Spin Labeling Methodologies for Measuring Precise Protein Distance Constraints

by

Timothy F. Cunningham

B.S. Chemistry, Mercyhurst College, 2008

Submitted to the Graduate 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

2015

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Timothy F. Cunningham

It was defended on

July 21st, 2015

and approved by

David Waldeck, Ph.D., Professor, Department of Chemistry

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

Andrea Berman, Ph.D., Assistant Professor, Department of Biological Sciences

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

Spin Labeling Methodologies for Measuring Precise Protein Distance Constraints

Timothy F. Cunningham, PhD

University of Pittsburgh, 2015

Protein spin labeling at helical and loop sites that yields the nitroxide based R1 side chain is a powerful method to measure protein dynamics and structure by electron spin resonance (ESR). This thesis addresses the lack of foundational work of R1 in β-sheets that has extremely limited the use of R1 in this secondary structure environment. This work provides the first essential steps for understanding R1 rotameric preferences in β-sheets through the use of various ESR experiments, X-ray crystallography, and molecular modeling. The results presented here indicate that R1 at internal β-strand sites display rotameric preferences previously not observed and extracting backbone dynamics information by current methods is not straight forward. Two distinct edge sites were also explored, again showing rotameric preferences unique to each site. This work highlights the need for new models that appropriately account for R1 motion in the highly complex and diverse β-sheet environment.

The second focus of this thesis is to develop strategies to site specifically incorporate Cu^{2+} ions into proteins as spin labels. Presented here is the first use of two Cu^{2+} chelating tags for use in ESR distance measurements. The results show that a cyclen based tag yields distance distributions comparable to that of R1 indicating that the tag is a reasonable alternative for R1. Additionally, given R1 instability in cells, the Cu^{2+} tag is likely well suited for in cell measurements.

The most severe limitation of the R1 label is its flexibility. We introduce a labeling procedure to site specifically incorporate Cu^{2+} ions. Notably, the labeling is achieved by exploiting naturally occurring amino acids and does not require post-expression synthetic modification. The Cu^{2+} motion is significantly restricted by coordination and therefore the resultant distances are remarkably precise with distance distributions widths that are five times more narrow when compared to R1. This development constitutes a decisive improvement in labeling methodology that is not only simple, but also capable of providing unambiguous, highly relevant, protein structural constraints.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

The last seven years of my life have been a difficult yet rewarding road. Not only did I find my calling in teaching but I found a confidence in myself that will keep with me throughout my life. First and foremost, I'd like to thank my advisor, Prof. Sunil Saxena for his guidance and help along the way. In his lab, I was pushed to levels I did not realize were achievable and I am absolutely proud of what I have accomplished is the lab over the years. Additionally I would like to thank him for his enthusiastic support of my passion for teaching throughout my time in graduate school.

I'd also like to thank my committee, Prof. Andrea Berman, Prof. David Waldeck, and Prof. W. Seth Horne for helping me with this integral step in my career. I'd especially like to thank Prof. Horne for his countless years of collaboration on a variety of projects.

Next I'd like to thank the many professors who I worked with for my teaching duties including Prof. Carol Fortney, Prof. George Bandik, and Prof. Ericka Huston. All of you were a huge help throughout my time here and are continuingly essential resources in the early stages of my career. In particular I'd like to thank Ericka for all of her time as a colleague and friend in the many years organizing the chaos that is undergraduate organic labs.

I'd also like to thank the many coworkers and friends in my lab, past and present. This includes Prof. Jessica Sarver, Prof. Sharon Ruthstein, Dr. Ming Ji, Dr. K. Ishara Silva, Matt

Lawless, and Austin Jarvi. All of you have made working in lab a wonderful experience and it has been fantastic getting to know all of you. I also want to thank the many undergraduates that worked with me in lab including Matt Shannon, Miriam Putterman, Kimi Shah, Astha Desai, and Amit Shimshi. I thoroughly enjoyed teaching all of you and I am so appreciative for all of your help on my projects. It was great working with and getting to know each of you and I wish you all the best of luck with your future endeavors.

I would also like to thank all my friends I have made in my time in Pittsburgh, in the department and out of it. Chris, Lacy, Mark, Jared, Robin, Tom, Rob, Amy, Ryan, Yang, Zach, and too many more; all of you made my graduate career more enjoyable and thank you so much for all of your support throughout the years. Especially I would like to thank Ashley though only witness to the tail end of my career, the support you provided in those end times helped me stay focused and confident through the finish.

Finally I want to thank my family. Mom, Dad, Jon, all of you have shown me love and support that has kept me going throughout the years. This was a difficult road at times but all of you helped me finish strong. Thank you so much for absolutely everything your support means the world to me.

1.0 INTRODUCTION

Electron spin resonance (ESR) is a powerful biophysical tool for elucidating a protein's dynamics and structure. In order for ESR to be used, unpaired electrons are required. As these are uncommon to many protein systems, they must be incorporated using the technique known as site directed spin labeling (SDSL). Currently, the most common spin label is the nitroxide based R1 side chain. With a single R1, site specific backbone dynamics information can be acquired through continuous wave (CW) ESR. With two or more R1 labels, long range distance constraints can be acquired with such pulsed ESR experiments as double electron electron resonance (DEER) spectroscopy.

Currently, site specific backbone dynamics can only be resolved when the R1 side chain is located at solvent-exposed α-helix sites or on a loop. This is possible because CW lineshape and crystallographic observations of R1 suggest a stabilizing interaction with the backbone that putatively restricts internal motions. This interaction is believed to couple the motion of the spin to the motion of the backbone and these motions can be detected in the CW lineshape. Despite the importance of this systematic analysis within α-helices, there is a significant lack of analogous work in solvent-exposed β-sheets despite the prevalence of β-sheets in proteins (Figure 1-1a). Chapter 2 directly addresses this lack and presents the first solvent exposed, internal strand β-sheet R1 rotamers observed through crystallography. Additionally the rotamers

are compared with DEER and variable temperature CW to gain further insight into rotameric preferences and motional dynamics of R1 at this site. Furthermore, two distinct solvent exposed edge β-strand environments are investigated and the results of these findings are discussed in Chapter 3. The sites are investigated with crystallography and DEER measurements and the resolved rotamers display contextual preferences unlike any other solvent-exposed environment. Additionally, common computational models used for predicting R1 based distances are assessed and it is clear that new models are necessary for appropriate modeling of R1 at these sites.

While R1 is the most widely used spin label, site specific incorporation of paramagnetic metals as alternate spin labels is becoming more prevalent. Chapter 4 utilizes two different site specific tags that specifically chelate Cu^{2+} for initial use in ESR protein distance measurements. In particular, a cyclen based tag yields distributions remarkably similar to R1 distance measurements. This result indicates the cyclen tag as a valuable alternative to R1. As nitroxides are not stable in cell, the cyclen tag shows potential promise for use in this environment.

Perhaps the most critical limitation of R1 is its flexibility. R1 contains 5 rotatable bonds that can place the nitroxide ring at a range of locations compared to the backbone attachment site (Figure 1-1b). Therefore, the distance measurements between two labeled sites are dominated by the conformations of the label at each site. In efforts to simplify DEER distance interpretation, a novel means of spin labeling is introduced in Chapter 5. Through two natural amino acid substitutions and the introduction of a metal salt and a small metal chelator, a highly rigid spin label self assembles in solution and provides distance distributions five times more narrow as compared to an analogous R1 distribution. This work displays the power of this method for measuring unambiguous distance contraints using ESR.

