^{2+}$ BASED DISTANCE MEASUREMENTS IN DNA BY PULSED ESR

*This work, written in collaboration with Jessica L. Sarver and Sunil Saxena, has been published in Angewandte Chemie International Edition, 2017, volume 56, pages 2115-2117. The thesis author collected and analyzed all ESR and CD data, performed all modeling and simulations, and prepared the manuscript.*

## 2.1 INTRODUCTION

We report a site-specific Cu$^{2+}$ binding motif within a DNA duplex for distance measurements by ESR. This motif utilizes a commercially available 2,2'-Dipicolylamine (DPA) phosphormadite easily incorporated into any DNA oligonucleotide during initial DNA synthesis. The method only requires the simple post-synthetic addition of Cu$^{2+}$ without the need for further chemical modification. Notably, the label is positioned within the DNA duplex, as opposed to outside the helical perimeter, for an accurate measurement of duplex distance. A distance of 2.7 nm was measured on a doubly Cu$^{2+}$-labeled DNA sequence, which is in exact agreement with the expected distance from both DNA modeling and molecular dynamic simulations. This result suggests that with this labeling strategy the ESR measured distance directly reports on backbone

DNA distance, without the need for further modeling. Furthermore, the labeling strategy is structure and nucleotide independent.

Nanometer range distance measurements using pulsed ESR methods provide useful insight into the conformation and dynamics of DNA and RNA[79-83] as well as on protein-DNA interactions[84-87]. For such experiments, nitroxide based spin labels have been incorporated into both single stranded and duplex DNA and RNA by attachment to specific nucleotides[80,88-94], sugar-phosphate backbone[95-97] or termini[98]. In many cases, the flexibility of the spin label can complicate the interpretation of distances in terms of DNA/RNA structure. Therefore, there has been a continuous effort to create more rigid labels for less ambiguous distance measurements[88,90,91]. Thus far, these rigid labels have come at the expense of complicated synthesis. Furthermore, due to limited spectral excitation, multiple experiments are often required to resolve a single distance. There has also been a large effort to develop spin labels that require less post-synthesis modification[88,90,92]. However, these labels position the nitroxide away from the duplex, further complicating data analysis. Alternatively, site specific tags, which can chelate, either $Gd^{3+}$ or $Mn^{2+}$, have been attached to DNA and RNA for ESR distance measurements[99]. These tags are typically longer than nitroxide spin labels. $Cu^{2+}$-$Cu^{2+}$ DEER distance in DNA has also been reported, using $Cu^{2+}$ binding porphyrins bound to quadruplexes containing multiple G-quartets[100]. This method requires an extremely specific DNA sequence to produce the necessary g-quadruplex structure. There still remains need for a labeling strategy that can place the label closer to the helix in order to incisively measure DNA structure and flexibility.

Herein, we report a $Cu^{2+}$ binding motif, for DEER measurements in DNA, which utilizes a 2,2'-Dipicolylamine (DPA) phosphormadite. Inorganic chelators have been used to create

metal mediated base pairing in DNA in order to increase the data storage capabilities of DNA[101-103]. Short range distance measurements (~3 Å) through dipolar coupling have also been performed with such DNA metal chelators[104]. The volume of the DPA-Cu$^{2+}$ motif is ~15% larger than guanine. The DPA containing DNA double helix used in this paper will be abbreviated as DPA-DNA, and is shown in Figure 2-1. The opposing strand uses a dSpacer, a commercially available tetrahydrofuranyl residue which is a nucleotide absent sugar-phosphate backbone.

## 2.2    MATERIALS AND METHODS

### 2.2.1   Oligonucleotides

2,2'-Dipicolylamine phosphoramidite was obtained from Glen Research. The sequence used was as follows: 5'-CAG(X)AGATATC(Z)TGA-3', as well as its complementary strand, where (X) represents the 2,2'-Dipicolylamine and (Z) represents the dSpacer. The DPA-DNA oligonucleotides were synthesized, purified by high performance liquid chromatography, characterized by mass spectroscopy and annealed at TriLink Biotechnologies. The finished DPA-DNA sequence was then lyophilized to a solid for resuspension. The control DNA oligonucleotide with a sequence of 5'-CAGTAGATATCATGA-3', and control DNA II 5'-GGGAATTAATTGGG-3', both purified by standard desalting, were purchased from Integrated DNA Technologies (IDT).
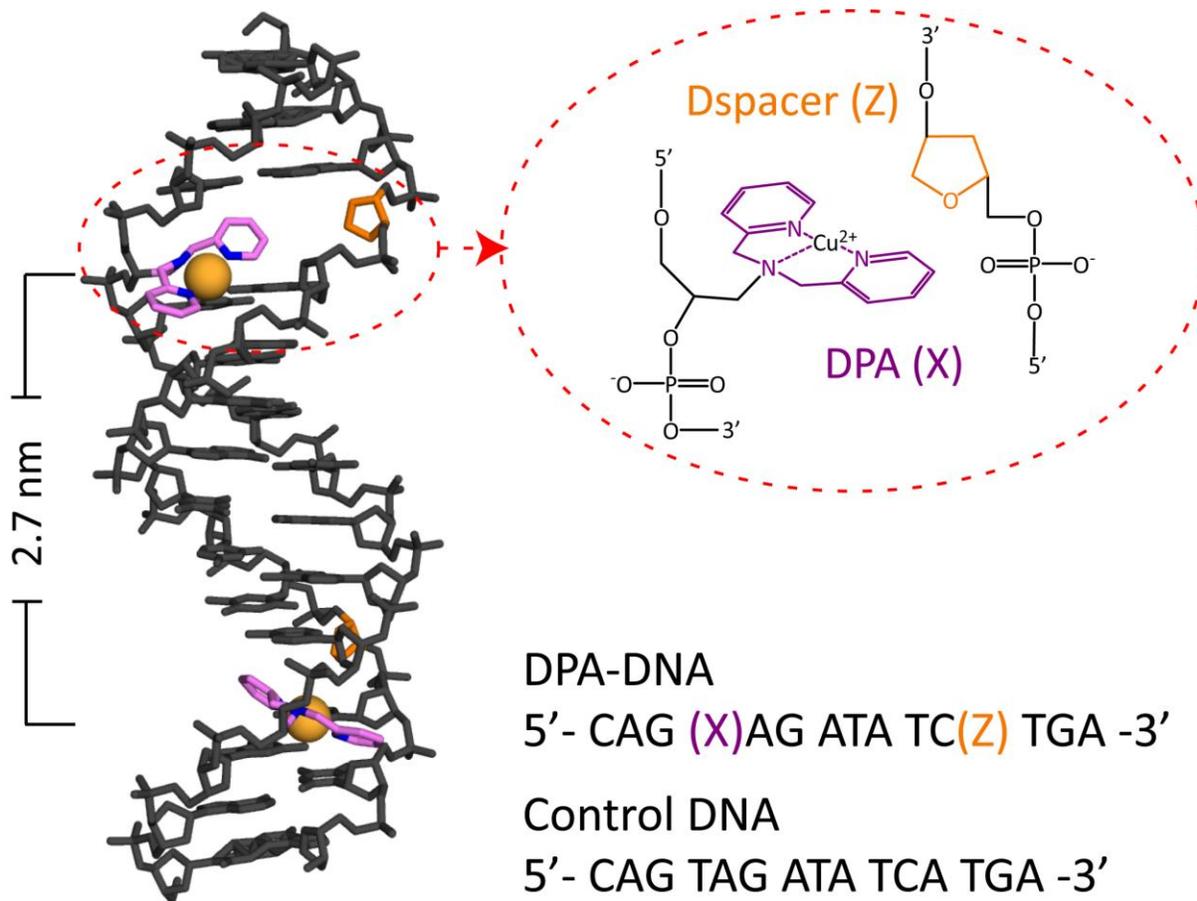
Figure 2-1. (Left) The DPA-DNA duplex used in this study. The inset shows the chemical structure of the Cu$^{2+}$ binding site. DPA binding sites are colored in purple and the dspacer is colored in orange. (Bottom right) The sequence of both DPA-DNA as well as control DNA are shown. These were paired with their respective complementary strands.

## 2.2.2   Sample preparation

All ESR samples were prepared at a concentration of 100 µM in DNA and 400 µM in $Cu^{2+}$. After addition of $Cu^{2+}$, DPA-DNA was annealed 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 10 minutes, 30°C for 5 minutes and then progressively cooled to 4°C. Samples were prepared in 50 µM NEM buffer at pH 7.4 in order to ensure that free excess $Cu^{2+}$ is ESR silent $Cu(OH)_2$[21]. For ESR experiments 20% v/v glycerol was added as a cryo-protectant.

## 2.2.3   Circular Dichroism and Melting Curves

CD spectra were collected using an Olis DSM17 Circular Cichroism Spectrometer. Samples were 100 µM DNA duplex prepared in 50 mM NEM (pH 7.4) buffer, and the data were collected in 2 mm quartz cells. Scans were completed at a constant temperature of 25 °C. The spectra were recorded from 200 nm to 335 nm in 1 nm increments with a 2 nm bandwidth. Melting curves were performed at 275 nm from 20 °C to 90 °C with 1 °C increments. The data was normalized to 1 for direct comparison.

## 2.2.4   ESR Measurements

Continuous wave (CW) spectroscopy experiments were performed on a Bruker ElexSys E580 CW/FT X-band spectrometer using a Bruker ER4118X-MD5 resonator at 80 K. All CW

experiments were collected at a center field of 3100 G with a sweep width of 2000 G for a total of 1024 data points. A modulation amplitude of 4 G, a modulation frequency of 100 kHz and a conversion time of 20.48 ms were used. Spectra were simulated using Bruker Simfonia. The CW data was acquired at 80 K.

Three-pulse electron spin echo envelope modulation (ESEEM) spectroscopy experiments were performed at X-band frequencies on a Bruker ElexSys E680 CW/FT X-band spectrometer equipped with a Bruker EN4118X-MD4 resonator. The $\pi/2 - \tau - \pi/2 - T - \pi/2 -$ echo pulse sequence was used with a $\pi/2$ pulse length of 16 ns. The first pulse separation, $\tau$, was kept fixed at 144 ns. The second pulse separation, T, was set to 288 and was progressively stepped out using a step size of 16 ns. Spectra were collected at a magnetic field of 3360 G and at a temperature of 80 K. The time for data acquisition ranged was 3 hours. The raw data was phase and baseline corrected, and zero filled to 2048 points. The data was then fast Fourier transformed.

The four-pulse two dimensional hyperfine sub-level correlation (HYSCORE) spectroscopy experiment was performed at X-band frequencies on a Bruker ElexSys E680 CW/FT spectrometer equipped with a Bruker EN4118X-MD4 resonator at 80 K. The $\pi/2 - \tau - \pi/2 - \tau1 - \pi - \tau2 - \pi/2 -$ echo pulse sequence with a $\pi/2$ pulse length of 16 ns and $\pi$ pulse length of 32 ns was used. Both $\tau1$ and $\tau2$ were stepped out by 16 ns. The data was collected at 3360 G, the field of maximum echo intensity. Data acquisition took 15 hours. The raw data was phase corrected and baseline corrected, and zero filled to 256 points in both dimensions. The data was then fast Fourier transformed and reported as a contour plot.

The four-pulse double electron-electron resonance (DEER) spectroscopy experiments were performed at X-band using a Bruker EN4118X-MD4 resonator. The experiments were

performed at 20 K. The following pulse sequence was used $(\pi/2)\nu1 - \tau1 - (\pi)\nu1 - T - (\pi)\nu2 - \tau2 - (\pi)\nu1 - \tau2 - \text{echo}$[48]. The observer pulse lengths, $(\pi/2)\nu1$ and $(\pi)\nu1$, were 16 ns and 32 ns, respectively, while the pump pulse length, $(\pi)\nu2$, was 16 ns. The delay, T, was incremented by a step size of 8 ns for 128 points. The pump frequency, $\nu2$, was placed at the maximum of the echo detected field swept nitroxide spectrum and the observer frequency, $\nu1$, was offset 150 MHz downfield. The raw time domain DEER data were analyzed using DeerAnalysis2013[105] via Tikhonov regularization. The data acquisition time was 3 hours.

## 2.3    RESULTS AND DISCUSSION

This DPA-DNA sequence is non-palendromic which avoids any possible self-folding. As a control, ESR experiments were also performed on 5'- CAGTAGATATCATGA -3' along with its complementary strand which omits the DPA-$Cu^{2+}$ binding site and dspacer. This DNA sequence will be referred to as the control DNA. Note, that both the DPA-DNA and the control DNA contain several guanine bases that can potentially bind $Cu^{2+}$.

### 2.3.1 DPA Does Not Affect DNA Helical Structure

Circular dichroism (CD) and melting temperature indicates that insertion of the DPA ligand does not disrupt the B-handed helical nature of the small DNA duplex. CD data was

collected on both control DNA and DPA-DNA to determine whether the addition of the DPA-$Cu^{2+}$ binding site had any effect on the DNA duplex structure. Shown in Figure 2-2, both the DPA-DNA and the control DNA CD spectrum showed negative ellipticity at 240 nm and positive ellipticity at 270 nm which is characteristic of a right handed B-DNA[106,107]. The DPA-DNA has a less negative band, which has also shifted by 5 nm. This shallower, shifted negative band is consistent with a decrease in A+T base pairing[108]. This result is expected because the DPA-dSpacer pairs were substituted with A+T pairs in the control DNA. The melting temperature of the both DPA-DNA and control DNA were examined as shown in Figure 2-3. The melting temperatures are 57±6 °C and 61±5 for DPA-DNA and control DNA respectively. This amount of lower melting temperature of DPA-DNA is expected because two A+T pairs have been replaced by the DPA binding site[109].

## 2.3.2  $Cu^{2+}$ Preferentially Binds to DPA Binding Site

$Cu^{2+}$ binding was first ascertained by collecting the CW spectrum of the DPA-DNA and the control DNA shown in Figure 2-4. Figure 2-4 shows the continuous wave (CW) $Cu^{2+}$-ESR spectra of DPA-DNA and control DNA. The CW spectrum of control DNA is likely due to $Cu^{2+}$ coordinating to guanine[110] (cf. Figure 2-4).  A 4:1 $Cu^{2+}$: DNA ratio was used to demonstrate that even at this high concentration ratio, non-specific binding would not affect DEER measurements. Since the samples were prepared in the NEM buffer at pH 7.4, any unbound $Cu^{2+}$ would precipitate into ESR silent $Cu(OH)_2$[21]. Therefore, any signal present must come from some form of bound $Cu^{2+}$.
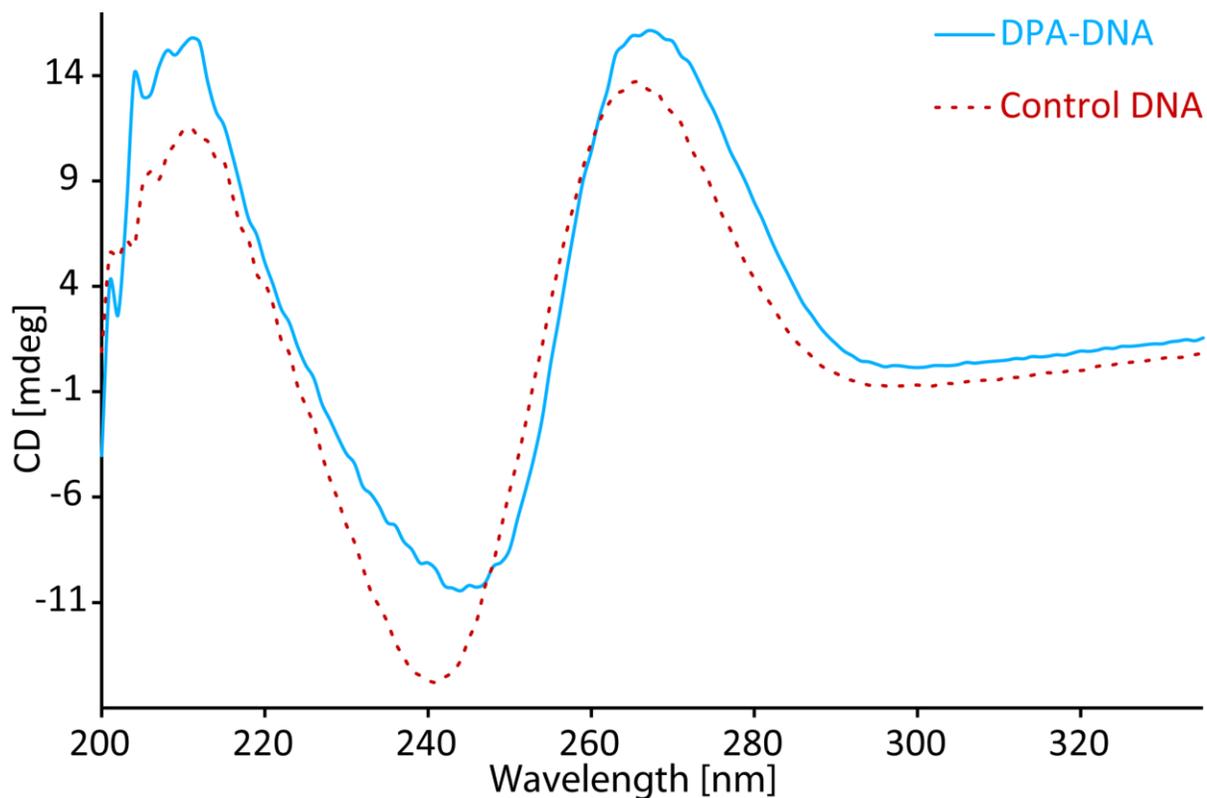
Figure 2-2. Circular dichroism spectra of DPA-DNA and control DNA. Both spectra show negative elipticty at 240 nm and positive elipticity at 270 nm. The pattern is the characteristic spectrum of right handed B-DNA double helix. Thus, the insertion of the $Cu^{2+}$ binding DPA into the DNA does not interfere with the typical DNA double helical structure.
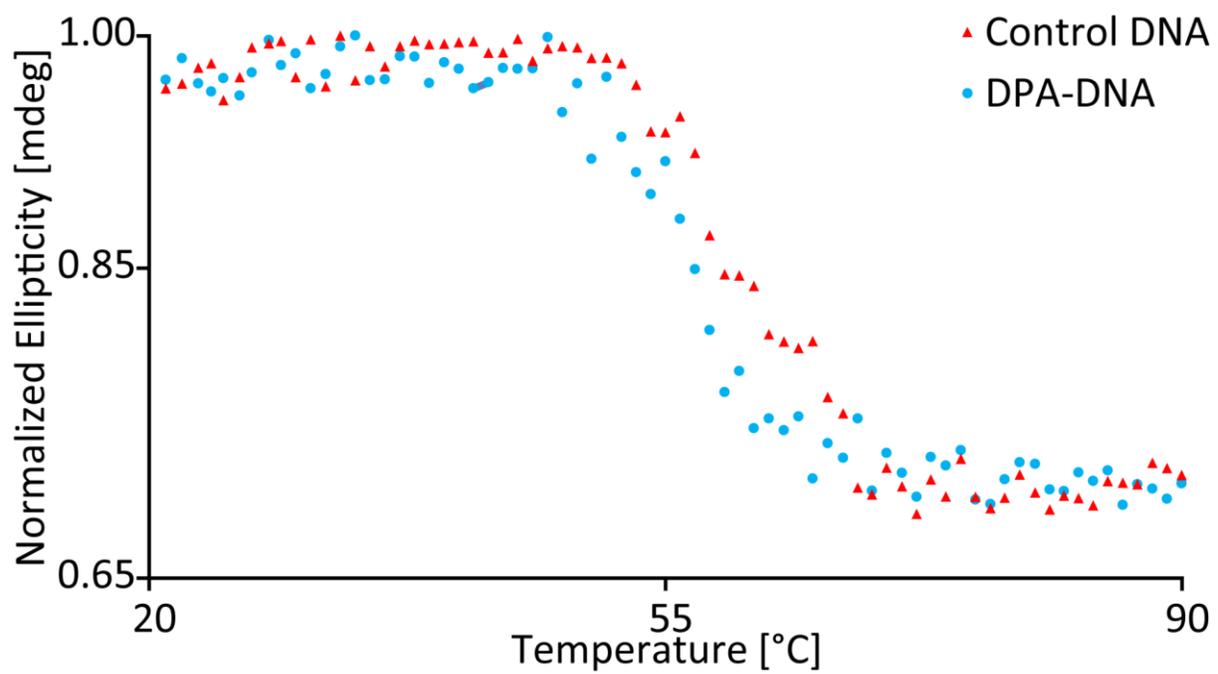
Figure 2-3. Melting curves for both DPA-DNA (blue dots) and control DNA (red triangles).
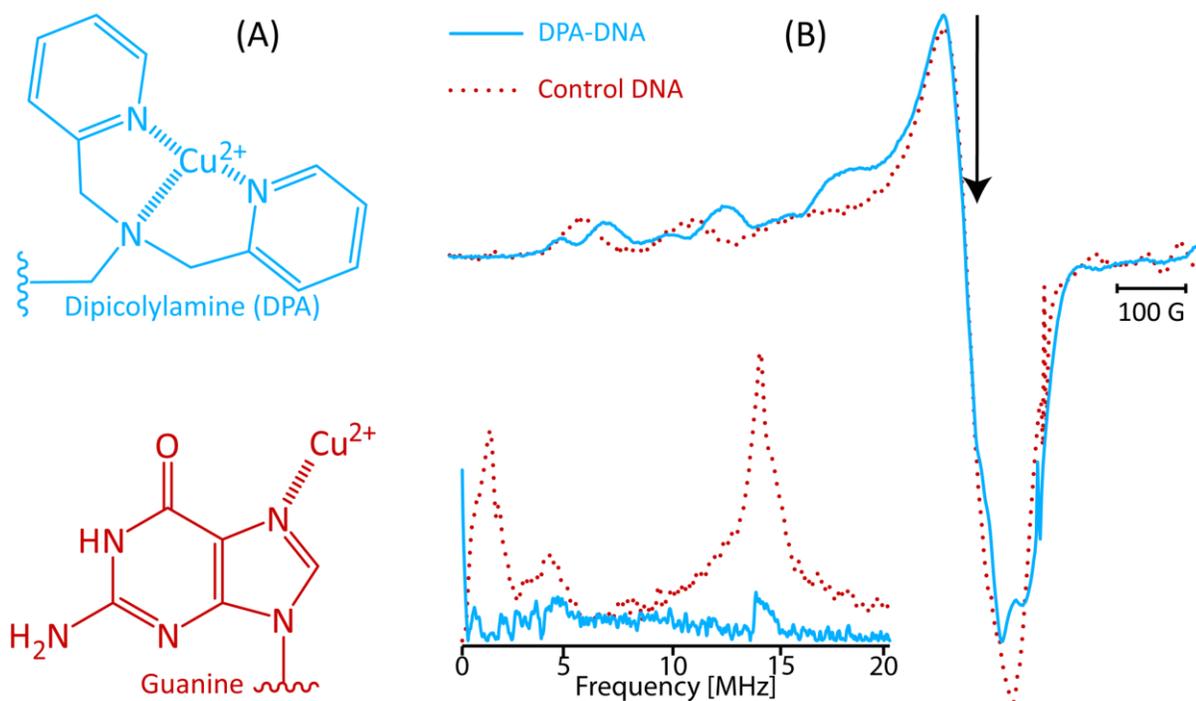
Figure 2-4. A) Coordination environment of both $Cu^{2+}$ with DPA (blue) and the guanine (red). B) CW spectra of DPA-DNA (blue) and control DNA (dotted red) clearly shows differences in both $g_{\parallel}$ value and hyperfine splitting. The Inset shows ESEEM spectra of DPA-DNA and control DNA. The magnetic field at which ESEEM was acquired is indicated by the black arrow. The peaks below 4 MHz are only present in control DNA indicative of imidazole coordination. The DPA-DNA has no such peaks as expected.

To determine that the DPA $Cu^{2+}$ binding site was populated by $Cu^{2+}$ rather than unspecific binding, both CW spectra were simulated to measure $g_\parallel$ and $A_\parallel$ values. Simulated CW spectra are shown in Figure 2-5 with corresponding parameters found in Table 2-1. The $g_\parallel$ and $A_\parallel$ values report on the number of directly coordinated nitrogen and/or oxygen atoms. The DPA-DNA spectrum shows two components each with $g_\parallel$ and $A_\parallel$ values typical of that of a $Cu^{2+}$ complex equatorially coordinated by four ligands, known as a Type II $Cu^{2+}$ center. The $g_\parallel$ and $A_\parallel$ values are consistent with either 3 nitrogen/1 oxygen or 2 nitrogen/2 oxygen equatorial ligands respectively[15]. The control DNA has $g_\parallel$ and $A_\parallel$ values consistent of a Type II $Cu^{2+}$ complex coordinated to 1 nitrogen/3 oxygen. This data shows that $Cu^{2+}$ preferentially binds to DPA over non-specific DNA coordination.

CW titrations were performed in order to determine the binding of $Cu^{2+}$ to DNA. By comparing the integrated intensity of the CW spectrum, which is directly proportional to the number of spins, to a calibration curve, we determined how many $Cu^{2+}$ ions are bound to the control DNA per equivalent of $Cu^{2+}$. The data is shown in Figure 2-6 on the control DNA. These results indicate that there are ~1 $Cu^{2+}$ per DNA in the control DNA and 0.5 $Cu^{2+}$ per DPA-DNA duplex. With 5 guanines in the DPA-DNA sequence there exist 0.1 $Cu^{2+}$ ions per guanine and full occupancy at the DPA sites.

Electron spin echo envelope modulation results further suggest the preferential binding of $Cu^{2+}$ to the DPA binding site over non-specific DNA coordination. Figure 2-4 shows the spectra for control DNA and DPA-DNA. The data was acquired at a magnetic field of 3360 G. The spectrum of control DNA, shows peaks below 4 MHz, that is consistent with coordination to the imidazole[32,40,58] of guanine. The spectrum of DPA-DNA is markedly different. Comparing these

Table 2-1. CW ESR Simulation Parameters for DPA-DNA and DNA

| Spectrum | $g_\parallel$ | $A_\parallel$ (G) | % |
|---|---|---|---|
| DPA Component 1 | 2.244 | 170 | 70 |
| DPA Component 2 | 2.305 | 157 | 30 |
| DNA | 2.280 | 163 | * |
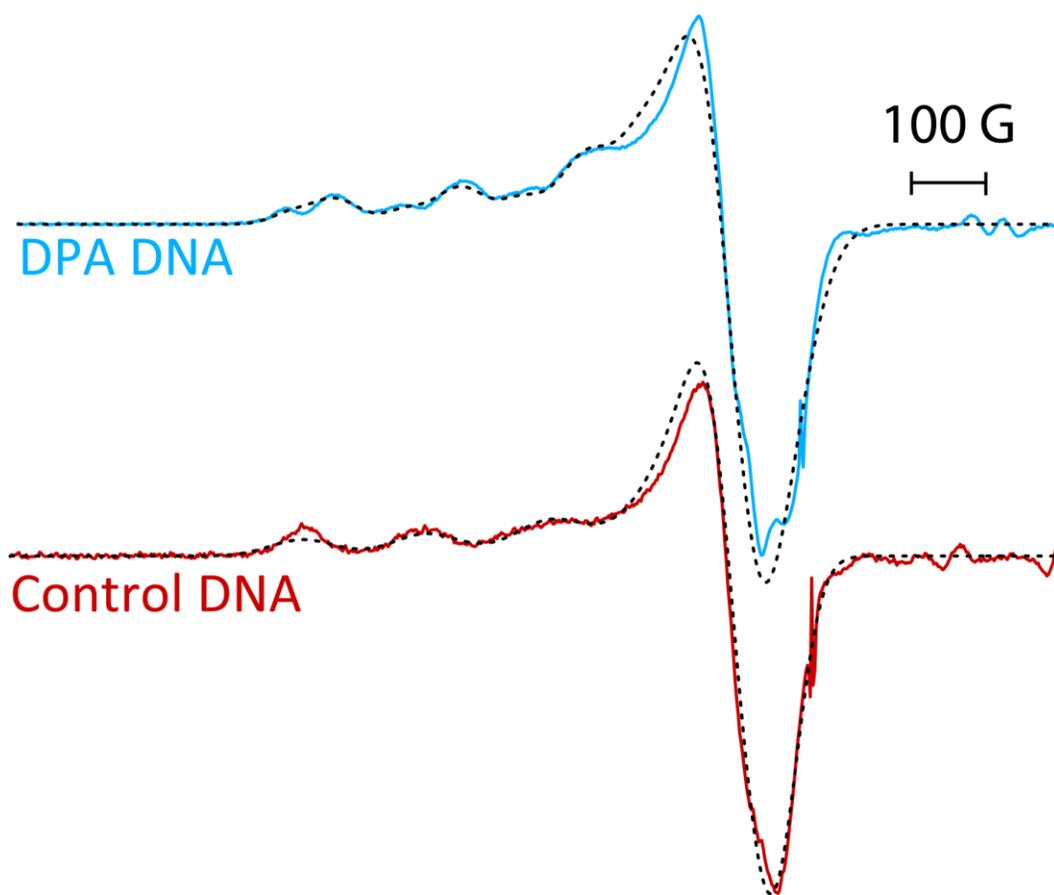


Figure 2-5. CW ESR spectra of DPA DNA (top, blue) and Control DNA (bottom red) with their respective simulations fit by Bruker Simfonia (black dotted).
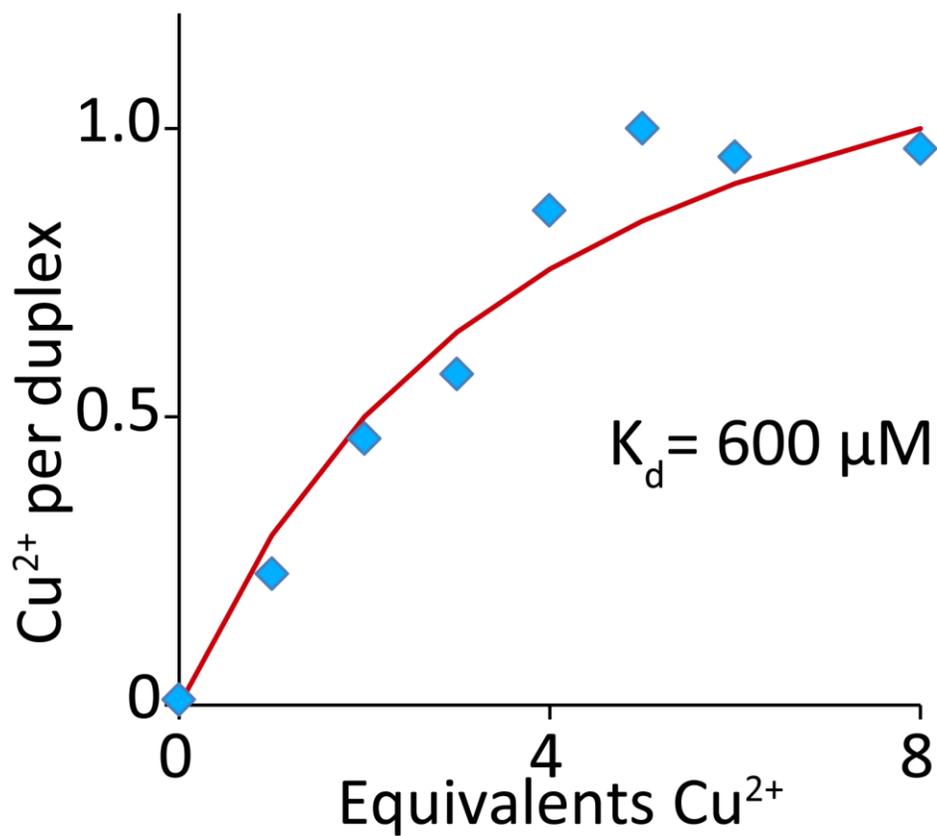
Figure 2-6. Titration curves illustrating sequential additions of equivalents of $Cu^{2+}$ to the control DNA. For control DNA, at 2 equivalents of free $Cu^{2+}$, there would be on average 0.1 $Cu^{2+}$ per guanine.

two spectra suggests that the $Cu^{2+}$ preferentially binds to the DPA binding site. There is only 0.1 $Cu^{2+}$ ion per guanine in DPA-DNA as per Figure 2-6. The ESEEM from guanine coordination is weak because the signal is dominated by $Cu^{2+}$ bound to DPA.

To further investigate the coordination of the DPA $Cu^{2+}$ binding site, we used the HYSCORE experiment. The Fourier transformed, frequency domain HYSCORE spectrum represented as a contour plot is shown in Figure 2-7. There are no features that would indicate any distal nitrogen coordination such as imidazole nitrogen, consistent with our expected binding environment shown in Figure 2-4. The peaks at 14 MHz indicate hydrogen coordination. This hydrogen signal is comprised of cross-peaks from 9 MHz to 19 MHz which are consistent with equatorially coordinated water molecules, suggesting the fourth equatorially coordinated ligand is $H_2O$. Once again, this is consistent with expected binding environment depicted in Figure 2-4.

### 2.3.3   DPA-DNA Reports Directly on Backbone-Backbone Measurement

Double Electron-Electron Resonance (DEER) experiments were performed on the DPA-DNA and control DNA. The data is shown in Figure 2-8. The control DNA shows no visible dipolar modulation, which suggests that non-specific binding should not impact $Cu^{2+}$-based DEER, especially given the higher affinity of DPA. DEER data on DPA-DNA was acquired at $g_\perp$ and $g_\parallel$, in order to examine whether limited spectral excitation leads to the data becoming sensitive to relative orientations of the g-tensors (i.e. orientational selectivity)[88,91].The background subtracted time domain data and the resultant distance distribution obtained using Tikhonov regularization[105]  are shown in Figure 2-8. The raw time domain DEER data and are shown in Figure 2-9.
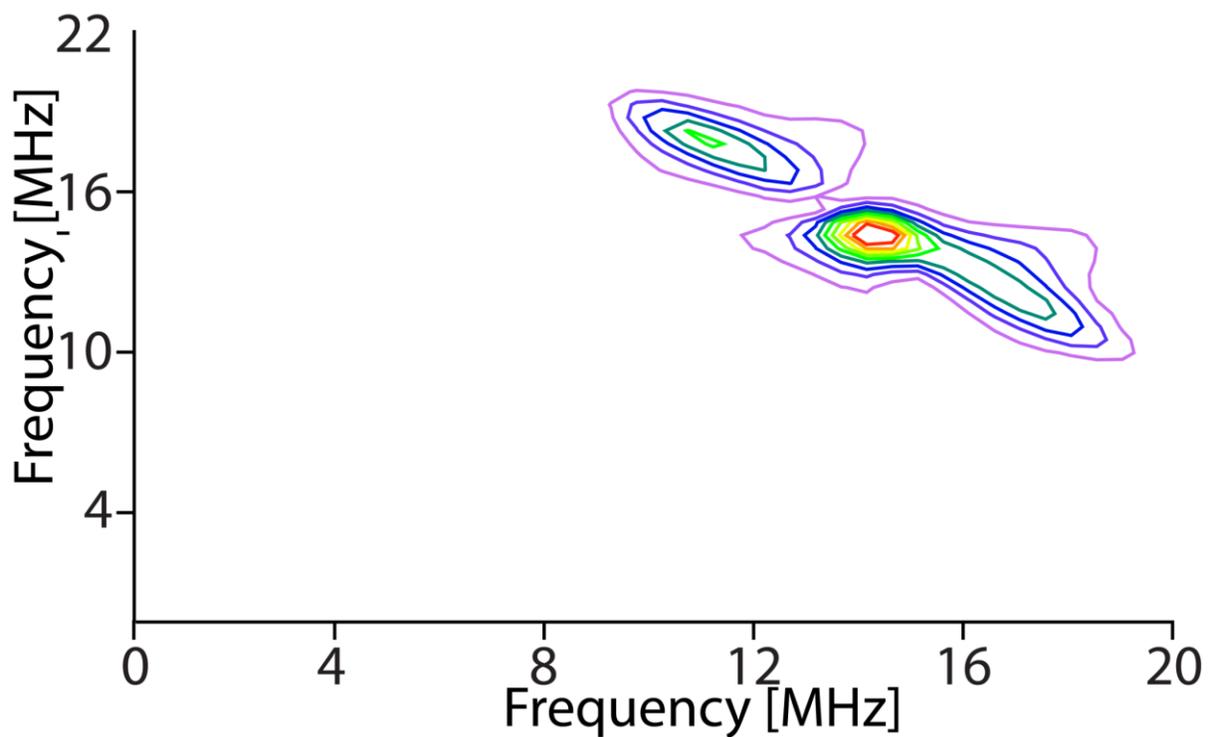
Figure 2-7. HYSCORE spectrum of DPA-DNA in the presence in $Cu^{2+}$. The major feature at 14 MHz comes from the interaction with nearby Hydrogen nuclei. The cross peaks visible between 9 MHz and 19 MHz are indicative of equatorially coordinated water molecules which are consistent with the expected binding environment of a single DPA chelator. The signals below 4 MHz can be attributed to noise.
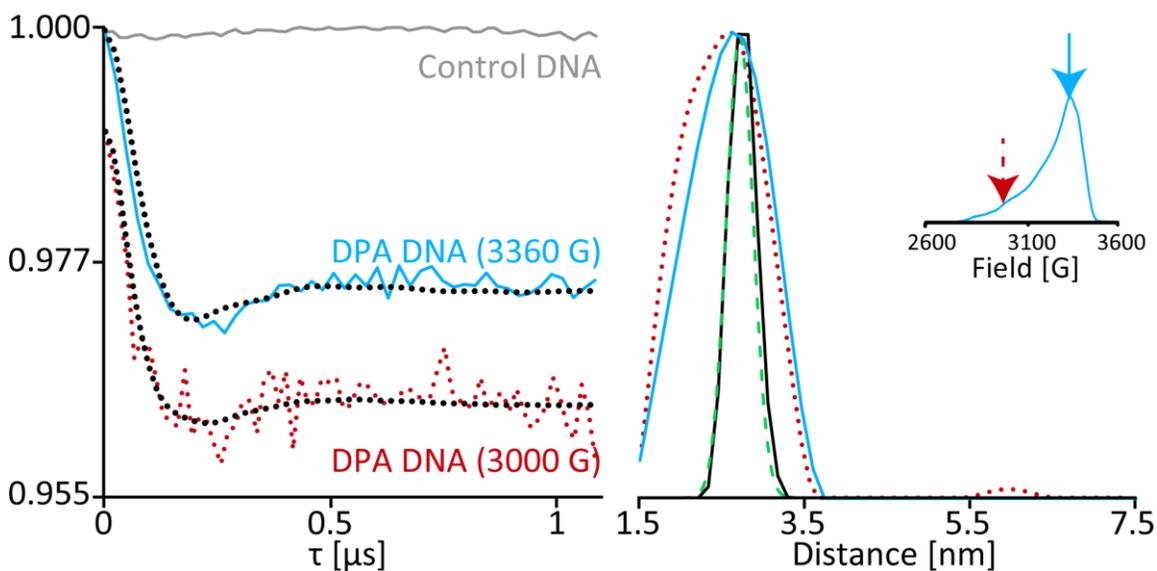
Figure 2-8. DEER data. (Left) Background subtracted time domain of DPA-DNA at two different fields and control DNA. The DPA-DNA showed no change in modulation period between the two differing fields suggesting no effects of orientational selectivity. It is clear that the control DNA had no observable dipolar modulations. Signals were normalized and offset for clarity. (Right) Resultant distance distribution of DPA-DNA at 3360 G and 3000 G. Inset shows the field swept spectrum and positions of the pump pulses 3360 G and 3000 G, blue solid arrow and red dashed arrow respectively. Simulated DEER distance distribution shown by dashed green line. MD simulation of DNA distribution shown by black solid line.

Furthermore, we have used validation techniques in DeerAnalysis2006 to produce error estimates of the reported DEER distance distributions. The DEER experiment was collected twice at the field of highest intensity (3360 G), and validation is shown in Figure 2-10. The validation of both spectra shows that the distance distribution is both reproducible and unaffected by differing background subtractions.

The DEER data and distance distributions at the two magnetic fields are similar, which suggests that orientational effects are negligible, likely due to a distribution in relative orientation of the g-tensors[1,8,75]. These data suggest that distance measurements using DPA within DNA duplexes can be achieved via a single experiment with the pump pulse positioned at the field of highest intensity. The most probable distance of 2.7 nm measured by DEER is in excellent agreement with the expected distance of 2.7 nm for B-DNA[79,111].

The DEER distance distribution can be analyzed by comparison to a model of distribution in DNA duplex length $(r)$[79,111,112]. For short DNA strands, the variance in duplex length $(\sigma_r)$ can be best modeled by a varying helix radius $(\sigma_R)$ which is called the "breathing" model[79]. For short DNA duplexes, $\sigma_R$ can be taken as 0.65 Å where the average DNA radius is taken to be 5.85 Å[79]. The work of Matthew-Fenn and coworkers[111] states:

$$\sigma_r^2 = 0.041n^2 + const \tag{2-1}$$

The first term in Eq 1 takes DNA mobility into account. In this equation 0.041 is a constant associated with $\sigma_R$, $n$ is the number of basepairs and the term labeled "const" accounts for label mobility. Taking $const = 0$, an end-to-end distance distribution of the DNA duplex is estimated. To model the expected backbone distribution for our experiments, $n$ was taken to be 9 basepairs. Using these values to calculate $\sigma_r$ with eqn. (1), we can produce a distribution of distances to model the inherent DNA mobility[79] using:
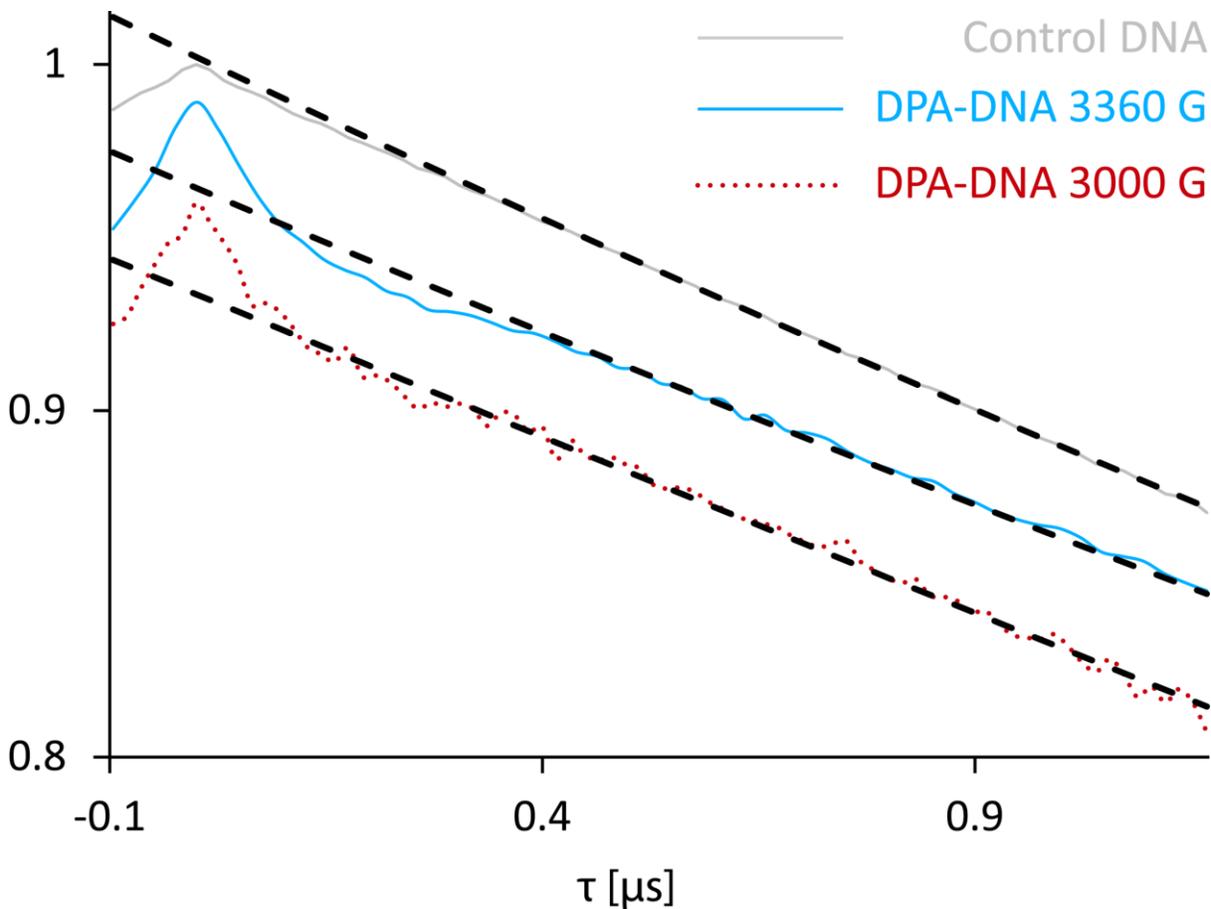
Figure 2-9. Raw time domain dipolar modulations of control DNA (grey), DPA DNA at 3360 G (blue) and 3000 G (red dotted) as well as their respective background fits (black dashed). Data are displayed at differing scales and offset for clarity. No modulations are apparent in the time domain of the control DNA meaning non-specific binding does not contribute to the DEER signal. Both DPA time domain signals are remarkably similar suggesting no orientional effects.
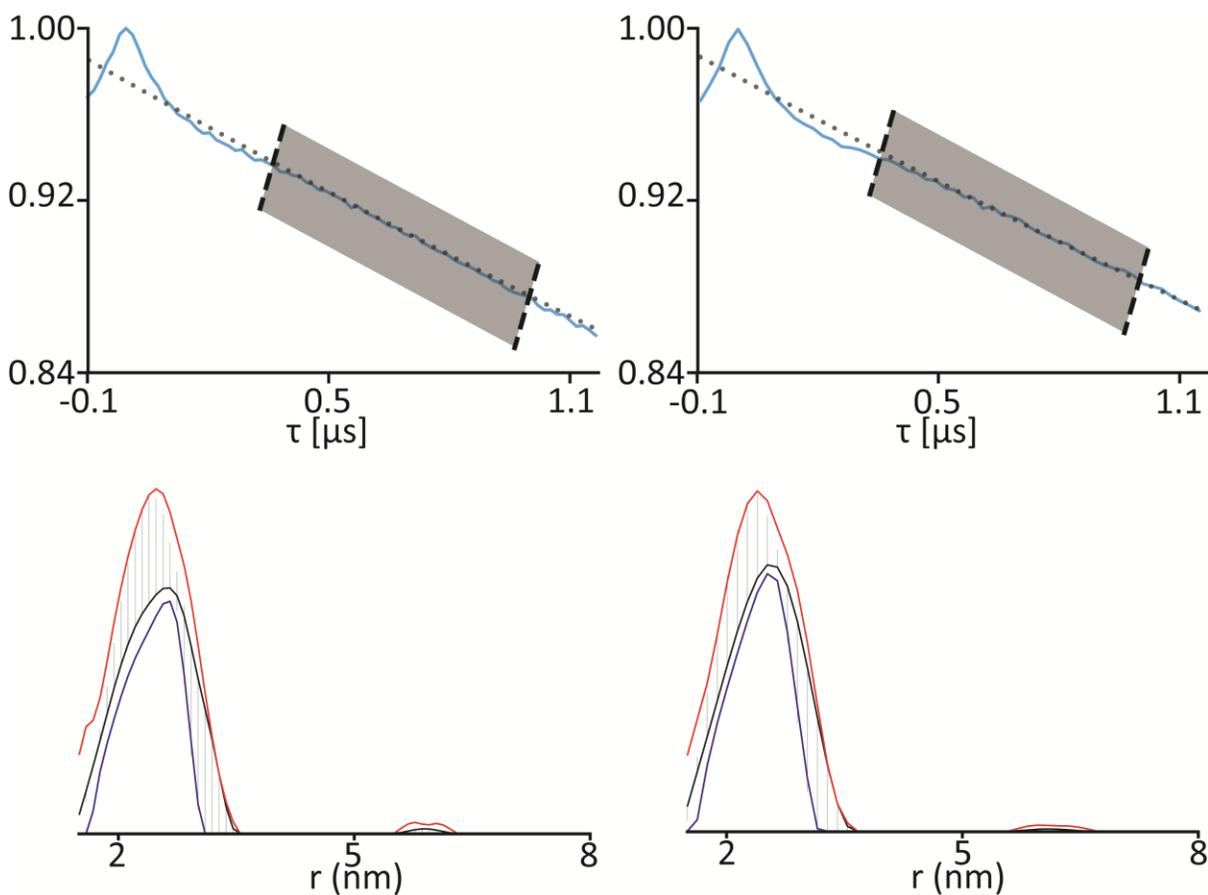
Figure 2-10. (Top) Raw time domain dipolar modulations of the DPA-DNA (blue line) collected twice (48 scans, left; 250 scans, right) along with the exponential background decay fit (black dotted). The shaded region is the range of the starting point for varied background subtraction used for validation. (Bottom) Resultant distance distributions (black) from validation techniques including upper (red) and lower (blue) acceptable limits.

$$f(r)d\Delta r = \frac{1}{\sqrt{2\pi\sigma_r^2}}\exp\left(\frac{(r-\langle r\rangle)^2}{2\sigma_r^2}\right)dr \tag{2-2}$$

Where $r$ is a small change in duplex length and $\langle r\rangle$ is the average duplex length determined by the number of base pairs and the average length per base pair, 3.4 Å.  The average length, $\langle r\rangle$, can be calculated by the following:

$$\langle r\rangle = n * 3.4 \tag{2-3}$$

Figure 2-8 shows the calculated backbone distribution for the control DNA using this breathing mode model[79,111] (green dashed line). The inherent DNA backbone distance distribution, calculated from a 33 ns MD simulation of the control DNA, is also shown in Figure 2-8. In both cases, the most probable distance is in exact agreement with the experimental results. The result is significant because it suggests that ESR directly measures the DNA backbone distance. With current labels, the reported distance can be up to 1-1.5 nm larger than expected. While the precise DNA backbone distance can be extracted, modeling is required. On the other hand, the $Cu^{2+}$-$Cu^{2+}$ DPA-DNA distribution is broader than what is expected for the control DNA (cf. Figure 2-8). The larger width is likely due to DPA motion in the plane perpendicular to the helical axis.  Such motion can be reduced by coordination to an atom in the opposing strand[101-104].

## 2.4    CONCLUSION

In summary, we report the first nucleotide and structure independent site specific $Cu^{2+}$-$Cu^{2+}$ distance measurement. The distance accurately reports on the DNA-backbone distance without the need for analysis of label conformations. Furthermore, only a single measurement is

required due to the lack of orientational effects. The distribution width can be further improved by including a coordinating atom on the opposing DNA strand[101,102]. The DPA-DNA scheme is likely to be useful for the measurement of DNA structure and dynamics, as well for understanding the role of DNA structure in mediating protein-DNA interactions. The ability to label DNA with $Cu^{2+}$ is also useful for orthogonal labeling for distance measurements[113,114].

## 2.5     ACKNOWLEDGEMENTS

# 3.0 ON THE USE OF THE CU$^{2+}$-IMINODIACETIC ACID COMPLEX FOR DOUBLE HISTIDINE BASED DISTANCE MEASUREMENTS BY PULSED ESR

*This work, written in collaboration with Shreya Ghosh, Timothy F. Cunningham, Amit Shimshi and Sunil Saxena, has been adapted from Ref. with permission from the PCCP Owner Societies. The thesis author collected a majority of the ESR data, analyzed the data and prepared the manuscript.*

## 3.1 INTRODUCTION

Electron spin resonance (ESR) distance measurements using paramagnetic metal ions is increasingly important for determining structure and dynamics[8,55,68] of biological macromolecules[5]. Efforts using paramagnetic metal centers include both intrinsic metal centers and metal-based spin labeling using metal ions such as Cu$^{2+}$,[1,12,58,64-67,69,70,74,78,115,116] Gd$^{3+}$,[53,55,117-125] and Mn$^{2+}$[126-131]. Once in place, the four pulse double electron-electron resonance (DEER) experiment can be performed utilizing these paramagnetic metals[6,48]. DEER measures the most probable distance as well as the distribution of distances between the two metals and thus can provide key insights into the flexibility and structure of the biological molecule of interest.

66

$Cu^{2+}$ based ESR distance measurements are of interest for many reasons. First, many proteins contain natural $Cu^{2+}$ binding sites that provide excellent rigidity for the unpaired electron as compared to the more common nitroxide sidechain, R1[8,73-76,132]. In addition, there are a number of newly developed strategies for site specific labeling that incorporate $Cu^{2+}$ with inherent flexibility comparable to R1[78]. One particular strategy that shows even greater promise is the double histidine (dHis) motif[133-139], in which two strategically placed histidine side-chains are used to chelate $Cu^{2+}$. For α-helical sites, an i and i+4 arrangement of two His residues allows for simultaneous cis-coordination of the $Cu^{2+}$, while in β-sheets, an i and i+2 double His mutation is necessary to place the two histidine on the same face of the sheet for metal binding. Once in place, the dHis motif can site specifically bind $Cu^{2+}$. However, a principal limitation is poor $Cu^{2+}$ selectivity for the dHis site, as $Cu^{2+}$ can coordinate unspecifically to unwanted protein sites.

Recently, we showed that $Cu^{2+}$ chelated to iminodiacetic acid (IDA) increases the selectivity of the $Cu^{2+}$ to the dHis site and reduces non-specific coordination[58]. IDA coordinates to the $Cu^{2+}$ by occupying two neighboring equatorial positions and one axial position, leaving two equatorial sites open for cis-coordination to dHis (cf. Figure 3-1A). The final axial position is occupied by solvent water[58]. In turn, this rigid motif provides extremely narrow distance distributions that are up to five times narrower as compared to similar distance measurements performed by R1-DEER[58]. In addition to the precise distance measurement, the entire dHis-$Cu^{2+}$-IDA binding pocket self assembles in solution, requires no post-expression covalent modification, and provides data that can be easily related to the backbone-backbone distance[58].
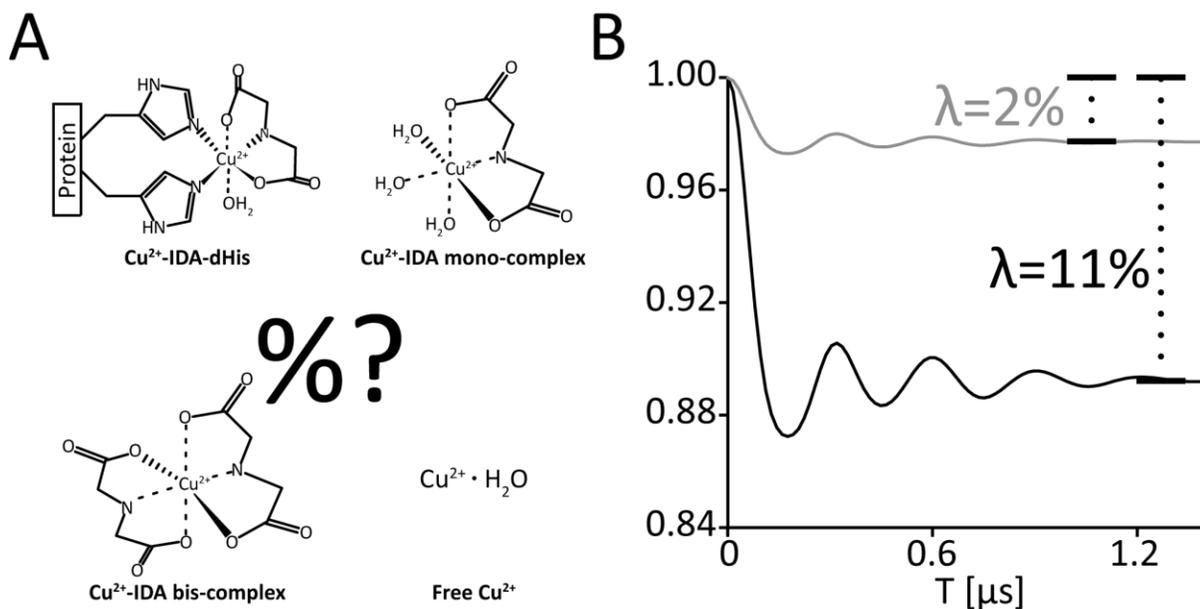
Figure 3-1. A) Possible $Cu^{2+}$ coordination environments in solution include the dHis-$Cu^{2+}$-IDA complex, $Cu^{2+}$-IDA mono-complex, $Cu^{2+}$-IDA bis-complex, and free $Cu^{2+}$. B) Simulated data illustrating DEER measurements with a modulation depth ($\lambda$) of 12% ( as expected for a 16 ns pump pulse for a $Cu^{2+}$-DEER experiment)[1] compared to the 2% achieved in the initial dHis measurement.

Despite the early promise of the dHis-$Cu^{2+}$-IDA approach, a concern from the initial results was the relatively shallow dipolar modulations[58]. The DEER experiment measures the inter-spin distance by measuring the dipolar frequency between two-coupled spins. The experiment generates a modulating signal and when the experiment exhibits deep modulations, the relevant distance information can be more easily separated from the unwanted noise and nuclear modulations[1]. A quantitative measure of these dipolar modulations is called the modulation depth ($\lambda$, cf. Figure 3-1B), which is defined as the difference from the maximum of the normalized background-corrected DEER signal to where the signal becomes constant, as shown in Figure 1B. The modulation depth depends on a number of variables including the length of the pump pulse, frequency offset between the pump and observer pulses, and the spectral lineshape[7]. In addition, the modulation depth is sensitive to the number of coupled spins in a system, where more spins per molecule produces a larger modulation depth[47,56]. A large modulation depth is highly desirable in a DEER experiment, since the sensitivity is directly proportional to the modulation depth[3,7,47,140].

Initial experiments utilizing the dHis-$Cu^{2+}$-IDA motif in the B1 immunoglobulin-binding domain of protein G (GB1) produced a modulation depth of ~2% whereas a modulation depth of 10-12% was anticipated from prior theoretical and experimental work on other $Cu^{2+}$-systems[1]. While the number of measured modulation periods still allowed for the extraction of relevant distance information, maximizing the modulation depth is essential in order to increase applicability[141]. Therefore, a systematic understanding of the causes of the low modulation depth is needed to improve the utility of the dHis motif as a spin label.

We hypothesize that the primary causes for the low modulation depth is due to inefficient dHis-$Cu^{2+}$-IDA loading and incomplete $Cu^{2+}$-IDA complex formation. Poor dHis-$Cu^{2+}$-IDA

loading produces a smaller modulation depth simply by providing less number of spins per protein. At the same time, unbound $Cu^{2+}$-IDA complex (cf. Figure 3-1A) will also contribute to the background, leading to further decrease of the modulation depth. In addition, incomplete $Cu^{2+}$-IDA formation could result in inaccurate estimates of proper $Cu^{2+}$-IDA:dHis ratios essential for proper sample loading. Since $Cu^{2+}$-IDA-imidazole coordination is highly dependent upon such factors as secondary structures, buffer salt concentration, and pH,[142] establishing the optimal conditions to promote dHis-$Cu^{2+}$-IDA formation is essential.

Herein, we perform a systematic investigation of these potential causes of the low modulation depth of the $Cu^{2+}$-DEER experiment when the $Cu^{2+}$ is coordinated at dHis sites. We use CW ESR to quantify the ratios of $Cu^{2+}$, $Cu^{2+}$-IDA mono-complex and $Cu^{2+}$-$IDA_2$ bis-complex in solution. We then quantitatively examine $Cu^{2+}$-IDA loading to α-helical and β-sheets by continuous wave (CW) ESR and Electron Spin Echo Envelope Modulation (ESEEM) spectroscopy. We develop methodology to combine CW_ESR and DEER results to extract populations of different $Cu^{2+}$-complexes in solution. These efforts are valuable in explaining the low modulation depth observed previously[141]. Finally, the wealth of information accrued is used to optimize the DEER experiment on the dHis-$Cu^{2+}$-IDA motif. With these optimized conditions, we are able to double the sensitivity of the DEER experiment.

## 3.2 MATERIALS AND METHODS

### 3.2.1 GB1 Expression, Purification and Cu$^{2+}$-IDA Loading

ESR experiments were performed on the B1 immunoglobulin-binding domain of protein G, called GB1. The mutation, expression, purification and Cu$^{2+}$-IDA loading of the 28H/32H, 6H/8H and 6H/8H/28H/32H mutant GB1 constructs were performed as previously described.[141] All samples for ESEEM and DEER were prepared in 50 mM of N-ethylmorpholine (NEM) buffer at pH 7.4 to exclude any signal originating from unbound Cu$^{2+}$[21].

### 3.2.2 EPR Measurements

For ESR measurements, a 120 µL aliquot of sample was placed in a quartz sample tube (3 mm inner diameter, 4 mm outer diameter). All samples were prepared in 20% v/v glycerol which is used as a cryoprotectant. All protein samples were prepared in 50 mM NEM buffer (pH 7.4) in 20% v/v glycerol, unless otherwise noted. Continuous Wave (CW) ESR, Electron Spin Echo Envelope Modulation (ESEEM) and DEER experiments were all 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. CW and ESEEM experiments were performed at 80 K while DEER experiments were performed at 20 K.

CW ESR experiments were performed at X-band frequencies. 1024 data points were collected over a sweep width of either 2000 G or 200 G with a modulation amplitude of 4 G or 1 G, respectively. All CW ESR data used a modulation frequency of 100 kHz, a time constant of 10.24 ms and a conversion time of 20.48 ms. Spectra were simulated and fit using the Bruker Simfonia software.

Three-pulse ESEEM experiments were performed at X-band frequencies using the $\pi/2 - \tau - \pi/2 - T - \pi/2 - $ echo pulse sequence with a $\pi/2$ pulse length of 16 ns. The first pulse separation, $\tau$, was set to 144 ns. The second pulse separation, T, was incremented by a step size of 16 ns from an initial value of 288 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[143,144]. Data acquisition took ~12 hours.

The DEER experiments were performed at 20 K at X-band frequencies. The four pulse sequence used was as follows: $(\pi/2)v1 - \tau1 - (\pi)v1 - T - (\pi)v2 - \tau2 - (\pi)v1 - \tau2 - $ echo[48]. The observer pulses, $(\pi/2)v1$ and $(\pi)v1$, were 16 ns and 32 ns respectively. The pump pulse, $(\pi)v2$, was 16 ns. The delay, T, was incremented using a step size of 10 ns for a total of 128 points. The pump frequency, $v2$, was placed at the maximum of the echo detected field swept spectrum with the observer frequency, $v1$, offset 150 MHz downfield. The raw time domain DEER data were analyzed by DeerAnalysis2013[105] via Tikhonov regularization. Distance distribution was corrected using proper g-values[115]. Data acquisition took between 5 and 12 hours.

### 3.2.3 Modulation Depth Calculations

The intramolecular signal $V_p(t)$ of the DEER experiment for a system with multiple $Cu^{2+}$ is[145]:

$$V_p(t) = \frac{1}{N}\sum_{j=1}^{N}\left\{\prod_{\substack{k=1\\k\neq j}}^{N}\left(1 - p_b(1 - \cos(D_{jk}t))\right)\right\} \tag{3-1}$$

where j and k are the different spins, $D_{jk}$ is the dipolar interaction between them, $p_b$ is the fraction of spins excited by the pump pulse, t is the time the pump pulse is applied and N is the number of spins. At t=∞, the $\cos(D_{jk}t)$ term is approximately zero[145]. Further simplification of Eq. (3-1) for each spin system gives the well-known[54,56,57]:

$$V_p(t = \infty) = (1 - p_b)^{N-1} \tag{3-2}$$

where modulation depth is related to $V_p(t)$ by:

$$V_p(t = \infty) = 1 - \lambda \tag{3-3}$$

In our 6H/8H/28H/32H-GB1 tetramutant sample, we have five possible systems in solution: fully loaded GB1 where a $Cu^{2+}$-IDA complex is loaded at both the 6H/8H site and also the 28H/32H site. We refer to this species as dHis2. In addition we have the half loaded GB1 where only one of the two dHis sites contains $Cu^{2+}$-IDA. Finally we have the free mono-complex ($Cu^{2+}$-IDA), bis-complex ($Cu^{2+}$-IDA$_2$), and free $Cu^{2+}$. Here, the only species containing two spins is dHis2. These different possible $Cu^{2+}$ systems, placed into Eq. (3-1) gives:

$$\lambda = 1 - f_{dHis2}(1 - p_b)^1 + f_{dHis}(1 - p_b)^0 + f_{IDA}(1 - p_b)^0 + f_{bis}(1 - p_b)^0 + f_{free}(1 - p_b)^0 \tag{3-4}$$

which simplifies to:

$$\lambda = 1 - [f_{dHis2}(1 - p_b)^1 + f_{dHis} + f_{IDA} + f_{bis} + f_{free}] \qquad (3\text{-}5)$$

where $f_{dHis2}$ is the fraction of $Cu^{2+}$-IDA bound to both dHis sites in the 6H/8H/28H/32H-GB1 tetramutant, $f_{dHis}$ is the fraction of $Cu^{2+}$-IDA bound to only a single dHis site in the tetramutant, $f_{IDA}$ is the fraction of $Cu^{2+}$-IDA mono-complex free in solution, $f_{bis}$ is the fraction of $Cu^{2+}$-IDA$_2$ bis-complex, and $f_{free}$ is the fraction of free $Cu^{2+}$. Finally, $f_{IDA}$ is the fraction of unbound $Cu^{2+}$-IDA complex in solution.

## 3.3     RESULTS AND DISCUSSION

Two different mutants of the B1 immunoglobulin-binding domain of protein G, called GB1, were used to understand $Cu^{2+}$-loading to dHis sites. The 28H/32H-GB1 mutant contains an α-helical dHis site, while the 6H/8H-GB1 mutant contains a β-sheet dHis site. Experimentation with GB1 is useful because the protein contains both α-helical segments and β-strands and because we have vast experience with its structural characterization[2,141,146].

### 3.3.1   Incomplete Cu²⁺-IDA complex formation revealed by CW ESR spectroscopy

We first use CW ESR spectroscopy to determine the amount of $Cu^{2+}$-IDA complex formed in solution. The CW ESR spectrum of the $Cu^{2+}$/IDA solution, where $Cu^{2+}$ was

held constant at 500 μM with pH 4.8-5.2 with differing amounts of acidic IDA, shown in Figure 3-2A (solid), was collected in water in order to detect any signal coming from free $Cu^{2+}$. The complexation of $Cu^{2+}$-IDA remains constant from pH 5 and pH 7.4[147], thus the titrations performed at pH ~5 hold true at pH 7.4. Two components are clearly visible in the CW ESR spectrum and the spectrum is simply generated (Figure 3-2A, dotted) by the ion of two individual CW ESR spectra: $Cu^{2+}$:IDA complex in NEM (inset black) and free $Cu^{2+}$ in water (inset grey). Note, that in NEM buffer free $Cu^{2+}$ forms insoluble $Cu(OH)_2$ and is ESR silent[21]. From the percentage of integrated intensities between these two components, the percentage of $Cu^{2+}$-IDA complex in solution can be determined. We collected the CW spectrum at varying $Cu^{2+}$:IDA ratios and measured the percent of $Cu^{2+}$-IDA complex in solution versus equivalents of IDA. The data is shown in Figure 3-2B. As evident, at a 1:1 ratio there exists ~65% $Cu^{2+}$-IDA complex in solution, which will lead to inefficient $Cu^{2+}$-loading to the dHis sites.

Although higher equivalents of IDA can reduce free $Cu^{2+}$, a potential complication is the formation of the $Cu^{2+}$-$(IDA)_2$ bis-complex, which cannot coordinate to dHis sites. In fact, the stability of the bis-complex is higher than that of the mono-complex at physiological pH, when no other coordinating group is present[148]. Figure 3-3 shows the CW spectra of $Cu^{2+}$:IDA at two different ratios: 1.0:0.5 (solid) and 1.0:2.0 (dotted). At the 1.0:0.5 ratio, there should be no bis-complex formed. As evident in Figure 3-3, with the addition of the excess IDA, the $g_{||}$ value decreases while the $A_{||}$ increases (shown by shift from dashed grey lines to solid lines). These shifts are consistent with an increase in nitrogen coordination[15]. Therefore, the data suggests that at the 1:2 ratio suggested by Figure 3-2, there is predominantly $Cu^{2+}$-$IDA_2$
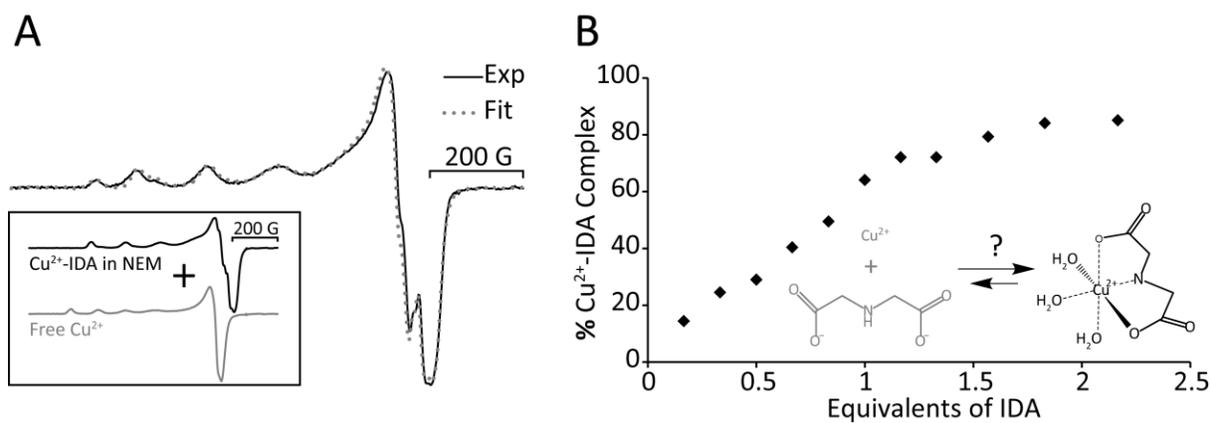
Figure 3-2. A) CW ESR spectrum of $Cu^{2+}$ with one equivalent of IDA in water (solid) which clearly shows two components. Spectrum was simulated (dotted) by the addition of two individual spectra: $Cu^{2+}$:IDA in NEM (inset black) and free $Cu^{2+}$ in water (inset grey). B) The percentage of the bound $Cu^{2+}$-IDA complex versus equivalents of added IDA. At 1 equivalent of IDA, only ~65% of $Cu^{2+}$ is bound to IDA. The concentration of $Cu^{2+}$ was 500 μM.
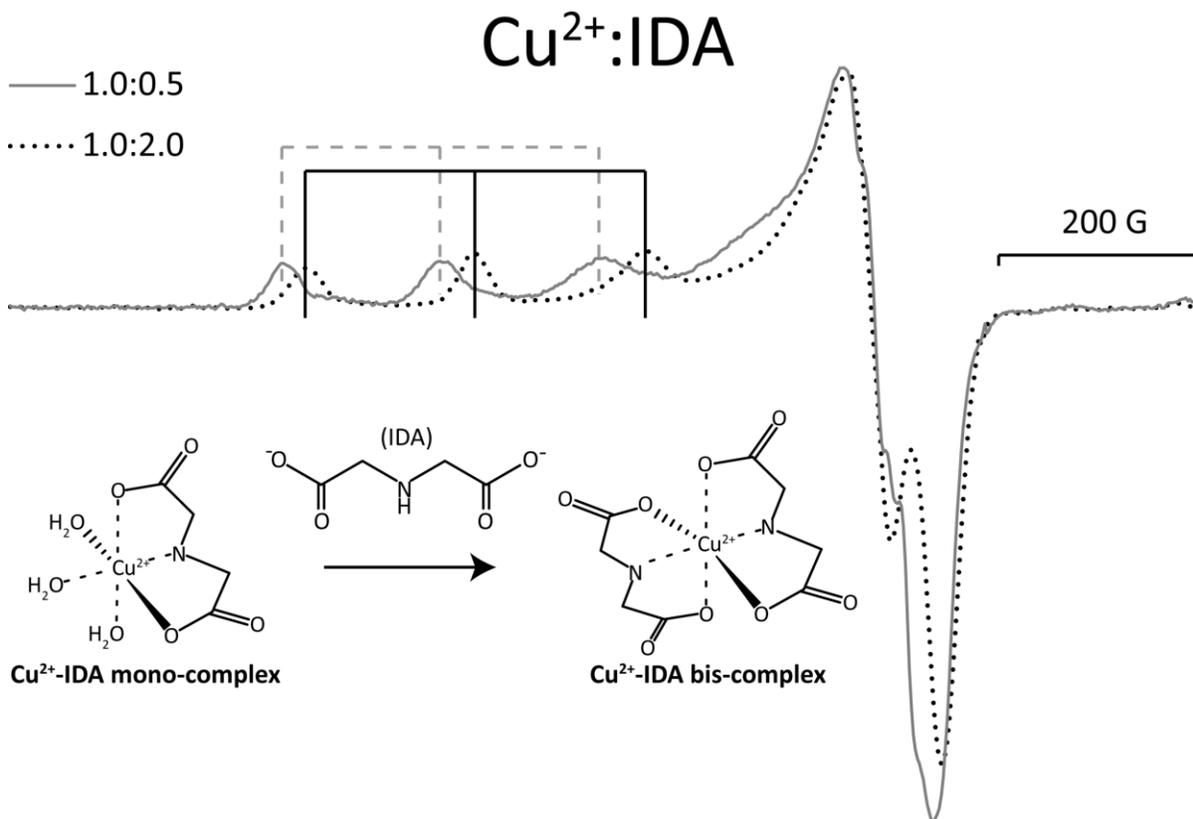
Figure 3-3. The CW ESR spectra of $Cu^{2+}$:IDA at two different ratios: 1.0:0.5 (solid) and 1.0:2.0 (dotted). The 1.0:0.5 ratio should consist of entirely of $Cu^{2+}$-IDA mono-complex. Upon addition of excess IDA, the lineshape shifts. The $g_\parallel$ decreases while $A_\parallel$ increases as shown by shift from dashed lines to solid lines. This shift is consistent with increased nitrogen coordination. This data suggests that at the 1.0:2.0 ratio of $Cu^{2+}$:IDA, the bis-complex is predominant in solution. Mono-complex to bis-complex formation shown in inset.

bis-complex in solution. In the bis-complex, all four equatorial coordinating ligands are occupied[148] which would inhibit dHis coordination. The mono-complex to bis-complex formation is shown in the inset in Figure 3-3.

If another coordinating ligand, such as an imidazole, is available, the imidazole-$Cu^{2+}$-IDA complex is heavily favored[148,149]. It has been documented that in the presence of a coordinating ligand, the $Cu^{2+}$-$IDA_2$ bis-complex is inhibited[149]. Therefore, we use CW ESR in order to detect if the bis-complex formation is inhibited by the presence of the coordinating ligand of the dHis motif: imidazole. Figure 3-4 shows the CW spectra of $Cu^{2+}$ in the presence of both IDA and imidazole at two different ratios of $Cu^{2+}$:IDA:Imidazole, 1:1:1 (solid) and 1:2:1 (dotted). As evident from Figure 3-4, the addition of excess IDA does not alter the lineshape. This data suggests that at the $Cu^{2+}$:IDA ratio of 1:2 suggested by Figure 3-2, the bis-complex formation will be inhibited in the presence of imidazole. The imidazole-Cu2+-IDA inhibition of the bis-complex is shown in the inset of Figure 3-4. Further supporting these results are ESEEM spectra shown in Figure 3-5. Figure 3-5 shows the Fourier transformed ESEEM spectra of both $Cu^{2+}$:IDA and $Cu^{2+}$:IDA:imidazole (1:2 and 1:2:1 respectively). While both samples show a similar shift in CW making analysis difficult, the ESEEM spectrum of $Cu^{2+}$:IDA:imidazole (grey line) clearly shows the presence of imidazole coordination as compared to the $Cu^{2+}$:IDA ESEEM spectrum (black dashed) which is largely featureless. In the presence of imidazole, the CW lineshape does not change upon the addition of excess IDA.