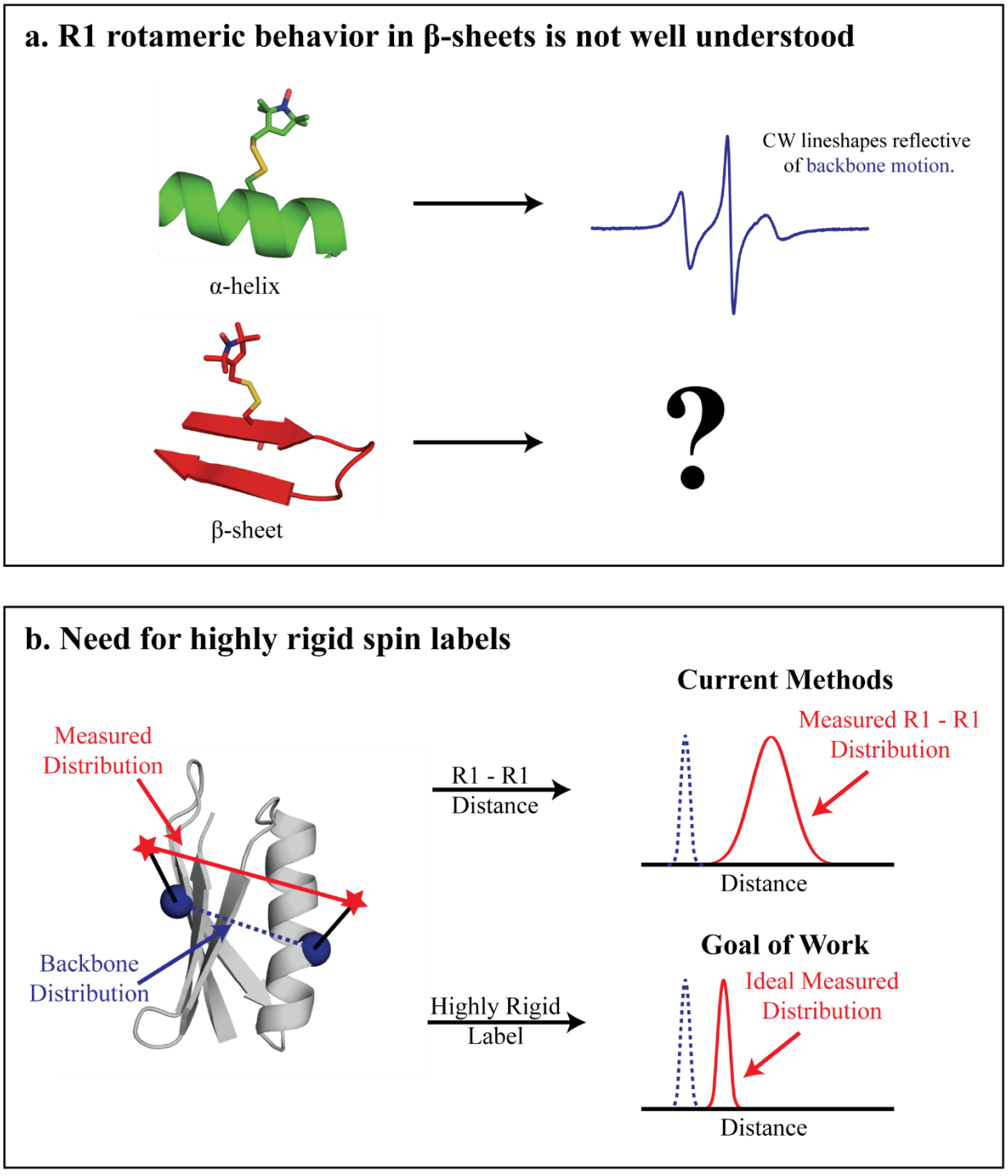

Figure 1-1. Overview of the two main problems addressed in this thesis.

This introductory chapter will provide an overview of work performed previously in the field to illustrate the motivations driving this thesis.

1.1 PROTEIN SPIN LABELING

Electron spin resonance (ESR) is a valuable biophysical tool that is commonly used to elucidate dynamics and structural information on a number of biologically relevant complexes.[1, 2] The sole requirement for using ESR to acquire such information is the presence of an unpaired electron within the system. An unpaired electron, or spin, can come in the form of either a paramagnetic metal ion or an organic radical. While many proteins do bind paramagnetic metals naturally, this is not the case for all systems. Additionally, natural metal binding sites are restricted to specific locations thus limiting regions of the protein that can be investigated. Therefore, to overcome these limitations, the technique known as site directed spin labeling (SDSL) was developed to provide a means to incorporate unpaired electrons into any protein site of interest.[3]

SDSL began with the development of a method which resulted in the nitroxide side chain R1 (Figure 1-2).[4] The basic concept behind SDSL is to create a specific labeling target within a protein and then incorporate the label at the target site. In the case of R1 labeling, the target is the amino acid cysteine (Figure 1-2). Cysteine is commonly used for site specific protein labeling due to the high specific reactivity associated with the free thiol. Site-directed mutagenesis can be used to incorporate cysteines at desired locations within the encoding DNA and in turn, the protein sequence. Naturally occurring cysteines must be removed if these

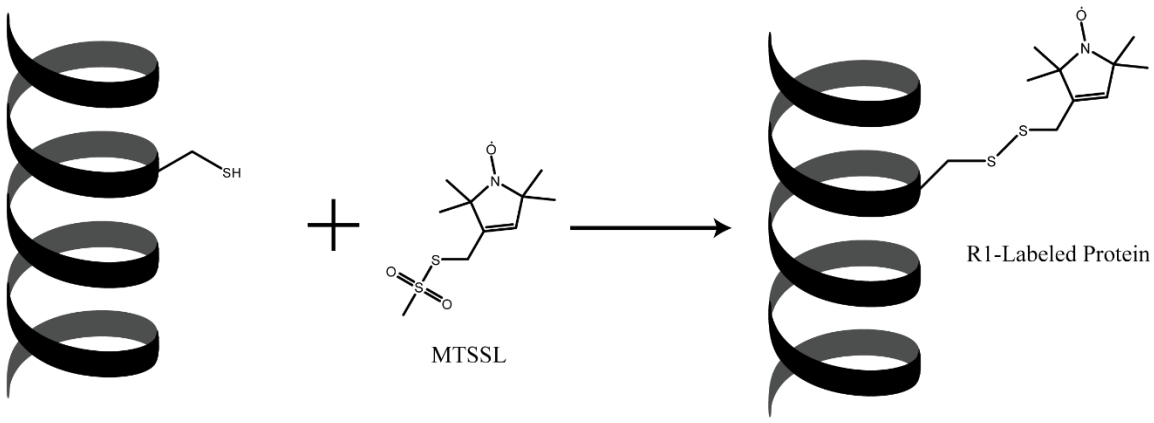

Figure 1-2. The most commonly used protocol for spin labeling proteins. The cysteine residues are inserted with mutagenesis then reacted with MTSSL to generate the R1 side chain.

labeling sites are undesired to create a Cys-null mutant of the protein. However this procedure is only required for cysteines that are not involved in disulfide bridges or are not solvent exposed. Once the cysteine-containing protein is overexpressed and purified, labeling can be achieved through introduction of the methanothiosulfonate spin label (MTSSL) reagent (Figure 1-2). The reagent reacts specifically with the free thiol of the introduced cysteines, yielding the R1 side chain at the desired location(s).

1.2 R1 RESOLVES PROTEIN DYNAMICS AND STRUCTURE

A number of different experiments can be performed to gain insight into a protein labeled with R1. While the vast majority of these experiments are ESR based, R1 labeling has also been used in conjunction with NMR to generate ~ 10 Å to 20 Å long distance constraints using the paramagnetic relaxation enhancement experiment.[5] Within ESR, two of the most common and most powerful experiments performed on R1 labeled proteins are continuous wave (CW) and double electron electron resonance (DEER). The focus of this section will be the information that can be elucidated from these two experiments.

1.2.1 Continuous wave ESR can resolve protein backbone dynamics

The simplest ESR experiment performed on R1 labeled proteins is CW. CW ESR detects magnetic interactions between the unpaired electron and an applied magnetic field. These interactions are dependent upon the orientation of the nitroxide relative to the applied field and

thus any motions that reorient R1 on the timescale of the measurement affect the CW lineshape. The most common ESR spectrometers use X-band (~9.5 GHz) frequencies and CW spectra collected on these instruments are sensitive to any nanosecond to microsecond motions. As such, motions that influence the CW lineshape include molecular tumbling, internal motions of the R1 side chain and, most interestingly, fluctuations of the backbone near where R1 is attached.

Resolving site specific backbone motion is highly useful as structural flexibility in proteins may play a large role in such molecular processes as ligand binding and protein recognition. While information on backbone motion is embedded within the CW lineshape, extracting this information requires the effects from molecular tumbling and internal motions be minimal. Tumbling is the simplest to address by slowing rates such that tumbling is frozen out on the ESR timescale. Proteins larger than 50 kDa or located within a lipid membrane experience tumbling rates too slow to affect the lineshape at X-band frequencies.[6] Tumbling rates of smaller complexes can be slowed through addition of a viscous agent to the solvent medium.[7] However accounting for R1 internal motions is not as simple.

The internal motions of R1 can be defined by the rotameric sampling of the five bonds that connect the nitroxide ring to the protein backbone (Figure 1-3a). The five dihedral angles, referred to as χ_1 through χ_5, are rapidly fluctuating leading to motion of the ring relative to the protein. While these motions can be affected by sterics or other interactions with the local environment, thorough analysis of R1 motion has been performed to elucidate a model of internal motion such that backbone fluctuations can be extracted from CW lineshape analysis (Figure 1-3b).

Pioneering work performed to isolate the backbone motion from internal R1 motions was performed by Hubbell and coworkers.[6-11] Using the protein T4 lysozyme, R1

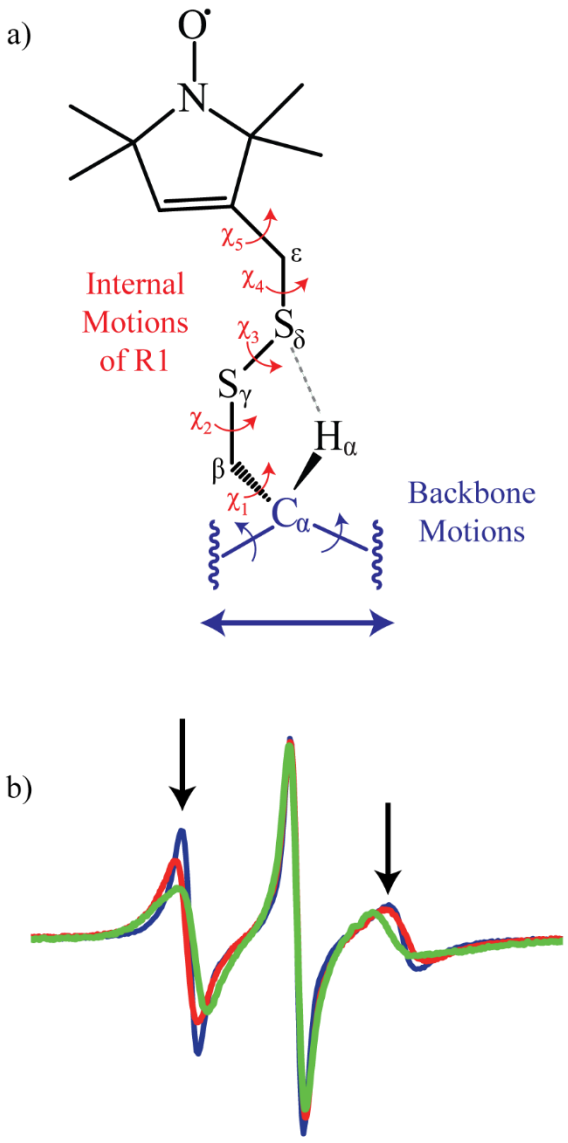

Figure 1-3. (a) The R1 spin label and the molecular motions that contribute to movement of the nitroxide ring. Additionally the dotted gray line represents the potential stabilizing hydrogen bond that occurs between the H_α and S_δ atoms. (b) Room temperature X-band CW spectra for various protein sites. The slight changes indicated by the arrows are indicative of differences in the molecular motions shown in (a)