The pKd of 10.63 was measured under the experimental condition of 293K[148] while our experimental condition was 80K. Also, using the enthalpy value for $Cu^{2+}$-IDA complex formation[150], we can calculate the expected $pK_d$ for $Cu^{2+}$-IDA at 80K. Using the van't Hoff
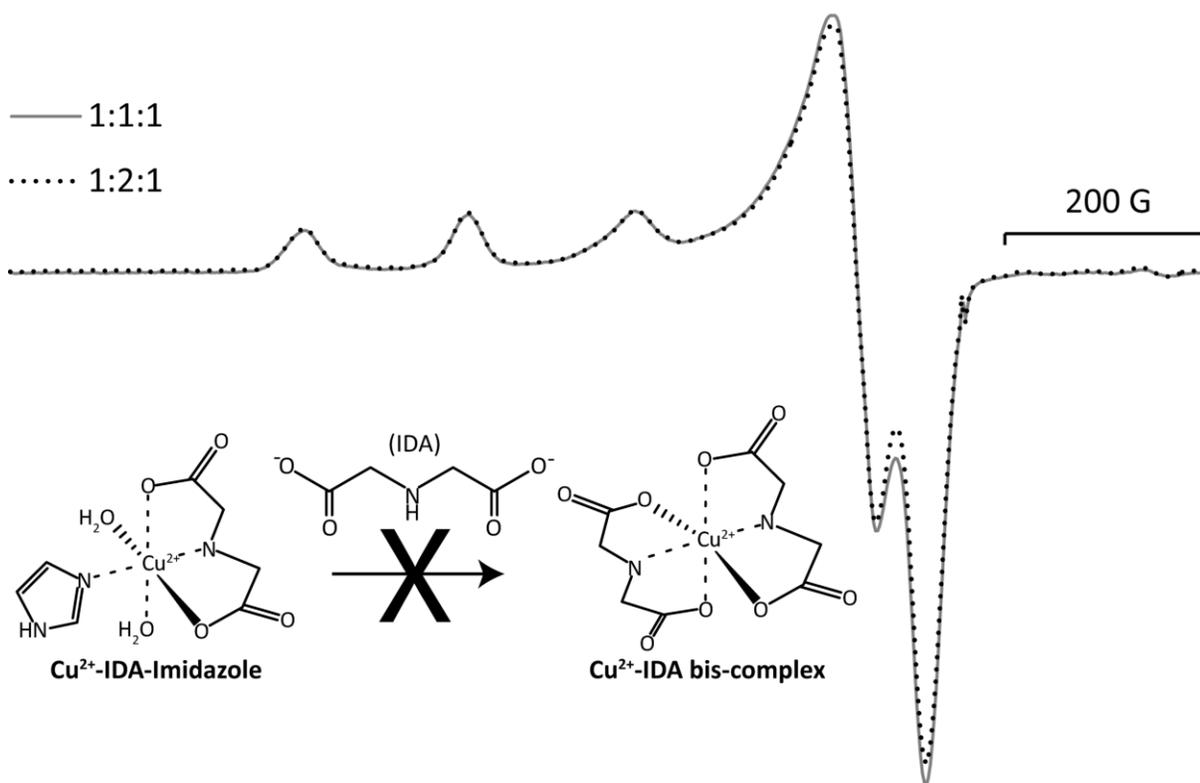
Figure 3-4. The CW ESR spectra of $Cu^{2+}$ in the presence of both IDA and imidazole at two different ratios of $Cu^{2+}$:IDA:Imidazole, 1:1:1 (solid) 1:2:1 (dotted). The lineshape does not change upon the addition of excess IDA. This suggests that the imidazole coordinating ligand inhibits the formation of the $Cu^{2+}$-$IDA_2$ bis-complex. This inhibition is shown in the inset.
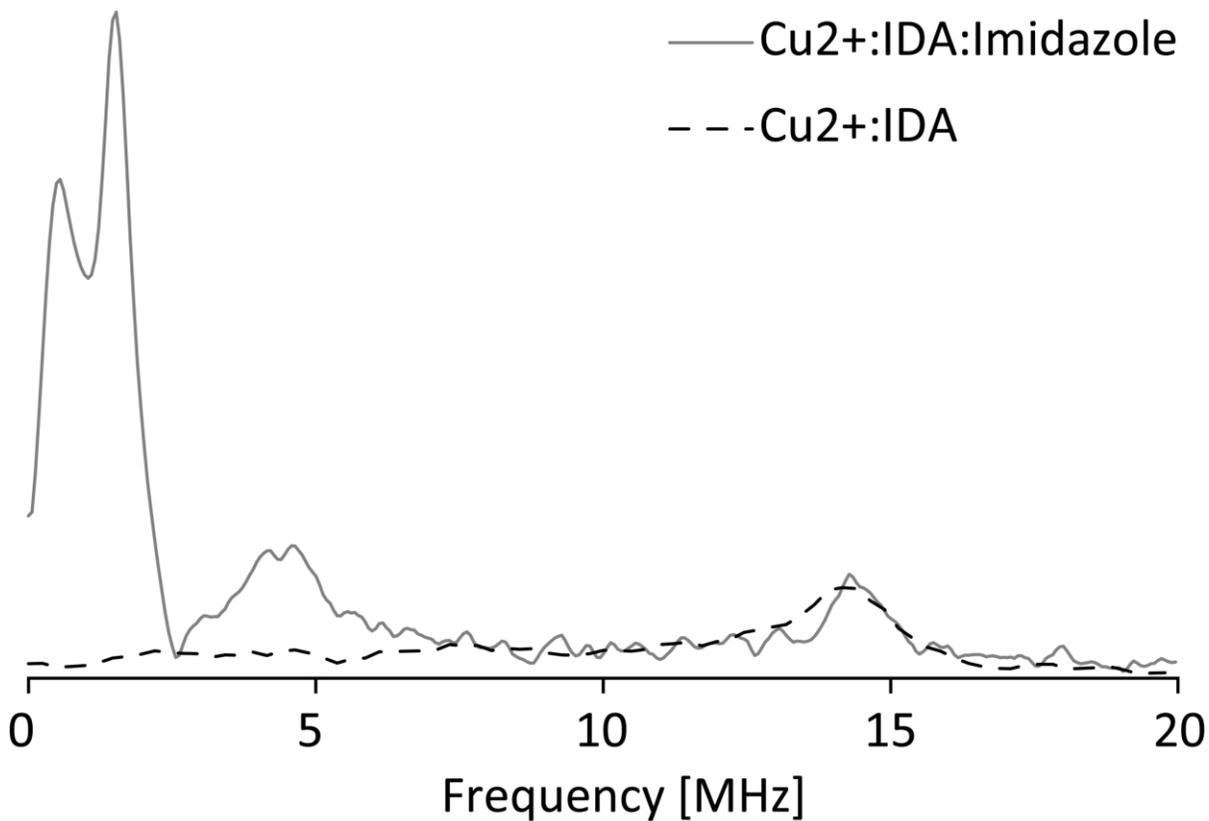
Figure 3-5. Fourier transformed ESEEM spectra of $Cu^{2+}$:IDA (black dashed) and $Cu^{2+}$:IDA:imidazole (grey line) solutions (1:2 and 1:2:1 respectively) in NEM buffer at pH 7.4. Even in the presence of two fold IDA, $Cu^{2+}$ complexes with the imidazole nitrogen as evident from the features below 2 MHz and at 4 MHz. This further supports that introduction of a coordinating ligand inhibits biscomplex formation.

equation and the above reported values, we get a $pK_d$ of 4.3 for $Cu^{2+}$-IDA at 80K, which is close to the experimental obtained value of 5.1, as shown in Figure 3-6. This model assumes the simplest case of 1:1 binding. Keeping in mind that theliterature values were reported at certain ionic strength, the expected and experimental pKd values differ slightly as our experimental conditions had zero ionic strength. Furthermore, the binding of $Cu^{2+}$-IDA to imidazole on proteins has been shown to be optimized at pH 7.5[142].

Therefore, the data suggests that in the presence of the dHis motif (two coordinating imidazoles), the bis-complex formation is likely inhibited.


### 3.4.2   CW ESR spectroscopy reveals amount of $Cu^{2+}$-IDA binding to protein


We next analysed the population of bound $Cu^{2+}$-IDA present for both α-helical and β-sheet dHis sites, under conditions used in prior DEER work[141]. In these experiments, an equimolar $Cu^{2+}$ to IDA solution was added to the protein. Figure 3-7A shows the CW ESR spectrum of $Cu^{2+}$-IDA added to the α-helix dHis site in 28H/32H-GB1 at two different $Cu^{2+}$-IDA:dHis ratios of 0.75:1 and 2.00:1 (dashed and solid respectively). The data were collected in NEM buffer where free $Cu^{2+}$ is ESR silent[21]. Upon addition of excess $Cu^{2+}$-IDA, a second component arises in the $g_{\parallel}$ region as shown in the inset of Figure 3-7A.

The data were simulated by using two components and a representative fit is shown in Figure 3-7B. The first component (cf. dashed line in the inset of Figure 3-7A) is consistent with $Cu^{2+}$-IDA coordinated to the dHis site[141]. The second component (cf. solid
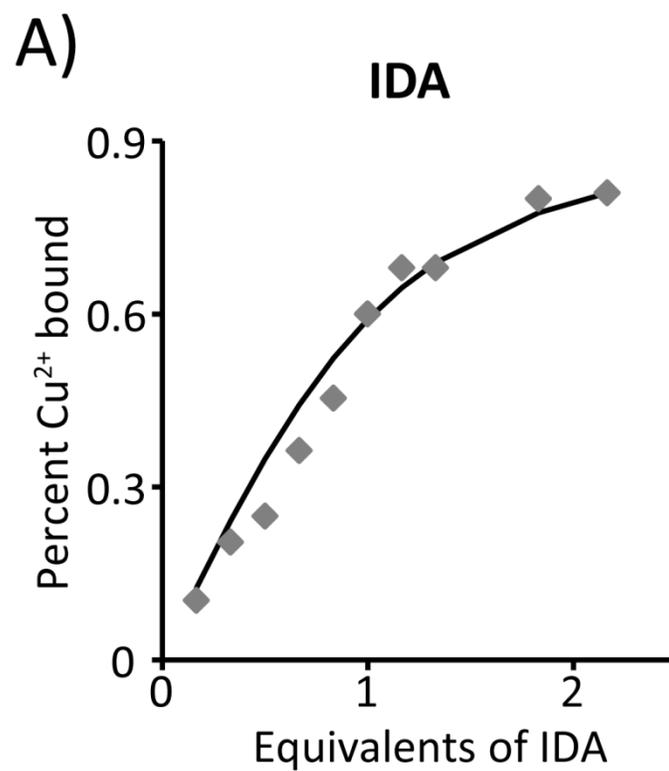
Figure 3-6. A) Apparent dissociation constant (Kd) fitting of $Cu^{2+}$:IDA titration resulting in

$K_d = 70$ μM. The fit assumes the simplest case of 1:1 binding.
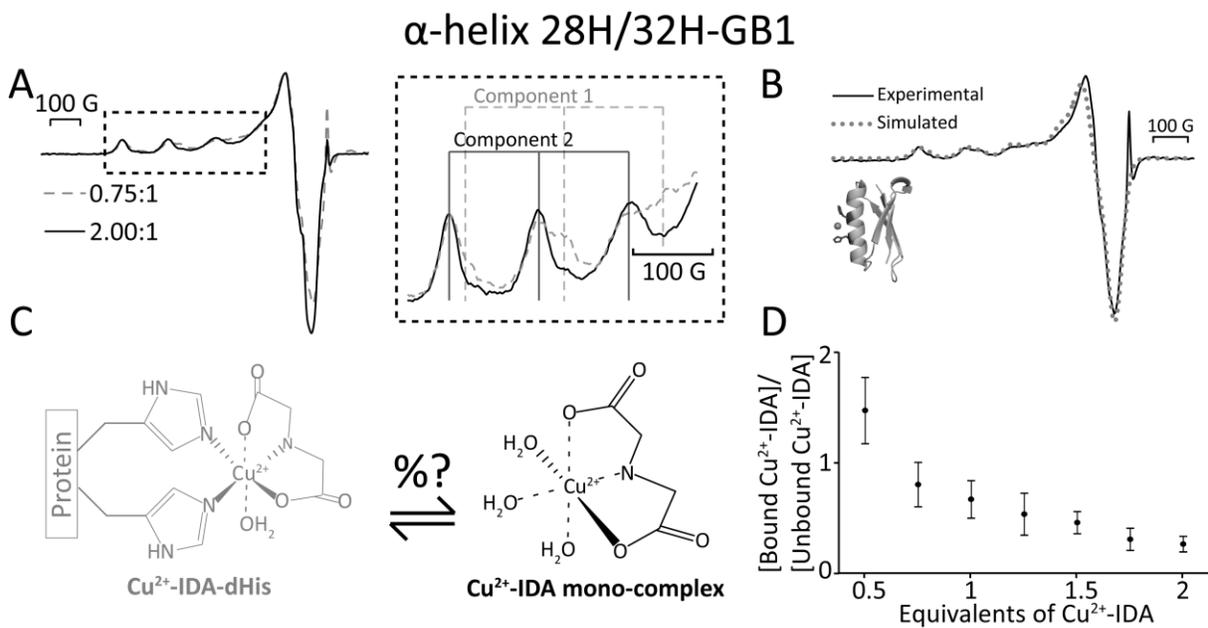
Figure 3-7. A) $Cu^{2+}$ CW ESR spectrum of 28H/32H-GB1 at two $Cu^{2+}$-IDA:dHis ratios of 0.75:1 (dashed) and 2.00:1 (solid), respectively. The CW spectrum showed two components apparent in the $g_\parallel$ region, indicated by the dashed box. The $g_\parallel$ region is magnified in the inset to better illustrate components 1 and 2 (dashed and solid, respectively). Component 2 is consistent with unbound $Cu^{2+}$-IDA complex in solution. B) The CW spectrum (solid) of 28H/32H-GB1 with 0.75 equivalents of $Cu^{2+}$-IDA with simulated two-component fit (dotted). C) The two species present in solution. D) Plot of the number of dHis-bound per unbound $Cu^{2+}$-IDA at different equivalents of $Cu^{2+}$-IDA. The error is determined from simulations. Upon comparison to a calibration curve, this data suggests that at a 1 equivalent of $Cu^{2+}$-IDA only 30% of dHis sites are loaded. The protein concentration was 500 μM.

line in the inset of Figure 3-7A) has $g_\parallel$ and $A_\parallel$ values consistent with unbound $Cu^{2+}$-IDA complex in solution. The CW spectrum of unbound $Cu^{2+}$-IDA complex in solution overlays precisely on the second component when the spectra are superimposed shown in Figure 3-8. The spectra were collected in NEM buffer thereby silencing any free $Cu^{2+}$ in solution. As noted, above, a $Cu^{2+}$-$(IDA)_2$ bis-complex is not anticipated in the presence of the dHis motif. Figure 3-7C shows the two different species (i.e. dHis-bound $Cu^{2+}$-IDA and unbound $Cu^{2+}$-IDA) that are present in solution. Accordingly, a titration was performed to monitor the amount of bound dHis-$Cu^{2+}$-IDA at varying equivalents of $Cu^{2+}$-IDA to dHis. The number of dHis-bound per unbound $Cu^{2+}$-IDA complex in solution, determined using simulations, is shown in Figure 3-7D. Since the doubly integrated CW ESR spectrum is directly proportional to the number of spins, we can also extract the concentration of the dHis bound $Cu^{2+}$-IDA complex by comparison to a calibration curve. Furthermore, comparison to the known protein concentration provides the percentage of dHis sites occupied by $Cu^{2+}$-IDA. This data is shown in Figure 3-9. The data indicates that at an equimolar ratio of $Cu^{2+}$-IDA solution to protein only 30% of the α-helical site is loaded (c.f. Figure 3-9). This is likely due to two reasons. First, in a 1:1 $Cu^{2+}$:IDA solution only about 65% of the metal ion exists as a $Cu^{2+}$-IDA complex (cf. Figure 2). Second, the $Cu^{2+}$-dHis binding affinity has been shown to be relatively low with apparent dissociation constants reported from 200 to 2 μM for various α-helical sites[137]. This data suggests that the dHis site within an α-helix has poor affinity for the $Cu^{2+}$-IDA complex. The apparent dissociation constant for the 28H/32H-GB1 dHis site is 1210 μM shown in Figure 3-10.
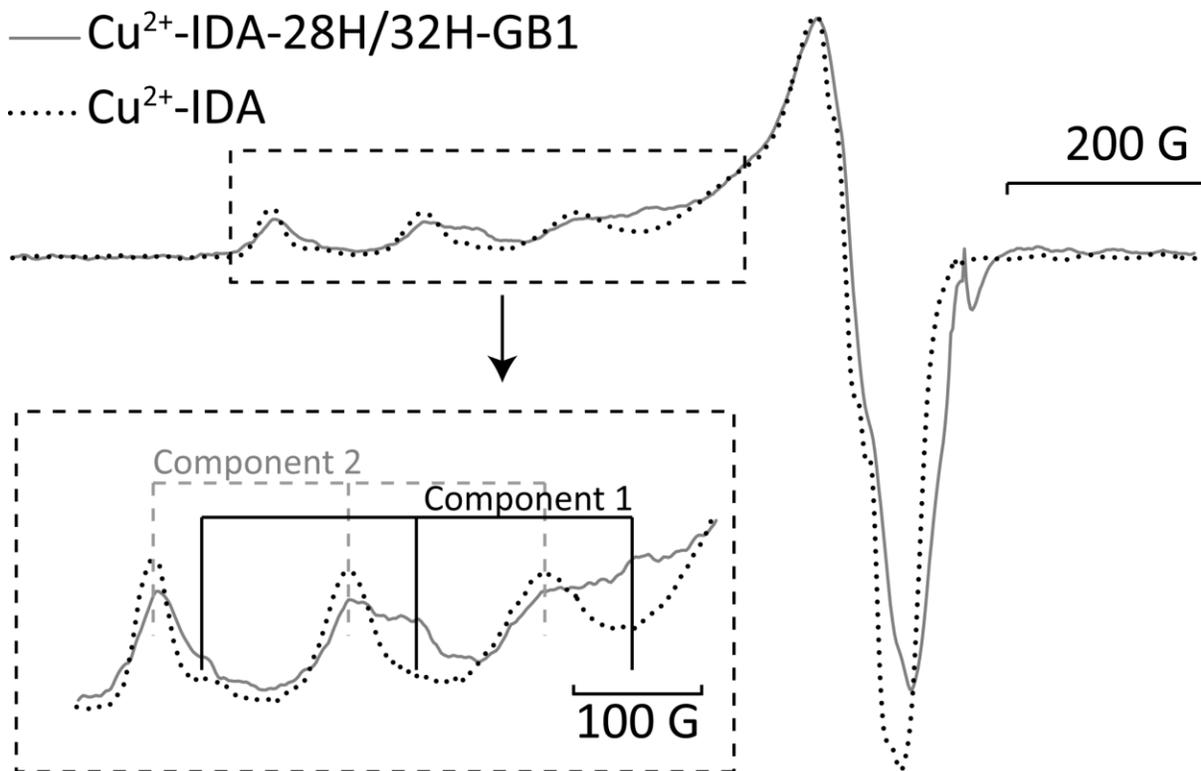
Figure 3-8. The CW ESR spectrum of $Cu^{2+}$-IDA in the presence of 28H/32H-GB1 (solid) overlaid with the CW ESR spectrum of unbound $Cu^{2+}$-IDA in solution (dotted). The spectrum $Cu^{2+}$-IDA in the presence of 28H/32H-GB1 clearly has two components, as evident in the $g_{\parallel}$ region indicated by dashed box. The inset is a magnified view of the $g_{\parallel}$ region showing component 1 (solid lines) and component 2 (grey dashed lines). Component 2 overlays exactly with the spectrum of unbound $Cu^{2+}$-IDA in solution. Therefore, the data suggests that the second component is due to free $Cu^{2+}$-IDA in solution.
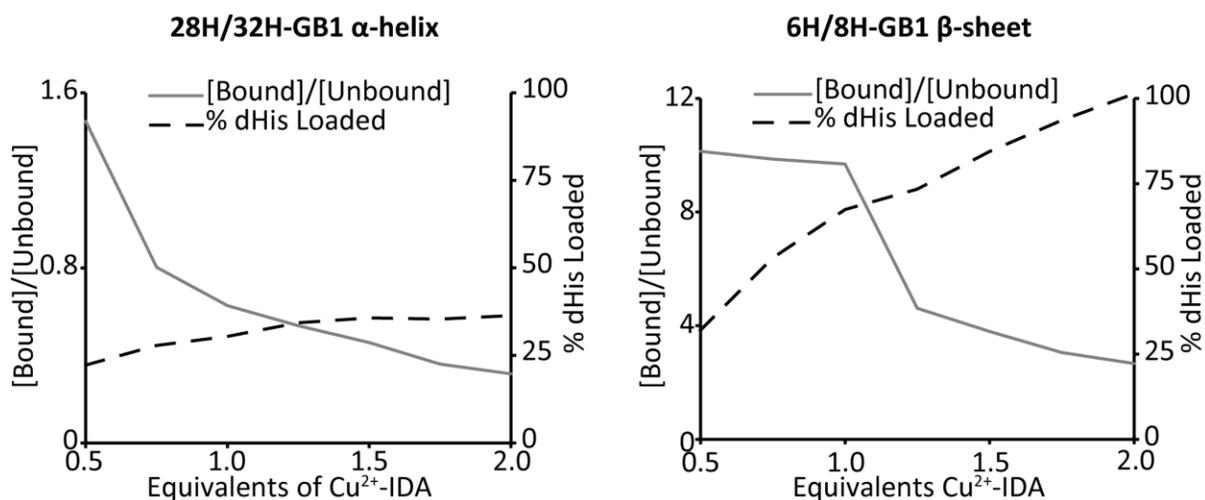
Figure 3-9. The amount of bound $Cu^{2+}$-IDA per unbound $Cu^{2+}$-IDA free in solution with respect to the equivalents of $Cu^{2+}$-IDA added (solid gray line) for both α-helix (left) and β-sheet (right). It is evident that the β-sheet has much higher selectivity than the α-helical dHis binding site. Also shown is the percent of dHis loaded (dashed black line). The β-sheet loads far quicker than the α-helix and is capable of full loading.
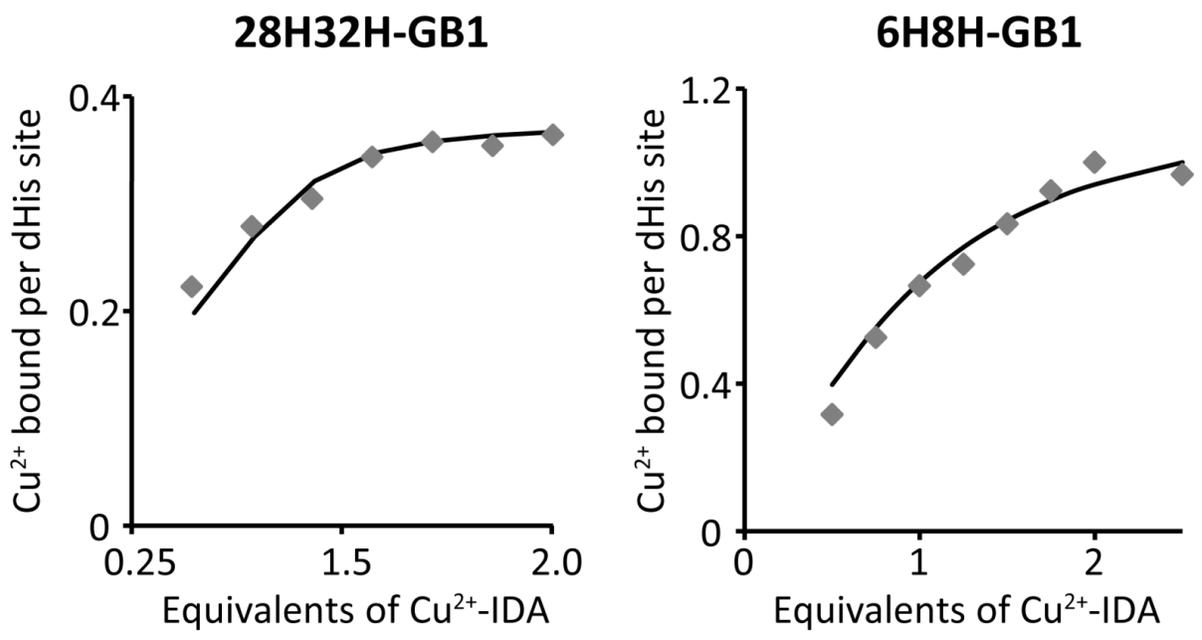
Figure 3-10. Apparent dissociation constant fits of titration curves of 28H32H-GB1 (left) and 6H8H-GB1 (right) with respect to equivalents of $Cu^{2+}$-IDA added.

Figure 3-11A shows the CW ESR spectrum of $Cu^{2+}$-IDA added to the β-sheet dHis site in 6H/8H-GB1 at two $Cu^{2+}$-IDA:dHis ratios of 0.75:1 and 2.00:1 (dashed and solid respectively). A second component arises in the $g_\perp$ with increased addition of $Cu^{2+}$-IDA as shown in Figure 3-11A (dashed box). To resolve this shoulder feature more clearly, we repeated the titration at only the $g_\perp$ region using a smaller modulation amplitude shown in Figure 3-11A (inset). Changes in the $g_\parallel$ region are not noticeable in this case due to such a low amount of unbound $Cu^{2+}$-IDA present in solution.

CW simulations (Figure 3-11B, dotted) using two components best fit this data. As with the α-helical dHis site, the first component is consistent with $Cu^{2+}$-IDA bound to dHis[141], albeit a narrower linewidth was observed. As more $Cu^{2+}$-IDA is added, a larger percent of the overall signal is due to the second component. The second component is consistent with unbound $Cu^{2+}$-IDA in solution. Figure 3-11C shows the difference in coordination between dHis-bound $Cu^{2+}$-IDA and unbound $Cu^{2+}$-IDA in solution. Figure 3-11D shows the amount of bound $Cu^{2+}$-IDA per unbound $Cu^{2+}$-IDA in solution. As evident in Figure 3-11D, after 1 equivalents of $Cu^{2+}$-IDA added, there is a drastic rise in the amount of unbound $Cu^{2+}$-IDA complex in solution. Upon comparison to a calibration curve, the dHis site within a β-sheet is ∼ 67% occupied when 1 equivalent of $Cu^{2+}$-IDA added (c.f. SI Figure 3-9). This result is remarkable because under these conditions only about ∼65% of $Cu^{2+}$-IDA complex exists in solution (cf. Figure 3-2). Therefore, up to 1 equivalent of $Cu^{2+}$-IDA added, the dHis site in the β-sheet loads stoichiometrically. This data suggests that the i, i+2 β-sheet dHis motif has an increased selectivity toward the $Cu^{2+}$-IDA complex with an apparent dissociation constant of 30 μM.
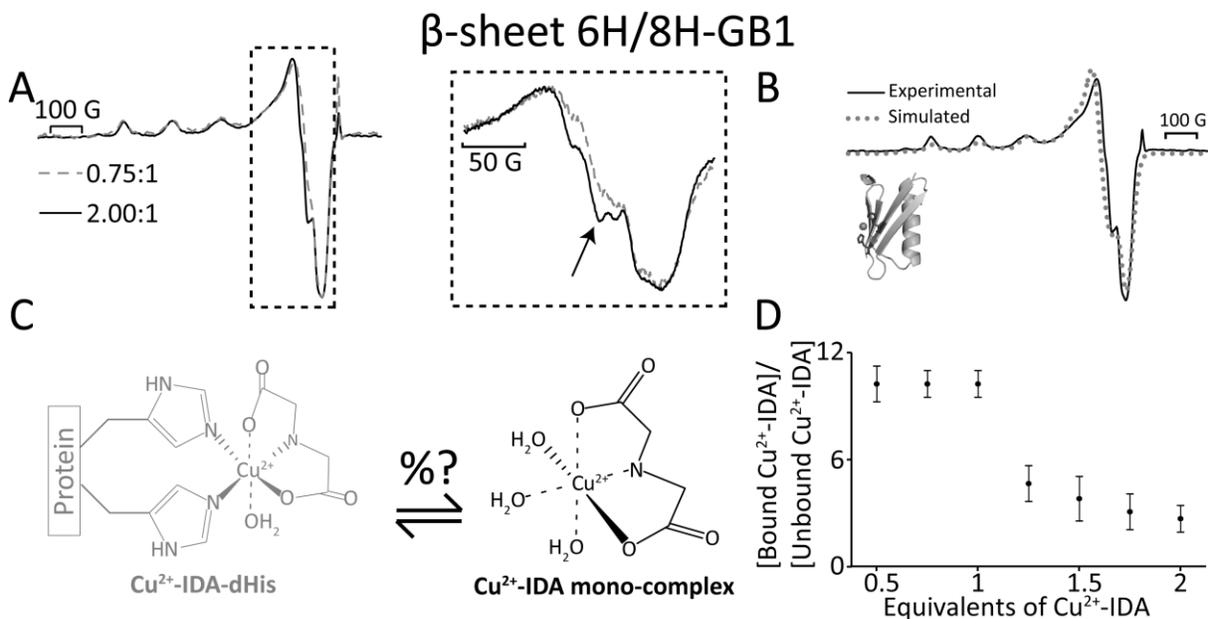
Figure 3-11. A) $Cu^{2+}$ CW-ESR spectrum of 6H/8H-GBI at two different $Cu^{2+}$-IDA:dHis ratios, 0.75:1 (dashed) and 2.00:1 (solid). The β-sheet CW spectrum showed two components apparent by a shoulder in the $g_\perp$ region, indicated by dashed box. The $g_\perp$ region was recollected with a smaller modulation amplitude for better resolution. B) CW ESR spectrum (solid) with simulated fit (dotted) at 1.5 equivalents of $Cu^{2+}$-IDA per dHis. C) The two species present in solution. D) Plot of the number of dHis-bound per unbound $Cu^{2+}$-IDA as determined by two-component fits at different equivalents of $Cu^{2+}$-IDA. The error is determined from the simulations. This data suggests that an equimolar ratio of $Cu^{2+}$-IDA per dHis minimizes unbound $Cu^{2+}$-IDA while maintaining enough $Cu^{2+}$-IDA for proper dHis loading. The protein concentration was 500 μM.

### 3.4.3 ESEEM spectroscopy provides insight on $Cu^{2+}$-IDA binding to dHis

To support the CW results, we use ESEEM spectroscopy to probe dHis-$Cu^{2+}$-IDA coordination. The coordination of $Cu^{2+}$ to the imidazole of the His yields a characteristic ESEEM signal due to the interaction of the unpaired electron spin with distal nitrogen of imidazole[25,32,38,151]. Figure 3-12 shows an example of the ESEEM signal for $Cu^{2+}$-IDA bound to the dHis site. The ESEEM data were collected for both α-helix and β-sheet dHis sites at varying equivalents of $Cu^{2+}$-IDA (c.f. Figure 3-13). From these data we quantified the relative modulation depth parameter, k,[152] using values from the second modulation period[46] as shown in Figure 3-12. The relative modulation depth parameter, k, increases with distal coordination of $^{14}N$ nuclei[25,151]. As dHis coordination increases, k should increase due to increased imidazole coordination. Conversely, as the fraction of unbound $Cu^{2+}$-IDA complex increases the k value should decrease.

Figure 3-12B shows the value of k as a function of $Cu^{2+}$-IDA equivalents, for the α-helical site. The value of k decreases as the equivalents of $Cu^{2+}$-IDA per dHis site is increased. This observation is in general agreement with CW results of Figure 3D, which show that the amount of unbound $Cu^{2+}$-IDA increases as the equivalents of $Cu^{2+}$-IDA are added. As more unbound $Cu^{2+}$-IDA is in solution, the modulations due to imidazole coordination become less apparent. This data confirms the CW results that the α-helical dHis site shows poor affinity for the $Cu^{2+}$-IDA complex.

Figure 3-12C shows the value of k as a function of $Cu^{2+}$-IDA equivalents, for the β-sheet site. The values of k are roughly similar until an equimolar ratio of $Cu^{2+}$-IDA
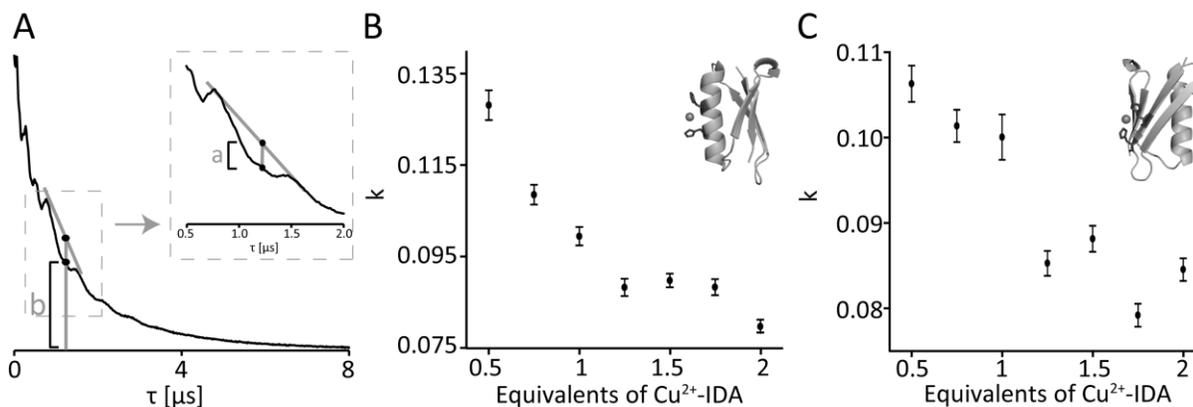
Figure 3-12. A) Raw ESEEM time domain data of dHis site loaded with $Cu^{2+}$-IDA within an α-helix. The modulation depth parameter ($k = a/(a+b)$) can be calculated by analyzing the second modulation period as shown in the Figure and the inset..B) Plot of k at different equivalents of $Cu^{2+}$-IDA per dHis site for an α-helix. C) Plot of k at different equivalents of $Cu^{2+}$-IDA per dHis site within a β-sheet. The protein concentration was 500 μM for both data-sets. The value of k decreases as the amount of unbound $Cu^{2+}$-IDA complex increases, since this species produces a modulation free ESEEM signal effectively washing out modulations due to His coordination.
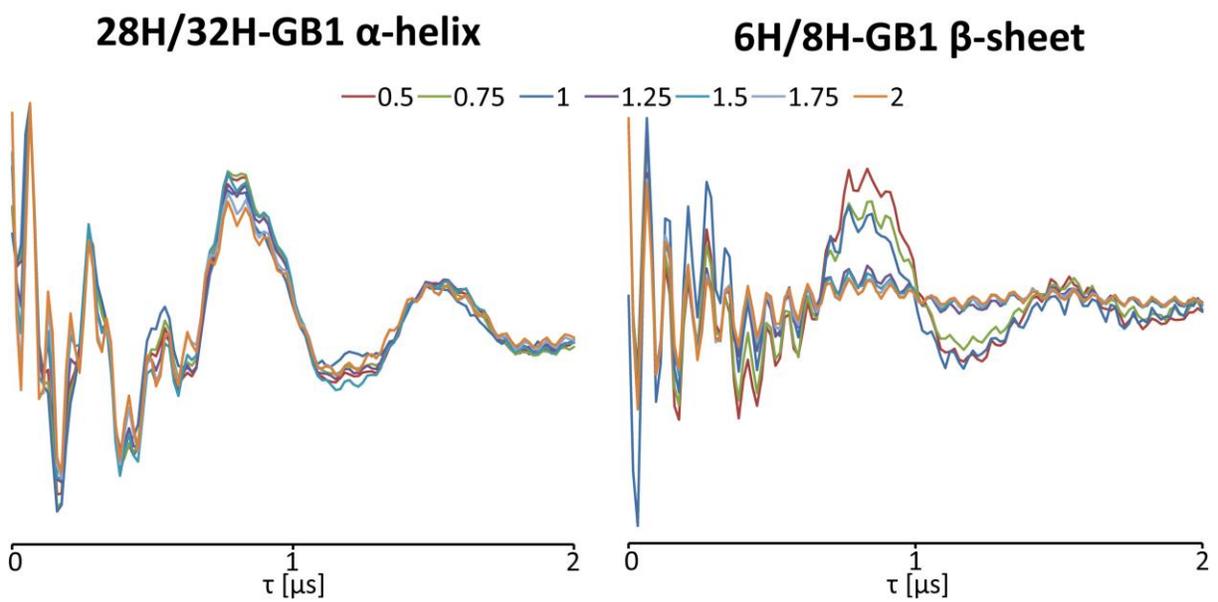
Figure 3-13. ESEEM signal on dHis site located within an α-helix (left) and β-sheet (right) at seven different equivalents of $Cu^{2+}$-IDA. Data was normalized to the point of maximum intensity.

solution to protein ratio is reached. At higher ratios of the $Cu^{2+}$-IDA the modulation depth is smaller. Again these observation are consistent with CW-ESR results of Figure 4D, which indicate that the fraction of the free $Cu^{2+}$-IDA complex increases at greater than 1 equivalent of $Cu^{2+}$-IDA. This data supports the conclusion that this β-sheet dHis site is highly selective toward the $Cu^{2+}$-IDA complex.

### 3.4.4 Analysis of the modulation depth of DEER

A tetramutant 6H/8H/28H/32H-GB1, which combines the α-helix and a β-sheet dHis sites, was used for DEER distance measurements. We have earlier confirmed through CD, DEER, and crystallography that this tetramutant retains native structure[141]. Figure 3-14A shows the baseline corrected DEER data obtained on an equimolar solution containing 0.5 mM $Cu^{2+}$, 0.5 mM IDA and 0.25 mM 6H/8H/28H/32H-GB1 (note: [$Cu^{2+}$] is twice [GB1] since the tetramutant contains two dHis sites). The raw data is shown in the Supporting Information (Figure 3-15). Data was collected on two different samples to ensure reproducibility as shown in Figure 3-16 (grey solid). The baseline subtracted DEER data, which produced a signal to noise ratio per scan of 0.8, indicates that the modulation depth is ~2% (cf. Figure 3-14, dashed line).