was systematically incorporated throughout the protein into a variety of different environments.[6-9] Initially, sterics with local residues were thought to be the primary influence on R1 motion, and mutational analysis was performed at those sites where neighboring bulky residues were substituted for smaller residues.[7] The CW lineshapes that changed upon mutation indicated that contact with the local environment was influencing the motion of R1. However, even for solvent-exposed α-helix R1 sites that exhibited no steric interactions, site specific changes in lineshape were still observed. This was rationalized as likely resulting from differences in protein backbone motion. This hypothesis was supported by observed differences in crystallographic thermal factors for each of the sites.[7]

In order to investigate why changes in backbone motions are observed in the CW lineshapes from these sites, the R1 rotameric preferences were investigated through solving the structure of spin labelled mutants of T4 lysozyme by X-ray crystallography.[12] These results displayed R1 conformations that place the S_δ atom in close proximity to the H_α atom, suggesting the presence of a stabilizing hydrogen bond (Figure 1-3a). This observation, coupled with the known slow rotational rates of disulfide bonds (χ_3),[6] suggests that the majority of R1 internal motion is due to rotations about χ_4 and χ_5. This model of internal motion of R1 became known as the χ_4/χ_5 model.[12] Accordingly, within the solvent exposed α-helical environment where this model was formulated, the internal motions of R1 are restricted and thus motion of the nitroxide ring is directly coupled to backbone motions. This allows for lineshape changes at these sites to be directly interpreted as changes in backbone motions.[13]

In order to illustrate the power of this model, Columbus et al explored the relative backbone dynamics of an α-helical protein dimer using CW lineshape analysis.[10] The spectra

were fit with the microscopic order macroscopic disorder (MOMD) model that was initially created to describe nitroxide motion in liquid crystals and membranes.[14] This model was used here in a protein environment to quantify contributions from the backbone motion. Specifically, the site specific differences in rotational correlation times and ordering potentials were directly attributed to changes in backbone dynamics. Performing this analysis along a protein helix allowed a map of relative backbone motion to be constructed.[10] These results were in remarkable agreement with previous NMR relaxation results[15] indicating the validity of this method for extracting backbone motions.

The CW work discussed thus far illustrates the power of CW lineshape analysis for resolving backbone dynamics. However, the χ_4/χ_5 model of R1 motion has only been shown to be valid in solvent-exposed α-helices. Another common secondary structure environment that warrants the same level of attention is the β-sheet. β-sheets are prevalent in protein structures and offer a diverse environment. Despite this, information on how R1 is influenced by the β-sheet environment is quite limited. The primary effort thus far has been a CW dynamics study of R1 labeled mutants of cellular retinol-binding protein by Lietzow et al in 2004.[11] A variety of β-sheet environments were investigated including pleated sheet, twisted sheet, inner strand, and outer strand. Additionally, simple modeling was performed on a crystal structure of the labeled protein to find the conformations of R1 that would be spatially allowed. The results indicate that if the individual sites spatially allow for the preferred R1 conformations, then the χ_4/χ_5 model may hold true here as well. While this work does provide the essential first steps for understanding R1 motion in various β-sheet contexts, a more thorough analysis of R1 in this environment is essential for potentially extracting backbone motions from these sites as well as understanding rotameric preferences to aid in relating DEER results to protein structure.

1.2.2 Double electron-electron resonance provides protein structural constraints

Another important ESR experiment is double electron electron resonance (DEER).[16, 17] DEER is able to resolve distances within a protein system by measuring the distance-dependent dipolar interaction between two unpaired electrons. Thus, to perform a DEER experiment, a doubly spin labeled protein is required. Through a variety of R1 label pairs, a series of long range (15 Å – 100 Å) distance constraints can be generated and provide essential structural details about the protein.[18] Many complex protein systems that are difficult to investigate may be accessible with DEER if the system can tolerate two R1 labels. Accordingly DEER has been heavily utilized within many biologically complicated environments, in particular the membrane environment, with great success.[1, 2, 19]

Within a doubly labeled protein sample, each R1 site experiences rotameric sampling. Though DEER experiments are performed at cryogenic temperatures, flash freezing captures the spatial sampling of the label leading to a wide variety of R1 rotameric states. The distance is measured between the nitroxides and the locations of the nitroxide rings relative to the backbone attachment sites play a large role in the resultant distance distribution. Figure 1-4 illustrates this point. As the R1 rotameric state varies drastically at both labeling sites, the measured distances also vary leading to a broad distribution. As the desired result is the distance between the protein backbones relative to one another, i.e. the distance between the C_α atoms, this aspect of R1 based DEER measurements is highly problematic as it complicates data interpretation. While the distribution between the backbone sites is encoded in the DEER results, this information is difficult to extract.

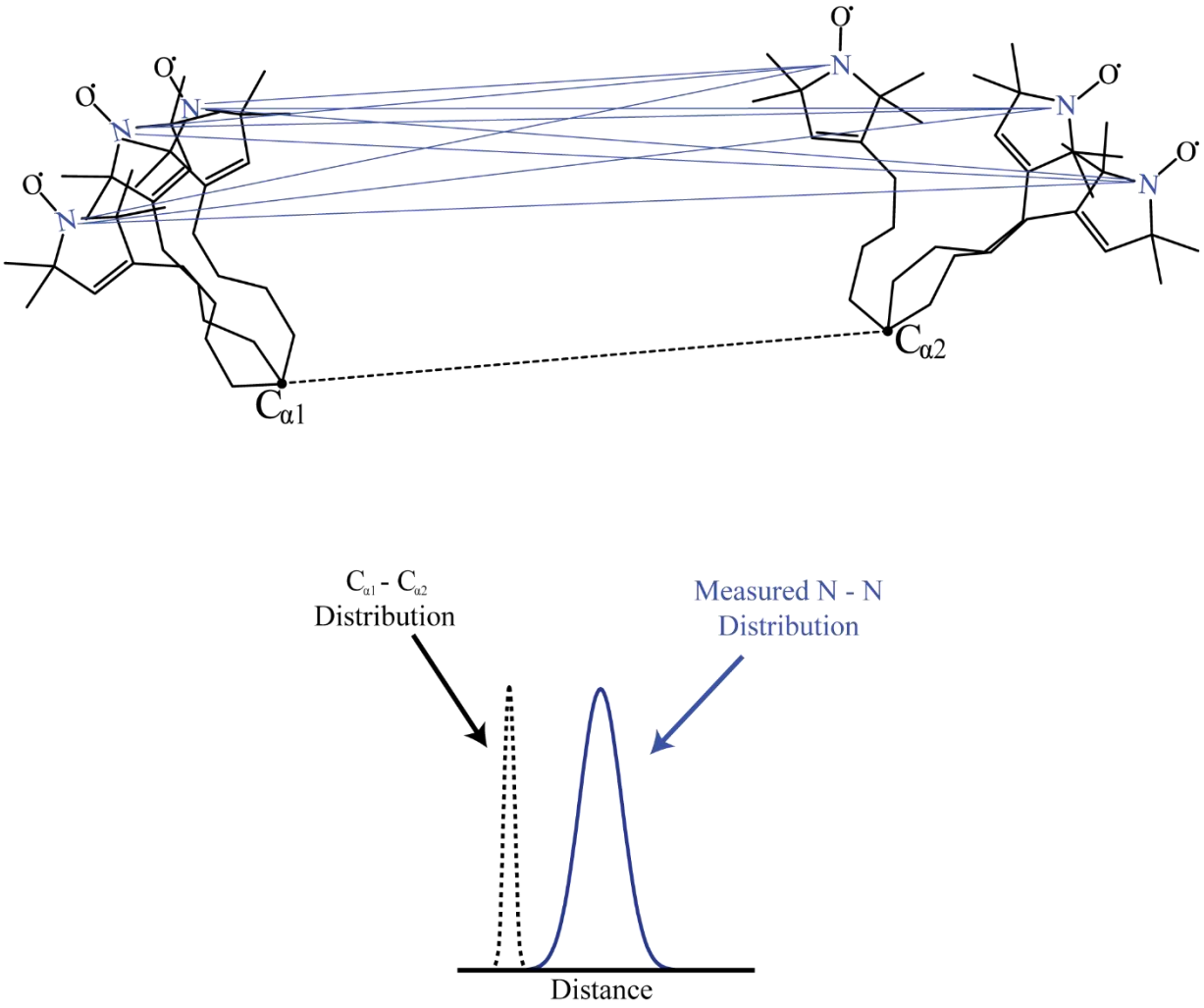

Figure 1-4. Cartoon representing the role R1 rotameric sampling plays in the interpretation of DEER distance distributions.