To analyze this modulation depth, λ, we must account for five separate contributions (c.f. Equations 3-1 – 3-5[54,56,57,145] in Experimental section):

$$\lambda = 1 - [f_{dHis2}(1-pb)^1 + f_{dHis} + f_{IDA} + f_{bis} + f_{free}] \qquad (3\text{-}5)$$

$$f_{dHis2} + f_{dHis} + f_{IDA} + f_{bis} + f_{free} \qquad (3\text{-}6)$$
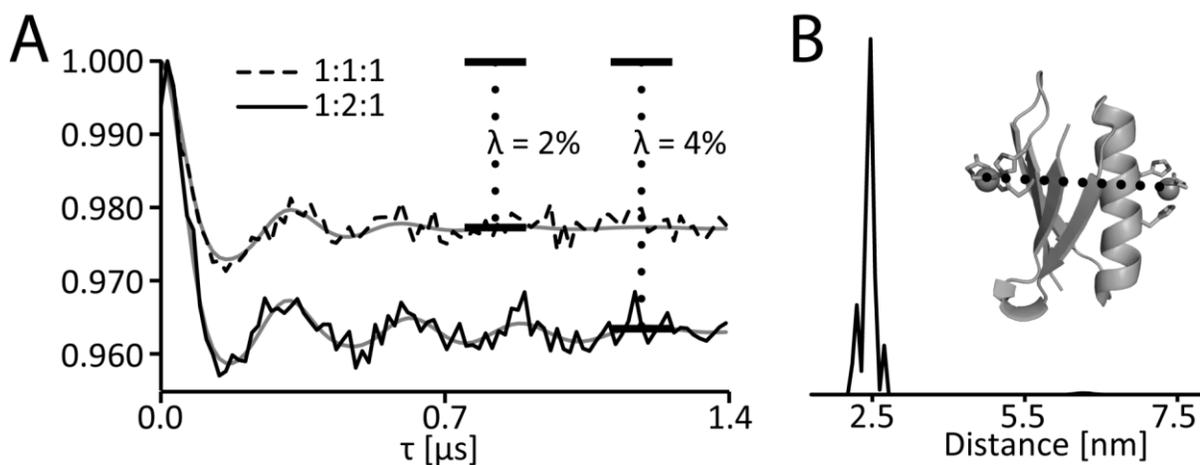
Figure 3-14. A) The background subtracted DEER signal for 6H/8H/28H/32H-GB1. The DEER data was obtained for $Cu^{2+}$:IDA:dHis ratios of 1:1:1 (dashed, number of scans = 900) and 1:2:1 (solid, number of scans =500), respectively. The protein concentration was 350 μM. B) The distance distribution analyzed via Tikhonov regularization showing the narrow distance distribution. The inset shows the tetramutant fully loaded with $Cu^{2+}$-IDA.
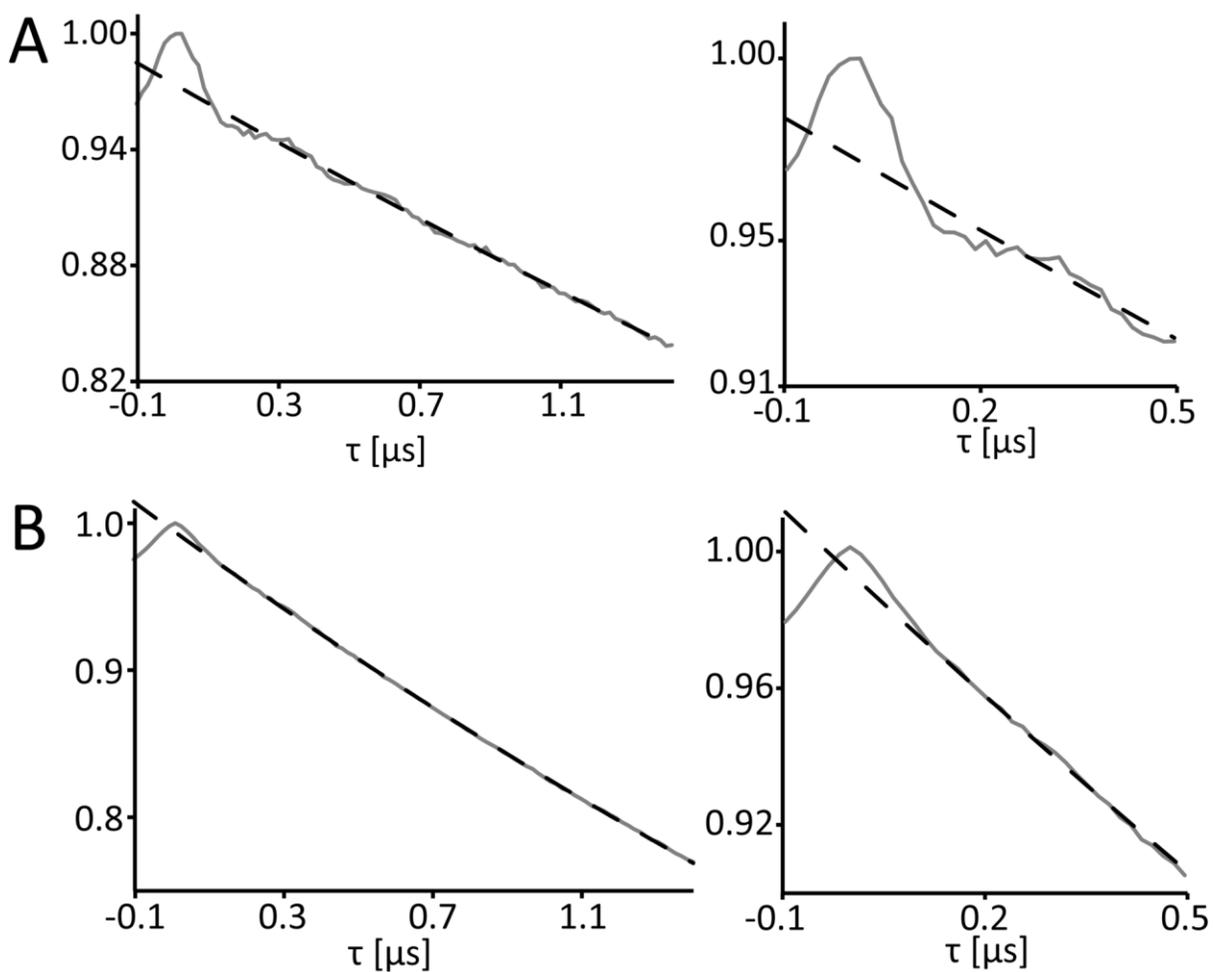
Figure 3-15. Left) Raw time domain DEER data for 6H/8H/28H/32H-GB1 at two different Cu$^{2+}$:IDA:dHis ratios: A) 1:2:1, B) 1:1:1. Right) Magnified raw time domain DEER signal to clearly see dipolar modulations for both samples.
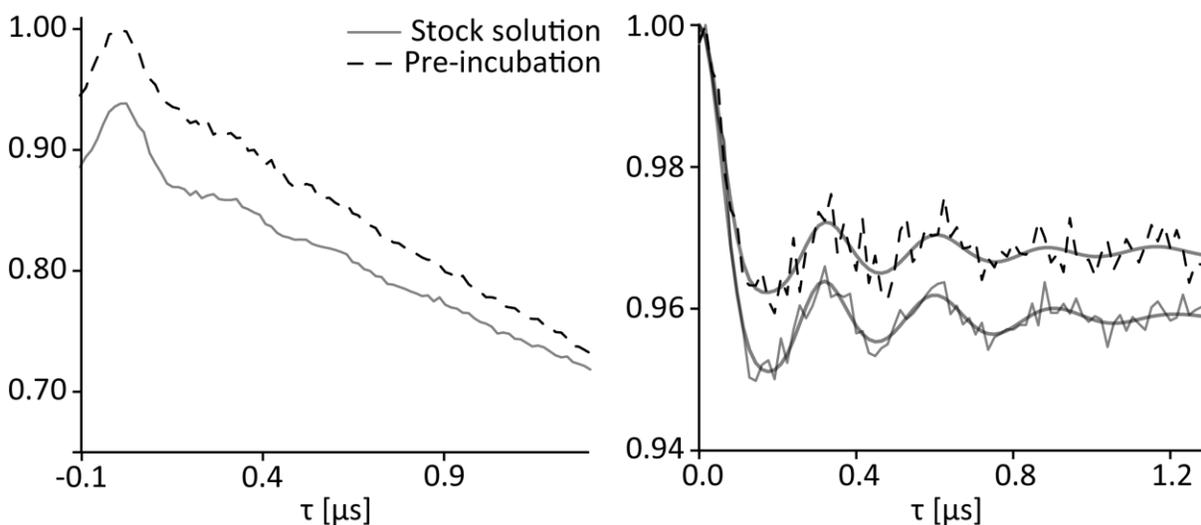
Figure 3-16. Left) Raw DEER time domain data for 6H/8H/28H/32H-GB1 prepared in two different methods: addition of $Cu^{2+}$-IDA (1:2 ratio) stock solution (dashed black line) and prepared via pre-incubation of protein with IDA followed by addition of $Cu^{2+}$ (grey solid). Data was offset for clarity. Right) Background subtracted DEER time domain data illustrating the improvement of modulation depth by pre-incubation of protein with IDA followed by the addition of $Cu^{2+}$ (~3.2% to ~4.0%).

where $f_{dHis2}$ is the fraction of $Cu^{2+}$-IDA bound to <u>both</u> dHis sites in the 6H/8H/28H/32H-GB1 tetramutant, $f_{dHis}$ is the fraction of $Cu^{2+}$-IDA bound to only a *single* dHis site in the tetramutant, $f_{IDA}$ is the fraction of $Cu^{2+}$-IDA mono-complex free in solution, $f_{bis}$ is the fraction of $Cu^{2+}$-IDA$_2$ bis-complex, and $f_{free}$ is the fraction of free $Cu^{2+}$. All contributions are fractions of the total signal and thus must add to one. The value for $p_b$ is 0.11 based upon the echo detected field swept spectrum and calculated as demonstrated by Yang et. al[1].

Since the solution buffer is NEM at pH 7.4, $f_{free}$ is near zero. Second, from Figure 3-4, we have shown that bis-complex formation is greatly inhibited presence of imidazole. Indeed the CW-ESR spectrum of the tetramutant does not show any features from the $Cu^{2+}$-IDA bis-complex (cf. Figure 3-17). Therefore, $f_{bis}$ can also be taken as zero. Finally, $f_{IDA}$ is the fraction of unbound $Cu^{2+}$-IDA complex in solution. This fraction can be estimated from the percentage of the total CW ESR spectrum due to the unbound $Cu^{2+}$-IDA complex in the data shown in Figures 3-7D and 3-11D. Furthermore:

$$f_{dHis2} = [P_\alpha P_\beta / (P_\alpha P_\beta + (P_\alpha P_{N\beta} + P_{N\alpha} P_\beta))] * P_{sig} \qquad (3\text{-}7)$$

$$f_{IDA} = 1 - (f_{dHis2} + f_{dHis}) \qquad (3\text{-}8)$$

$$P_{sig} = 1 - f_{IDA} \qquad (3\text{-}9)$$

where $P_\alpha$ and $P_\beta$ are the probabilities of $Cu^{2+}$-IDA binding a α-helical and β-sheet dHis site, respectively. Similarly, $P_{N\alpha}$ and $P_{N\beta}$ are the probabilities of $Cu^{2+}$-IDA not binding a α-helical and β-sheet dHis site, respectively and $P_{sig}$ is the fraction of the DEER signal contributed by all bound $Cu^{2+}$-IDA.
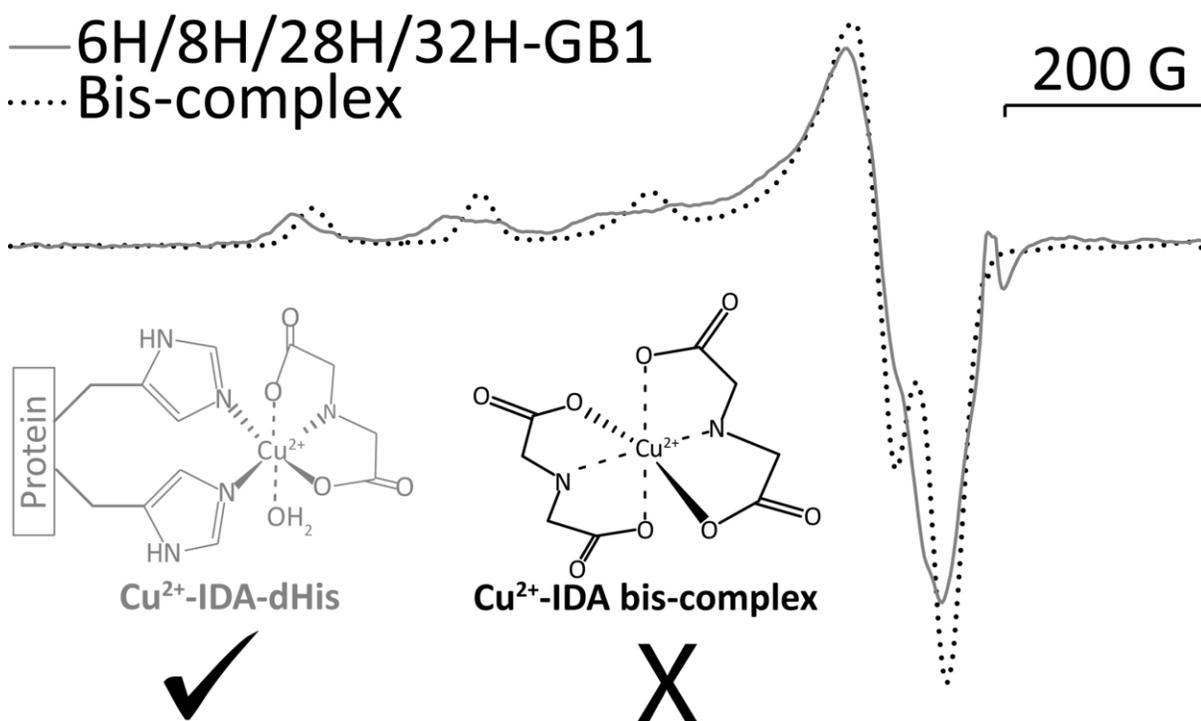
Figure 3-17. The CW ESR spectrum of 6H/8H/28H/32H-GB1 (gray solid lines) overlaid with the spectrum of the bis-complex (black dotted lines). It is evident that the bis-complex is not present in solution for the 6H/8H/28H/32H-GB1 sample. Inset shows the favored dHis coordination compared to the bis-complex.

With respect to Equations 3-7 – 3-9, $f_{dHis2}$ and $f_{dHis}$ can be determined from Figures 3-7D and 3-11D. In particular, $P_\alpha$ and $P_\beta$ can be determined by the concentration of bound $Cu^{2+}$-IDA divided by the known protein concentration. Furthermore, $P_\alpha$ and $P_{N\alpha}$ must add to 1, and likewise for $P_\beta$ and $P_{N\beta}$. $P_{sig}$ is comprised of two values: the probability that there are two $Cu^{2+}$-IDA per GB1 and the probability that there is only one $Cu^{2+}$-IDA per GB1. Therefore, we take a ratio of the two probabilities which both must sum to $P_{sig}$. From this methodology, we can experimentally determine $f_{dHis2}$, $f_{dHis}$ and $f_{IDA}$ to produce a theoretical value of $\lambda$.

From Figure 3-7D, we determine that $P_\alpha$ and $P_{N\alpha}$ are 0.3 and 0.7, and from Figure 3-11D we determine $P_\beta = 0.65$ and $P_{N\beta} = 0.35$, under these conditions. From these values, and using Equations 3-7 and 3-8, we calculate that $f_{IDA}$ is 0.09, $f_{dHis}$ is 0.75 and $f_{dHis2}$ is 0.14 (c.f. Equations 7-9). These values from Figures 3-7 and 3-11 allow us to calculate a theoretical $\lambda$ of 1.8% ± 0.3% (error is propagated from CW simulation fits from Figures 3-7 and 3-11), which is in reasonable agreement with the observed value.

Together these results indicate that the low modulation depth that was observed in DEER is likely due to two reasons. First incomplete complex formation leads to only 65% $Cu^{2+}$-IDA complex (cf. Figure 3-2) available for binding to dHis sites, which in turn reduces $Cu^{2+}$ loading to the protein. Second, the α-helical site exhibits a much smaller affinity for $Cu^{2+}$. Due to a combination of these factors the fraction of protein that contains $Cu^{2+}$ at both sites is small.

### 3.4.5   Improvement in sensitivity of DEER

In order to improve this modulation depth, it is necessary to increase $f_{dHis2}$ and lower $f_{dHis}$, which can be done by increasing the amount of $Cu^{2+}$-IDA complex formed in solution. Taking our results from Figure 3-2, we prepared a solution containing 0.5 mM $Cu^{2+}$, 1 mM IDA, and 0.25 mM protein. These conditions increase the formation of the $Cu^{2+}$-IDA complex to ~80% while decreasing free $Cu^{2+}$ in solution to ~20% (cf. Figure 2).  In addition, the IDA was added to the protein prior to the addition of $Cu^{2+}$, in order to inhibit the formation of the $Cu^{2+}$-IDA bis-complex. The increase in the amount of the $Cu^{2+}$-IDA complex should raise $f_{dHis2}$. The DEER data, under these conditions, is shown in Figure 3-14A (solid line).  The modulation depth, $\lambda$, of the dHis-DEER improved from ~2% to ~4%. This greater modulation depth increases the signal to noise ratio per scan to 1.8 from 0.8, effectively doubling the  DEER sensitivity (note that the number of scans is ~500 and ~900 for 1:2:1 and 1:1:1 samples, respectively). DEER experiments were also performed by adding a 1:2 $Cu^{2+}$:IDA stock solution to the protein.  As shown in Figure S10 this approach led to a lower modulation depth.

We support these results with CW spectra of the 6H/8H-GB1 and 28H/32H-GB1 obtained in a 1:2 $Cu^{2+}$:IDA (c.f. SI Figure 3-18). As noted above, we can determine $P_{\alpha}$, $P_{\beta}$, $P_{N\alpha}$ and $P_{N\beta}$ from the CW results in Figure 3-18. From Figure 3-18 we determine that, under these conditions, $P_{\alpha}$ and $P_{N\alpha}$ are both ~0.5. According to our previous results (c.f. Figures 3-2) ~80% of the metal ion exists as a $Cu^{2+}$-IDA complex under these conditions. Note that there is increased super hyperfine features in the $g_{\perp}$ which is indicative of a more homogeneous binding environment. Furthermore, this superhyperfine structure was simulated
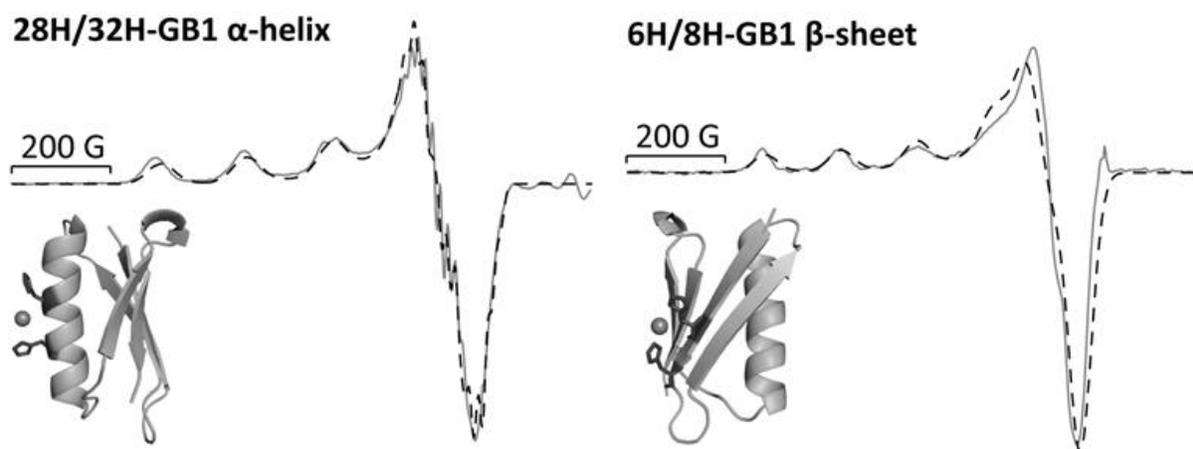
Figure 3-18. The CW ESR spectra (gray solid lines) and respective simulations (black dashed lines) 28H/32H-GB1 (left) and 6H/8H-GB1 with one equivalent of $Cu^{2+}$ and two equivalents of IDA. Both spectra were fit with two components, the second using parameters consistent with unbound $Cu^{2+}$-IDA in solution. The percent of dHis loaded increased for both α-helix and β-sheet.

using parameters of three equatorial coordinated nitrogen. This superhyperfine structure is not evident in Figure 3 most likely due to inconsistent coordination, due to the possibility of IDA-free Cu2+ occupying the dHis site. Given that that $Cu^{2+}$-IDA adds to the β-sheet stoichiometrically (cf. Figure 3-11) $P_β$ and $P_{Nβ}$ are 0.8 and 0.2, respectively. Indeed CW ESR data, collected under these conditions and shown in Figure 3-18, support these values. From these values, we calculate that $f_{IDA}$ is 0.07, $f_{dHis}$ is 0.60 and $f_{dHis2}$ is 0.33 (c.f. Equations 7-9). These values allow us to estimate a λ of 3.8% ± 0.4%, which is in reasonable agreement with our experimental value.

Taken together, the DEER results suggest that the primary limitation of the dHis motif is the poor selectivity of the α-helical dHis site. While it is possible to fully load the β-sheet site, albeit at substantial amounts of $Cu^{2+}$-IDA, we never experienced over ~50% α-helical sites loaded. Taking this into account, where $P_α$ and $P_β$ are 0.5 and 1.0 respectively, our preliminary data and calculations suggest a maximum modulation depth of ~4.6%. Indeed, using the range of reported affinities for $Cu^{2+}$ to α-helical dHis sites[137] and our results on binding of $Cu^{2+}$ to β-sheet dHis site, we anticipate that modulations depths of ~2-4% are achievable for systems with two α-helical dHis sites, up to 12% for systems with two β-sheet dHis sites and finally ~4-7% for the heterogeneous case as presented in this work. Such values of modulation depth allow for quality distance measurements, especially given that high quality DEER measurements using paramagnetic metal ions are routinely reported with modulation depths of less than 1%[115,128,140,153].

## 3.4    CONCLUSIONS

Herein, we have provided a comprehensive analysis of the distance measurements using the $Cu^{2+}$-IDA complex and the dHis motif. We have shown that $Cu^{2+}$-IDA complex formation is inefficient and at most, 80% complexation is achieved at physiological pH. In order to improve loading of the metal ion to the protein, an excess of IDA is needed in order to minimize free $Cu^{2+}$ and maximize $Cu^{2+}$-IDA complex formation. In addition, it is important to add the IDA to the protein, prior to the addition of $Cu^{2+}$ in order to inhibit the formation of the $Cu^{2+}$-$IDA_2$ bis-complex, as shown in Figure 3-17. CW lineshape analysis and ESEEM data shows that $Cu^{2+}$-IDA binding to a β-sheet dHis site is highly efficient, while α-helical dHis site exhibits poor selectivity towards $Cu^{2+}$-IDA. ESEEM data further supports our CW results. The DEER data is therefore, limited by the poor complexation and weak affinity of $Cu^{2+}$-IDA to α-helical dHis site. Nevertheless, by adequately understanding all these effects, we are able to double the DEER sensitivity.

## 3.5    ACKNOWLEDGEMENTS

# 4.0 ESR SHOWS THAT THE C-TERMINUS OF HUMAN GLUTATHIONE S-TRANSFERASE A1-1 EXISTS IN TWO DISTINCT HELICAL CONFORMATIONS IN THE LIGAND-FREE STATE

*This work, written in collaboration with John R. Pettersson, Frederick Lanni, Gordon S. Rule and Sunil Saxena has been submitted and is under revision in Biophysical Journal, 2017. The thesis author collected and analyzed all ESR data, performed all modeling and simulations, and prepared the manuscript.*

## 4.1 INTRODUCTION

Glutathione S-transferases (GSTs) are a family of detoxification enzymes that play an important role in cellular defense in a broad range of species from plants to animals[154]. GSTs catalyze the conjugation of glutathione (GSH) to various endogenous and exogenous electrophiles. This results in a deactivation of the reactive centers of such electrophiles, providing protection from a wide range of xenobiotics[155]. GST's have also been used as prodrug activators[156-158].

Currently, there are eight known classes of dimeric, soluble human GSTs (Alpha, Kappa, Pi, Mu, Theta, Zeta, Sigma and Omega)[159]. Each class also contains multiple isoforms, which typically catalyze several different reactions. This broad substrate specificity of GSTs can be described by a flux across a four-state pathway that is determined by several factors such as ligand concentration[160]. Two limiting cases of this flux are as follows; either the enzyme alters its conformation upon substrate binding (induced-fit model) or the unliganded enzyme may exist in multiple conformers and each conformer conjugates GSH to a different substrate (conformation selection model)[161]. It has been suggested that human alpha GST (hGSTA1-1) samples a range of states in the absence of substrates, and therefore is more accurately described by the latter model[162].

hGSTA1-1 contains 222 residues, which form two distinct domains, shown in Figure 4-1. The first, N-terminal domain (domain 1, residues 3-83) is a thioredoxin-like fold. The second domain (domain 2, residues 85-207) is made up of 5 α-helices and ends with a 15-residue sequence at the C-terminal that contains an 11 residue helix (210-220) referred to as α9[163-165]. Helix α9 is of particular interest as it is believed that the conformation of this C-terminal helix is affected by ligand binding[164,165]. Based on crystallographic[164,166-168] and NMR data[165], the C-terminal region of the Alpha-class GSTs is a well ordered helix when the active site is occupied by glutathione, products, or inhibitors[169,170]. In the absence of the ligand, α9 is believed to be folded but is delocalized[164,165,169-171]. It was shown that α9 in hGSTA1-1 likely exists in two conformations as indicated by ligand induced differences in ensemble fluorescence energy transfer[172], however it was difficult to extract exact distances from the energy transfer data because of the unknown orientation of donor and acceptor transition dipoles. The substrate-free
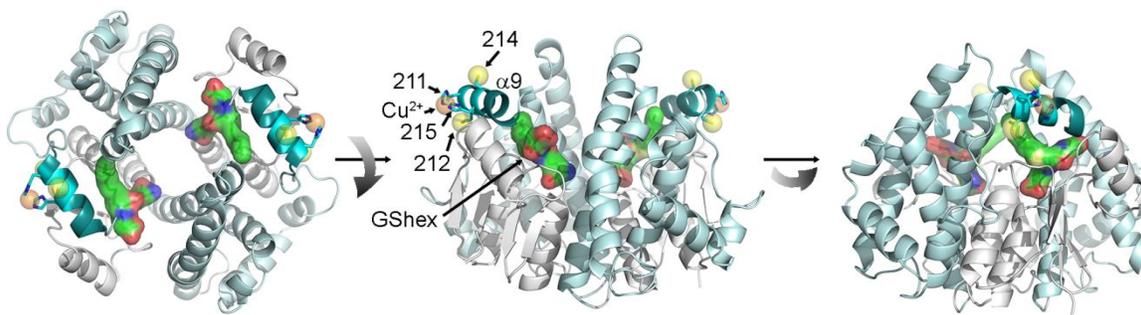
Figure 4-1. A) Looking down the 2-fold symmetry axis of the hGSTA1-1 dimer (PDB: 1K3L), while B) and C) were generated by 90 degree rotations around the x- and y-axis, respectively. In this image, residues 212 and 214 were converted to Cys residues and residues 211 and 215 were converted to His residues for illustrative purposes. The C-terminal helix is colored blue. The bound ligand, GS-hex, is highlighted with a green surface. The attachment points for R1 on either 212 or 214 are shown as yellow spheres. The $Cu^{2+}$ that interacts with the sidechains of H211 and H215 is represented by an orange sphere between the two His sidechains.

state is suggested to exist in a broad conformational ensemble with barrier-less transitions between states[162]. The existence of multiple conformations in hGSTA1-1 is supported by the two-step binding mechanism, wherein the initial binding event is followed by a closing of α9[172,173].

While the structure of the ligand bound state of hGSTA1-1 has been resolved by NMR and X-ray crystallography[164-168], the conformation of the delocalized α9 in the ligand-free state has yet to be elucidated. Not only is the conformation of α9 unknown, but the dynamics and equilibrium between the conformational states remain to be determined. The conformation of α9 critically influences the catalytic activity of the enzyme as a truncated wild-type hGSTA-1 that is missing the majority of α9 is essentially inactive[164]. Given that catalytic functions are highly dependent upon the C-terminal region[165,171,174-179], determination of the structure, conformation, and dynamics of the C-terminal α9 of hGSTA1-1 will provide key insights into how the enzyme achieves such breadth of substrate selectivity.

In this work, we use electron spin resonance (ESR) to shed light on the structure, conformation, and dynamics of the C-terminal α9 of hGSTA1-1. ESR has no molecular size limitation, requires no crystallization, and can probe highly dynamic areas. Techniques such as continuous wave (CW) ESR, double electron-electron resonance (DEER)[48,54,180] and double quantum coherence (DQC)[59,61,181] coupled with site-directed spin labeling (SDSL)[5] have enabled distance measurements and provided dynamical information within proteins and nucleic acids. These techniques have proven effective both in vitro and within the cell[182]. The most common approach of SDSL is the incorporation of the nitroxide functionalized side-chain, R1, created from the reaction of the thiol specific methanothiosulfanate spin label (MTSSL) with cysteine
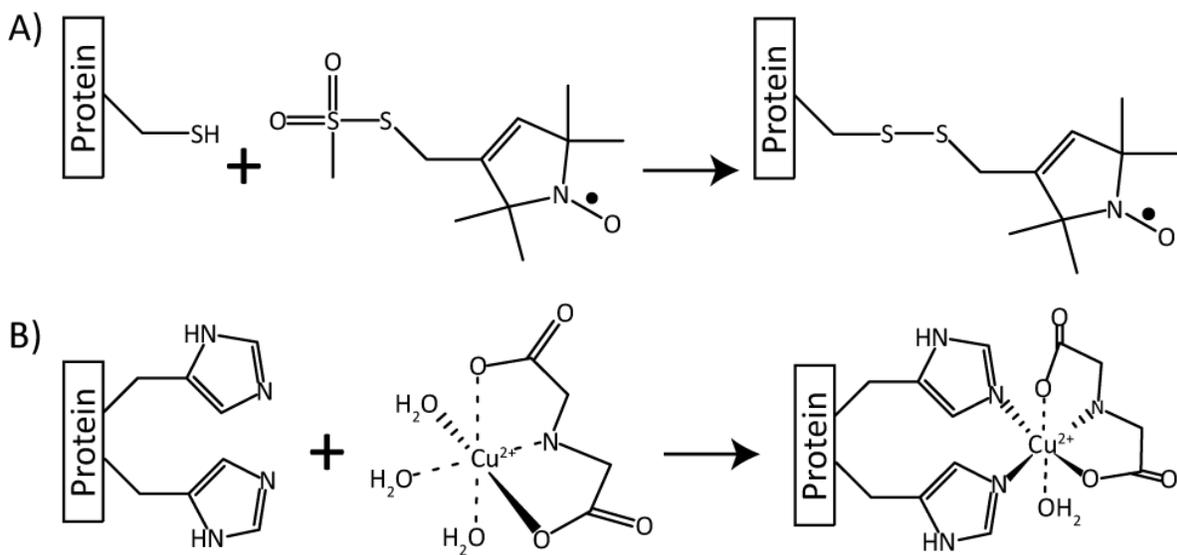
Figure 4-2. A) Reaction scheme of MTSSL with cysteine thiol to produce the R1 sidechain.

B) Schematic of the $Cu^{2+}$-binding dHis motif.

residues, as shown in Figure 4-2A. The R1 sidechain is capable of providing site-specific information on dynamics[183], long-range distance measurements within proteins[6,62,93], and information on conformational equilibria[4].

Aside from nitroxide based spin-labeling, SDSL using paramagnetic metal centers have also been shown to be suitable reporters of distance constraints and dynamics[8,55,68]. In particular, there are a number of newly developed strategies that incorporate $Cu^{2+}$ into specific sites in proteins[68,78,184]. Of these, the double histidine (dHis) motif strategy[52,58,135-139] places two histidine residues in close proximity so that they chelate $Cu^{2+}$, creating an extremely rigid reporter. Introduction of the $Cu^{2+}$ as a complex with the chelating agent iminodiacetic acid (IDA) increases the selectivity for $Cu^{2+}$ binding to the dHis site, and reduces nonspecific binding to native $Cu^{2+}$ binding sites[58]. The dHis motif using IDA is shown in Figure 4-2B.

Herein we use a combination of R1 and dHis SDSL to report on the conformation of the C-terminal α9 of hGSTA1-1 in the unliganded state and in complex with an S-alkyl glutathione derivative. R1-based DEER distance measurements are performed to create structural constraints, which are then applied to generate a model of the conformation of C-terminal α9 of hGSTA1-1 in the unliganded state. dHis-based DEER measurements are then implemented to provide an unbiased validation of the newly proposed model. Furthermore, the CW and DEER data reveal new insights into the effect of ligand binding on the conformation of the two states of the C-terminal α9 of hGSTA1-1. Surprisingly, a significant population of the disordered helix persists in the presence of bound ligand.

## 4.2  MATERIALS AND METHODS

### 4.2.1  Expression of hGSTA1-1

ESR experiments were performed on the human glutathione S-transferase A1-1 (hGSTA1-1). The hGSTA1-1 used in this study differs from the wild-type sequence in two aspects: i) cysteine 112 has been replaced with serine to generate a cys-null protein for R1 labeling, ii) the presence of an amino-terminal biotin tag sequence (GLNDIFEAQKIEWHE) for single-molecule studies[172]. Neither of these changes affects the activity of the enzyme[165,172]. For R1 labeling, the S212C and E214C mutants of hGSTA1-1 were prepared as previously described[172]. Protein concentration was 0.15-0.30 mM (monomer) and 1 mM DTT was added to reduce cysteines. The protein was washed in 25 mM HEPES, 125 mM KCl, pH 7.8 to remove DTT before labeling. Each hGSTA1-1 mutant was labeled with 10-fold excess S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSSL), obtained from Toronto Research Chemicals, using 100-fold excess of tris(2-carboxyethyl)phosphine (TCEP) in the same buffer as a reducing agent. Protein was labeled at room temperature for 4 hours followed by 4°C over-night. The protein was then washed and stored in DTT free buffer. Ellman's assay was used to confirm that there were no free thiols present. The activity of S212C and its R1 derivative were not significantly different from the wild-type activity. In the case of the E214C mutant and its R1 derivatives, their activity was slightly reduced from the wild-type activity.

For dHis DEER studies the following pairs of dHis mutants were constructed using the QuikChange mutagenesis kit (Agilent) and confirmed using Sanger DNA sequencing:

E210H/E214H, K211H/E215H, and E214H/K218H. Of these, only the K211H/E215H was active (equivalent to wild-type), the other two dHis mutants were inactive (<5% of wild-type). GST activity was determined as previously described using chlorodinitrobenzene as the substrate [172]. The n-hexyl conjugate of GSH (GS-hex) was purchased from Sigma-Aldrich. To ensure that the enzyme was saturated with GS-hex, it was added to each mutant in a 10-fold excess, approximately 10,000 times its KD of 70 nM[185]. Although one would expect some change in the affinity due to these mutations, this effect is expected to be relatively minor because the enzymatic activity of S212C, S214C, and dHis 211/215 are essentially the same as the wild-type enzyme. Given that the KD for GS-hex binding to the wild-type enzyme is 70 nM and that the concentration of GS-hex that was used in our experiments was 104 times this value, all proteins would be saturated with GS-hex, even with substantial changes in KD.

### 4.2.2   ESR Measurements and Analysis

ESR measurements were performed at three different temperatures, 294 K, 80 K and 20 K. Samples were prepared by drawing 8-10 µL aliquots into Pyrex capillary sample tubes for experiments performed at 294 K. All 294 K samples were run on a Bruker ElexSys 580 X-band FT/CW spectrometer equipped with an ER4122 SHQE-W1 high resolution resonator. Both 80 K and 20 K samples were prepared as 120 µL aliquots which were placed in either a Pyrex sample tube or a quartz sample tube for R1-based or $Cu^{2+}$-based measurements, respectively. Sample tube dimensions were 3 mm inner diameter by 4 mm outer diameter. For both 80 K and 20 K, samples were prepared in 20% v/v glycerol which is used as a cryoprotectant. Glycerol does not affect GST binding affinities or stability up to 30% v/v[186]. All 80 K and 20 K experiments were

performed on either a Bruker EleXsys 580 X-band FT/CW spectrometer or a Bruker EleXsys E680 CW/FT X-band spectrometer equipped with a Bruker ER4118X-MD5 or a Bruker ER4118X-MD4 resonator, respectively.

Continuous Wave (CW) ESR experiments were performed at X-band frequencies for both R1- and Cu2+-based samples. We collected 1024 data points over a sweep width of either 200 G or 2000 G, for R1 or $Cu^{2+}$, respectively. All experiments used a modulation frequency of 100 kHz, a time constant of 10.24 ms and a conversion time of 20.48 ms. R1-based samples used a modulation amplitude of 1 G while $Cu^{2+}$-based samples used a modulation amplitude of 4 G. Spectra were simulated using the Bruker Simfonia software.

The CW-ESR spectra of R1 mutants were further simulated using the microscopic order macroscopic disorder (MOMD) model using the NLSL program[187]. This model describes the anisotropic motion of an R1 sidechain in an orienting potential. The MOMD model relates the R1 sidechain to the protein through a series of coordinate frames, the magnetic frame of the nitroxide and the rotational diffusion frame. The relationship between these frames can be described by three Euler angles, however only one, $\beta D$, set to 36°, was needed to fit the observed data[11,188]. The CW spectra were fit using an axially symmetric rotational diffusion frame with two rotational rates ($R_\parallel$ and $R_\perp$) where the average diffusion rate is given by $R = (R_\parallel R_\perp^2)^{1/3}$ and the asymmetry parameter is given by $N = (R_\parallel/R_\perp)$. The rotational correlation time of R1 ($\tau$) is given by $\tau = 1/6R$. The MOMD model also includes an ordering potential parameter ($S_{20}$) which is typically used for R1 sidechains within a solvent exposed $\alpha$-helical site.

The three-pulse electron spin echo envelope modulation (ESEEM) experiment was performed at X-band using the $\pi/2 - \tau - \pi/2 - T - \pi/2 - \tau - $ echo pulse sequence with four-step phase cycling[143,144] at 80 K. A $\pi/2$ pulse length of 16 ns was used and the separation after the first

112

pulse, τ, was set to 144 ns. The second pulse separation, T, was incremented by a step size of 16 ns starting with an initial value of 288 ns. The experiment was performed at the magnetic field corresponding to the highest intensity in the echo-detected field swept spectrum. Raw time domain data were phase and baseline corrected, and then Fourier transformed.

The four-pulse double electron-electron resonance (DEER) experiments were performed at X-band at either 80 K (for R1) or 20 K (for $Cu^{2+}$). The following pulse sequence was used $(\pi/2)v1 - \tau1 - (\pi)v1 - \tau1+T - (\pi)v2 - \tau2\text{-}T - (\pi)v1 - \tau2 - echo$[48]. Observer pulse lengths $(\pi/2)v1$ and $(\pi)v1$ were 16 ns and 32 ns, respectively, for all samples. The pump pulse length $(\pi)v2$ was 16 ns for all samples. The delay, T, was incremented by a step size between 10-22 ns (depending on the sample) for 128 data points. The pump frequency, v2, was placed at the maximum of the echo detected field swept spectrum and the observer frequency, v1, was offset by downfield 150 MHz for $Cu^{2+}$ samples and 70 MHz for R1 samples. The raw time domain DEER data were analyzed using DeerAnalysis2013. $Cu^{2+}$ distances were analyzed using proper g-values[115]. Distances were verified using the validation tool of DEER analysis which generates error estimates using a Monte Carlo algorithm for analyzing the influence of background correction parameters as well as noise[105].

R1 generated distance constraints were used to generate a proposed structure using the multiscale modeling of macromolecular systems (MMM) software[189,190]. This software uses a computational approach to generate constraints based on an elastic network model (ENM) and draws from a rotamer library of R1 conformations based on a coarse-grained representation of the space of R1[189,190]. The ENM was constructed using ed-ENM, ed-ENM-p and Jeschke anisotropic network model parametrizations[191-194]. MTSSLWizard[195] as well as MMM were used to predict distance distributions using previous crystal structures as well as the model proposed

by MMM. Both ed-ENM and ed-ENM-p failed to fit the experimental data for S212R1. Jeschke parametrization produced acceptable fits, and thus, was the chosen parametrization method.

## 4.3 RESULTS AND DISCUSSION

### 4.3.1 DEER distance measurements for unliganded glutathione S-transferase

Two different mutants of hGSTA1-1, S212C and E214C, were chosen for nitroxide spin labeling. Since hGSTA1-1 is a homodimer, single cysteine mutations provide two sites for spin labeling and distance measurements. These specific residues are close to, but not involved in glutathione (GSH) binding. The residues also do not engage in any inter-monomer interactions upon GSH binding. Residues past E215 (216-222) have a more direct involvement in protein functionality. In particular, altering F220 shows an increased off-rate of GSH by a factor of 3.2[176]. In addition, I219 significantly contributes to numerous tertiary contacts which help localize the C-terminal helix[185,196]. Any alteration of I219 leads to a repositioning of F220 and F222 which are required for GSH activation[185]. Altering F222 can possibly lead to destabilization of the C-terminal helix[165]. Mutation of either S212 or E214, as well as their subsequent labeling with R1, leads to minimal changes in the function of hGSTA1-1, based on enzymatic activity (see material and methods). These mutations also should not affect the thermodynamic stability of the enzyme[170,197,198]. After reaction of the individual cysteine mutants with MTSSL, creating S212R1 or E214R1, the labeling efficiency was determined to be

114

>95% via the doubly integrated intensity of the CW ESR spectrum when compared to a previously prepared standard calibration curve.