A common means of reducing the ambiguity associated with R1 distances is through triangulation. This is performed by choosing three labeling sites (for example A, B, and C), generating each possible combination of labeling sites into pairs (AB, AC, and BC), and measuring the distance distribution between each pair. Through this approach, R1 behavior at one site is assumed to be identical for both measurements (i.e. R1 at site A is behaving the same in the AB and AC measurements). Thus this approach can potentially decrease the ambiguity associated with the R1 sampling. Triangulation alone allowed for Jeschke et al. in 2005 to resolve conformational states of the highly flexible N-terminus of the light-harvesting chlorophyll a/b protein.[20] In 2006, Park et al. were impressively able to construct a model of the chemotaxis receptor-kinase complex using seven different labeling sites.[21] Also, Gaffney et al. in 2012 were able to triangulate a spin labeled lipid in soybean seed lipoxygenase-1 with five different R1 labeled sites.[22] Triangulation has also been combined with molecular modelling techniques to circumvent the issue of R1 flexibility. In 2008, Finiguerra et al. combined triangulation with simple steric modeling of R1 in a protein crystal structure of the protein azurin to gain insight into R1-distance based distribution widths.[23] Triangulation has also been combined with molecular dynamics (MD) modeling by Swanson et al in 2011[24] and by Sarver et al in 2012[25] to gain structural insight on large biomacromolecular complexes. Though this triangulation approach does provide a means of addressing the R1 rotameric flexibility, it commonly requires additional modeling to fully resolve the protein – protein distances.

Ideally, DEER should be able to report unambiguous distance distributions that can be simply interpreted and related to protein structure. One approach towards this goal has been to better understand R1 rotameric behavior. Through a variety of experimental and computational

techniques, R1 behavior has been thoroughly investigated in efforts to be better understand general behavior and rotameric preferences.

1.3 UNDERSTANDING R1 ROTAMERIC BEHAVIOR

Despite the many successful applications of R1 within protein systems, rotameric sampling of R1 continues to complicate the interpretation of DEER results. One approach to minimize this complication is to better understand R1 rotameric behavior. The secondary structure of the R1 labeled site can influence R1 rotameric sampling and thus different environments must be individually assessed. As past studies have excluded β-sheets from R1 analysis, this section focuses on the crystallography results and computational modeling methods that have been used to understand R1 at solvent-exposed sites on α-helices.

1.3.1 Observing R1 rotameric preferences with X-ray crystallography

X-ray crystallography has been used to explore rotameric preferences of the R1 spin label in a variety of protein environments.[12, 26-31] Rather than providing dynamics information, crystallography provides a means to directly observe rotameric preferences of R1. As is the case with the χ_4/χ_5 model, resolved R1 rotameric preferences may be indicative of dynamical modes of R1. Additionally, knowledge of R1 rotameric preferences can aid in relating R1 based DEER distance distributions to protein structure. Commonly, rotameric states of R1 are described by

the first two dihedral angles (χ_1 and χ_2) and the convention used to categorize these angles is the m (-60° ± 20°), p (+60° ± 20°), t (180° ± 20°) notation initially described by Lovell et al.[32]

Crystallography has thus far been used to observe R1 rotameric states in a variety of different protein contexts including both buried and solvent-accessible sites. At sites with R1 buried within the protein[26] or in the lipid environment,[29-31] the observed rotamers are highly dependent on the immediate environment. Additionally, there is adequate electron density to resolve full R1 structure at these buried sites suggesting a full immobilization of the side chain due to local sterics.

The majority of solved crystal structures that feature R1 in a solvent exposed environment have primarily been focused only on α-helical sites.[12, 27, 28] At these solvent-exposed α-helical sites, there is a distinct trend in the rotameric preferences of R1. With regards to χ_1 and χ_2, commonly either {m,m} or {t,p} is observed (Figure 1-5). Both of these conformations place the S_δ close to the H_α atom. Additionally, electron density beyond the S_δ was poorly resolved, which was rationalized as being due to rapid fluctuations about the final two dihedrals of the side chain, χ_4 and χ_5. These crystallographic observations led to the formulation of the χ_4/χ_5 model.[12] All other χ_1/χ_2 conformations result in the S_δ atom too far from the backbone thus precluding the presence of the hydrogen bond (Figure 1-5). However, these additional observed conformations provide essential information into site specific R1 rotameric preferences. Knowledge of rotameric preferences can be used to better relate DEER distance distributions to protein structure through modeling the rotamers into protein structures. While solvent-exposed α-helical contexts have been thoroughly investigated with crystallography, crystal resolved rotamers of R1 at solvent-exposed β-sheet sites is currently nonexistent.

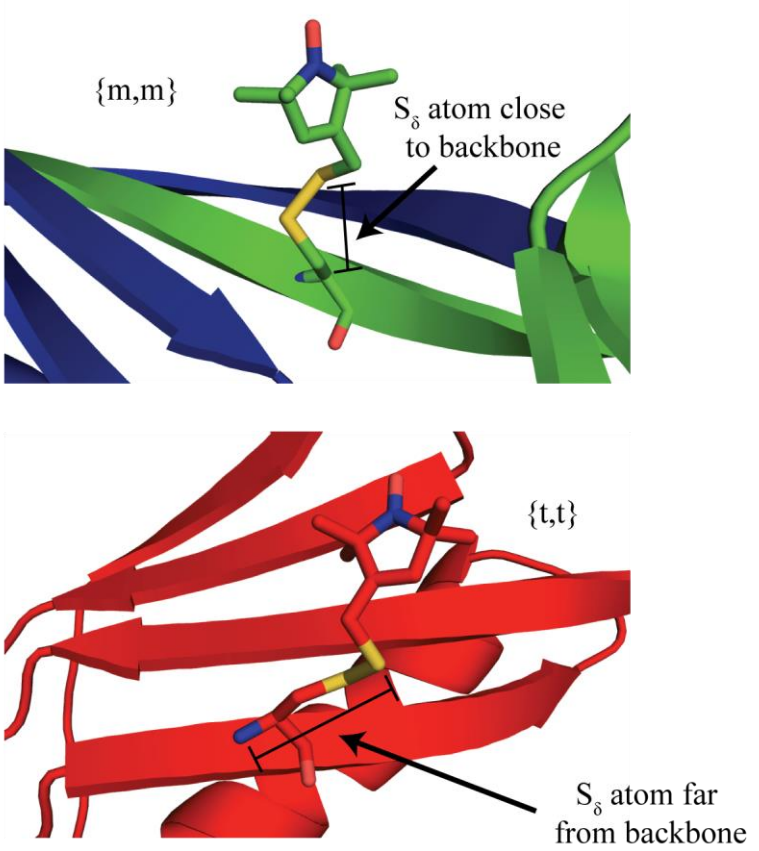

Figure 1-5. Crystallography allows for direct observation of R1 rotameric preferences in different contexts. The χ_1 = -60° and χ_2 = -60°, or {m,m}, conformation, is commonly observed and potentially allows for a hydrogen bond to form between the S_δ and H_α atoms. All other conformations, for example the χ_1 = 180° and χ_2 = 180°, or {t,t}, conformation, places these atoms too far apart for such an interaction to be present.

1.3.2 Computational efforts for predicting R1 rotameric behavior

In addition to crystallography, an alternate approach for exploring R1 rotameric preferences is through various computational methods. Ab initio[33] and density functional theory[34] calculations were performed on R1 in an α-helical environment to identify allowed conformations of R1 and calculate the energies of these rotameric states. Additionally, MD methods have been developed including a simulated scaling approach[35] and ab initio calculations were used to generate MD force fields for R1 at solvent-exposed α-helix sites.[36] These works provided theoretical comparisons of R1 that were directly compared with the results observed in the various crystals. While these efforts have provided essential information regarding R1 at solvent-exposed α-helix sites, most of these approaches are costly with regards to the necessary computation time. Accordingly, alternate computational approaches that can be performed quickly have been created for general use in predicting R1 rotameric preferences.

One such approach for simple prediction of R1 preferences is with the use of a rotamer library. The most common rotamer library is a collection of 200 of the most populated rotamers predicted in a 100 ns MD trajectory.[37] The library uses the relative populations of each rotameric state as calculated with MD. The rotamers are then modeled into selected sites of a crystal or NMR-determined structure. Using the relative populations and sterics with the environment, a population of viable rotamers is generated. When performed at two different sites, a distance distribution can be generated. For ease of use, this approach can be implemented with the Multiscale Modeling of Macromolecules (MMM) open source software package available for use with Matlab.[37] MMM is commonly used to model R1 DEER measurements and has been shown to be the method of choice when compared to MD due to

similar results and lower computational demands.[38] An alternative tool is MtsslWizard, a plugin available for use with the protein visualization software PyMol.[39] Rather than using conformational preferences to predict the location of the nitroxide nitrogen relative to the backbone C_α, MtsslWizard simply builds the spin label using randomly selected dihedral angles. If the created rotamer experiences no steric clashes, then the rotamer is allowed. The clash cutoff and the number of allowed rotamers generated can be adjusted by the user. Similar to MMM, if this modelling is performed at two locations, the resultant distance distribution can be calculated.

While these tools have yielded useful distance comparisons for numerous DEER measurements,[40-42] the methods once again focus primarily on R1 in solvent-exposed α-helix sites with the exception of MtsslWizard. The validity of using such modeling techniques in β-sheets has yet to be systematically tested.

1.4 DEVELOPMENT OF ALTERNATIVE PROTEIN SPIN LABELS

R1 rotameric uncertainty severely limits the interpretation of DEER results and relating those results to protein structure. No amount of experimental or computational work has been able to overcome this limitation in general. As such, many have focused their efforts on developing alternate spin labels that are thoughtfully designed to address these flexibility concerns. This section discusses the alternate spin labels that have been created thus far to address the limitations associated with R1.

1.4.1 Rigid spin labels for protein structure determination

The majority of alternate spin labels being created are labels that directly address the R1 flexibility concerns. The primary motivation of these labels is to either reduce the distance between the protein and the spin, reduce the flexibility of the spin relative to the protein, or both. Perhaps the most successful alternative spin label is the artificial amino acid TOAC (Figure 1-6a).[43] TOAC is arguably the most rigid commercially available nitroxide label with incorporation of both the C_α of the backbone and the nitroxide nitrogen within the same six membered ring. While this label has yielded numerous specific and unambiguous distance measurements, the label is currently restricted for use with synthesized peptides and proteins only.

Many alternative labels have been generated for general use with overexpressed proteins. Starting with the basic R1 structure, the label has been altered in many ways including the incorporation of bulky groups off of the nitroxide ring[6, 44] and changing the ring structure.[45] However, the most rigid spin label created for use with overexpressed systems is the bifunctional RX label (Figure 1-6b).[46] RX incorporates the same nitroxide ring but has two side chains coming off the ring that are connected to the protein via covalent bonding with two strategically placed cysteine mutations. While the label has shown great success in providing precise distance measurements in a number of protein systems,[47-51] incorporation of the label is experimentally more complicated. Though drastically more rigid as compared to R1, slight rotational freedom exists and as a result, some uncertainty in DEER interpretation remains.

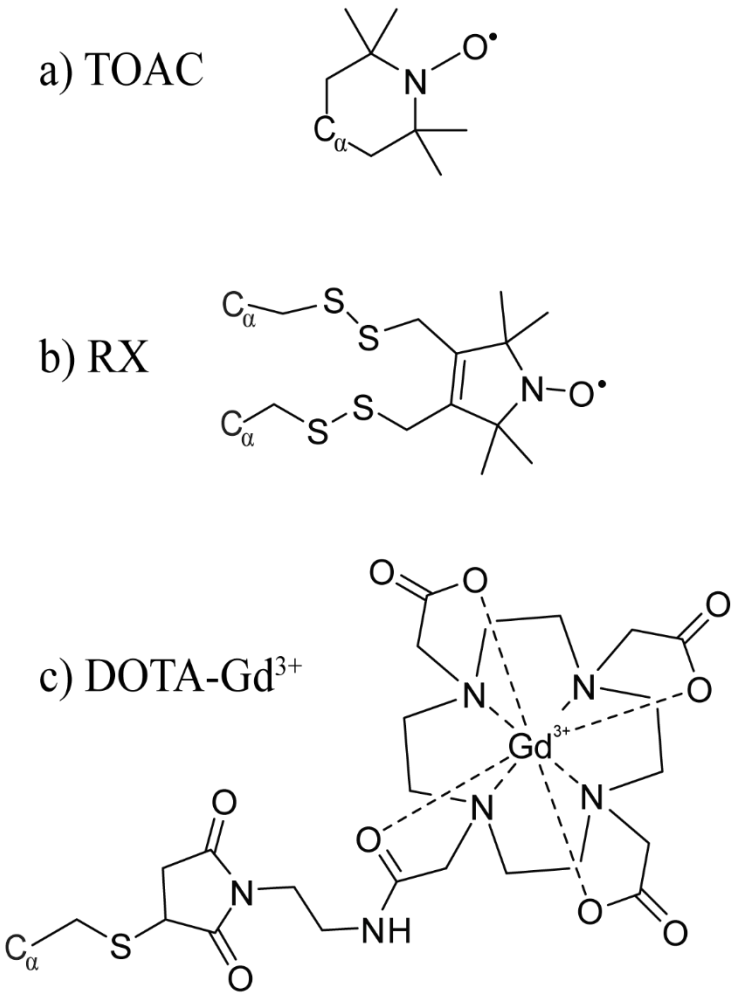

Figure 1-6. Three alternative spin labels used in place of R1. The most rigid is TOAC (a) but its use is limited to synthesized peptides. RX (b) is the most rigid that can be used in overexpressed proteins. Finally for use in cell is the DOTA-Gd^{3+} tag (c).

1.4.2 Paramagnetic metal tags for in-cell

An additional limitation of the R1 spin label has become apparent due to recent ESR efforts to measure protein distances within the cellular environment. Nitroxides are unstable within the cellular environment.[52] The nitroxide is known to have ~ 2 hour half life in some cells and thus any in cell work with nitroxides can be challenging.[53] The primary cause has been shown to be enzymatic degradation[53] and in turn introducing greater sterics surrounding the radical has led to greater stability in the cellular context.[54] Despite these efforts, it seems likely that the best spin label candidate for cellular studies are paramagnetic metals because of their well-known stability in this environment.[55] Primarily, the metal of choice has been Gd^{3+} which is commonly used as a contrast dye for magnetic resonance imaging in humans.[56] A number of site specifically incorporated chelating tags that chelate Gd^{3+}, such as the DOTA tag (Figure 1-6c),[57] have proven to be viable options for DEER distance measurements in solution,[58-61] in membranes,[57, 62] and in cells.[55, 63] While these tags do provide in cell stability which in turn allows for distance measurements to be performed, the flexibility concern remains. In many cases, the Gd^{3+} specific tags tend to exhibit greater flexibility as compared to R1.[59] Despite this, these metal chelating tags are making it possible for ESR to become increasingly used within the cellular environment.

1.4.3 Cu^{2+} as a paramagnetic probe for ESR distance measurements

Recently, our group[64-66] and others[67] have developed Cu^{2+} based DEER. Unlike Gd^{3+}, Cu^{2+} naturally binds to many proteins allowing for general use with these proteins.[68-72] For proteins that do not bind Cu^{2+}, binding motifs made up of natural amino acids can been

constructed through a series of point mutations.[73-76] Additionally, Cu^{2+} can be used on the much more common low field, X-band instruments while Gd^{3+} is better suited for the less common high field, W-band instruments. Cu^{2+} can also potentially be used to potentially resolve orientational differences between the two Cu^{2+} centers, thus providing additional biologically relevant information in beyond the distance distribution.[66, 70] In addition to natural and artificial Cu^{2+} binding sites, Cu^{2+}-specific chelating tags have also been used for NMR studies.[77-80] We have been inspired by these methodological advances to seek biochemical procedures for site specific attachment of Cu^{2+} labels in proteins.

2.0 HIGH-RESOLUTION STRUCTURE OF A PROTEIN SPIN-LABEL IN A SOLVENT-EXPOSED BETA-SHEET AND COMPARISON WITH DEER SPECTROSCOPY

This work, written in collaboration with Marshall S. McGoff, Ishita Sengupta, Christopher P. Jaroniec, W. Seth Horne, and Sunil Saxena, has been published in Biochemistry, 2012, V. 51, pages 6350-6359. The thesis author prepared samples, performed ESR experiments, analyzed data, and prepared the manuscript.

2.1 CHAPTER SYNOPSIS

X-ray crystallography has been a useful tool in the development of site-directed spin labeling by resolving rotamers of the nitroxide spin-label side chain in a variety of α-helical environments. In this work, the crystal structure of a doubly spin-labeled N8C/K28C mutant of the B1 immunoglobulin-binding domain of protein G (GB1) was solved. The double mutant formed a domain-swapped dimer under crystallization conditions. Two rotameric states of the spin-label were resolved at the solvent-exposed α-helical site, at residue 28; these are in good agreement with rotamers previously reported for helical structures. The second site, at residue 8 on an interior β-strand, shows the presence of three distinct solvent-exposed side-chain rotamers. One

of these rotamers is rarely observed within crystal structures of R1 sites and suggests that the H_α and S_δ hydrogen bond that is common to α-helical sites is absent at this interior β-strand residue. Variable temperature continuous wave (CW) experiments of the β-strand site showed two distinct components that were correlated to the rotameric states observed by crystallography. Interestingly the CW data at room temperature could be fit without the use of an order parameter which is consistent with the lack of the interaction between H_α and S_δ. Additionally, double electron electron resonance (DEER) spectroscopy was performed on the GB1 double mutant in its monomeric form and yielded a most probable interspin distance of 25 ± 1 Å. In order to evaluate the accuracy of the measured DEER distance, the rotamers observed in the crystal structure of the domain-swapped GB1 dimer were modeled into a high-resolution structure of the wild type monomeric GB1. The distances generated in the resulting GB1 structural models match the most probable DEER distance within ~ 2 Å. The results support the hypothesis that the rotameric states of R1 found in the crystal provide a very close match to the most probable distance measured by DEER.

2.2 INTRODUCTION

Site-directed spin labeling (SDSL) has emerged as an important biochemical technique that allows application of electron spin resonance (ESR)[13, 81-83] and nuclear magnetic resonance (NMR)[5, 84, 85] spectroscopy to explore protein structure and dynamics. A commonly utilized reagent in SDSL is the methanothiosulfonate nitroxide spin label (MTSSL), which reacts with cysteine residues to generate the side chain R1 (Figure 2-1).