Figure 4-3 shows the background-corrected DEER data and the corresponding distance distribution of the unliganded hGSTA1-1 S212R1 and E214R1 proteins. All of the raw DEER data are shown in Figure 4-4. These R1-based DEER distance measurements suggest two conformations of the C-terminal α-helix (α9) in the unliganded state of hGSTA1-1. In the case of S212R1, shown in Figure 4-3 (top), the distance distribution shows two distinct distances with the most probable distances of 3.7 ± 0.2 nm and 4.6 ± 0.2 nm (mean ± s.d.), respectively (validation shown in Figure 4-5). The bimodal nature of the distribution could be due to the α9 helix of hGSTA1-1 existing in two distinct conformations as suggested by recent NMR results [172]. However, it is possible that the two populations could be due to two distinct rotameric species of R1 present. We used MTSSLWizard to simulate expected R1-R1 distances based on previous x-ray crystal structure of the liganded state (PDB: 1K3L) (13). MtsslWizard is a PyMol plugin that can predict R1 based distance distributions using spatial sampling for R1 conformations free of steric clashes[195]. The predicted distance distribution is shown by dotted lines in Figure 4-3. This modeling indicates that the expected most probable R1-R1 distance for the liganded structure is 5.4 nm. This value is larger than the distance reported by DEER. The DEER data unambiguously show the presence of two distances, 3.7 and 4.6 nm. Modest occupation of the distance predicted by MtsslWizard (5.4 nm) is also supported by the DEER data, but with a higher degree ofuncertainty. Furthermore, since S212R1 is located near the N-terminus of the helix, the poor fit from MtsslWizard may be due to the static treatment of the protein environment. In solution, alternative R1 rotamers not sampled using the program may be possible[2]. For E214R1, shown in Figure 4-3 (bottom), the distance distribution shows a major

115
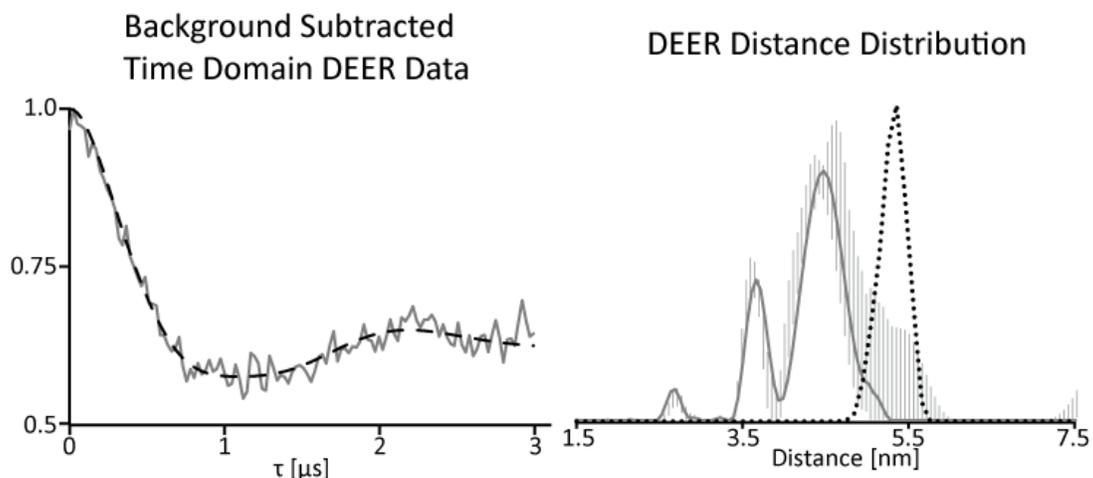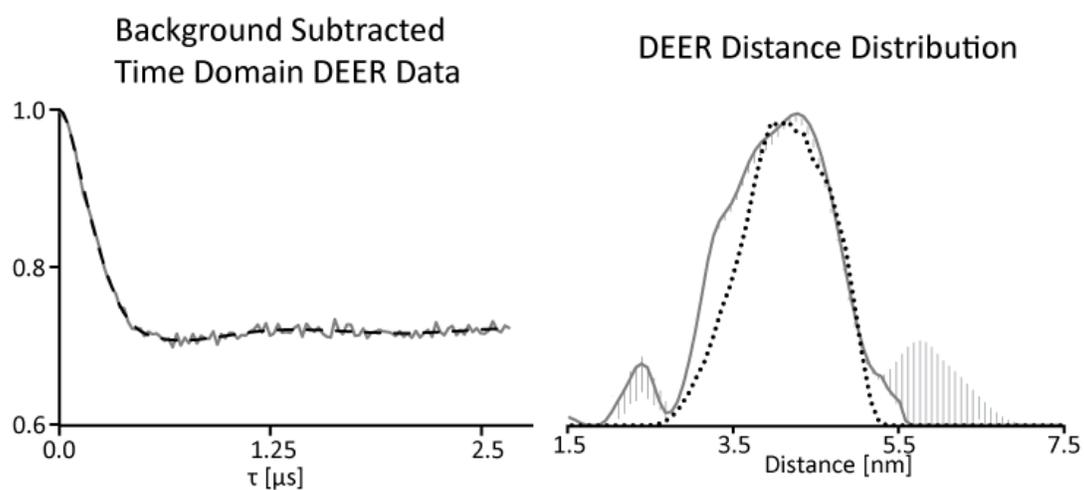
**Figure 4-3.** Background-corrected time domain DEER data (left) and resultant distance distribution (right) for both unliganded hGSTA1-1 mutants S212R1 and E214R1 (top and bottom, respectively). Fit modeled by DEERAnalysis 2013 shown as dashed line. The bimodal distance distributions are strong evidence that the unliganded hGSTA1-1 exists in equilibrium between two conformations. The distance distribution predicted by MTSSLWizard based on the liganded GST structure (PDB: 1K3L) is shown as dotted line.
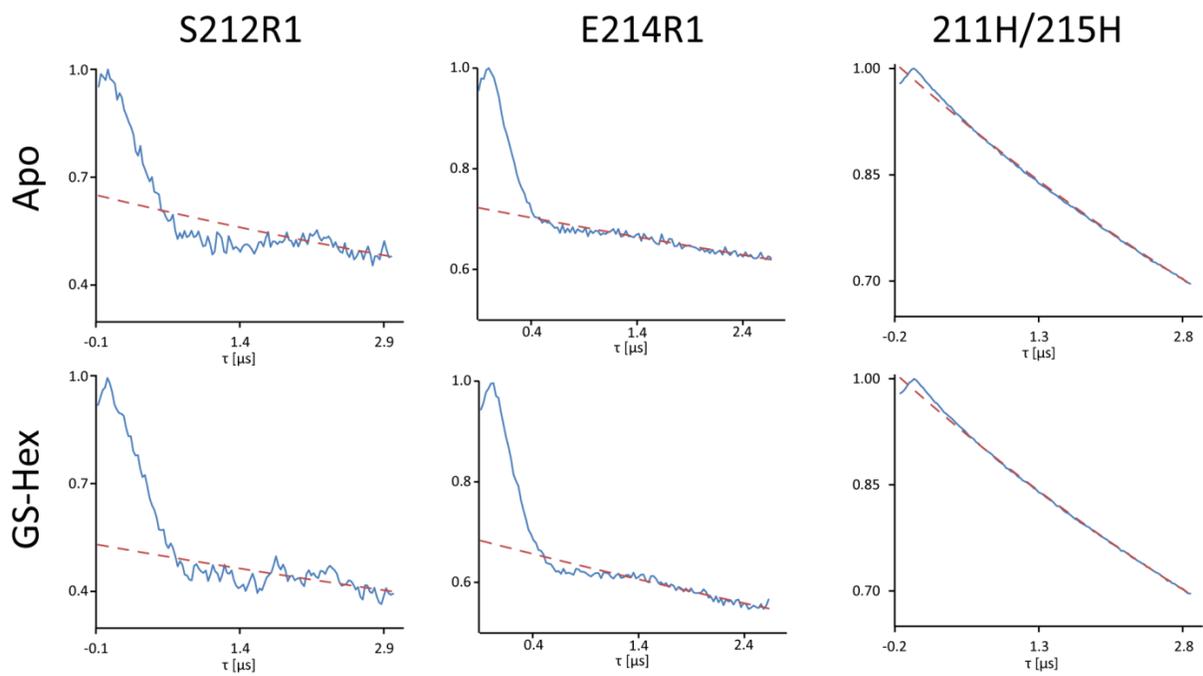
Figure 4-4. Raw time domain DEER data for S212R1, E214R1 and 211H/215H hGSTA1-1 (solid) with fitted background exponential (dashed).
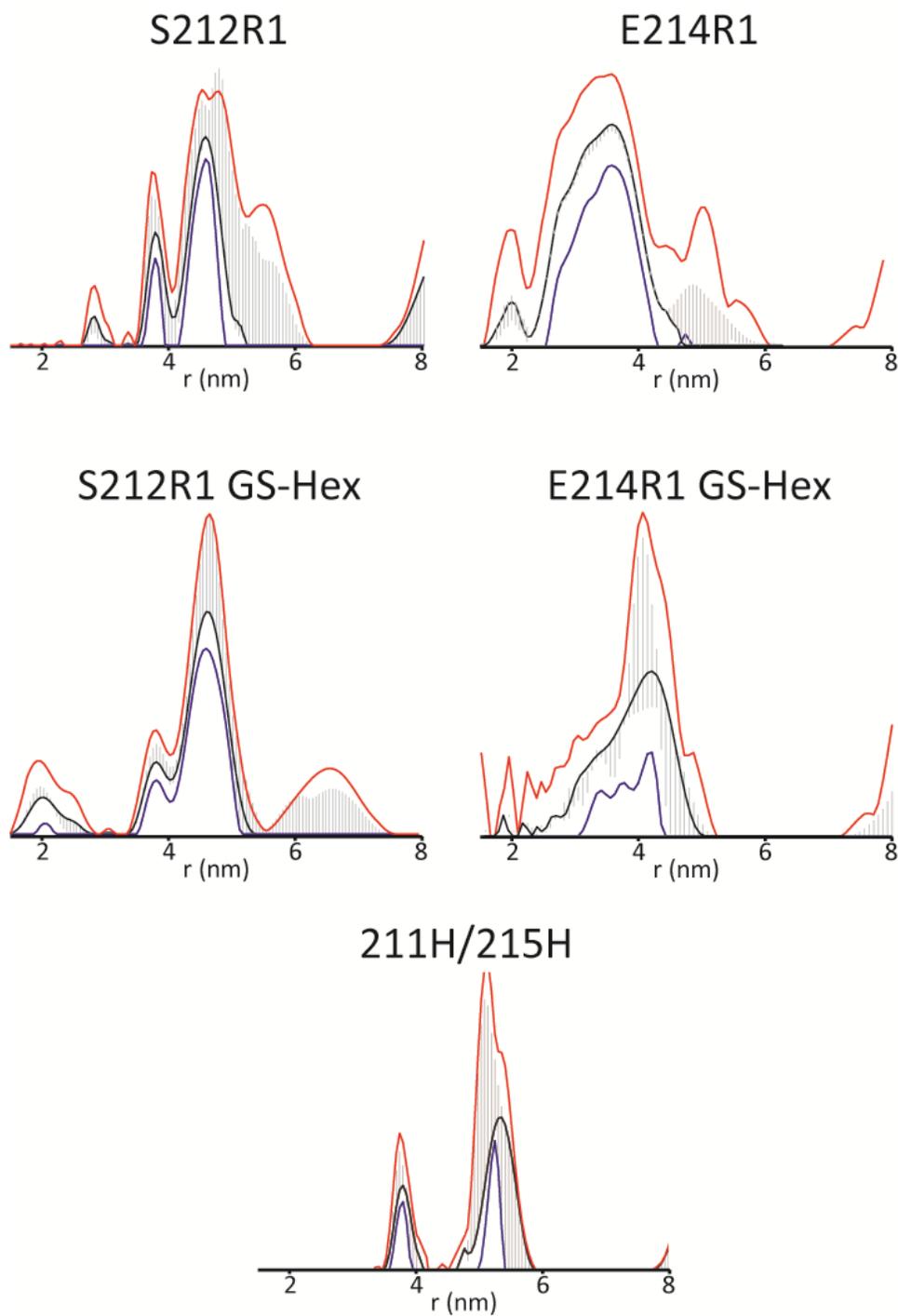
Figure 4-5.Validation of Tikhonov regularization based DEER distance distributions. The figures illustrate the average values (black), maximum (red) and minimum (blue) along with acceptable associated error (vertical lines).

component with a most probable distance 4.1 ± 0.3 nm (mean ± s.d.) and a possible shoulder representing a most probable distance of 3.6 ± 0.2 nm (mean ± s.d.). Validation of the distances is shown in Figure 4-5. To further assess the significance of two features in the E214R1 distance distribution, we have provided a dual Gaussian analysis as shown in Figure 4-6. The small shoulder may be an indication of the second conformation present in the unliganded hGSTA1-1 or due to a second conformer of the spin label. The expected R1-R1 distance of the liganded structure is 4.2 nm as simulated by MTSSLWizard. This value is in excellent agreement with the 4.1 nm distance reported by DEER. Therefore, we conclude that the larger distance represents the conformation of α9 assumed in liganded hGSTA1-1.

In order to proceed, we first assumed that both of the second distances measured by DEER in S212R1 and E214R1 represent the second conformer of the α9 helix. The presence of a second conformation is in accord with recent NMR and fluorescence data (10, 19). We used the most probable distances determined by DEER spectroscopy to generate a constraint-based model of the C-terminal α-helix confirmation for Apo-hGSTA1-1. We then performed independent $Cu^{2+}$-based DEER measurements and modeling to verify this model.

To model the unliganded α9, we started with the full length liganded GST (PDB: 1K3L) wherein α9 is fully resolved. This template provides us with the secondary structure of α9. We used the two inter-monomer R1 DEER constraints along with four inter-monomer Cα-Cα constraints (210-210, 211-211, 213-213 and 215-215) taken from a structure generated via x-ray crystallography (PDB: 1PKZ) where the unliganded GST C-terminal α-helix was partially resolved (residues 2-215 on one monomer and 2-219 on the other)[164]. All parameters used are listed in Table 4-1. This crystal structure, although truncated, provides the only available

Table 4-1. Constraints used for MMM generated model (PDB: 1K3L)

|              | Distance | Std. Dev. |
|--------------|----------|-----------|
| **DEER**     |          |           |
| (A)214-(B)214 | 3.6     | 0.4       |
| (A)212-(B)212 | 3.7     | 0.4       |
| **Direct**   |          |           |
| (A)210-(B)210 | 4.0     | 0.1       |
| (A)211-(B)211 | 4.5     | 0.1       |
| (A)213-(B)213 | 3.6     | 0.1       |
| (A)215-(B)215 | 4.0     | 0.1       |
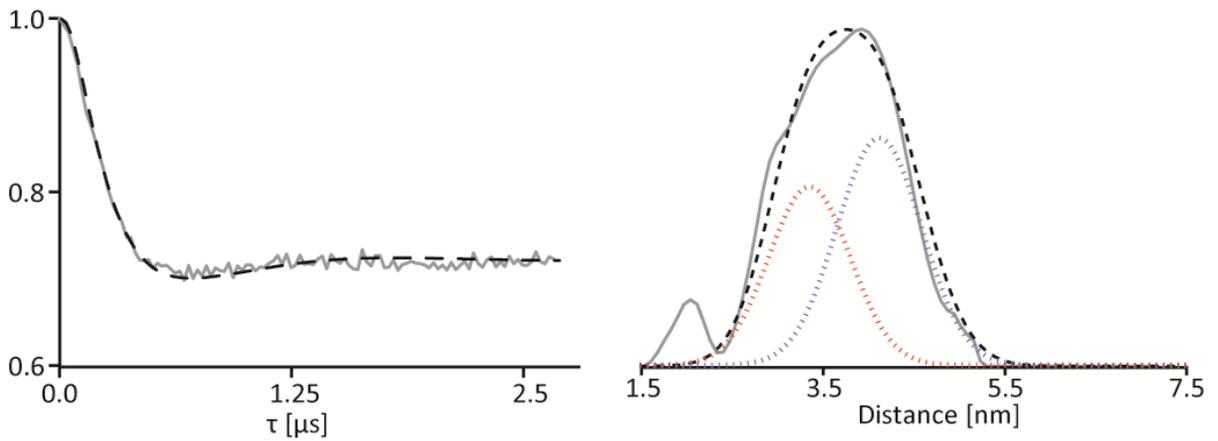


Figure 4-6. E214 R1 data analyzed via a dual Guassian fit. Left panel shows the experimental data (solid) overlaid with the dual Gaussian fit (dashed). Right panel shows the Gaussian distribution (dashed) overlaid on the Tikhonov fit (solid) for comparison. Both Gaussian components are shown (slotted, red and blue).

structural information on the unliganded C-terminal α-helix of hGSTA1-1. Since only the orientation of the α-helix was in question, the small number of constraints was sufficient to define its orientation. The model was generated using the elastic-network model implemented in the multiscale modeling of macromolecular systems using the MMM2015 software[189,190]. Figure 4-7 shows suggested ligand-free conformations of the C-terminal. As evident in Figure 4, both α9 helices undergo a ~30% rotation. MMM's quoted ~4 Å error can account for the slight dissimilarity with regard to helical translation. The similarity between the helical shifts within the dimer is exemplified by the side view shown in Figure 4-7.

We performed both MtsslWizard and MMM simulations to show that the DEER data agrees reasonably with our proposed unliganded conformation. Figure 4-8 shows the distance distribution predicted by MTSSLWizard for both S212 and E214 mutants (dashed) overlaid on the experimental data (solid). Notably, the distance based on our proposed structure overlays well with the shorter distance of both experimental distributions of S212 and E214 (see Figure 4-8). These data also support the presence of two distinct helical conformations of α9. Distance distributions were similarly predicted using MMM2015 and are shown in Figure 4-8A (shaded). Both the S212 and E214 results by MMM2015 agree well with both experimental data and MTSSLWizard simulations.

Next, we performed $Cu^{2+}$-based DEER measurements to provide further validation that the C-terminal α-helix of hGSTA1-1 exists in two conformations, and that both conformations are helical. The mutant K211H/E215H creates a dHis site in the C-terminal α-helix, shown in Figure 4-9A. $Cu^{2+}$-IDA was attached to the K211H/E215H dHis site in each subunit. The dHis-$Cu^{2+}$-IDA motif is a highly rigid, site-specific spin label useful for precise ESR distance measurements in proteins[13,58]. Importantly, the coordination to $Cu^{2+}$ restricts the conformations
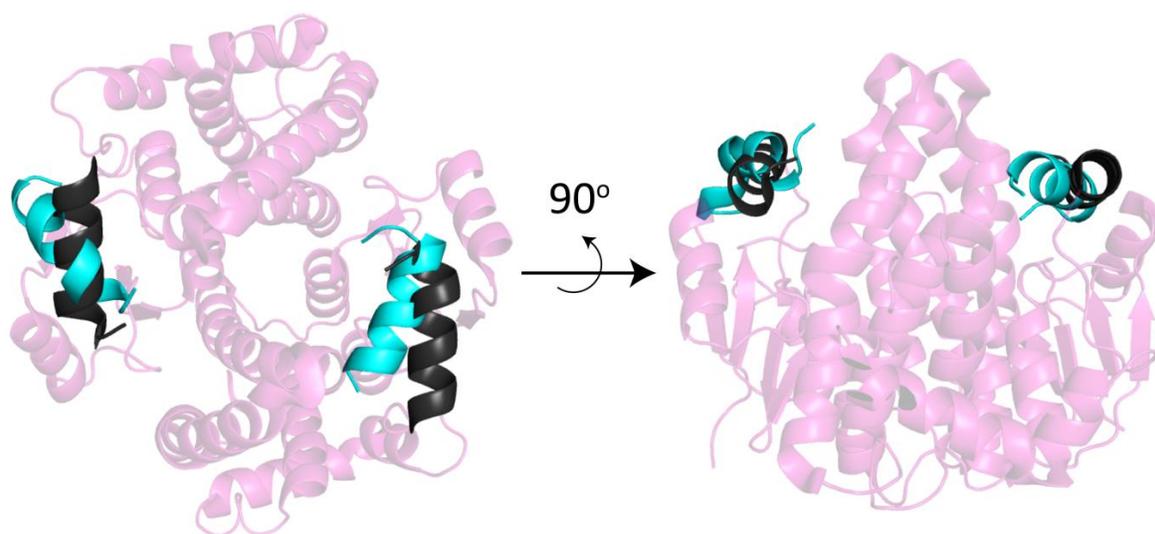
Figure 4-7. Comparison of known α9 conformation of liganded hGSTA1-1 shown in black (PDB: 1K3L) with the MMM generated model of α9 of the second unliganded conformation hGSTA1-1 shown in blue. DEER data indicates that unliganded hGSTA1-1 exists in conformational equilibrium between these two helical conformations.
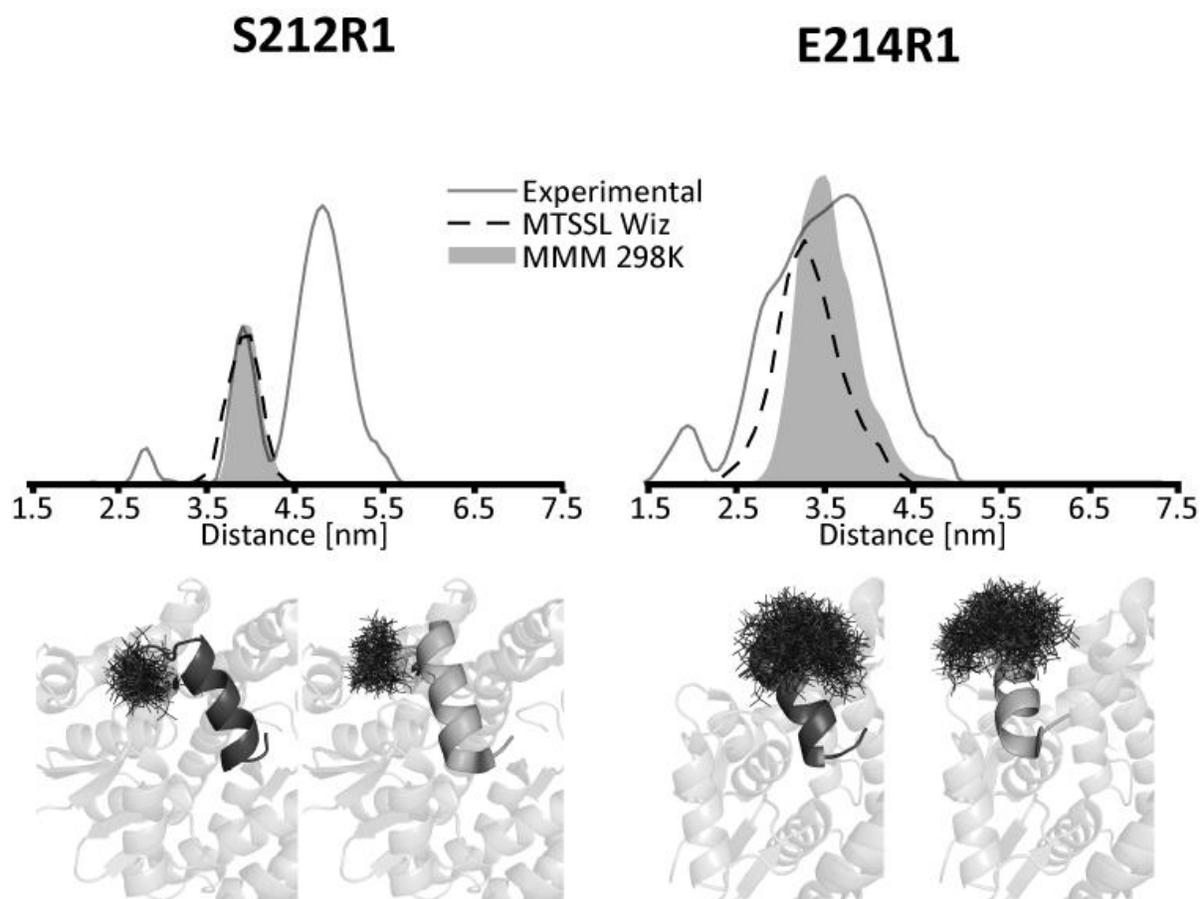
Figure 4-8. Simulated R1-R1 distance distributions of both S212 (left) and E214 (right) mutants of hGSTA1-1 produced by both MtsslWizard (dashed) and MMM (shaded) overlaid upon actual experimental data (solid). In the case of the S212R1 sample, the MtsslWizard curve is on top of the experimental curve for the peak at 3.7 nm. Insets shows orientation of R1 distribution in both ligand free model (dark grey) and ligand bound crystal structure (grey).
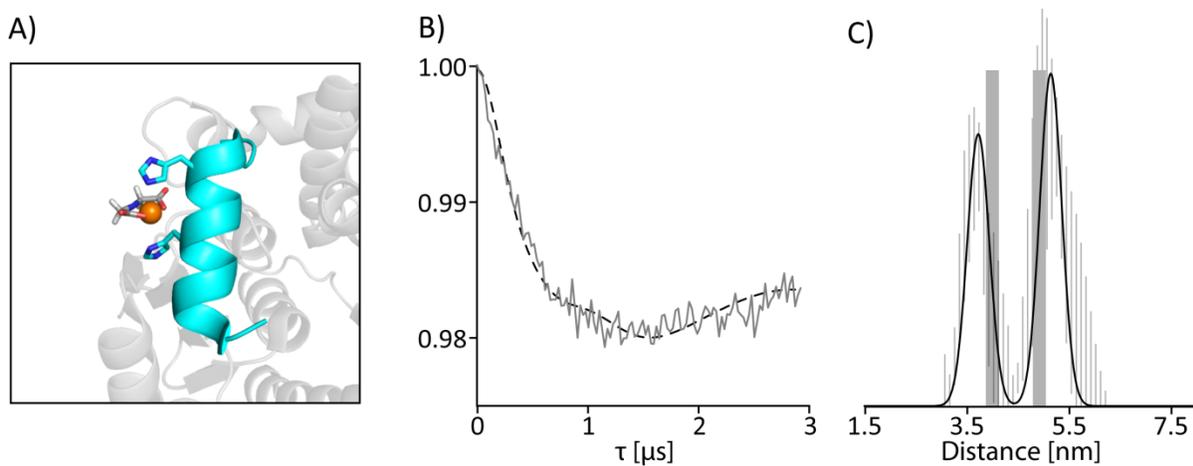
Figure 4-9. A) The C-terminal α-helix (blue) of hGSTA1-1 mutant K211H/E215H containing the $Cu^{2+}$-IDA binding dHis site. B) The background subtracted time domain DEER data (solid) and corresponding fit (dashed). C) The distance distribution showing two distinct distances. Expected distances shown by grey lines.

of the His side-chain and the motif provides narrow distance distributions that can be easily related to backbone conformation.  The dHis-$Cu^{2+}$-IDA coordination was confirmed via CW-ESR and ESEEM shown in Figures 4-10 and 4-11. DEER experiments were performed at both $g_\perp$ and $g_\parallel$ regions of the field-swept spectrum. Orientational effects were negligible as shown in Figure 4-12, see Supporting Information. The background-corrected time domain DEER data, as well as the distance distribution for unliganded K211H/E215H are shown in Figure 4-9 (B and C, respectively). Raw DEER data and validation are shown in Figures 4-4 and 4-5 respectively.

The bimodal distribution in Figure 4-9C strongly supports the above R1 studies as well as previous fluorescence data[172]. As shown in Figure 4-9, the distance distribution shows two components with most probable distances of $3.6 \pm 0.3$ nm (mean $\pm$ s.d.) and $5.1 \pm 0.3$ nm (mean $\pm$ s.d.). The histidine residues do not sample a wide range of rotamers when simultaneously bound to $Cu^{2+}$-IDA. Therefore, the bimodal distribution cannot be from rotameric preference and strongly supports that two distinct conformations of α-9 are indeed present. Using previously published rotamers of the dHis motif within an α-helix[58], expected distances can be generated for both conformations using crystal structures as well as our proposed helical conformation. These expected distances are shown by the grey vertical lines in Figure 4-9. The longer distance agrees with the ligand-bound crystal structure (PDB 1K3L)[168]. Expected $Cu^{2+}$-$Cu^{2+}$ distances were modeled using previously reported $Cu^{2+}$ bound to IDA and two imidazole ligands[199]. Likewise the shorter distance is in reasonable agreement with our model, as shown in Figure 4-9C. Note, that the dHis distance constraints were not used in MMM2015, providing an unbiased validation.

Previous data has suggested the secondary structure of the unliganded conformation of α9 is of "helical nature"[164,165]. The presence of dipolar modulations in the dHis-based DEER time domain data suggests that the secondary structure of the second conformation is indeed helical.
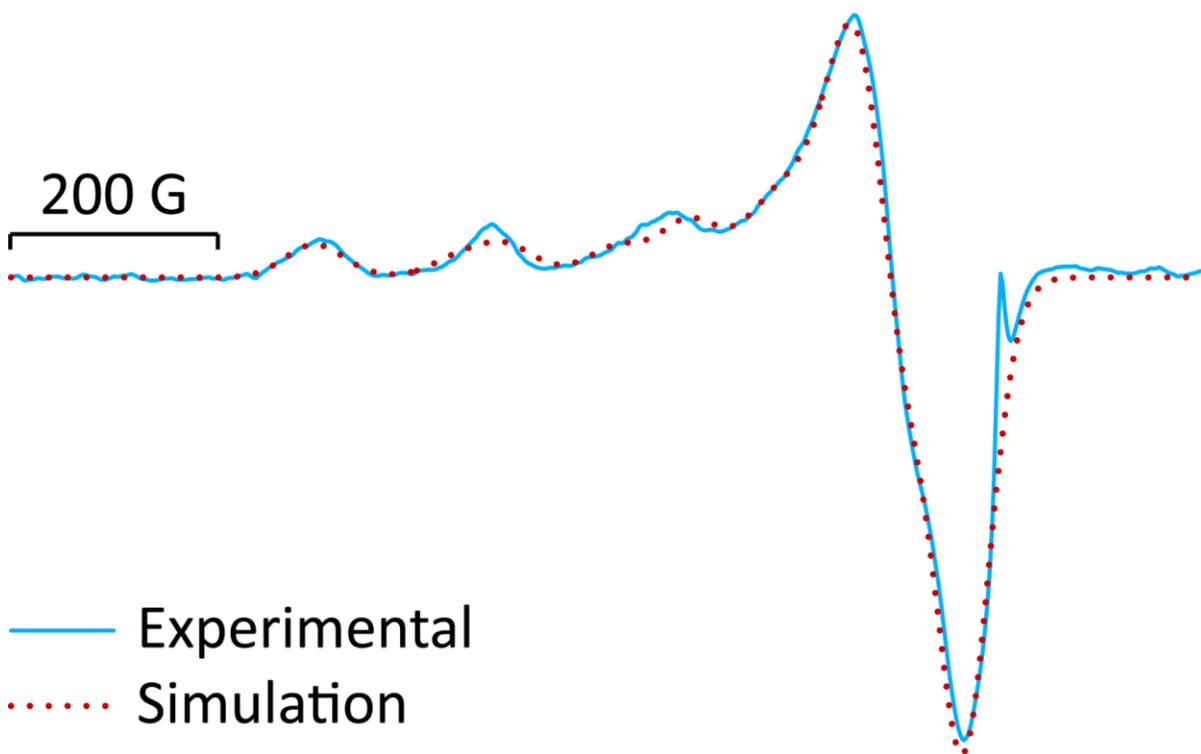
Figure 4-10. CW spectrum (solid) of 211H/215H hGSTA1-1 in the presence of $Cu^{2+}$-IDA with respective simulation. Two components were used, one consistent with $Cu^{2+}$-IDA bound to two histidine residues and one consistent with free $Cu^{2+}$-IDA in solution.
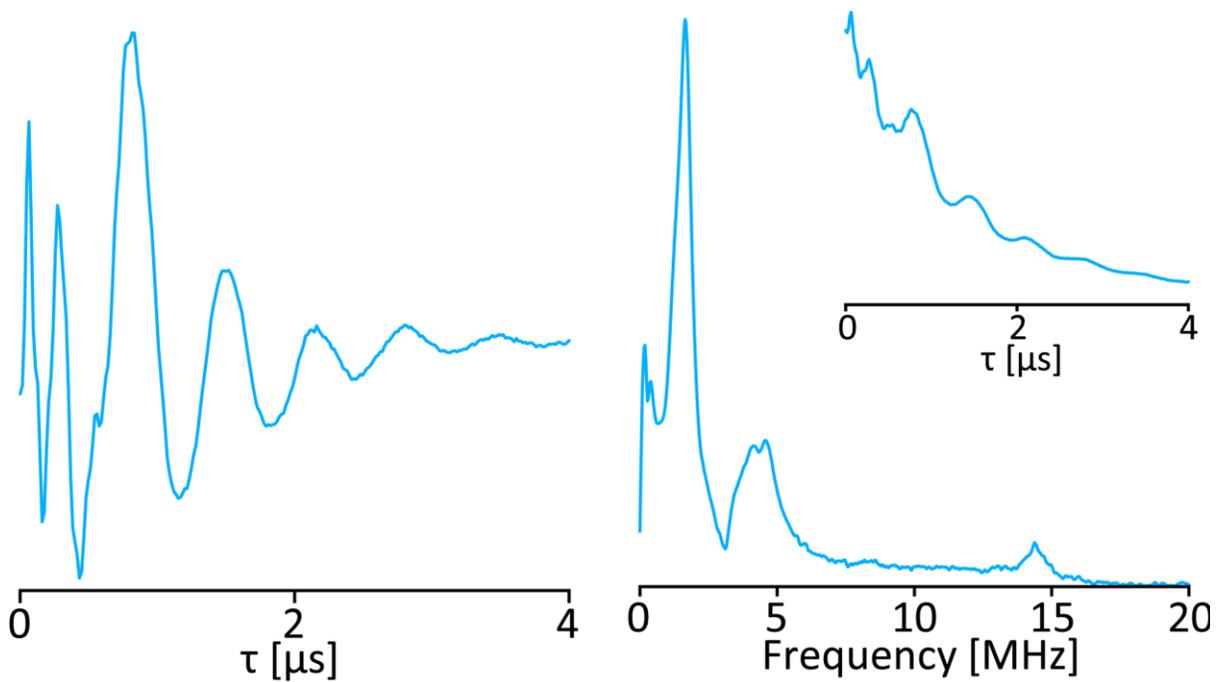
Figure 4-11. Left) Background corrected time domain ESEEM data of 211H/215H hGSTA1-1 showing strong nitrogen modulations. Right) ESEEM FFT showing DQ peak at 4 MHz and NQI peaks below 2 MHz consistent with imidazole coordination.
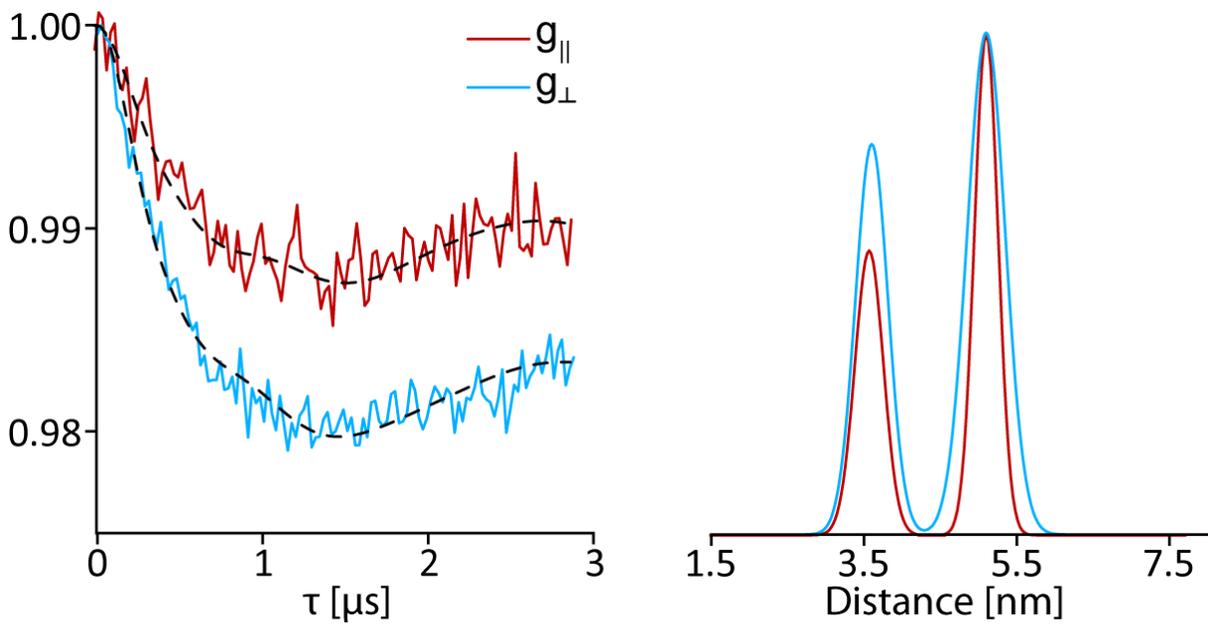
Figure 4-12. DEER data of 211H/215H hGSTA1-1 at two different magnetic fields, 3355 G for the $g_{\parallel}$ region (blue lines), and 3080 for the $g_{\perp}$ region (red lines). Left shows the time domain data overlaid to exhibit the similar dipolar modulation periods while the right shows the similar distance distributions. No major orientational effects are noted.

The dHis motif for an α-helix uses an i,i+4 pattern in order to position the two His residues on neighboring helical turns[138]. If the structure of the C-terminal α-helix was not helical, the His residues would not be in proximity to bind $Cu^{2+}$ and there would no observed distance in the DEER distribution. However, we note, that in the presence of $Cu^{2+}$, the dHis site may promote helical stabilization to some degree[58,136].

## 4.3.2  Populations within two conformations of the C-terminal helix of hGSTA1-1

We performed experiments to probe which conformation of hGSTA1-1 is favored in the liganded state. Upon addition of the ligand, the n-hexyl conjugate of glutathione (GS-hex), to hGSTA1-1, the relative populations of the DEER distance distributions shifted to favor the conformation associated with the longer distance. Figure 4-13 shows the DEER distance distributions for all three proteins, both unliganded (solid) and in the presence of GS-hex (dotted). The populations of each of these conformations can be obtained by comparing integrated areas of each component of the bimodal distance distribution[200]. The enhancement of the conformation with the longer distance is consistent with all previous work that suggests that, upon ligand addition, the C-terminal α9 of hGSTA1-1 adopts a single conformation[163-165]. In contrast, our DEER data suggests that in the presence of GS-hex, there are still two conformational states, however one conformation is heavily favored. The shift in the populations due to GS-hex binding were supported by continuous wave (CW) ESR. Figure 4-14 shows the nitroxide CW ESR spectra with and without the presence of GS-hex for both S212R1 and E214R1. The data were collected at 294 K to measure biologically relevant dynamics. The CW
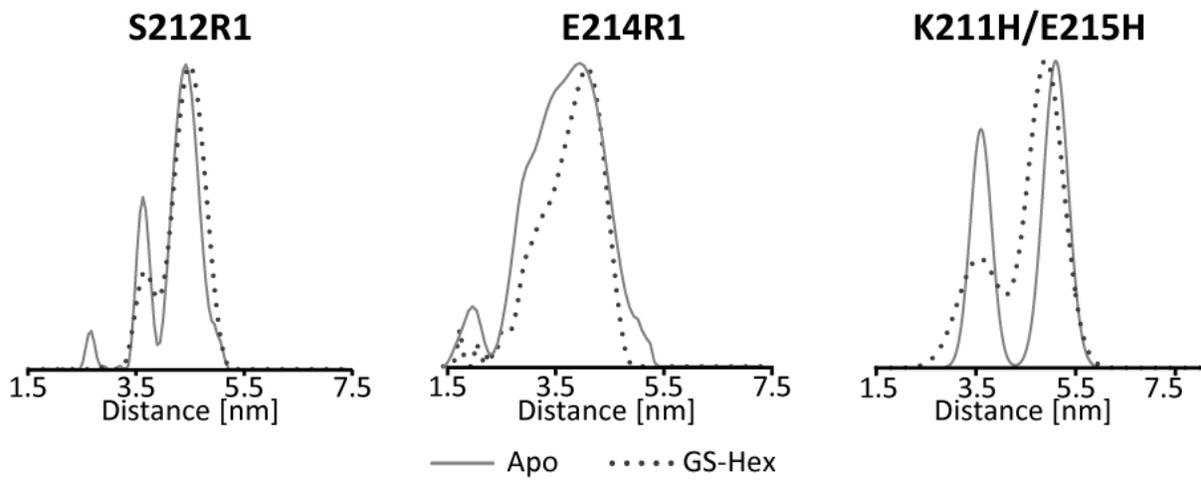
Figure 4-13. DEER distance distributions with (dotted) and without (solid) the presence of GS-hex of hGSTA1-1 S212R1, E214R1 and K211H/E215H. Upon addition of GS-hex, the population of the conformation with the larger distance occurs for all three sites. This larger distance coincides with expected distances of the liganded conformation based on X-ray crystallography (PDB: 1K3L).
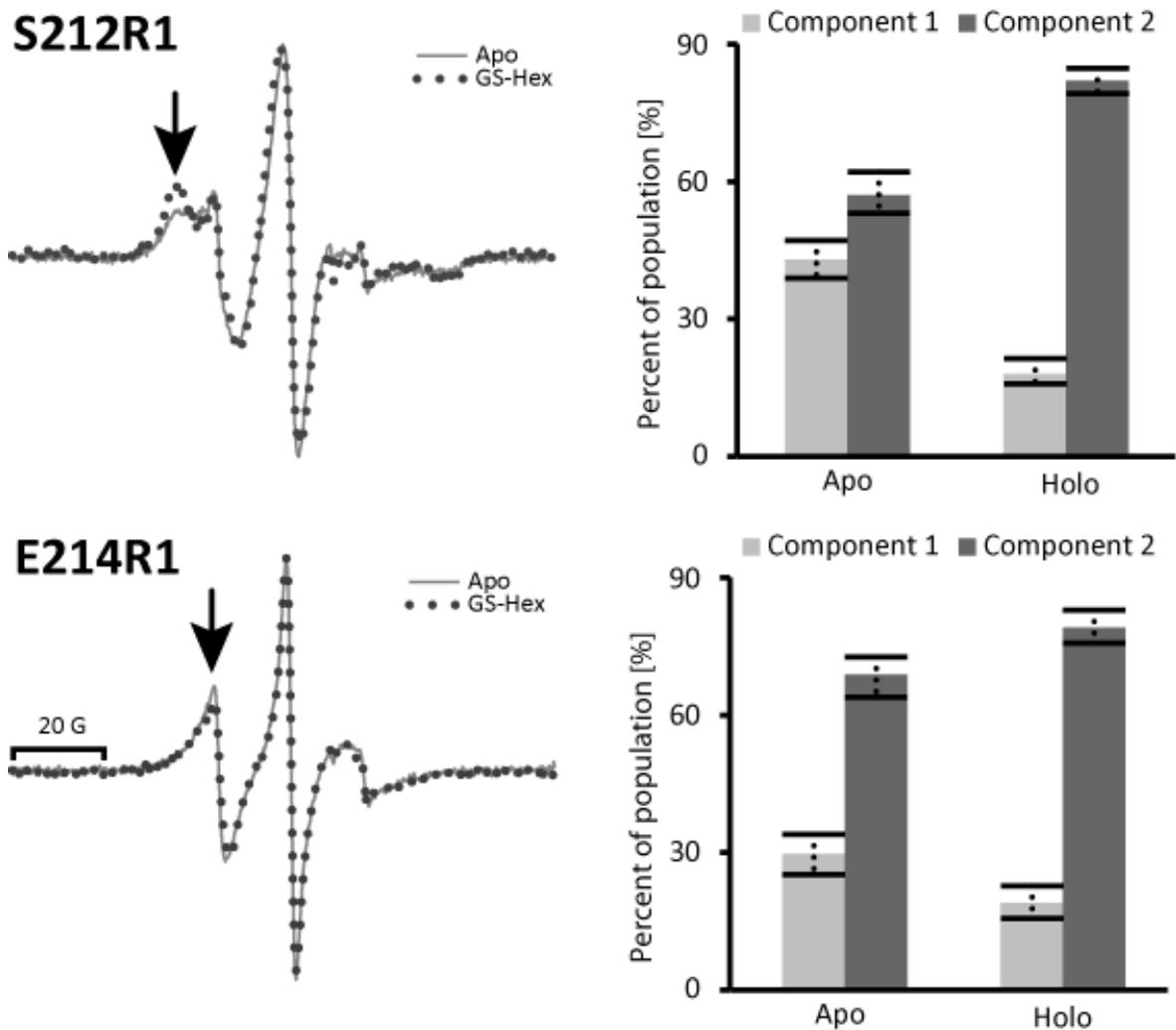
Figure 4-14. 294 K nitroxide CW spectra of S212R1 (top) and E214R1 (bottom) both unliganded (solid) and in the presence of GS-hex (dotted). Spectral changes are highlighted by the black arrows. Spectra were simulated using the MOMD model and differences upon ligand addition could be accounted for by changing only the relative population between components. This change in population is shown by the bar graph (black lines represent error in fits).

spectra consist of two components of different mobility. The CW spectra were best fit with two components using the microscopic order macroscopic disorder (MOMD) model where three parameters were allowed to change, the rotational correlation rate ($\tau$), the order parameter ($S_{20}$) and an asymmetry parameter (N)[187]. Simulations of all CW spectra are shown in Figures 4-15 and 4-16. Respective parameters used in simulations are listed in Table 4-2. Each spectrum contained a mobile component (small $\tau$) and an immobile component (large $\tau$). In both cases, for S212 and E214, the positioning of the nitroxide did not result in a difference in orientation or steric interference, as shown in Figure 4-8. Therefore, we can claim the differences in component mobility could be resultant of shifts in backbone dynamics. Based on the relative populations of these two components, the immobile component likely represents the helix as found in the ligand containing X-ray based structures. We can then infer that the mobile component represents the unliganded state of α9. The higher mobility suggests that the α9 helix in the unliganded state is highly dynamic. This inference is supported by earlier solution NMR studies[165].

For the S212R1, upon addition of GS-hex, there is a large increase in the slow component as indicated by the black arrow in Figure 4-14 (top). Both spectra were fit without changing any parameter other than their relative population. These data agree well with DEER data which showed a shift in population toward the longer distance upon the addition of the ligand (see Table 4-3). The DEER population distribution shifted from 34/66% to 13/87% (short/long conformations). The CW ESR simulations suggest a shift from 40/60% to 18/82% (component 1/component 2).

In the case of E214R1, there is also a shift in population toward the slower component, as indicated by the arrow in Figure 4-14 (bottom). The changes are smaller than those for S212R1, which is in agreement with DEER data above (c.f. Figure 4-13, see Table 4-3). For E214R1, the

Table 4-2. Parameters used for MOMD model simulations of S212R1 and E214R1

| | | S20 | τ (ns) | N | Occupancy (%) |
|---|---|---|---|---|---|
| S212R1 Apo | Component 1 | 0.15 | 0.6 | 25 | 40 |
| | Component 2 | 0.5 | 10.5 | 40 | 60 |
| S212R1 GS-Hex | Component 1 | 0.15 | 0.6 | 25 | 18 |
| | Component 2 | 0.5 | 10.5 | 40 | 82 |
| E214R1 Apo | Component 1 | 0.1 | 0.5 | 25 | 29 |
| | Component 2 | 0.2 | 3.3 | 35 | 71 |
| E214R1 GS-Hex | Component 1 | 0.1 | 0.5 | 25 | 22 |
| | Component 2 | 0.2 | 3.3 | 35 | 78 |

Table 4-3. Summary of Populations as Detected by R1, $Cu^{+2}$ DEER and CW ESR

| | Apo | GS-hex |
|---|---|---|
| Sample-Technique | % conf I/% conf II | % conf I/% conf II |
| S212R – R1 DEER | 34/66 | 13/87 |
| S212R - CW | 40/60 | 18/82 |
| E214R – R1 DEER | 40/60 | 25/75 |
| E214R - CW | 29/71 | 22/78 |
| K211H/E215H -dHis DEER | 40/60 | 25/75 |
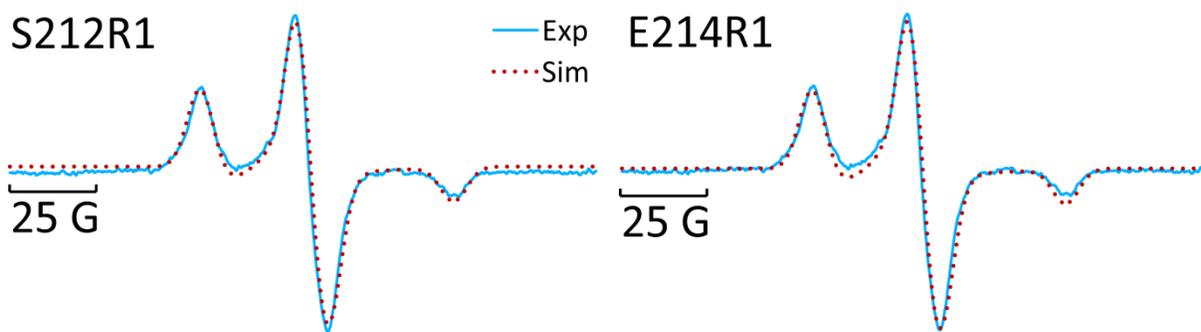
Figure 4-15. 80 K CW spectra (solid) of both S212R1 (left) and E214R1 (right) hGSTA1-1 with respective simulations (dotted).
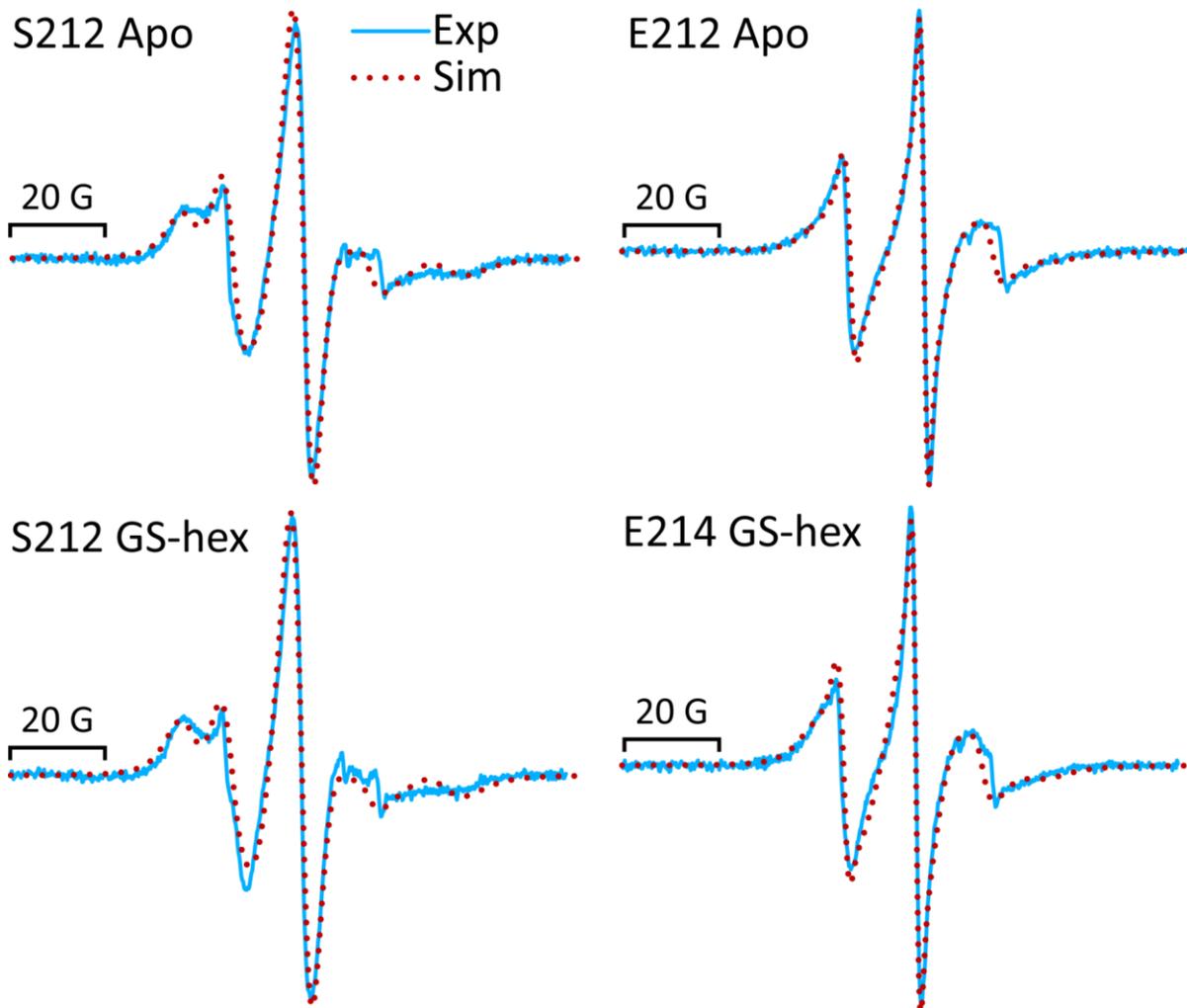
Figure 4-16. CW simulations of both S212R1 (left) and E214R1 (right) in the unliganded (top) and liganded (bottom) states. Experimental data is shown as solid blue line while simulated data is shown as red dotted.

DEER population distribution shifted from 30/70% to 20/80% while the CW data indicate a shift from 29/71% to 22/78%. Both data sets suggest that even upon addition of GS-hex, the C-terminal α-helix of hGSTA-1 still exists in two distinct conformations. However, both DEER and CW data suggest that in the presence of GS-hex, the C-terminal α-helix does shift its relative population to highly favor the conformation that is observed in crystal structures of hGSTA-1 ligand complexes. The difference in the response to ligand binding by S212R1 versus E214R1 suggest that modification of these residues, by mutagenesis and nitroxide labeling, has led to different effects on the ability of the bound ligand to favor one conformation. The difference in populations for the E214R1 site may also be reflected in a slight decrease in enzyme activity for this protein (see Figure 4-17).

Taking the agreement between populations revealed by DEER and CW-ESR data into account, we propose that the immobile component represents the liganded conformation for two reasons. First, the population of the immobile component increases upon addition of GS-hex in a similar manner as to the population increase of the DEER distance representative of the liganded conformation. For S212, the immobile component increases 22% upon addition of GS-hex while the DEER population increases 23%. In the case of E214, the immobile component increases 9% upon addition of GS-hex while the DEER population increases 10%. Second, the larger $\tau$ value is indicative of a nitroxide conformation that is more highly constrained as opposed to a conformation that is sampling space (low $\tau$)[11,183,201]. Furthermore, the higher $S_{20}$ of the immobile component is also consistent with a more rigid protein conformation[11,183,201]. Therefore, we can conclude that the immobile component represents the liganded conformation while the mobile component represents the unliganded conformation.
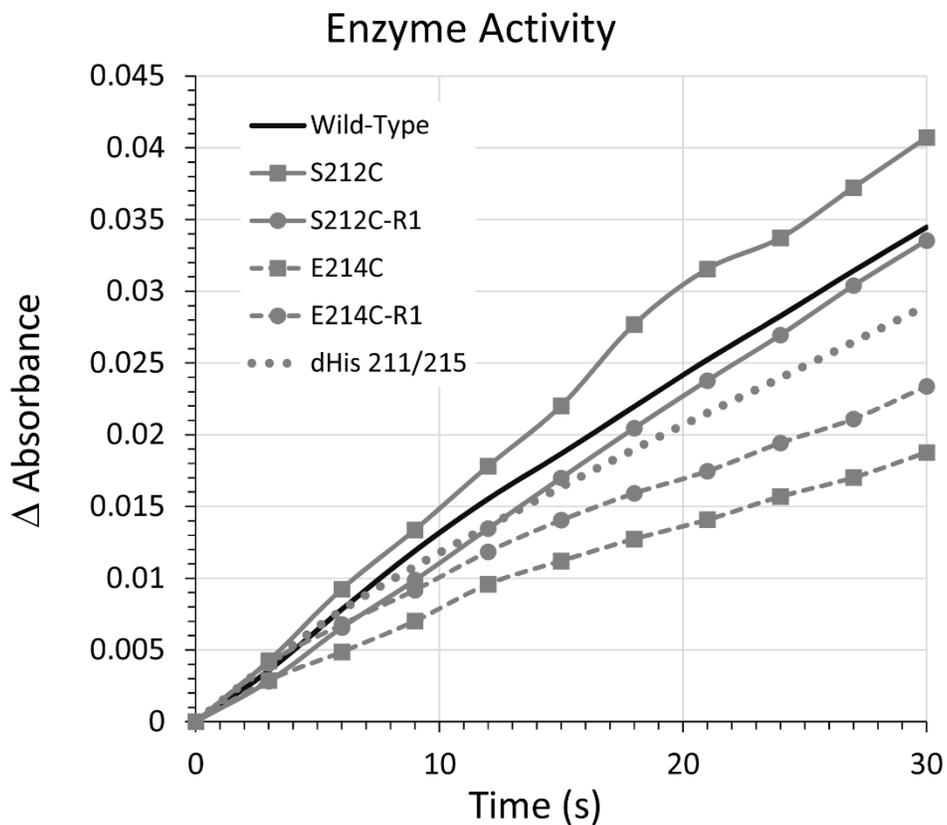
Figure 4-17. Enzymatic activity of Mutant and R1 labeled proteins. The change in absorbance as a function of time is shown for the wild-type enzyme (solid black), S212C (solid grey, squares),S212C-R1 (solid grey, circles), E214C (dashed, squares), E214C-R1 (dashed, circles), and dHIs 211/215 dotted. Each trace is the average of two separate experiments. Initial rates are not significantly different for wild-type, S212C, S212C-R1, and dHis21/215 (T statistic < 1). The activity of S214C and S214C-R1 are lower than wild-type at the 86 and 77 % confidence limits, respectively.

137

## 4.4    CONCLUSION

In summary, we have demonstrated that the terminal helix in hGSTA1-1 exists in two conformations, both of which appear to be helical. One conformation, which is more highly populated when the active site is occupied, is very similar to that observed in X-ray derived structures of liganded hGSTA1-1. The second conformation, although still helical and localized, is more dynamic in solution under physiological conditions. We also demonstrate the enhanced resolution and utility of dHis-Cu$^{2+}$ DEER measurements to measure protein structure and conformational changes.

## 4.5    ACKNOWLEDGEMENTS

# 5.0      AN ANALYSIS OF NITROXIDE BASED DISTANCE MEASUREMENTS *IN CELL-EXTRACT* AND *IN-CELL* BY PULSED ESR SPECTROSCOPY

*This work, written in collaboration with Amit Shimshi, Timothy F. Cunningham, Monica N. Kinde, Pei Tang and Sunil Saxena has been published in ChemPhysChem, 2017, volume 18, pages 1653-1660. The thesis author collected most ESR data, analyzed all data and prepared the manuscript.*

## 5.1      INTRODUCTION

Electron spin resonance (ESR) techniques such as double electron-electron resonance (DEER)[54,202] and double quantum coherence (DQC)[8,59,61,181,203,204] coupled with site-directed spin labeling (SDSL)[205] have been used to measure distances to study structure and dynamics in both proteins[4,6,8,140,205-208] and nucleic acids[79,82,83,89,209-211] *in vitro*. In recent years, there has been a considerable effort to further the ESR field *in-cell* in order to measure the structure and or conformation of biological macromolecules in their native environment[212-215]. Inside the cell, it is estimated that the concentration of biological macromolecules is in the range of 80-400 mg/mL[216] which creates a "crowded" environment[217]. This "crowded" environment may not only

affect the structure of biological macromolecules, but also various biological equilibria such as protein folding, aggregation, binding of small molecules, enzymatic activity and protein-protein interactions[217-219]. Despite various efforts, this crowded environment cannot be accurately replicated *in vitro*[214,220]. Accordingly, several ESR studies on biological macromolecules have been carried out in various types of cells including *Xenopus laevis* oocytes[221-224], *Escherichia coli* cells[225-229] as well as human cells[53,230]. ESR is an ideal technique for determination of structure and dynamics of biological macromolecules *in-cell* because it is non-destructive, highly sensitive when compared to similar techniques such as NMR, and can operate over a wide range of temperatures.

The bottleneck of *in-cell* distance measurements by ESR is that the cellular environment causes a drastic loss in the DEER signal over time for nitroxide spin labels[221,231-233]. To make progress in the field, it is necessary to understand the significant contributions to the loss of signal. One potential source of signal loss is cleavage of the disulfide bond of the typical methanothiosulfonate nitroxide side chain (called R1) in the cellular environment[221,233]. Cleavage of the R1 side chain results in decreased protein labeling without a loss in overall signal which in turn increases background noise, hindering the DEER experiment. Issues of cleavage can be minimized by the use of different nitroxide linkage methods such as maleimides[53,221]. However, R1 labeling for *in-cell* DEER and CW analysis is common in order to extract information on protein dynamics and structure[226,227,229]. In addition to disulfide cleavage, signal loss can also result from reduction of the nitroxide to an ESR silent hydroxylamine[232]. Both these factors lead to a rapid loss of DEER signal upon incubation in-cell. While nitroxide based distances have been reported upon incubation, in-cell DEER experiments show much clearly resolved dipolar modulations when cells are flash frozen immediately after the protein was introduced[221]. While

immediate flash freezing does allow for maximum signal intensity, it does not allow time for the protein to equilibrate in the cellular milieu, bringing into question the biophysical significance of these measurements[53].

As an alternative, $Gd^{3+}$ labels have been found to be useful for distance measurements *in-cell*[53,121,234]. $Gd^{3+}$ labels have enabled DEER distance distributions to be obtained after 5 hours of incubation *in-cell*. This method uses an unreducible maleimide linkage, though many $Gd^{3+}$ tags still rely on a reducible disulfide bond. However, in comparison to nitroxide labels, $Gd^{3+}$ requires high fields such as W-band (95 GHz) in order to extract distance distributions. Recently, sterically shielded trityl radicals show improved resistance to *in-cell* radical reduction[235-238], and also allow for room temperature[239,240] and *in-cell* distance measurements[241,242]. These trityl radicals show tremendous promise, especially for orthogonal labeling[242]. In comparison to the typical R1 sidechain, the trityl radicals are large in size. Many of the  trityl labeling methods still employ thiol based reactions creating disulfide bonds which could be subject to reductive disulfide cleavage. However, trityl labels that attach without the use of a disulfide bond[241] have recently been developed and these are likely to be useful for in-cell work.

An alternative to the *in-cell* approach is using cytoplasm extracted from *Xenopus laevis* oocytes as the sample buffer (i.e. *in cell-extract*). However, it is not clear whether the *in cell-extract* environment is an accurate mimic of the *in-cell* environment due to a lack of organelles and membrane. Thus, differences in nitroxide reduction, disulfide bond cleavage and protein structure/dynamics may be different *in cell-extract* when compared to *in-cell*.

Herein, we have examined the contribution of R1 sidechain disulfide reduction to the loss of DEER signal both *in-cell* and *in cell-extract*.  To better understand the behavior of the nitroxide radical within a cellular environment, we first report a quantitative look at the

suggested use of the oxidizing agent $K_3Fe(CN)_6$[232,243]. This oxidizing agent is added in order to extend the lifetime of the nitroxide radical within *Xenopus laevis* oocytes. Using the extended nitroxide radical lifetime, we report a comparison of nitroxide signal reduction versus modulation depth decrease both *in cell-extract* and *in-cell* ESR spectroscopy using R1 labeled GB1, the immunoglobulin binding domain of protein G. Our results indicate disulfide bond cleavage contributes to the loss of DEER signal. Also, it is evident that both nitroxide reduction and disulfide cleavage occur at a much faster rate *in-cell*, likely due to organelle and membrane associated factors. Data also suggest that there are no further changes in local concentration past 30 minutes indicating that 30 minutes is enough time for this protein to equilibrate *in-cell* post-microinjection. Taken together, these data provide essential knowledge for improving ESR methods *in-cell*.

## 5.2    MATERIALS AND METHODS

### 5.2.1    Reagents and Cellular Extract

3-carboxy-2,2,5,5,-tetramethylpyrrolidinyl-1-oxy (PCA) nitroxide was obtained from Sigma-Aldrich and dissolved into sterile, ultra-filtered water. Potassium ferricyanide ($K_3Fe(CN)_6$) was purchased from Sigma-Aldrich.

Healthy *Xenopus laevis* oocytes (stages V-VI) were selected for cytosol preparation. They were maintained at 15 °C in a modified Barth's solution containing (in mM): 88 NaCl, 1 KCl, 2.4 $NaHCO_3$, 15 HEPES, 0.3 $Ca(NO_3)_2$, 0.41 $CaCl_2$, 0.82 $MgSO_4$, and 10µgml−1 sodium

penicillin, 10μgml−1 streptomycin sulphate, 100μgml−1 gentamycin sulphate, pH 6.7. Oocytes were removed from the Barth's solution and homogenized using a Scilogex D160 tissue homogenizer. Cellular debris was pelleted and cytosol was collected by ultracentrifugation at 100,000xg for 30 min at 4°C using a Beckman Coulter TL-100 ultracentrifuge.

### 5.2.2   GB1 Expression, Purification and MTSSL Labeling

The expression, purification and MTSSL labeling of the 15R1/28R1 mutant GB1 construct were carried out as described previously[2,146,244].

### 5.2.3   Microinjection of Oocytes

Stage V-VI *Xenopus laevis* oocytes were selected and maintained as described above. Every oocyte was manually injected with 50 nL of 500 μM spin-labeled GB1 in a sodium phosphate buffer (pH) containing 2.5 mM $K_3Fe(CN)_6$ using a Nanoject II microinjector by Drummond Scientific. Approximately 60 oocytes were individually placed into a Pyrex sample tube (4 mm OD, 3 mm ID) and were gently settled by gravity. The oocytes were subsequently flash frozen by immersion in a bath of liquid nitrogen either immediately (on average it took 10 minutes to inject 60 oocytes), or after 30 minutes and 60 minutes of incubation at 15°C in the modified Barth's solution.

143

### 5.2.4 ESR measurements

Both CW and time scan experiments were performed on a Bruker ElexSys E580 CW/FT X-band spectrometer using a Bruker ER4122 SHQE-W1 high resolution resonator. Samples for room temperature time scan experiments were prepared in a total volume of 40 µL with either 200 µM or 500 µM PCA and varying concentrations of $K_3Fe(CN)_6$ while keeping the total cytosol amount constant at 36 µL. 8-10 µL were then drawn into Pyrex capillary sample tubes. All CW experiments were run at a center field of 3440 G with a sweep width of 200 G with a modulation amplitude of 1 G and a modulation frequency of 100 kHz for a total of 1024 data points using a conversion time of 20.48 ms. Time scan experiments used the same experimental parameters as CW, except a total of 8192 data points were collected over the course of approximately 6 hours. The static magnetic field was set to the maximum of the center peak of the nitroxide triplet. Experiments were performed at 294 ± 1 K. *in cell-extract* CW experiments were also performed at 80 K. These samples consisted of 120 µL of cytosol containing 250 µM doubly nitroxide labeled GB1 with 20% v/v glycerol which was flash frozen in liquid methylacetylene-propadiene propane. The same parameters were used as above and 10 scans were collected.

Four pulse DEER experiments were performed on a Bruker ElexSys E680 CW/FT X-band spectrometer equipped with a Bruker EN4118X-MD4 resonator. The *in-cell* samples consisted of approximately 60 microinjected *Xenopus laevis* oocytes. The *in cell-extract* samples consisted of 120 µL of cytosol containing 250 µM doubly nitroxide labeled GB1 with 20% v/v glycerol which was flash frozen in liquid methylacetylene-propadiene propane. The experiments were completed at 80 K controlled using an Oxford ITC503 temperature controller and an

Oxford CF935 dynamic continuous flow cryostat connected to an Oxford LLT 650 low-loss transfer tube. The following pulse sequence was used: $(\pi/2)_{v1} - \tau1 - (\pi)_{v1} - T - (\pi)_{v2} - \tau2 - (\pi)_{v1} - \tau2 - echo^{202}$. The observer pulses, $(\pi/2)_{v1}$ and $(\pi)_{v1}$, were 16 ns and 32 ns respectively while the pump pulse, $(\pi)_{v2}$, was 16 ns. T was incremented by a step size of either 8 ns or 10 ns for 128 points. The pump frequency, $v2$, was placed at the maximum of the echo detected field swept nitroxide spectrum and the observer frequency, $v1$, was offset 70 MHz downfield. The time domain DEER data were analyzed using DeerAnalysis2013[245] via Tikhonov regularization. Data acquisition ranged from five hours to approximately three days.

**5.2.5 Theoretical Modulation Depth Calculations.**

The theoretical modulation depths ($\lambda_T$) are calculated using the integrated intensities of the CW spectra. The integrated intensity of the CW spectrum is directly proportional to the number of spins present. By dividing the integrated intensities of all three incubation times (A, A+30 and A+60) by the integrated intensity at time A, we can calculate the percent nitroxide signal remaining. Here we make the assumption that 95% signal remains at time A due to the time required for microinjection. The equation for modulation depth is given as follows[54,56,57]:

$$V_p = (1 - p_b)^{N-1} \tag{5-1}$$

Where:

$$V_p = 1 - \lambda_T \tag{5-2}$$

Such that $p_b$ is the percent of spins excited by the pump pulse (calculated to be 0.36) and N is the number of spins where N=2 describes 100% nitroxide signal and N=1 describes 50% nitroxide

signal. Next we convert the percent nitroxide signal remaining, evaluated from the CW integrated intensities, into the theoretical N:

$$(percent\ nitroxide\ signal) * 2 = N_{theoretical} \tag{5-3}$$

For instance, 76% nitroxide signal would yield N=1.52. By inserting theoretical N from equation (5-3) into equation (5-1), we calculate $V_p$ which is related to $\lambda_T$ by equation (5-2).

## 5.3 RESULTS AND DISCUSSION

3-carboxy-proxyl (PCA) was used as a R1 nitroxide mimic for measuring the reduction of the nitroxide radical to a hydroxyl amine *in cell-extract*[232]. This reduction, *in cell-extract*, is believed to be enzyme mediated[232,246]. The reduction of PCA, as well as the typical ESR spin probe, R1, can be seen in Figure 5-1. Cytosol was extracted from stage V-VI *Xenopus laevis* oocytes and used to mimic the *in-cell* environment.

First, we explored the use of an oxidizing agent ($K_3Fe(CN)_6$ in order to extend the lifetime of the nitroxide radical in the cellular milieu. $K_3Fe(CN)_6$ has been reported to successfully extend the lifetime of a nitroxide radical *in vivo* for ESR imaging and radical detection, but has never been quantitatively measured[243,247]. In order to assess the effect of $K_3Fe(CN)_6$, continuous wave (CW) time scan ESR experiments were carried out at room temperature (294 ± 1 K). By using the time scan experiment, the signal intensity (which is directly proportional to the number of ESR active species present) at a static magnetic field can be monitored over time. Time scans were then completed with differing aliquots of $K_3Fe(CN)_6$ shown in Figure 5-2. The error associated with the decay curves is smaller than the actual line

146

Figure 5-1. A) 3-carboxy-proxyl (PCA) nitroxide radical being reduced to the ESR silent hydroxyl amine. The forward reaction is believed to be enzyme mediated while the back reaction can be facilitated by a strong reducing agent, in this case, $K_3Fe(CN)_6$. B) The typical R1 nitroxide spin label covalently attached to a cysteine residue in a protein. C) The protein GB1 (PDB ID: 5BMG)[2] double R1 labeled at positions 15 and 28 (red spheres).
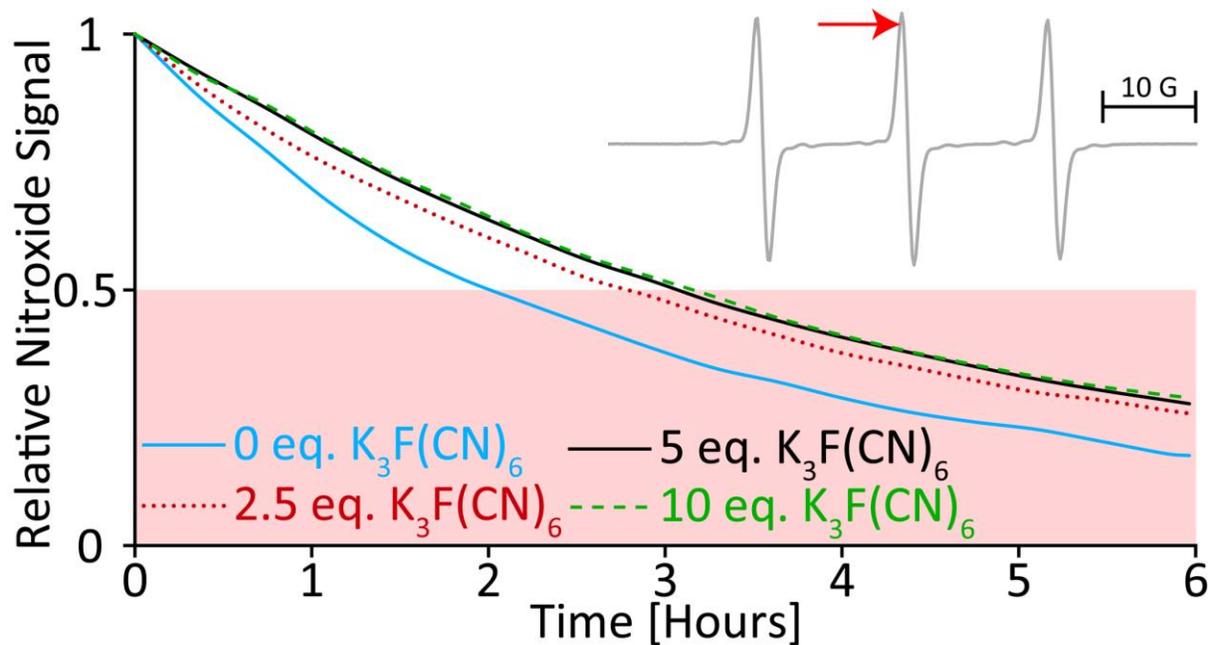
Figure 5-2. Time scan decay curves of 200 µM PCA in cytosol with differing equivalents of $K_3Fe(CN)_6$. The shaded region indicates the cut off of 50% where, on average, there is only one spin per system, and measurement of distances by DEER becomes difficult. It is clear that the decay rate ceases to decrease past 5 equivalents of $K_3Fe(CN)_6$. Inset shows the CW spectrum of PCA in cytosol where the red arrow represents the position of the static magnetic field for the time scan experiments.

represented in Figure 5-2. The rate of nitroxide decay is decreased from $0.96 \pm 0.03$ μM/min to $0.64 \pm 0.02$ μM/min by the addition of 5 equivalents of $K_3Fe(CN)_6$. Increasing the $K_3Fe(CN)_6$ concentration past 5 equivalents led to no further decrease of the decay.

First, we explored the use of an oxidizing agent ($K_3Fe(CN)_6$ in order to extend the lifetime of the nitroxide radical in the cellular milieu. $K_3Fe(CN)_6$ has been reported to successfully extend the lifetime of a nitroxide radical *in vivo* for ESR imaging and radical detection, but has never been quantitatively measured[243,247]. In order to assess the effect of $K_3Fe(CN)_6$, continuous wave (CW) time scan ESR experiments were carried out at room temperature ($294 \pm 1$ K). By using the time scan experiment, the signal intensity (which is directly proportional to the number of ESR active species present) at a static magnetic field can be monitored over time. Time scans were then completed with differing aliquots of $K_3Fe(CN)_6$ shown in Figure 5-2. The error associated with the decay curves is smaller than the actual line represented in Figure 5-2. The rate of nitroxide decay is decreased from $0.96 \pm 0.03$ μM/min to $0.64 \pm 0.02$ μM/min by the addition of 5 equivalents of $K_3Fe(CN)_6$. Increasing the $K_3Fe(CN)_6$ concentration past 5 equivalents led to no further decrease of the decay.

While five equivalents of $K_3Fe(CN)_6$ is the most needed even for larger concentrations of nitroxide, it has been shown that an increase in [PCA] will increase the rate of nitroxide reduction[232]. We next tested to see if more than 5 equivalents of $K_3Fe(CN)_6$ would be necessary in order to reach the lowest decay possible with an increased [PCA]. Time scan decay curves of 500 μM PCA in cytosol with different equivalents of $K_3Fe(CN)_6$ are shown in Figure 5-3. As evident from Figure 5-3, no more than 5 equivalents of $K_3Fe(CN)_6$ are required to achieve the maximum nitroxide radical lifetime *in cellulo*. These results are consistent with the same
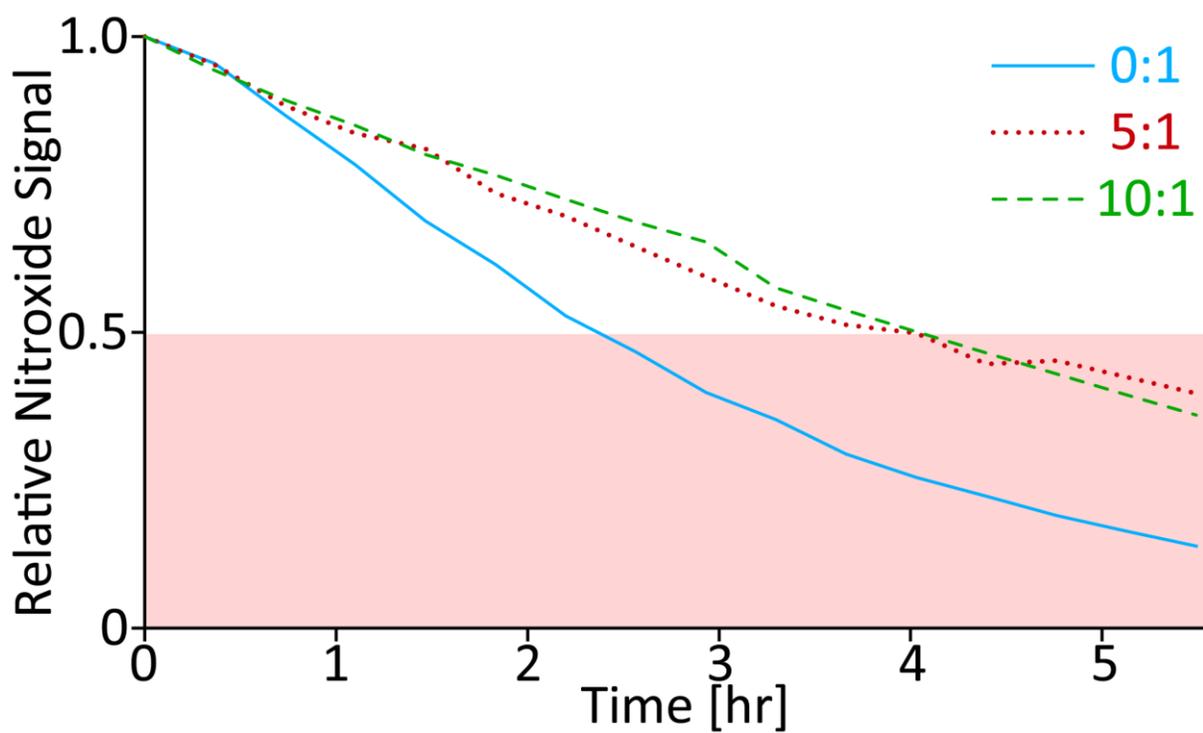
Figure 5-3. Time scan decay curves of 500 μM PCA in cytosol with differing equivalents of K$_3$Fe(CN)$_6$. The shaded region indicates the cut off of 50% nitroxide signal necessary for the DEER experiment. As consistent with 200 μM PCA, no more than 5 equivalents of K$_3$Fe(CN)$_6$ are required to achieve the maximum nitroxide radical lifetime.

experiment repeated with 200 μM PCA (Figure 2). Note, that the spectral shape remains constant over time as shown in Figure 5-4.