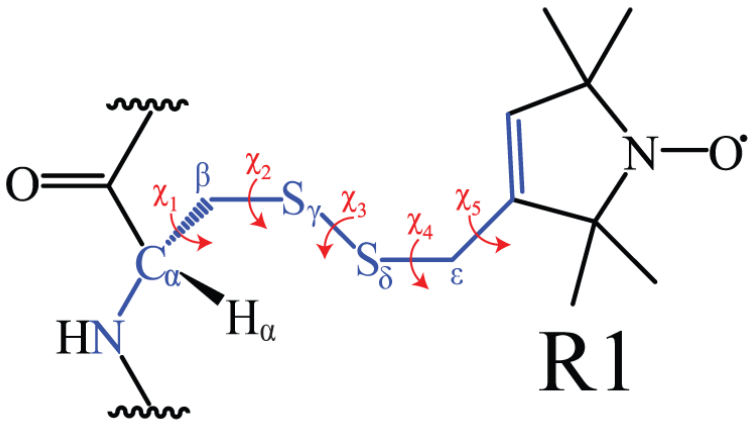

Figure 2-1. The structure of the R1 residue with side-chain torsions labeled.

The continuous wave (CW) ESR lineshape of R1 is used to infer information about the identity of protein secondary structure[7] and is often sensitive to the dynamics of the protein backbone.[6, 8] However, the deconvolution of spin label dynamics from protein backbone dynamics typically requires sophisticated analysis.[86-88] SDSL is also essential in the pulsed ESR technique, double electron electron resonance (DEER).[89, 90] DEER is used to explore protein structure and conformational changes through the measurement of distance distributions between two spin-labeled sites. However, this experiment does not directly report on the critical distance between the backbone C_α atoms and thus, a systematic method to derive this distance from the measured interspin distance is needed. In both CW and DEER ESR, complications arise due to the dynamics and/or conformation of the side chain itself. As a result, much work has been done to better understand the conformational preferences of the R1 side chain.

The conformation of the R1 spin label is defined by the five dihedral angles between the protein backbone and the nitroxide ring, denoted χ_1 through χ_5 (Figure 2-1). The preferred conformations of R1 have been investigated by using X-ray crystallography[12, 26-29] and various computational methods.[36, 91] Thus far, the focus has primarily been on spin-label rotamers in surface-exposed α-helical environments. One significant finding from these efforts is the existence of an interaction between H_α and S_δ (Figure 2-1),[12, 26-29] which is thought to reduce rotational freedom about the first two bonds (χ_1 and χ_2). Since rotation about the disulfide bond (χ_3) is known to be slow in solution, the majority of the R1 mobility has been accredited to rotation about χ_4 and χ_5 [this is collectively referred to as the χ_4/χ_5 model].[6] This model is further supported by the rotamers observed by crystallography. In solvent-exposed cases where local interactions are minimal, only χ_1, χ_2 and occasionally χ_3 are typically measured, as electron density is not often resolved beyond S_δ, suggesting a higher mobility about χ_4 and χ_5.[12, 27, 28]

Crystal structures of proteins with nitroxide rings containing bulky substituents are also available (PDB 1ZWN and 1ZUR). The χ_4/χ_5 model has been particularly useful in the interpretation of continuous wave (CW) ESR spectra. Given that the χ_4/χ_5 model suggests that the internal motions of the spin label are similar at all surface exposed helical sites, the site to site spectral changes in surface exposed helical sites can be directly attributed to fluctuations of the protein backbone. This was shown to be the case with the DNA binding protein, GCN4-58.[10] In cases where R1 is buried, local interactions dominate leading to more restricted motion and thus it is common that the entire side chain is fully resolved.[26, 29] Since χ_1 and χ_2 are the only consistently measureable angles throughout these studies, the conformation of **$\{\chi_1/\chi_2\}$** is commonly presented in the notation m(-60°±20°), p(+60°±20°), or t(180°±20°).[32]

As noted above, the focus of published R1 conformational studies to date has been on α-helices, including solvent-exposed,[27, 28, 92] buried,[26] and membrane-embedded sites.[29] In contrast, the spin-label conformations in β-sheet structures have been relatively unexplored. A previous dynamics study by Lietzow et al.[11] investigated R1 in a variety of solvent-exposed β-sheet environments (pleated vs. twisted sheet, interior vs. edge strand, etc.) using CW ESR. While thorough, many of the important conclusions concerning the **$\{\chi_1/\chi_2\}$** conformations used to interpret the ESR results were obtained by in silico modeling rather than by direct crystallographic observation. Additionally, for interior strands, it was concluded that the χ_4/χ_5 model would only hold true if the degree of strand twist was enough to sterically allow the H_α and S_δ hydrogen bonding to occur. Recently, the first crystal structure of a protein containing R1 in a β-sheet was published by Freed et al.[93] This work provides a thorough investigation of membrane effects on R1 when located on a membrane-embedded β-barrel. Interestingly, when R1 is present within the lipid environment, the spin label strongly interacts with hydrophobic

pockets on the protein surface which influences the resultant CW spectra. There is, however, no direct structural examination of R1 in solvent-exposed β-sheets, making it difficult to draw direct comparisons with the data set from Lietzow.[11]

The goals of this work are to investigate the conformations of R1 side chains in a solvent-exposed β-sheet, use these results to interpret distance distributions obtained by DEER spectroscopy, and assess the dynamics of R1 located at the β-sheet site using variable temperature CW measurements. To this end we used the 56 residue B1 immunoglobulin-binding domain of protein G (GB1). GB1 is commonly used as a model system in protein folding studies [94-96] as well as a variety of other applications[77, 97-99] due to its high stability[100, 101] and well understood structure.[100, 102-105] The GB1 mutant utilized here includes two spin labels, one on the α-helix and one on the β-sheet (Figure 2-2). The crystal structure of this mutant was solved and provided multiple rotameric states of R1 at both locations. DEER spectroscopy was used to measure the interspin distance for the doubly labeled GB1 mutant, allowing for a comparison with the crystal structure. This work provides the initial characterization of R1 rotamers in a solvent-exposed, interior strand, twisted β-sheet. Variable temperature CW analysis of the β-sheet site was also performed to assess the internal motion of R1 at the β-sheet site in light of the resolved rotameric states of R1 at this site. Additionally, a remarkable correspondence between the most probable DEER distance and the predicted distances obtained by modeling the resolved R1 rotamers into a high-resolution WT GB1 structure was observed. This constitutes a promising step towards a general approach for the elucidation of backbone $C_\alpha - C_\alpha$ distances in proteins by DEER spectroscopy.

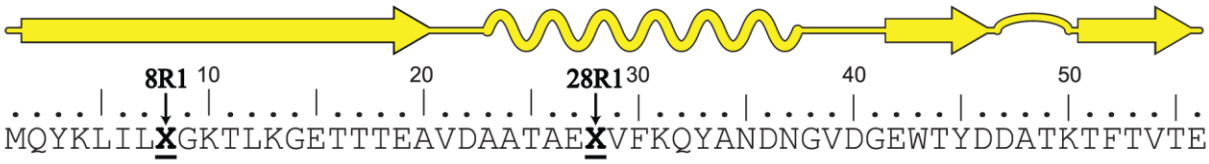

Figure 2-2. The secondary structure map and sequence of the spin labeled N8C/K28C GB1 mutant used in this study. The sites labeled with X (8 on a β-sheet and 28 on the α-helix) designate sites mutated to Cys and subsequently labeled with MTSSL.

2.3 METHODS

2.3.1 N8C/K28C GB1 Expression and Purification

The wild type GB1 plasmid was obtained as a gift from Prof. Angela M. Gronenborn (University of Pittsburgh) and used to create a plasmid encoding for the N8C single mutant and the N8C/K28C double mutant by previously described methods.[5] The plasmid was transformed into BL21(DE3) *E. coli* cells which were subsequently grown in 1 L of LB media with 100 μg/mL of ampicillin. The culture was grown at 37 °C until OD_{600} reached 0.6 to 0.8 at which point IPTG was added to a final concentration of 500 μM. Growth was continued for 4 hrs and the cells were pelleted by centrifugation. The pellet was resuspended in 20 mM Tris-HCl, pH 8.5, 5 mM NaCl (buffer A) with 5 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mg/mL lysozyme, 11 units/mL DNAse I, 1.3 units/mL RNase A and 0.1 % w/v Triton X-100. The resuspended cells were mixed on ice for 1 hr, then sonicated 3 x 30 s. The mixture was centrifuged at 20,000 rpm in a Sorvall SS-34 rotor for 30 min and the pellet was discarded. The supernatant was placed in an 80 °C water bath for 10 min, then again centrifuged. The supernatant was sterile-filtered (0.22 μm) and loaded onto a GE Healthcare HiTrap Q HP column equilibrated in buffer A. The [NaCl] was increased to 500 mM over 40 mL elution and individual fractions were checked with SDS-PAGE. The fractions containing GB1 were combined, treated with 10 mM TCEP, concentrated with Amicon 3,000 MWCO centrifugal filter, and further purified by size exclusion chromatography on a GE Healthcare Sephacryl S-100 26/60 GFC column equilibrated with 150 mM NaCl and 50 mM sodium phosphate, pH 6.5 (buffer B). The GB1 fractions were combined and concentrated to

160 μM (1 mg/mL) with 20 % (v/v) glycerol and 10mM TCEP then flash frozen in liquid nitrogen.