After the addition of 5 equivalents, no further decrease in decay was observed. The rate of nitroxide decay is decreased from $1.68 \pm 0.03$ μM/min to $1.17 \pm 0.12$ μM/min by the addition of 5 equivalents of $K_3Fe(CN)_6$. Interestingly, at 200 μM PCA the decay was minimized by 33% while at 500 μM PCA the decay was minimized by 30%. This indicates that 5 equivalents of $K_3Fe(CN)_6$ is sufficient for concentrations typically used in pulsed EPR spectroscopy at X-band ( <500 μM). The extended lifetime of a nitroxide radical might also be important for *in vivo* biomedical assays which require long nitroxide incubation times[243,247-250]. With the addition of $K_3Fe(CN)_6$, we have extended the feasible incubation time for DEER data by 50%. Note, that if nitroxide reduction was the only contribution to the loss in the DEER signal, then DEER should be feasible after two hours even without $K_3Fe(CN)_6$. Thus far, nitroxide based distance measurements have been made either by flash freezing immediately after microinjection or flash freezing after a more limited time of incubation[221]. Therefore, we must consider that nitroxide reduction is not the only contributing factor toward loss of DEER signal.

The B1 domain of the immunoglobulin-binding protein G (GB1) was doubly spin labeled at positions 15 and 28. This mutant will be called 15R1/28R1-GB1. 250 μM 15R1/28R1-GB1 with 5 equivalents of $K_3Fe(CN)_6$ was then introduced *in cell-extract* or microinjected *in-cell* for both CW and pulsed ESR spectroscopy. Microinjection of the 60 cells required for adequate sample size took approximately 10 minutes. We will call this time A. After time A, samples were either immediately frozen, or allowed to incubate at 15°C for either A+30 or A+60 minutes.
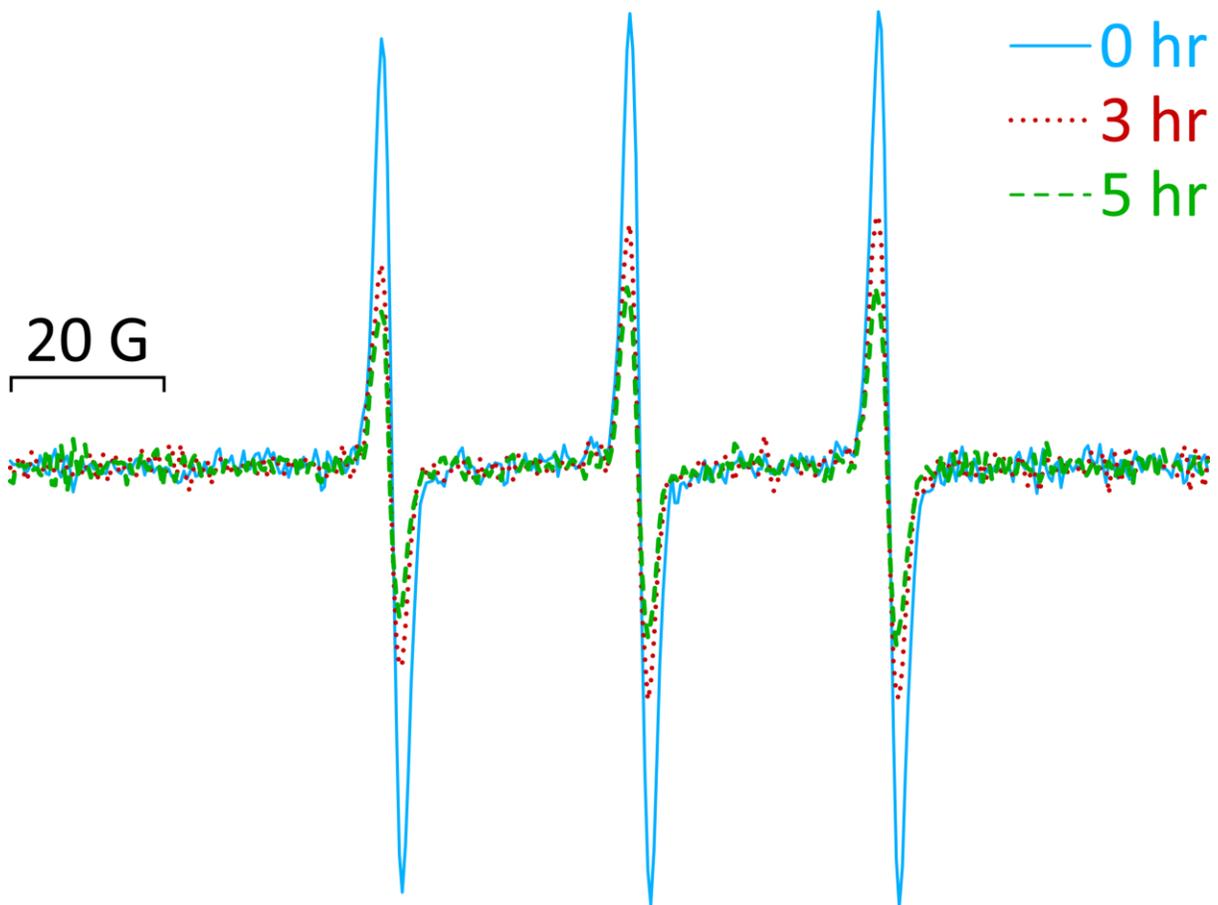
Figure 5-4. CW spectrum of 200 µM PCA in cytosol after 0 hours (solid), 3 hours (dotted) and 5 hours (dashed) incubation times showing consistent spectral line shape.

*Xenopus laevis* oocytes are used for microinjection. As opposed to other cell types, *Xenopus laevis* oocytes can be directly injected with nitroxide labeled proteins or moleculeswithout prejudice[224,251]. Compared with other cell lines, *Xenopus laevis* allow for long incubation periods[252] which makes them ideal for studies involving slow folding kinetics. Additionally, these "giant cells" allow for functional studies on proteins either expressed directly within or that have been microinjected into oocytes[253].

We next investigated if the decay rate of a doubly R1 labeled protein is consistent with the decay of the free nitroxide, PCA. The CW spectra of 15R1/28R1-GB1 at three different 15°C incubation times is shown in Figure 5-5. From analysis of the time scan of 500 μM PCA (S.I.), we can predict the expected percent nitroxide signal for a designated incubation time. The doubly integrated intensity of the CW spectrum, which is directly proportional to the number of active ESR species, shows a comparable rate of decay for 15R1/28R1-GB1 *in cell-extract* for the first A+30 minutes when compared to the expected results. By A+60 minutes, the 15R1/28R1-GB1 signal had decayed 8% faster than expected. However, the nitroxide signal decays much faster during incubation *in-cell*. After A+60 minutes incubation, the nitroxide radical decayed 21% faster than expected as compared to PCA reduction *in cell-extract*. The increased *in-cell* nitroxide reduction suggests the possibility of other cofactors, such as ascorbates[254-256] and glutathione[257,258], contributing to nitroxide reduction outside of those within the cytosol only.

DEER data was acquired for both *in-cell* and *in cell-extract* 15R1/28R1-GB1 samples at three different incubation times. Using the DEER data, first, we compared the rates of molecular diffusion using local concentrations. The *in vitro* DEER signal ($V(\tau)$), shown in Figure 5-6 (top), is the product of two signals stemming from the intramolecular dipolar interaction ($V_{intra}(\tau)$) and the numerous intermolecular dipolar interactions ($V_{inter}(\tau)$). $V_{inter}(\tau)$ is given as follows[50,206,259]:
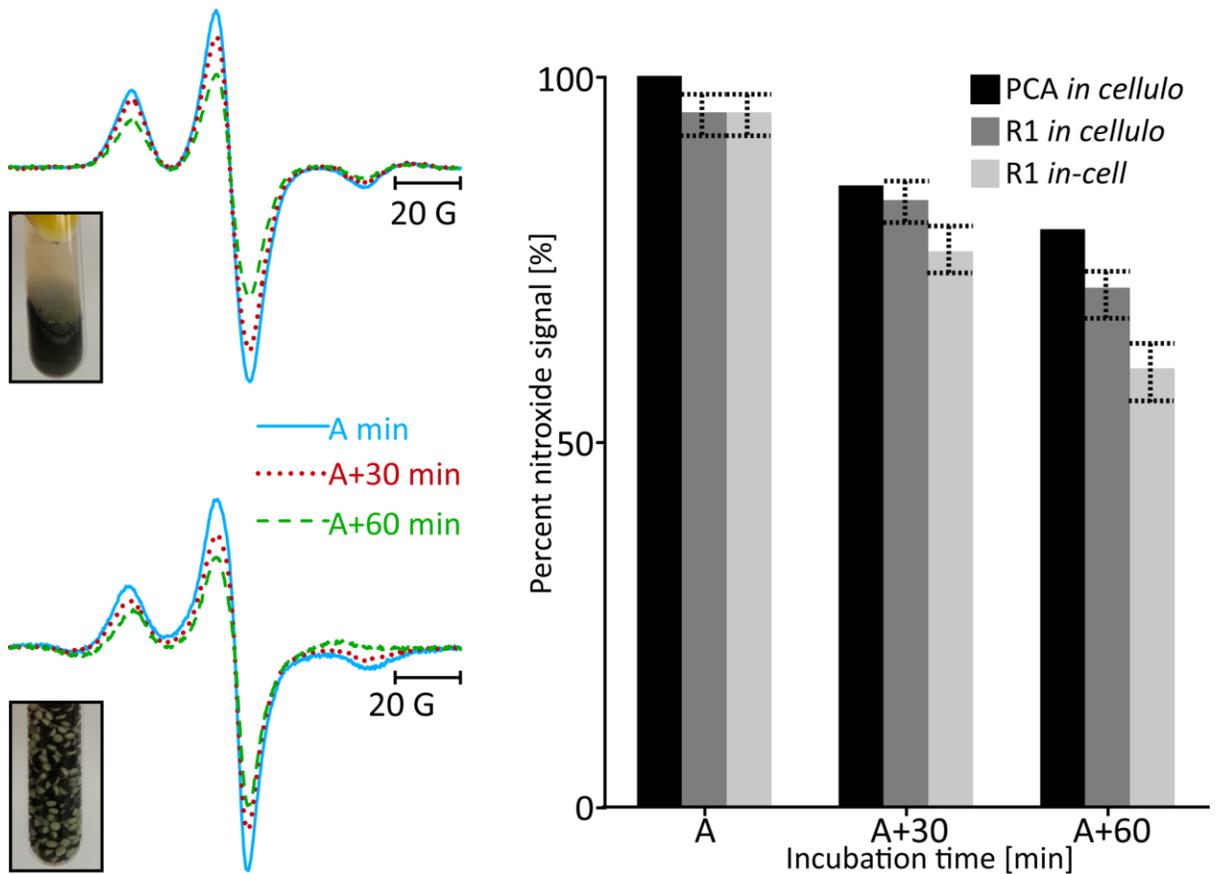
Figure 5-5. (Left) 80 K CW spectra of 15R1/28R1-GB1 *in cell-extract* (top) and *in-cell* (bottom) at three different 15°C incubation times; A (blue solid), A+30 (red dotted) and A+60 minutes (green dashed). (Right) Bar graph visualization of nitroxide decay from PCA (black), 15R1/28R1-GB1 *in cell-extract* (dark grey) and 15R1/28R1-GB1 *in-cell* (light grey). Dotted lines represent error. The error for PCA is less than 1%.
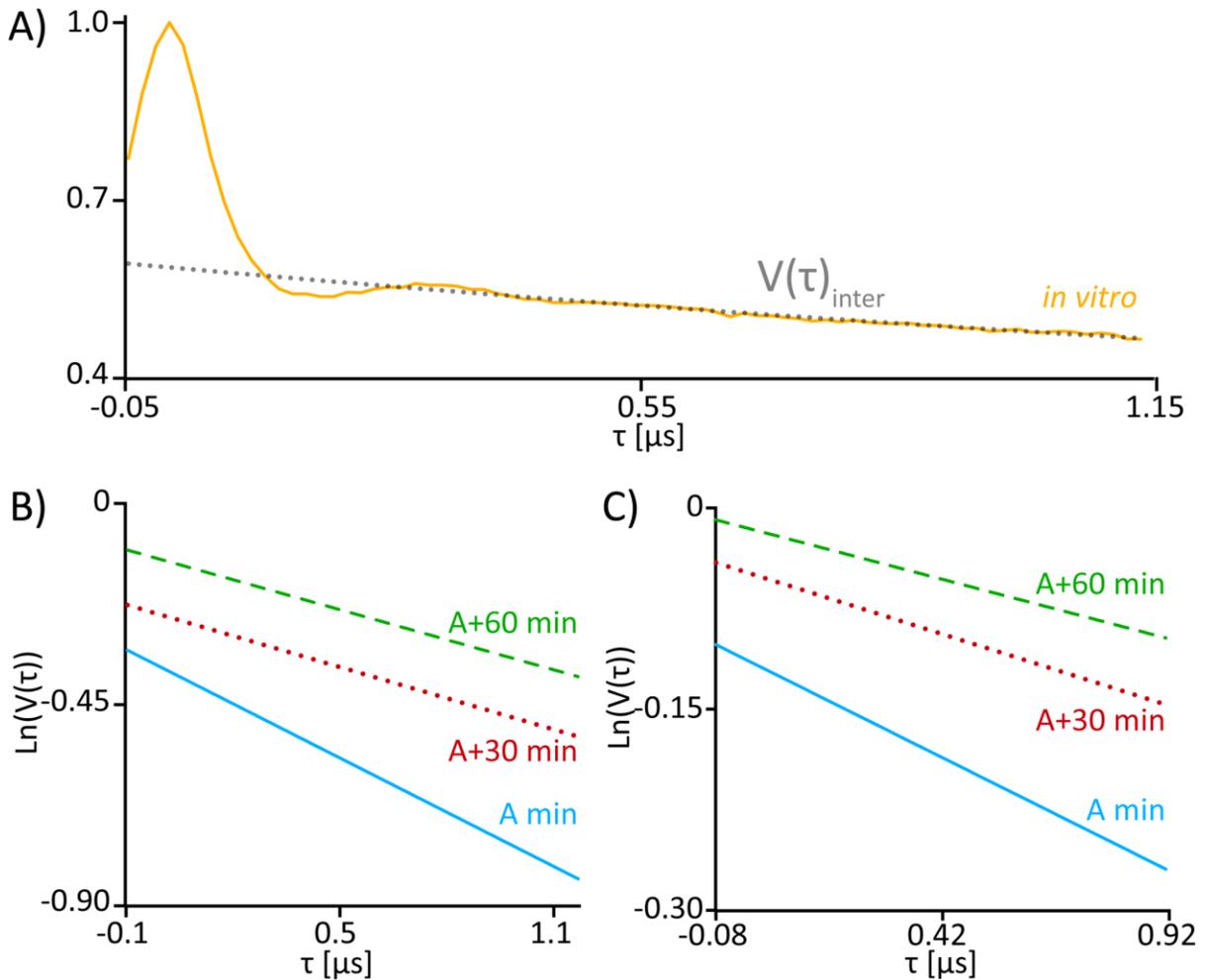
Figure 5-6. A) 15R1/28R1-GB1 *in vitro* raw DEER time domain illustrating contributions from the intermolecular signal, $V_{inter}(\tau)$, and the intramolecular signal, $V_{intra}(\tau)$. $V_{inter}(\tau)$ consists of an exponential decay shown by the grey dotted line. $V_{inter}(\tau)$ is dependent upon the local concentration which is indicative of the overall nitroxide signal. B) Logarithmic scale of the DEER background decay signal *in cell-extract*. C) Logarithmic scale of the DEER background decay signal *in-cell* (B).

$$V_{inter}(\tau) = \exp(-kt) \tag{4}$$

with the decay rate, k, as [51]:

$$k^{-1} = 1.0027 \frac{10^{-3}}{pC} \tag{5}$$

where p is the probability of excitation of the B spins and C is the molar concentration. Fitting of the background decay results in the extraction of local concentration of the sample. This exponential decay is shown as the black dotted line in Figure 5-6(A).

Figure 5-7 shows $V_{inter}(\tau)$ of all varying incubation times both *in cell-extract* (B) and *in-cell* (C) on a logarithmic scale where the negative slope is proportional to the local concentration. The change in local concentration can indicate whether the protein has equilibrated within the sample. Local concentrations of the *in cell-extract* samples are $1.13 \pm 0.03$ mM, $0.65 \pm 0.06$ mM and $0.63 \pm 0.08$ mM for A minutes, A+30 minutes and A+60 minutes of incubation respectively. The local concentrations of the in-cell samples were as follows: $0.47 \pm 0.06$ mM, $0.29 \pm 0.08$ mM and $0.26 \pm 0.13$ mM for A minutes, A+30 minutes and A+60 minutes of incubation respectively Both *in cell-extract* and *in-cell* samples showed no further decrease in local concentration after A+30 minutes of incubation. The constant slope suggests that A+30 minutes is ample time for molecular diffusion of a small globular protein both prepared *in cell-extract* and microinjected *in-cell*. Note that dipolar modulations in the DEER experiment are still clearly visable after A+30 minutes incubation for both sample environments (cf. below).

We next used the background subtracted time domain DEER data in order to ascertain whether R1 bond cleavage is a contributor to the decrease DEER signal. The modulation depth
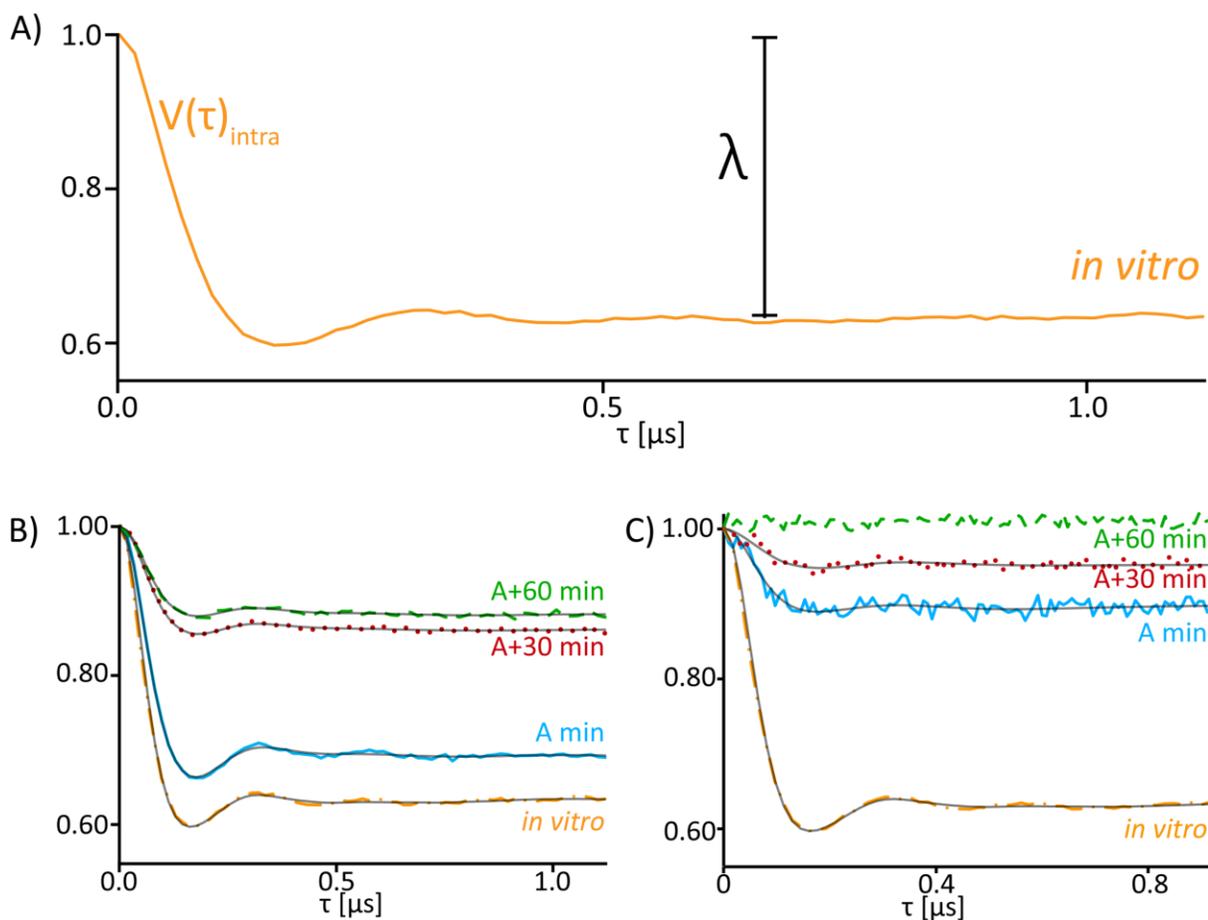
Figure 5-7. A) Background subtracted time domain DEER data, $V_{intra}(\tau)$, of *in vitro* 15R1/28R1-GB1 illustrating the modulation depth parameter ($\lambda$). B) Background subtracted time domain of 15R1/28R1-GB1 with *in cell-extract* (B) and *in-cell* (C) incubation times of A minutes (blue solid), A+30 minutes (red dotted) and A+60 minutes (green dashed) as compared to *in vitro* (yellow dashed/dotted). Fits via Tikhonov regularization shown as gray line.

157

($\lambda$), described by *in vitro* 15R1/28R1-GB1 data shown in Figure 5-7 (A), is indicative of how many spins are present per system [54]. Analyzing the modulation depths of the *in cell-extract* and *in-cell* time domain signals, we see that both show decreased modulation depth over time, as clearly evident in Figure 5-7 (B,C). These modulations depths are significantly lower than expected if only nitroxide reduction occurs. A visual representation of the expected versus actual modulation *in cell-extract* is seen in Figure 5-8 (left). The expected modulation depths were calculated using the percent nitroxide signal present from the CW experiments (see Materials and Methods.). The discrepancy between the modulation depth loss compared to the total nitroxide signal loss over time suggests that the nitroxide reduction is not the only contribution to the loss of DEER signal. Cleavage of the R1 side chain by reduction of the disulfide bond[221] is a possible explanation for the difference in signal between CW intensity and DEER modulation depth. If the R1 sidechain is cleaved while the nitroxide radical is not reduced, the CW signal would not change while the modulation depth would decrease. Even with such rapid loss of modulation depth, the DEER time domain signal still has visible dipolar modulations even after A+60 minutes *in cell-extract*.

Furthermore, the modulation depth decreases at an even faster rate upon microinjection *in-cell*. The *in-cell* background subtracted time domain signals are shown in Figure 5-7 (C). *in-cell* DEER results indicate that the disulfide bond is cleaved at a much faster rate than *in cell-extract*. No visible dipolar modulations were present after A+60 minutes of incubation, thus we can assume the number of attached and intact nitroxides has fallen below 50%. This discrepancy suggests that *in cell-extract* is not a sufficient representation of the true *in-cell* environment with respect to disulfide reduction. The faster reduction and cleavage of the disulfide bond can likely be attributed to the presence of organelle and membrane associated factors such as glutathione-
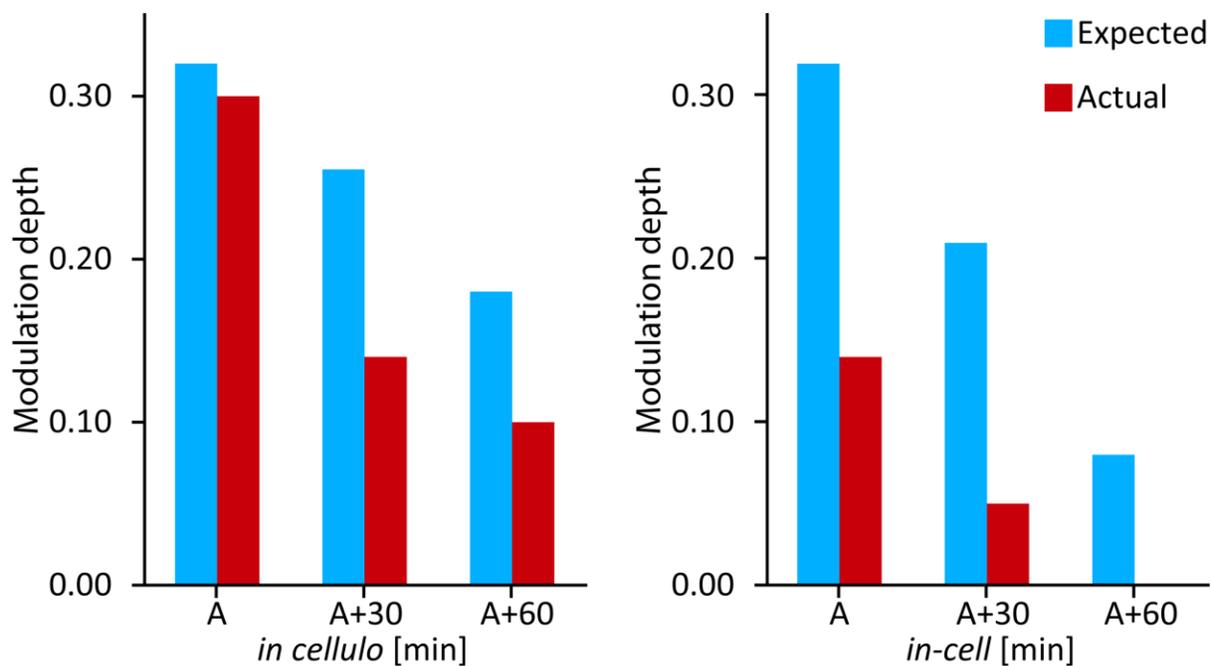
Figure 5-8. Bar graph visual representation of both *in cell-extract* (left) and *in-cell* (right) of the expected (blue) vs actual (red) modulation depths at three different incubation times. The expected modulation depth is derived from the CW spectra which only takes nitroxide reduction into account. The discrepancy between the expected and actual is evidence for cleavage of the R1 disulfide bond.

reductase[260], disulfide-isomerase[261] and thioredoxin[262]. These enzymes are powerful reducers that specialize in disulfide reduction of proteins *in-cell*. Organelles, membranes and most other organelle and membrane associated factors are separated and pelleted out of the cytosol along with other cellular debris during preparation of the cytosol.

Finally, the distance distribution did not change between *in vitro*, *in cell-extract* and *in-cell* DEER experiments as shown in Figure 5-9. All time domains were background subtracted and analyzed via Tikhonov regularization. The narrow distance distributions were achievable due to the ability to resolve clear dipolar modulations *in cell-extract*, for 60 minutes, and *in-cell*, for up to 30 minutes. Figure 5-10 shows the background subtracted time domain DEER data of 15R1/28R1-GB1 *in-cell* after incubation times at 15°C of A and A+30 to better demonstrate the clearly visible dipolar modulations.

## 5.4    CONCLUSION

Herein we have shown that $K_3Fe(CN)_6$ is capable of extending the feasible incubation times for DEER experiments by up to 50%. Using this extended nitroxide radical lifetime, first, using analysis of local concentrations, we show that 30 minutes is enough time to allow for molecular diffusion for a small globular protein such as GB1. Second, we have shown that the cleavage of the R1 sidechain contributes to the loss of DEER signal both *in-cell* and *in cell-extract*. This cleavage is pertinent to the expansion of ESR *in-cell* because many modern ESR probes still use disulfide linkages to covalently attach to biological macromolecules. DEER and CW results suggest that both nitroxide redution and R1 cleavage occur at a faster rate *in-cell* than
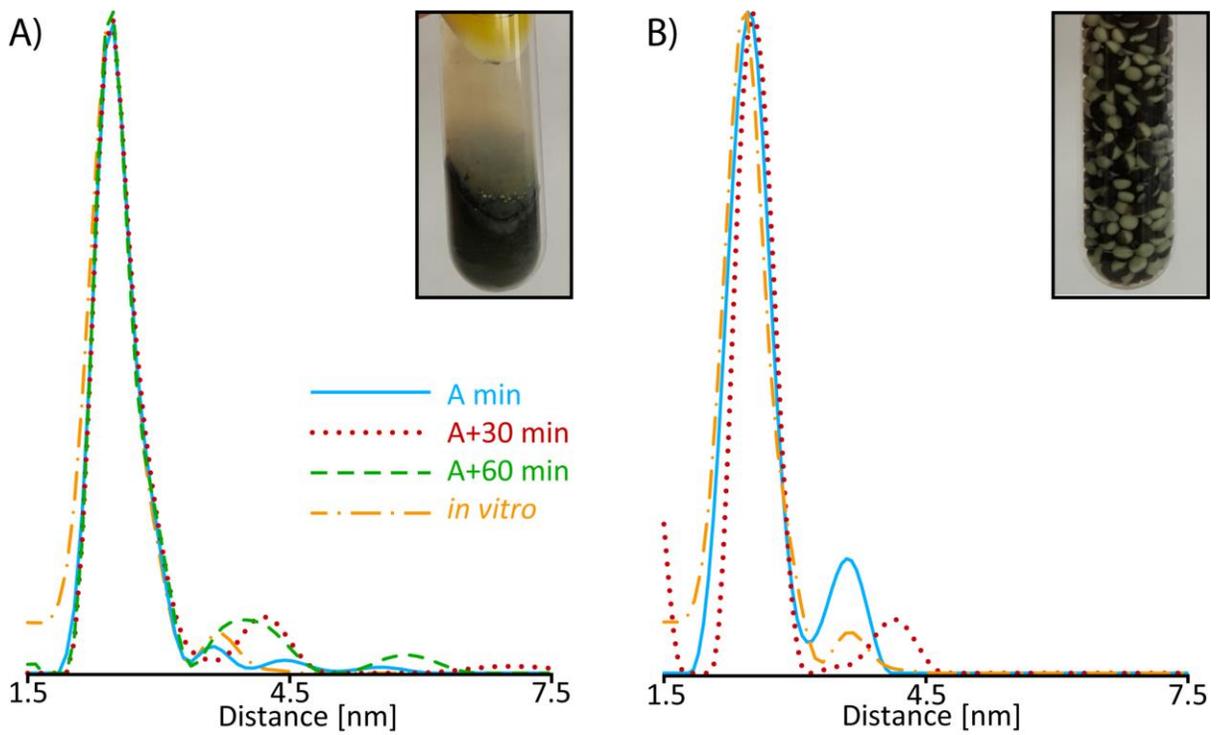
Figure 5-9. DEER distance distribution showing the same most probable distance as well as the same distance distribution width between *in cell-extract* (A) and *in-cell* (B) for three different incubation times: A (blue solid), A+30 minutes (red dotted) and A+60 minutes (green dashed) as well as *in vitro* (yellow dashed/dotted).
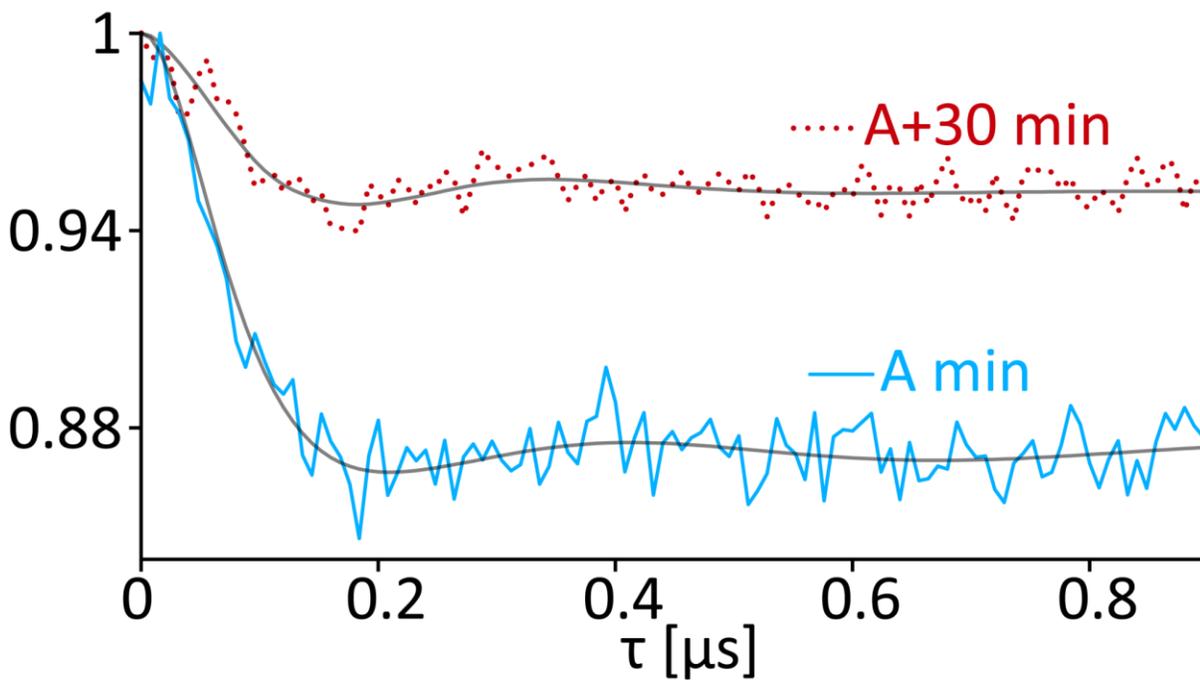
Figure 5-10. Background subtracted time domain DEER data of 15R1/28R1-GB1 *in-cell* after incubation times of A (blue solid) and A+30 (red dotted). Fits via Tikhonov regularization are shown as grey solid lines.

*in cell-extract*. Furthermore, 15R1/28R1-GB1 did not experience any changes in structure, at least at this site, among *in vitro*, *in cell-extract* and *in-cell*. The ability to resolve such distance distrubitons illustrates the ability of *in-cell* nitroxide DEER to produce clear dipolar modulations.

## 5.5    ACKNOWLEDGEMENTS

# 6.0    SUMMARY OF MAJOR ACHIEVMENTS


The work presented here is stark progress in the field of site-directed spin labeling in biological macromolecules using the $Cu^{2+}$ metal ion. This work presents the first nucleotide and structure independent labeling strategy to incorporate $Cu^{2+}$ within the DNA duplex. The so called DPA-DNA labeling method positions the spin label within the interior of the DNA duplex minimizing ambiguity in distance measurements. Furthermore, this technique reports on DNA backbone-backbone distances without the need of modeling. $Cu^{2+}$ labeling in DNA presents a new way to incorporate spin labels within nucleic acids and allows for orthogonal labeling with nitroxide labeled protein. The combination can lead to multiple distance constraints generated from the same sample.

$Cu^{2+}$ labeling in proteins has proven to report concise distance distributions using the double histidine motif. In this work, we measure, by through spectroscopic analysis, optimum conditions to increase the ESR signal from this labeling strategy. Using the optimized protocol, we show a two-fold increase in ESR signal allowing for a minimization of time required for data acquisition.

This work also presents the first example of the utility of the double histidine motif in answering questions about protein structure and function. Using a combination of nitroxide- and $Cu^{2+}$-labeling, we investigate the functionally important terminal helix of the human glutathione S-transferase enzyme. We show that the terminal helix exists in two distinct conformations, in agreement with previous NMR and fluorescence data. The most probable DEER distances are

164

then used to generate a full length structure of the protein for the first time. This model structure is validated by further $Cu^{2+}$-based distance constraints.

The efforts of my work in the University of Pittsburgh have created the following publications:

Harris, R. D., Auletta, J. T., Mohaghegh S. A., **Lawless, M. J.**, Perri, N. M., Saxena, S., Weiland, L. M., Waldeck, D. H., Clark, W. W., Meyer, T. Y. "Chemical and electrochemical manipulation of mechanical properties in stimuli-responsive copper-cross-linked hydrogels." *ACS Macro Lett.* **2013**, 2, 1095-1099.

Kinde, M. N., Chen, Q., **Lawless, M. J.**, Mowrey, D. D., Xu, J. Saxena, S., Xu, Y., Tang, P. "Conformational changes underlying desensitization of the pentameric ligand-gated ion channel ELIC." *Structure*, **2015**, 23, 995-1004.

De Santis, E., Minicozzi, V., Proux, O., Rossi, G., **Lawless, M. J.**, Silva, K. I., Stellato, F., Saxena, S., Morante, S. "Cu(II)-Zn(II) cross-modulation in amyloid-beta peptide binding: an X-absorption spectroscopy study." *J. Phys. Chem. B.* **2015**, 119, 15813-15820.

**Lawless, M. J.**, Sarver, J. L., Saxena, S. "Nucleotide-independent copper(II)-based distance measurements in DNA by pulsed ESR spectroscopy." *Angew. Chem. Int. Ed.* **2017**, 56, 2115-2117.

**Lawless, M. J.**, Shimshi, A., Cunningham, T. F., Kinde, M. N., Tang, P., Saxena, S. "An analysis of nitroxide based distance measurements in cell extract and in-cell by pulsed ESR spectroscopy." *ChemPhysChem*, **2017**, 18, 1653-1660.

**Lawless, M. J.**, Ghosh, S., Cunningham, T. F., Shimshi, A., Saxena, S., "On the use of the Cu(II)-iminodiacetic acid complex for double histidine based distance measurements by pulsed ESR." *Phys. Chem. Chem. Phys.*, **2017**, 19, 20959-20967.

Ghosh, S., **Lawless, M. J.**, Rule, G. S., Saxena, S., "The $Cu^{2+}$-nitrilotriacetic acid complex improves loading of a-helical double histidine site for precise distance measurements by pulsed ESR." *J. Magn. Reson.* **2018**, 286, 163-171.

**Lawless, M. J.**, Pettersson, J. R., Rule, G. S., Lanni, F., Saxena, S., "ESR shows that the c-terminus of human glutathione S-transferase A1-1 exists in two distinct helical conformations in the ligand-free state." *Biophys. J.* **2018**, accepted.

# BIBLIOGRAPHY

(1)     Yang, Z.; Ji, M.; Saxena, S. *Appl. Magn. Reson.* **2010**, *39*, 487.

(2)     Cunningham, T. F.; Pornsuwan, S.; Horne, W. S.; Saxena, S. *Prot. Sci.* **2016**, *25*, 1049.

(3)     Borbat, P. P.; Freed, J. H. In *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centers in the Biosciences. Structure and Bonding.*; Timmel, J. H. a. C., Ed.; Springer: New York, USA, 2012, p 1.

(4)     Cafiso, D. S. *Acc. Chem. Res* **2014**, *47*, 3102.

(5)     Hubbell, W.; Lopez, C.; Altenbach, C.; Yang, Z. *Curr. Opin. Struct. Biol.* **2013**, *23*, 725.

(6)     Jeschke, G. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419.

(7)     Jeschke, G.; Y., P. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895.

(8)     Ji, M.; Ruthstein, S.; Saxena, S. *Acc. Chem. Res.* **2014**, *47*, 688.

(9)     Reginsson, G. W.; Schiemann, O. *Biochem. J.* **2011**, *434*, 353.

(10)    Borbat, P.; Freed, J. H. In *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centres in the Biosciences. Structure and Bonding*; Springer: Berlin, 2013; Vol. 152.

(11)    Columbus, L.; Kalai, T.; Jeko, J.; Hideg, K.; Hubbell, W. *Biochemistry* **2001**, *40*, 3828.

(12)    Lawless, M. J.; Sarver, J.; Saxena, S. *Angew. Chem. Int. Ed.* **2017**, *56*, 2115.

(13)    Lawless, M. J.; Ghosh, S.; Cunningham, T. F.; Shimshi, A.; Saxena, S. *Phys. Chem. Chem. Phys.* **2017**, *19*, 20959.

(14)    Mims, W. B. *Phys. Rev. B.* **1972**, *5*, 2409.

(15)    Peisach, J.; Blumberg, W. E. *Arch. Biochem. Biophys.* **1974**, *195*, 691.

(16)    Weil, J. A.; Bolton, J. R. *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*; 2 ed.; John Wiley & Sons, Inc.: Hoboken, New Jersey, 2007.

(17)    Savelieff, M. G.; Wilson, T. D.; Elias, Y.; Nilges, M. J.; Garner, D. K.; Lu, Y. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7919.

(18)    In *Biomolecular EPR Spectroscopy*; Hagan, W. R., Ed.; CRC Press: Boca Raton, Florida, 2009, p 67.

(19)    Slichter, C. P. In *Principles of Magnetic Resonance*; Fulde, P., Ed.; Springer-Verlag: 1990, p 503.

(20)    Aasa, R.; Vanngard, T. *J. Magn. Res.* **1975**, *19*, 308.

(21)    Syme, C. D.; Nadal, R. C.; Rigby, S. E. J.; Viles, J. H. *J. Biol. Chem.* **2004**, *279*, 18169.

(22)    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.

(23)    Ghosh, S.; Lawless, M. J.; Rule, G. S.; Saxena, S. *J. Magn. Res.* **2017**, *10.1016/j.jmr.2017.12.005*.

(24)    Ruthstein, S.; Stone, K. M.; Cunningham, T. F.; Ji, M.; Cascio, M.; Saxena, S. *Biophys.* **2010**, *99*, 2497.

(25)    Shin, B. K.; Saxena, S. *J. Phys. Chem. A* **2011**, *115*, 9590.

(26)     Silva, K. I.; Saxena, S. *J. Phys. Chem. B.* **2013**, *117*, 9386.

(27)     Mims, W. B.; Peisach, J. *J. Chem. Phys.* **1978**, *69*, 4921.

(28)     Mims, W. B.; Davis, J. L. *J. Chem. Phys.* **1976**, *64*, 4836.

(29)     Kevan, L.; Bowman, M. K.; Narayana, P. A.; Boeckman, R. K.; Yudanov, V. F.; Tsvetkov, Y. D. *J. Chem. Phys.* **1975**, *63*, 409.

(30)     Stoll, S.; Calle, C.; Mitrikas, G.; Schweiger, A. *J. Magn. Res.* **2005**, *177*, 93.

(31)     Dikanov, S. A.; Tsvetkov, Y. D.; Bowman, M. K.; Astashkin, A. V. *Chem. Phys. Lett.* **1982**, *90*, 149.

(32)     McCracken, J.; Peisach, J.; Dooley, D. M. *J. Am. Chem. Soc.* **1987**, *109*, 4064.

(33)     Burns, C. S.; Aronoff-Spencer, E.; Dunham, C. M.; Lario, P.; Avdievich, N. I.; Antholine, W. E.; Olmstead, M. M.; Vrielink, A.; Gerfen, G. J.; Peisach, J.; Scott, W. G.; Millhauser, G. L. *Biochemistry* **2002**, *41*, 3991.

(34)     Lee, H.-I.; Doan, P. E.; Hoffman, B. M. *J. Magn. Res.* **1999**, *140*, 91.

(35)     Flanagan, H. L.; Singel, D. J. *J. Chem. Phys.* **1987**, *87*, 5606.

(36)     Deligiannakis, Y.; Louloudib, M.; Hadjiliadis, N. *Coord. Chem. Rev* **2000**, *201*, 1.

(37)     Huffman, D. L.; Huyett, J.; Outten, F. W.; Doan, P. E.; Finney, L. A.; Hoffman, B. M.; O'Halloran, T. V. *Biochemistry* **2002**, *41*, 10046.

(38)     Jiang, F.; McCracken, J.; Peisach, J. *J. Am. Chem. Soc.* **1990**, *112*, 9035.

(39)     McCracken, J.; Pember, S.; Benkovic, S. J.; Villafranca, J. J.; Miller, R. J.; Peisach, J. *J. Am. Chem. Soc.* **1988**, *110*, 1069.

(40)     Silva, K. I.; Michael, B. C.; Geib, S. J.; Saxena, S. *J. Phys. Chem. B* **2014**, *118*, 8935.

(41)     Hernandez-Guzman, J.; Sun, L.; Mehta, A. K.; Dong, J.; Lynn, D. G.; Warncke, K. *ChemBioChem* **2013**, *14*, 1762.

(42)     Kosman, D. J.; Peisach, J.; Mims, W. B. *Biochemistry* **1980**, *19*, 1304.

(43)     McCracken, J.; Desai, P. R.; Papadopoulos, N. J.; Villafranca, J. J.; Peisach, J. *Biochemistry* **1988**, *110*, 4133.

(44)     Goldfarb, D.; Fauth, J. M.; Tor, Y.; Shanzer, A. *J. Am. Chem. Soc.* **1991**, *113*, 1941.

(45)     Rao, G.; Bansal, S.; Law, W. X.; O'Dowd, B.; Dikanov, S. A.; Oldfield, E. *Biochemistry* **2017**, *56*, 3770.

(46)     Carmieli, R.; Papo, N.; Zimmermann, H.; Potapov, A.; Shai, Y.; Goldfarb, D. *Biophys. J.* **2006**, *90*, 492.

(47)     Milov, A. D.; Maryasov, A. G.; Tsvetkov, Y. D. *Appl. Magn. Reson.* **1998**, *15*, 107.

(48)     Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiees, H. W. *J. Magn. Reson.* **2000**, *142*, 331.

(49)     Larsen, R. G.; Singel, D. J. *J. Chem. Phys.* **1993**, *98*, 5134.

(50)     Raitsimring, A. M.; Salikhov, R. M.; Umanskii, B. A.; Tsvetkov, Y. D. *Sov. Phys. Sol. State* **1974**, *16*, 493.

(51)     Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H. *J. Magn. Res.* **2012**, *216*, 69.

(52)     Lawless, M. J.; Shimshi, A.; Cunningham, T. F.; Kinde, M. N.; Tang, P.; Saxena, S. *ChemPhysChem* **2017**, *18*, 1653.

(53)     Martorana, A.; Bellapadrona, G.; Feintuch, A.; Di Gregorio, E.; Aime, S.; Goldfarb, D. *J. Am. Chem. Soc.* **2014**, *136*, 13458.

(54)     Milov, A. D.; Ponomarev, A. B.; Tsvetkov, Y. D. *Chem. Phys. Lett.* **1984**, *110*, 67.

(55)     Goldfarb, D. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9685.

(56)     Bode, B. E.; Margraf, D.; Plackmeyer, J.; Durner, G.; Prisner, T. F.; Schiemann, O. *J. Am. Chem. Soc.* **2007**, *129*, 6736.

(57)     Tsvetkov, Y. D.; Milov, A. D.; Maryasov, A. G. *Russ. Chem. Rev.* **2008**, *77*, 487.

(58)     Cunningham, T. F.; Putterman, M. R.; Desai, A.; Horne, W. S.; Saxena, S. *Angew. Chem. Intl. Ed.* **2015**, *54*, 6330.

(59)     Saxena, S.; Freed, J. H. *Chem. Phys. Lett.* **1996**, *262*, 17.

(60)     Martin, R. E.; Pannier, M.; Diederich, F.; Gramlich, V.; Hubrich, M.; Spiess, H. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 2833.

(61)     Borbat, P. P.; Freed, J. H. *Chem. Phys. Lett.* **1999**, *313*, 145.

(62)     Schiemann, O.; Prisner, T. F. *Q. Rev. Biophys.* **2007**, *40*, 1.

(63)     Fanucci, G. E.; Cafiso, D. S. *Cur. Opin. Struct. Biol.* **2006**, *16*, 644.

(64)     Becker, J. S.; Saxena, S. *Chem. Phys. Lett.* **2005**, *414*, 248.

(65)     Yang, Z.; Becker, J.; Saxena, S. *J. Magn. Reson.* **2007**, *188*, 337.

(66)     Yang, Z.; Kise, D.; Saxena, S. *J. Phys. Chem. B* **2010**, *114*, 6165.

(67)     Ruthstein, S.; Ji, M.; Mehta, P.; Jen-Jacobson, L.; Saxena, S. *J. Phys. Chem. B.* **2013**, *117*, 6227.

(68)     Yang, Z.; Ji, M.; Cunningham, T. F.; Saxena, S. *Methods. Enzymol.* **2015**, *563*, 459.

(69)    van Amsterdam, I. M.; Ubbink, M.; Canters, G. W.; Huber, M. *Angew. Chem. Int. Ed.* **2003**, *42*, 62.

(70)    Abdullin, D.; Florin, N.; Hagelueken, G.; Schiemann, O. *Angew. Chem. Int. Ed.* **2015**, *54*, 1827.

(71)    Bowen, A. M.; Jones, M. W.; Lovett, J. E.; Gaule, T. G.; McPherson, M. J.; Dilworth, J. R.; Timmel, C. R.; R., H. J. *Phys. Chem. Chem. Phys.* **2016**, *18*, 5981.

(72)    Evans, E. G. B.; Pushie, J. M.; Markham, K. A.; Lee, H.-W.; Millhauser, G. L. *Structure* **2017**, *24*, 1057.

(73)    Merz, G. E.; Borbat, P. P.; Pratt, A. J.; Getzoff, E. D.; Freed, J. H.; Crane, B. R. *Biophys. J.* **2014**, *107*, 1669.

(74)    van Wonderen, J. H.; Kostrz, D. N.; Dennison, C.; MacMillan, F. *Angew. Chem. Int. Ed.* **2013**, *52*, 1990.

(75)    Yang, Z.; Kurpiewski, M. R.; Ji, M.; Townsend, J. E.; Mehta, P.; Jen-Jacobson, L.; Saxena, S. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E993.

(76)    Kay, C. W. M.; El Mkami, H.; Cammack, R.; Evans, R. W. *J. Am. Chem. Soc.* **2007**, *129*, 4868.

(77)    Ji, M.; Tan, L.; Jen-Jacobson, L.; Saxena, S. *Mol. Phys.* **2014**, *112*, 3173.

(78)    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*, 2839.

(79)    Marko, A.; Denysenkov, V.; Margraf, D.; Cekan, P.; Schiemann, O.; Sigurdsson, S. T.; Prisner, T. F. *J. Am. Chem. Soc.* **2011**, *133*, 13375.

(80)    Schiemann, O.; Piton, N.; Y., M.; Stock, G.; W., E. J.; Prisner, T. F. *J. Am. Chem. Soc.* **2003**, *126*, 5722.

(81)     Sicoli, G.; Wachowius, F.; Bennati, M.; Hobartner, C. *Angew. Chem. Int. Ed.* **2010**, *49*, 6443.

(82)     Ward, R.; Schiemann, O. *Struct. Bond.* **2014**, *152*, 249.

(83)     Endeward, B.; Marko, A.; Denysenkov, V.; Sigurdsson, S. T.; Prisner, T. F. *Methods. Enzymol.* **2015**, *564*, 403.

(84)     Grytz, C. M.; Marko, A.; Cekan, P.; Sigurdsson, S. T.; Prisner, T. F. *Phys. Chem. Chem. Phys.* **2016**, *18*, 2993.

(85)     Reginsson, G. W.; Shelke, S. A.; C., R.; White, M.; Sigurdsson, S. T.; Schiemann, O. *Nucleic Acids Res.* **2013**, *41*, e11.

(86)     Reyes, C. V.; Tangprasertchai, N. S.; Yogesha, S. D.; Nguyan, R. H.; Zhang, X.; Rajan, R.; Qin, P. Z. *Cell Biochem. Biophys.* **2016**, DOI: 10.1007/s12013.

(87)     Stone, K. M.; Townsend, J. E.; Sarver, J.; Sapienza, P. J.; Saxena, S.; Jen-Jacobson, L. *Angew. Chem. Int. Ed.* **2008**, *47*, 10192.

(88)     Erlenbach, N.; Endeward, B.; Schops, P.; Gophane, D. B.; Sigurdsson, S. T.; Prisner, T. F. *Phys. Chem. Chem. Phys.* **2016**, *18*, 16196.

(89)     Frolow, O.; Endeward, B.; Schiemann, O.; Prisner, T. F.; W., E. J. *Nucl. Acids Symp. Ser.* **2008**, *52*, 153.

(90)     Gophane, D. B.; Endeward, B.; Prisner, T. F.; Sigurdsson, S. T. *Chem. Eur. J.* **2014**, *20*, 15913.

(91)     Gophane, D. B.; Sigurdsson, S. T. *Chem. Comm.* **2013**, *49*, 999.

(92)     Kerzhner, M.; Abdullin, D.; Wiecek, J.; Matsuoka, H.; Hagelueken, G.; Schiemann, O.; Famulok, M. *Chem. Eur. J.* **2016**, *22*, 12113.

(93)     Saha, S.; Jagtap, A. P.; Sigurdsson, S. T. *Chem. Comm.* **2015**, *51*, 13142.

(94)    Shelke, S. A.; Sigurdsson, S. T. *Angew. Chem. Int. Ed.* **2010**, *49*, 7984.

(95)    Nguyan, P. H.; Popova, A.; Hideg, K.; Qin, P. Z. *BMC Biophys.* **2016**, *8*, 6.

(96)    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. Protocols* **2007**, *2*, 2354.

(97)    Tangprasertchai, N. S.; Zhang, X.; Ding, Y.; Tham, K.; Rohs, R.; Haworth, I. S.; Qin, P. Z. *Methods. Enzymol.* **2015**, *564*, 427.

(98)    Fedin, M. V.; Shevelev, G. Y.; Pyshnyi, D. V.; Tormyshev, V. M.; Jeschke, G.; Yulikov, M.; Bagryanskaya, E. G. *Phys. Chem. Chem. Phys.* **2016**.

(99)    Wojciechowski, F.; Gross, A.; Holder, I. T.; Knorr, L.; Drescher, M.; Hartig, J. S. *Chem. Comm.* **2015**, *51*, 13850.

(100)   Donohue, M. P.; Szalai, V. A. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15447.

(101)   Atwell, S.; Meggers, E.; Spraggon, G.; Schultz, P. G. *J. Am. Chem. Soc.* **2001**, *123*, 12364.

(102)   Meggers, E.; Holland, P. L.; Tolman, W. B.; Romesberg, F. E.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 10714.

(103)   Tor, Y.; Weizman, H. *J. Am. Chem. Soc.* **2001**, *123*, 3375.

(104)   Clever, G. H.; Shionoya, M. *Coord. Chem. Rev* **2010**, *254*, 2391.

(105)   Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J. *Appl. Magn. Reson.* **2006**, *30*, 473.

(106)   Miyahara, T.; Nakatsuji, H.; Sugiuama, H. *J Phys Chem A* **2013**, *117*, 42.

(107)   Taboury, J. A.; Adam, S.; Taillandier, E.; Neumann, J.-M.; Tran-Dinh, S.; Huynh-Dinh, T.; Langlois d'Estaintot, B.; Conti, M.; Igolen, J. *Nucl. Acids Res.* **1984**, *12*, 6291.

(108)   Kypr, J.; Kejnovska, I.; Renciuk, D.; Vorlickova, M. *Nucl. Acids Res.* **2009**, *37*, 1713.

(109)   Freier, S. M.; Kierzek, R.; Jaeger, J. A.; Sugimoto, N.; Caruthers, M. H.; Neilson, T.; Turner, D. H. *Proc. Natl. Acad. Sci.* **1986**, *83*, 9373.

(110)   Theophanides, T.; Bariyanga, J. *J. Mol. Struct.* **1989**, *214*, 177.

(111)   Matthew-Fenn, R. S.; Das, R.; Harbury, P. A. *Science* **2008**, *322*, 446.

(112)   Matthew-Fenn, R. S.; Das, R.; Fenn, T. D.; Schneiders, M.; Harbury, P. A. *Science* **2009**, *325*, 538.

(113)   Garbuio, L.; Bordignon, E.; Brooks, E. K.; Hubbell, W.; Jeschke, G.; Yulikov, M. *J. Phys. Chem. B* **2013**, *117*, 3145.

(114)   Joseph, B.; Tormyshev, V. M.; Rogozhnikova, O. Y.; Akhmetzanov, D.; Bagryanskaya, E. G.; Prisner, T. F. *Angew. Chem. Int. Ed.* **2016**, *55*, 11538.

(115)   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.

(116)   Sarver, J. L.; Silva, K. I.; Saxena, S. *Appl. Magn. Reson.* **2013**, *44*, 583.

(117)   Abdelkader, E. H.; Feintuch, A.; Yao, X.; Adams, L. A.; Aurelio, L.; Graham, B.; Goldfarb, D.; Otting, G. *Chem. Commun.* **2015**, *51*, 15898.

(118)   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.

(119)   Barthelmes, D.; Granz, M.; Barthelmes, K.; Allen, K. N.; Imperiali, B.; Prisner, T. F.; Schwalbe, H. *J. Biol. NMR* **2015**, *63*, 275.

(120)   Doll, A.; Qi, M.; Wili, N.; Pribitzer, S.; Godt, A.; Jeschke, G. *J. Magn. Res.* **2015**, *259*, 153.

(121)    Mascali, F. C.; Ching, H. Y. V.; Rasia, R. M.; Un, S.; Tabares, L. *Angew. Chem.* **2016**, *128*, 11207.

(122)    Matalon, E.; T., H.; Hagelueken, G.; Graham, B.; Frydman, V.; Feintuch, A.; Otting, G.; Goldfarb, D. *Angew. Chem. Int. Ed.* **2013**, *125*, 1204.

(123)    Qi, M.; Gross, A.; Jeschke, G.; Godt, A.; Drescher, M. *J. Am. Chem. Soc.* **2014**, *136*, 15366.

(124)    Raitsimring, A. M.; Gunanathan, C.; Potapov, A.; Efremenko, I.; Martin, J. M.; Milstein, D.; Goldfarb, D. *J. Am. Chem. Soc.* **2007**, *129*, 14138.

(125)    Yulikov, M.; Lueders, P.; Warsi, M. F.; Chechik, V.; Jeschke, G. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10732.

(126)    Akhmetzyanov, D.; Plackmeyer, J.; Endeward, B.; Denysenkov, V.; Prisner, T. F. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6760.

(127)    Banerjee, D.; Yagi, H.; Huber, T.; Otting, G.; Goldfarb, D. *J. Phys. Chem. Lett.* **2012**, *3*, 157.

(128)    Ching, H. Y. V.; Demay-Drouhard, P.; Bertrand, H. C.; Policar, C.; Tabares, L.; Un, S. *Phys. Chem. Chem. Phys.* **2015**, *17*, 23368.

(129)    Kaminker, I.; Bye, M.; Mendelman, N.; Gislason, K.; Sigurdsson, S. T.; Goldfarb, D. *Phys. Chem. Chem. Phys.* **2015**, *17*, 15098.

(130)    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.

(131)    Martorana, A.; Yang, Y.; Zhao, Y.; Q., L.; Goldfarb, D. *Dalton Trans.* **2015**, *44*, 20812.

(132) Yang, Z.; Jimenez-Oses, G.; Lopez, C. J.; Bridges, M. D.; Houk, K. N.; Hubbell, W. L. *J. Am. Chem. Soc.* **2014**, *136*, 15356.

(133) Arnold, F. H.; Haymore, B. L. *Science* **1991**, *252*, 1796.

(134) Higaki, J. N.; Fletterick, R. J.; Craik, C. S. *Trends in biochemical sciences* **1992**, *17*, 100.

(135) Jung, K.; Voss, J.; He, M.; Hubbell, W. L.; Kaback, H. R. *Biochemistry* **1995**, *34*, 6272.

(136) Nicoll, A. J.; Miller, D. J.; Futterer, K.; Ravelli, R.; Allemann, R. K. *Journal of the American Chemical Society* **2006**, *128*, 9187.

(137) Regan, L. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 257.

(138) Todd, R. J.; Van Dam, M. E.; Casimiro, D.; Haymore, B. L.; Arnold, F. H. *Proteins* **1991**, *10*, 156.

(139) Voss, J.; Salwinski, L.; Kaback, H. R.; Hubbell, W. L. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12295.

(140) Goldfarb, D. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9685.

(141) Cunningham, T. F.; Putterman, M. R.; Desai, A.; Horne, W. S.; Saxena, S. *Angew. Chem. Int. Ed.* **2015**, *127*, 6330.

(142) Chen, W.; Wu, C.; Liu, C. *J. Colloid Interface Sci.* **1996**, *180*, 135.

(143) Fauth, J. M.; Schweiger, A.; Braunschweiler, L.; Forrer, J.; Ernst, R. R. *J. Magn. Res.* **1986**, *66*, 74.

(144) Gemperle, C.; Aebli, G.; Schweiger, A.; Ernst, R. R. *J. Magn. Res.* **1990**, *88*, 241.

(145) Milov, A. D.; Tsvetkov, Y. D.; Formaggio, F.; Crisma, M.; Toniolo, C.; Raap, J. *J. Am. Chem. Soc.* **2000**, *112*, 3843.

(146)    Cunningham, T. F.; McGoff, M. S.; Sengupta, I.; Jaroniec, C. P.; Horne, W. S.; Saxena, S. *Biochemistry* **2012**, *51*, 6350.

(147)    Hubicki, Z.; Jakowicz, A. *Desalination* **2003**, *155*, 121.

(148)    Lukes, I.; Smidova, I.; Vleck, A.; Podlaha, J. *Chem. Zvesti* **1982**, *38*, 331.

(149)    Yip, T.; Nakagawa, Y.; Porath, J. *Anal. Biochem.* **1989**, *183*, 159.

(150)    Roche, T. S.; Wilkins, R. G. *J. Am. Chem. Soc.* **1974**, *96*, 5082.

(151)    Silva, K. I.; Michael, B. C.; Geib, S. J.; Saxena, S. *J. Phys. Chem. B* **2014**, *118*, 8935.

(152)    Kevan, L. In *Time Domain Electron Spin Resonance*; John Wiley & Sons: New York, 1979, p 279.

(153)    Potapov, A.; Yagi, H.; Huber, T.; Jergic, S.; Dixon, N. E.; Otting, G.; Goldfarb, D. *J. Am. Chem. Soc.* **2010**, *132*, 9040.

(154)    Hayes, J. D.; Pulford, D. J. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 445.

(155)    Armstrong, R. N. *Chem. Res. Toxicol.* **1997**, *10*, 2.

(156)    Koumaravelou, K.; Shoaib, Z.; Adithan, C.; Charron, D.; Srivastava, A.; Tamouza, R.; Krishnamoorthy, R. *Clin. Clhim. Acta.* **2011**, *412*, 1465.

(157)    Liu, Y.; Hyde, A. S.; Simpson, M. A.; Barycki, J. J. *Adv. Cancer Res.* **2014**, *122*, 69.

(158)    Moden, O.; Mannervik, B. *Adv. Cancer Res.* **2014**, *122*, 199.

(159)    Hayes, J. D.; Flanagan, J. U.; Jowsey, I. R. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 51.

(160)    Hammes, G. G.; Chang, Y.-C.; Oas, T. G. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13737.

(161)   Eaton, D. L.; Bammler, T. K. *Toxicol. Sci.* **1999**, *49*, 156.

(162)   Honaker, M. T.; Acchione, M.; Zhang, W.; Mannervik, B.; Atkins, W. M. *J. Biol. Chem.* **2013**, *288*, 18599.

(163)   Frova, C. *Biomol. Eng.* **2006**, *23*, 149.

(164)   Grahn, E.; Novotny, M.; Jakobsson, E.; Gustafsson, A.; Grehn, L.; Olin, B.; Madsen, D.; Wahlberg, M.; Mannervik, B.; Kleywegt, G. J. *Acta. Crystallogr. D. Biol. Crystallogr.* **2006**, *62*, 197.

(165)   Zhan, Y.; Rule, G. S. *Biochemistry* **2004**, *43*, 7244.

(166)   Balchin, D.; Fanucchi, S.; Achilonu, I.; Adamson, R. J.; Burke, J.; Fernandes, M.; Gildenhuys, S.; Dirr, H. W. *Biochim. Biophys. Acta* **2010**, *1804*, 2228.

(167)   Karpusas, M.; Axarli, I.; Chiniadis, L.; Papakyriakou, A.; Bethanis, K.; Scopelitou, K.; Clonis, Y. D.; Labrou, N. E. *PLOS One* **2013**, *8*, e56337.

(168)   Le Trong, I.; Stenkamp, R. E.; Ibarra, C.; Atkins, W. M.; Adman, E. T. *Proteins* **2002**, *48*, 618.

(169)   Adman, E. T.; Le Trong, I.; Stenkamp, R. E.; Nieslanik, B. S.; Dietze, E. C.; Tai, G.; Ibarra, C.; Atkins, W. M. *Proteins* **2001**, *42*, 192.

(170)   Dirr, H. W.; Wallace, L. A. *Biochemistry* **1999**, *38*, 15631.

(171)   Nilsson, L. O.; Edalat, M.; Pettersson, P. L. M.; Mannervik, B. *Biochim. Biophys. Acta - Proteins Proteomics* **2002**, *1597*, 199.

(172)   Pettersson, J. R.; Lanni, F.; Rule, G. S. *Biochemistry* **2017**.

(173)   Nieslanik, B. S.; Ibarra, C.; Atkins, W. M. *Biochemistry* **2001**, *40*, 3536.

(174)   Allardyce, C. S.; McDonagh, P. D.; Lian, L.; Wolf, C. R.; Roberts, G. C. *Biochem. J.* **1999**, *343*.

(175)   Board, P. G.; Mannervik, B. *Biochem. J.* **1991**, *275*, 171.

(176)   Gustafsson, A.; Etahadieh, M.; Jemth, P.; Mannervik, B. *Biochemistry* **1999**, *38*, 16268.

(177)   Hou, L.; Honaker, M. T.; Shireman, L. M.; Balogh, L. M.; Roberts, A. G.; Nguyen, K.; Nathan, A.; Atkins, W. M. *J. Biol. Chem.* **2007**, *282*, 23264.

(178)   Nieslanik, B. S.; Atkins, W. M. *J. Biol. Chem.* **2000**, *275*, 17447.

(179)   Nieslanik, B. S.; Dietze, E. C.; Atkins, W. M.; Trong, I. L.; Adman, E. T. *Pac. Symp. Biocomput.* **1999**, *275*, 554.

(180)   Manukovsky, N.; Frydman, V.; Goldfarb, D. *J. Phys. Chem. B* **2015**, *119*, 13732.

(181)   Bonora, M.; Becker, J.; Saxena, S. *Journal of Magnetic Resonance* **2004**, *170*, 278.

(182)   Roser, P.; Schmidt, M. J.; Drescher, M.; Summerer, D. *Org. Biomol. Chem.* **2016**, *14*, 5468.

(183)   Columbus, L.; Hubbell, W. L. *Biochemistry* **2004**, *43*, 7273.

(184)   Sengupta, I.; Gao, M.; Arachchige, R. J.; Naduad, P. S.; Cunningham, T. F.; Saxena, S.; Schwieters, C. D.; Jaroniec, C. P. *J. Biol. Magn. Reson.* **2015**, *62*, 1.

(185)   Kuhnert, D. C.; Sayed, Y.; Mosebi, S.; Sayed, M.; Sewell, T.; Dirr, H. W. *J. Mol. Biol.* **2005**, *349*, 825.

(186)   Pabst, M. J.; Habig, W. H.; Jakoby, W. B. *J. Biol. Chem.* **1974**, *249*, 7140.

(187)   Budil, D. E.; Lee, S.; Saxena, S.; Freed, J. H. *Journal of Magnetic Resonance* **1996**, *120*, 155.

(188)   Columbus, L.; Hubbell, W. L. *Trends Biochem. Sci.* **2002**, *27*, 288.

(189)   Jeschke, G. *Prog. Nuc. Magn. Reson. Spec.* **2013**, *72*, 42.

(190)   Polyhach, Y.; Bordignon, E.; Jeschke, G. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356.

(191)   Hinsen, K.; Petrescu, A. J.; Dellerue, S.; Bellissent-Funel, M. C.; Kneller, G. R. *Chem. Phys.* **2000**, *261*, 25.

(192)   Jeschke, G. *Z. Phys. Chem.* **2012**, *226*, 1395.

(193)   Orellana, L.; Rueda, M.; Ferrer-Costa, C.; Lopez-Blanco, J. R.; Chacon, P.; Orozco, M. *J. Chem. Theor. Comput.* **2010**, *6*, 2910.

(194)   Zheng, W.; Brooks, B. R. *Biophys. J.* **2006**, *90*, 4327.

(195)   Hagelueken, G.; R., W.; Naismith, J. H.; Schiemann, O. *Appl. Magn. Reson.* **2012**, *42*, 377.

(196)   Mosebi, S.; Sayed, Y.; Burke, J.; Dirr, H. W. *Biochemistry* **2003**, *42*, 15326.

(197)   Wallace, L. A.; Dirr, H. W. *Biochemistry* **1999**, *38*, 16686.

(198)   Wallace, L. A.; Sluis-Cremer, N.; Dirr, H. W. *Biochemistry* **1998**, *37*, 5320.

(199)   Brandi-Blanco, M. P.; Dumet-Fernandes, B.; González-Pérez, J. M.; Choquesillo-Lazarte, D. *Acta. Cryst.* **2007**, *63*, m1598.

(200)   Blackburn, M. E.; Veloro, A. M.; Fanucci, G. E. *Biochemistry* **2009**, *48*, 8765.

(201)   Fleissner, M. R.; Cascio, D.; Hubbell, W. L. *Protein Sci.* **2009**, *18*, 893.

(202)   Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W. *J. Magn. Reson.* **2000**, *142*, 331.

(203)   Borbat, P. P.; Freed, J. H. In *Biological Magnetic Resonance*; Kluwer Academic/Plenum Publishers: New York, 2000; Vol. 19.

(204)   Borbat, P.; Mchaourab, H. S.; Freed, J. H. *J. Am. Chem. Soc.* **2002**, *124*, 5304.

(205)    Hubbell, W. L.; Lopez, C. J.; Altenbach, C.; Yang, Z. *Cur. Opin. Struct. Biol.* **2013**, *23*, 725.

(206)    Jeschke, G.; Polyhach, Y. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895.

(207)    Reginsson, G. W.; Schiemann, O. *Biochem. J.* **2011**, *434*, 353.

(208)    Borbat, P.; Freed, J. H. In *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centers in the Biosciences. Structure and Bonding*; Timmel, J. H. a. C., Ed.; Springer: 2014, p 1.

(209)    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*, 4722.

(210)    Zhang, X.; Tung, C.; Sowa, G. Z.; Hatmal, M. M.; Haworth, I. S.; Qin, P. Z. *J. Am. Chem. Soc.* **2012**, *134*, 2644.

(211)    Prisner, T. F.; Marko, A.; Sigurdsson, S. T. *J. Magn. Res.* **2015**, *252*, 187.

(212)    Dedmon, M. M.; Patel, C. N.; Young, G. B.; Pielak, G. J. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12681.

(213)    Schlesinger, A. P.; Wang, Y.; Tadeo, X.; Millet, O.; Pielak, G. J. *J. Am. Chem. Soc.* **2011**, *133*, 8082.

(214)    Monteith, A. C.; Pielak, G. J. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11335.

(215)    Luchinat, E.; Banci, L. *J. Biol. Chem.* **2016**, *291*, 3776.

(216)    Zimmermann, S. B.; Trach, S. O. *J. Mol. Biol.* **1999**, *222*, 599.

(217)    Minton, A. P. *Curr. Opin. Biotechnol.* **1997**, *8*, 65.

(218)    Eggers, D. K.; Valentine, J. S. *Protein Sci.* **2001**, *10*, 250.

(219)    Kuznetsova, I. M.; Turoverov, K. K.; Uversky, V. N. *Int. J. Mol. Sci.* **2014**, *15*, 23090.

(220)  Ellis, R. J. *Trends Biochem. Sci.* **2001**, *26*, 597.

(221)  Igarashi, R.; Sakai, T.; Hara, H.; Tenno, T.; Tanaka, T.; Tochio, H.; Shirakawa, M. *J. Am. Chem. Soc.* **2010**, *134*, 8228.

(222)  Azarkh, M.; Okle, O.; Singh, V.; Seemann, I. T.; Hartig, J. S.; Dietrich, D. R.; Drescher, M. *ChemBioChem* **2011**, *12*, 1992.

(223)  Krstic, I.; Hansel, R.; Romainczyk, O.; Engels, J. W.; Dotsch, V.; F., P. T. *Angew. Chem. Int. Ed.* **2011**, *50*, 5070.

(224)  Azarkh, M.; Singh, V.; Okle, O.; Seemann, I. T.; Dietrich, D. R.; Hartig, J. S.; Drescher, M. *Nature Protocols* **2013**, *8*, 131.

(225)  Schmit, M. J.; Borbas, J.; Drescher, M.; Summerer, D. *J. Am. Chem. Soc.* **2014**, *136*, 1238.

(226)  Dunkel, S.; Pulagam, L. P.; Steinhoff, H.-J.; Klare, J. P. *Phys. Chem. Chem. Phys.* **2015**, *17*, 4875.

(227)  Joseph, B.; Sikora, A.; Bordignon, E.; Jeschke, G.; Cafiso, D. S.; Prisner, T. F. *Angew. Chem. Int. Ed.* **2015**, *54*, 1.

(228)  Schmidt, M. J.; Fedozeev, A.; Bucker, D.; Borbas, J.; Peter, C.; Drescher, M.; Summerer, D. *ACS Chem. Biol* **2015**.

(229)  Benesh, J.; Sikora, A.; Cafiso, D. S. *J. Am. Chem. Soc.* **2016**, *138*, 1844.

(230)  Theillet, F. X.; Binolfi, A.; Bekei, B.; Martorana, A.; Rose, H. M.; Stuvier, M.; Goldfarb, D.; Selenko, P. *Febs J.* **2015**, *282*.

(231)  Shafer, A. M.; Bennett, V. J.; Kim, P.; Voss, J. C. *J. Biol. Chem.* **2003**, *278*, 34203.

(232) Azarkh, M.; Okle, O.; Eyring, P.; Dietrich, D. R.; Drescher, M. *J. Magn. Reson.* **2011**, *212*, 450.

(233) Roser, P.; Schmidt, M. J.; Drescher, M.; Summerer, D. *Org. Biomol. Chem.* **2016**, *14*, 5468.

(234) Qi, M.; Grob, A.; Jeschke, G.; Godt, A.; Drescher, M. *J. Am. Chem. Soc.* **2014**, *136*, 15366.

(235) Jagtap, A. P.; Krstic, I.; Kunjir, N. C.; Hansel, R.; Prisner, T. F.; Sigurdsson, S. T. *Free Radical Res.* **2015**, *49*, 78.

(236) Bobko, A. A.; Dhimitruka, I.; Zweier, J. L.; Khramstov, V. V. *J. Am. Chem. Soc.* **2007**, *129*, 7240.

(237) Liu, Y.; Villamena, F. A.; Sun, J.; Xu, Y.; Dhimitruka, I.; Zweier, J. L. *J. Org. Chem.* **2008**, *73*, 1490.

(238) Reginsson, G. W.; Kunjir, N. C.; Sigurdsson, S. T.; Schiemann, O. *Chem. Eur. J.* **2012**, *18*, 13580.

(239) Yang, Z.; Liu, Y.; Borbat, P.; Zweier, J. L.; Freed, J. H.; Hubbell, W. L. *J. Am. Chem. Soc.* **2012**, *134*, 9950.

(240) Meyer, V.; Swanson, M. A.; Clouston, L. J.; Boratynski, P. J.; Stein, R. A.; Mchaourab, H. S.; Rajca, A.; Eaton, S. S.; Eaton, G. R. *Biophys. J.* **2015**, *108*, 1213.

(241) Jassoy, J. J.; Berndhauser, A.; Duthie, F.; Kuhn, S. P.; Hagelueken, G.; Schiemann, O. *Angew. Chem. Int. Ed.* **2016**, *56*, 177.

(242) Joseph, B.; Tormyshev, V. M.; Rogozhnikova, O. Y.; Akhmetzyanov, D.; Bagryanskaya, E. G.; Prisner, T. F. *Angew. Chem. Int. Ed.* **2016**, *55*, 11538.

(243)   Wipf, P.; Xiao, J.; Jiang, J.; Belikova, N. A.; Tyurin, V. A.; Fink, M. P.; Kagan, V. E. *J. Am. Chem. Soc.* **2005**, *127*, 12460.

(244)   Cunningham, T. F.; Putterman, M. R.; Desai, A.; Horne, W. S.; Saxena, S. *Angew. Chem. Int. Ed.* **2015**, *127*, 1.

(245)   Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. *Appl. Mag. Res.* **2006**, *30*, 473.

(246)   Swartz, H. M.; Sentjurc, M.; Morse, P. D. *Biochim. Biophys. Acta* **1986**, *888*, 82.

(247)   Ji, J.; Kline, A. E.; Amoscato, A.; Samhan Arias, A. K.; Sparvero, L. J.; Tyurin, V. A.; Tyurina, Y. Y.; Fink, B.; Manole, M. D.; Puccio, A. M.; Okonkwo, D. O.; Cheng, J. P.; Alexander, H.; Clark, R. S. B.; Kochanek, P. M.; Wipf, P.; Kagan, V. E.; Bayir, H. *Nature Neuroscience* **2012**, *15*, 1407.

(248)   Kagan, V. E.; Wipf, P.; Stoyanovsky, D.; Greenberger, J. S.; Borisenko, G.; Belikova, N. A.; Yanamala, N.; Samhan Arias, A. K.; Tungekar, M. A.; Jiang, J.; Tyurina, Y. Y.; Ji, J.; Klein-Seetharaman, J.; Pitt, B. R.; Shvedovam, A. A.; Bayir, H. *Adv. Drug Deliv. Rev.* **2009**, *61*, 1375.

(249)   Yoshimura, T.; Yokoyama, H.; Fujii, S.; Takayama, F.; Oikawa, K.; Kamada, H. *Nature Biotechnology* **1996**, *14*, 992.

(250)   Weaver, J.; Burks, S. R.; Liu, K. J.; Kao, J. P. Y.; Rosen, G. M. *J. Magn. Res.* **2016**, *271*, 68.

(251)   Shafer, A. M.; Kalai, T.; Bin Liu, S. Q.; K., H.; C., V. J. *Biochemistry* **2004**, *43*, 8470.

(252)   Swick, A. G.; Janicot, M.; Cheneval-Kastelic, T.; McLenthian, J. C.; Lane, M. D. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1812.

(253)    Kinde, M. N.; Chen, Q.; Lawless, M. J.; Mowrey, D. D.; J., X.; Saxena, S.; Xu, Y.; Tang, P. *Structure* **2015**, *23*, 995.

(254)    Keana, J. F.; Pou, S.; Rosen, G. M. *Magn. Reson. Med.* **1987**, *5*, 525.

(255)    Swartz, H. M.; Timmins, G. S. In *Toxicology of the human environment: the critical role of free radicals*; Rhodes, C. J., Ed.; Taylor & Francis Inc.: London, New York, 2000, p 91.

(256)    Saphier, O.; Silberstein, T.; Shames, A. I.; Likhtenshtein, G. I.; Maimon, E.; Mankuta, D.; Mazor, M.; M., K.; Meyerstein, D.; Meyerstein, N. *Free Radic. Res* **2003**, *37*, 301.

(257)    Schafer, F. Q.; Buettner, G. R. *Free Radic. Biol. Med.* **2001**, *30*, 1191.

(258)    Kuppusamy, P.; Li, H.; Ilangovan, G.; Cardounel, A. J.; Zweier, J. L.; Yamada, K.; Krishna, M. C.; Mitchell, J. B. *Cancer Res.* **2002**, *62*, 307.

(259)    Borbat, P. P.; Freed, J. H. *Methods. Enzymol.* **2007**, *432*, 52.

(260)    Ziegler, D. M. *Ann. Rev. Biochem.* **1985**, *54*, 305.

(261)    Couet, J.; de Bernard, S.; Loosfelt, H.; Saunier, B.; Milgrom, E.; Misrahi, M. *Biochemistry* **1996**, *35*, 14800.

(262)    Rhee, S. G.; Z., C. H.; Kim, K. *Free Radic. Biol. Med.* **2005**, *38*, 1543.