2.3.2 MTSSL Labeling

(1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSSL) was purchased from Toronto Research Chemicals. Labeling reactions were performed such that the final MTSSL:GB1 was 20:1. Thus, to label 1 mg (0.16 μmol) of GB1, 0.85 mg (3.2 μmol) MTSSL was dissolved into 100 μL DMSO, and diluted to 1 mL with buffer B. The N8C/K28C GB1 was thawed and run through a series of 5 x 5 mL GE Healthcare HiTrap desalting columns equilibrated with buffer B to remove TCEP and glycerol. The eluted protein was added directly to the premade MTSSL solution and allowed to react overnight. The mixture was concentrated and again run through the desalting columns to remove excess label. The eluted GB1 was concentrated to 0.25 mM for ESR studies or concentrated to 2.7 mM (17 mg/mL) for crystallographic studies.

2.3.3 Size Exclusion Chromatography

ESR samples were further checked for the possible formation of domain swapped dimers using a GE Healthcare Superdex 75 10/300 GL equilibrated with buffer B. 100 μL of 50 μM spin-labeled N8C/K28C GB1 was injected to the column using a 100 μL loading loop at a flow rate of 0.5 mL/min. The remainder of the 50 μM labeled protein was stored at 4 °C and run under

identical conditions for three consecutive days. WT GB1 was run under similar conditions as a basis of comparison.

2.3.4 Crystallization, Data Collection, and Structure Determination

Crystals of the nitroxide-labeled N8C/K28C double mutant of GB1 were grown by hanging drop vapor diffusion. A 17 mg/mL solution of protein in water was mixed (1 μL + 1 μL) with 150 mM sodium acetate pH 4.5, 18% w/v PEG 3350; the drop was allowed to equilibrate at room temperature over a well containing the crystallization buffer. A single crystal was flash frozen in liquid nitrogen after cryoprotection in well buffer supplemented with 20% v/v glycerol. Diffraction data were collected using CuKα radiation on a Rigaku/MSC diffractometer (FR-E generator, VariMax optics, AFC-Kappa goniometer, Saturn 944 CCD detector) equipped with an X-Stream 2000 low temperature system operated at 100 K. Raw diffraction data were indexed, integrated, and scaled with d*TREK. The crystal diffracted to 2.0 Å resolution with C2 symmetry and unit cell dimensions a = 105.3, b = 35.8, c = 86.99, β = 126.4°.

Structure solution and refinement were carried out using the CCP4 software suite.[106] The structure was solved by molecular replacement using the program Phaser[107] and a model derived from a published structure of wild-type GB1 (PDB: 2QMT).[108] Refinement was performed by a combination of Refmac[109] for automated refinement, Coot [110] for manual model building, and ARP/wARP [111] for automated model building. NCS restraints were applied during refinement for all four chains in the asymmetric unit. Geometric restraints for the nitroxide spin label were constructed from a published high-resolution small molecule crystal structure (CSD: 710750).[112] The final model was validated using the MolProbity[113] server and scored in the 87th percentile as

compared to 12522 published structures of 2.00 Å ± 0.25 Å resolution. Coordinates and structure factors have been deposited in the Protein Data Bank (PDB: 3V3X). Data collection and refinement statistics are listed in Table 2-1.

2.3.5 ESR Measurement and Analysis

All ESR experiments were performed on a Bruker Elexsys 580 spectrometer with a Bruker ER4118X-MD5 resonator. The temperature for all experiments was controlled with an Oxford ITC503 temperature controller and an Oxford ER 4118CF gas flow cryostat.

Approximately 5 μL of spin-labeled GB1 in buffer B and 30% w/v Ficoll 70 was drawn into quartz capillary tubes (1.0 o.d. x 0.8 i.d.) for CW measurements. Spectra were collected at an incident microwave power of 0.1995 mW. The modulation frequency and amplitude were set to 100 kHz and 1 G respectively. Spectra were collected within a temperature range of 293 K to 253 K in increments of 10 K. Additionally, a low temperature (80 K) CW spectrum was also collected.

The 80 K CW spectrum was fit using the EasySpin [114] program to find the principal A and g tensor values, which for 8R1 are A_{xx}=6.4 G, A_{yy}=5.7 G, A_{zz}=37 G and g_{xx}=2.0080, g_{yy}=2.0060, g_{zz}= 2.0020. The higher temperature spectra were fit with the microscopic order macroscopic disorder (MOMD) model using the NLSL program described by Budil.[115] The MOMD model is used to describe the anistropic motion of the R1 side chain with respect to the protein using a series of coordinate frames. The first frame is the magnetic frame of the nitroxide (x_m, y_m, z_m) where x_m lies along the N – O bond, z_m lies along the p orbital of the N, and y_m completes the right-handed frame. The second frame is the rotational diffusion frame.

Data Collection	
Unit cell dimensions (Å, °)	a = 105.3, b = 35.8, c = 86.99 $\alpha = \gamma = 90$; β = 126.4
Space group	C2
Resolution (Å)	27.3–2.0 (2.07–2.00)
Total observations	60,218
Unique observations	17,743
Redundancy	3.4 (2.7)
Completeness (%)	98.8 (94.8)
I/σ	15.5 (3.0)
R_{merge} (%)	4.9 (27.2)
Refinement	
Resolution (Å)	25.0–2.0
R (%)	23.1
R_{free} (%)	28.8
Avg. B factor (Å^2)	36.0
RMSD	
Bonds (Å)	0.018
Angles (°)	2.2

Table 2-1. X-ray crystal data collection and refinement statistics for N8R1/K28R1 GB1.

Spectra were fit assuming an axially symmetric rotational diffusion frame ($R_{\parallel}$ and $R_{\perp}$) using the average rotational time $\bar{R} = (R_{\parallel} R_{\perp}^2)^{1/3}$ and the asymmetry parameter $N = R_{\parallel}/R_{\perp}$. The rotational correlation time of R1 (τ) is equal to $1/6\bar{R}$. Transformation between the magnetic frame and the rotational frame is defined by three Euler angles, α_D, β_D, and γ_D. Fitting of the spectra here, as seen previously, was found to depend primarily on the angle β_D. Over the specified range of temperatures, β_D was assumed to remain constant. Previous work with solvent-exposed α-helices[6, 10] and β-sheets[11] has shown that holding β_D at 36° provided sufficient fits, however for the 8R1 spectra presented here, higher quality fits resulted when β_D was held at 32°. Additionally, the MOMD model includes an ordering potential (S_{20}) that is commonly utilized to fit spectra of R1 at solvent-exposed α-helical sites. Therefore only $\bar{R}$, N, and the Gaussian inhomogeneous line width, gib0, were varied for all fits with the priority of fitting the width of the central line and the low field peak.

For the DEER experiment, the pulse sequence used was $(\pi/2)\nu_1$-τ_1-$(\pi)\nu_1$-T-$(\pi)\nu_2$-τ_2-$(\pi)\nu_1$-τ_2-echo [90]. The pump frequency (ν_2) was placed at the maximum of the nitroxide spectrum and the observer frequency (ν_1) was offset ~67 MHz at the maximum of the low field component. The length of the $(\pi/2)\nu_1$ and $(\pi)\nu_1$ pulses were set to 16 ns and 32 ns respectively while the $(\pi)\nu_2$ pulse was set to 16 ns. The parameters τ_1 and T were set to 200 ns and 160 ns respectively, and T was incremented by a stepsize of 10ns for 256 points; τ_2 was adjusted such that $T + \tau_2 = 2760$ ns. The raw time domain DEER spectrum was analyzed using *DEER Analysis 2008.*[116]

2.4 RESULTS

2.4.1 Design of the Double Mutant

Crystallographic structure determination and ESR experiments were performed on an MTSSL spin-labeled N8C/K28C GB1 mutant. The K28 site in the wild type (WT) protein resides on an α-helix which allowed us to compare our results with literature. The N8 site is located on the β-sheet and serves to provide new insight into R1 rotamers on solvent-exposed β-sheets. The double mutant enables the use of DEER spectroscopy to acquire distance constraints that are compared with the crystallographic results.

2.4.2 N8R1/K28R1 GB1 Crystal Structure

The spin-labeled N8C/K28C GB1 double mutant crystallized as a domain-swapped dimer (Figure 2-3a), which was refined to 2.0 Å (Table 2-1). The asymmetric unit is comprised of two dimers for a total of four crystallographically independent GB1 chains. The four different monomers are labeled A, B, C, and D (shown in green, red, yellow, and blue respectively) and the dimers are paired A/C and B/D. Each of the four GB1 monomers has two R1 residues (Figure 2-1), one at site 8 on the β-sheet and the other site 28 on the α-helix (Figure 2-2). For some R1 residues, two different partial-occupancy rotamers were visible in electron density. Stereo views of the electron density surrounding each of the spin label rotamers on chains A and B are shown in Figure 2-4 and on chains C and D in Figure 2-5. The notation utilized to differentiate between the various R1 rotamers in the asymmetric unit is illustrated in Figure 2-3b.

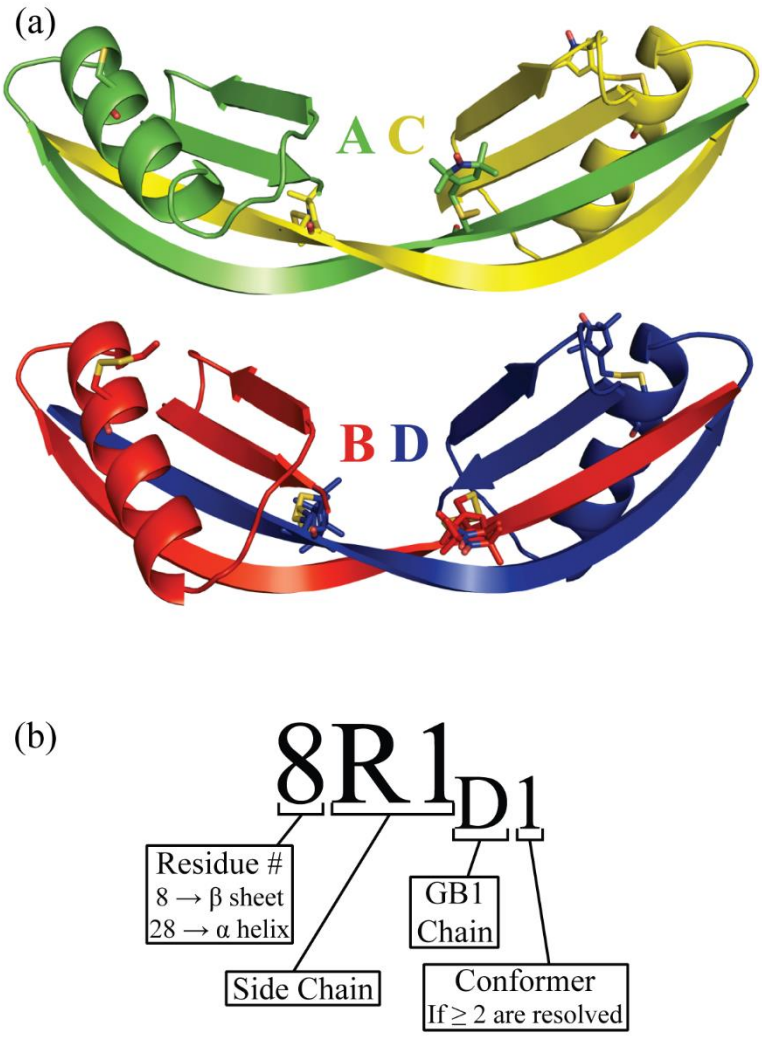

Figure 2-3. (a) Two crystallographically independent domain-swapped dimers observed in the crystal structure of the spin-labeled double mutant. Chains A, B, C, and D colored in green, red, yellow, and blue respectively. (b) Key for the notation used to identify R1 rotamers in the text.

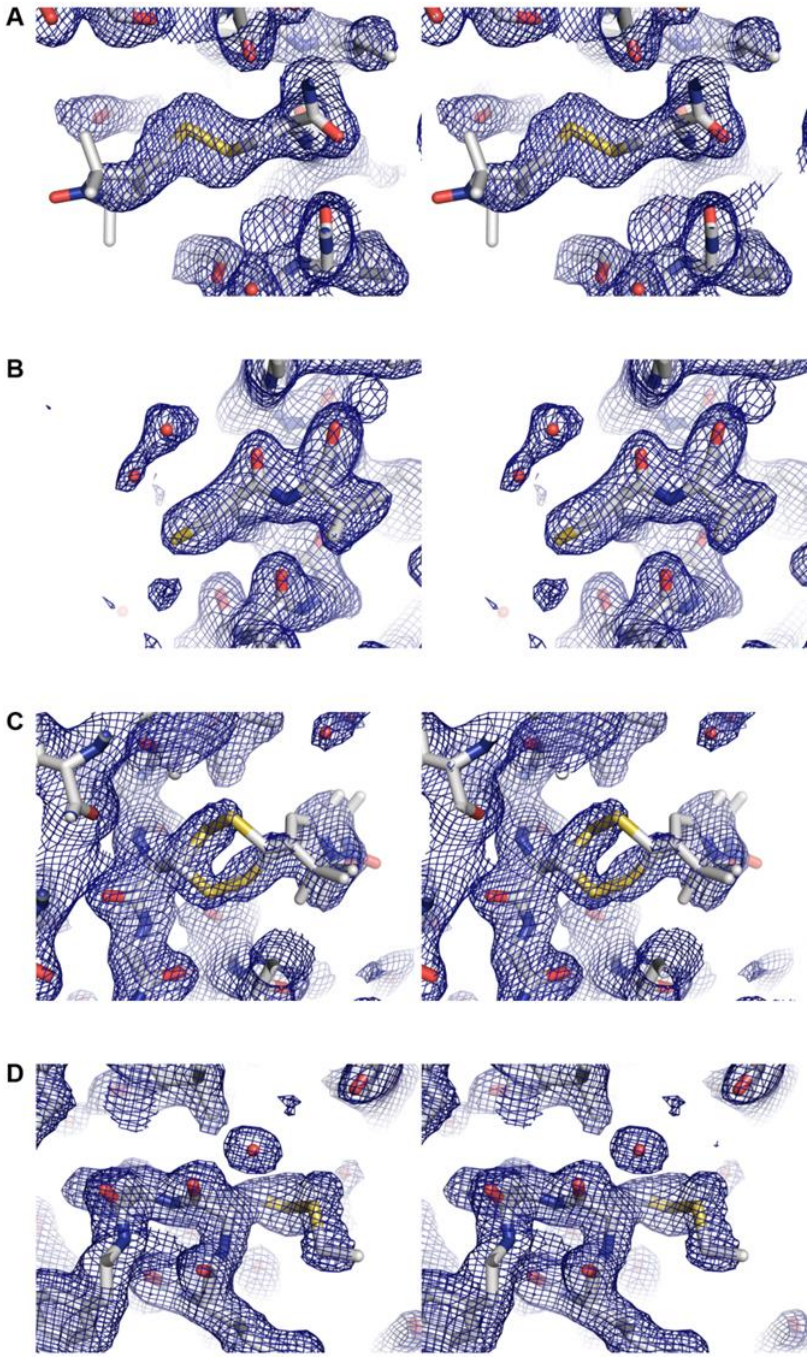

Figure 2-4. Stereo views of electron density around the nitroxide modified residues from the refined structure of the N8R1/K28R1 GB1 double mutant: (A) chain A, residue 8; (B) chain A, residue 28; (C) chain B, residue 8; (D) chain B, residue 28. The maps depict σa weighted 2Fo−Fc electron density contoured at 1.0 σ.

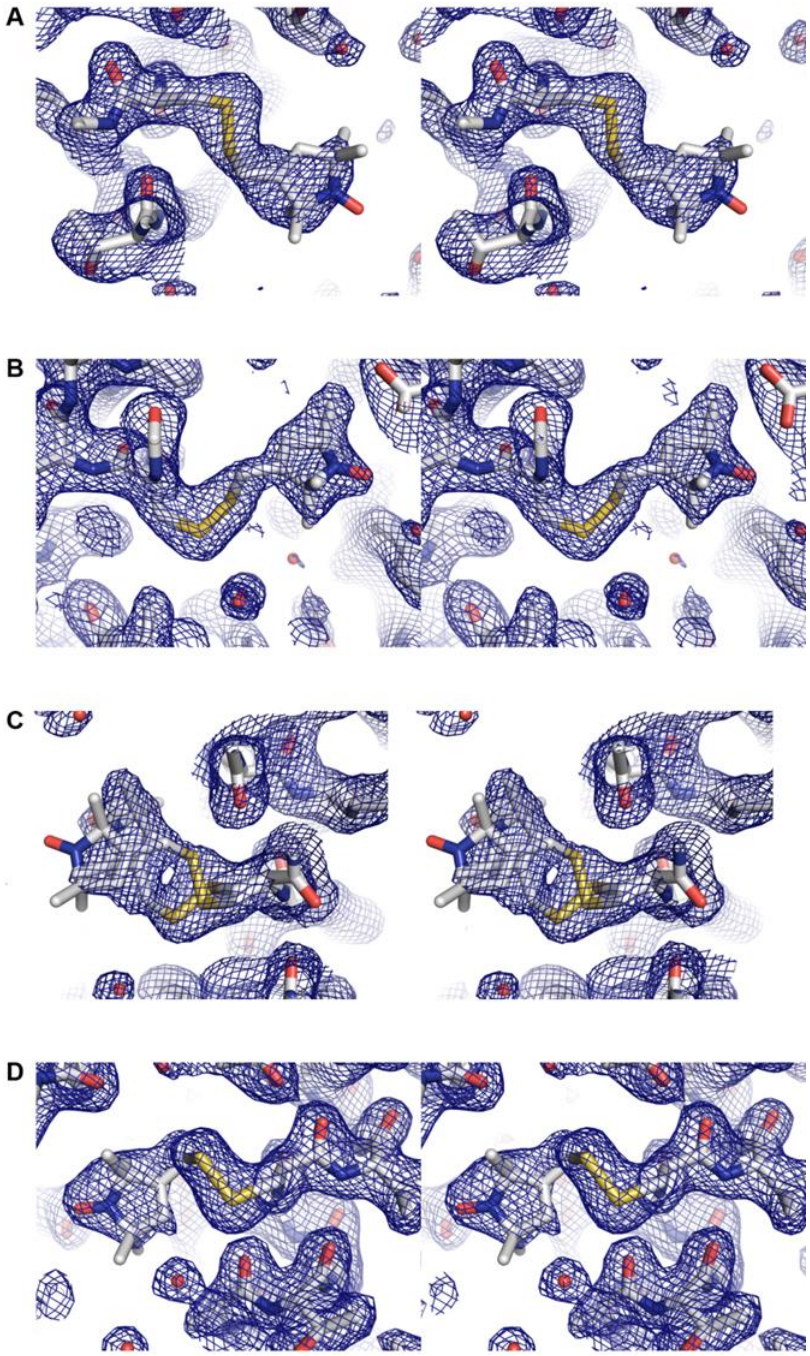

Figure 2-5. Stereo views of electron density around the nitroxide modified residues from the refined structure of the N8R1/K28R1 GB1 double mutant: (A) chain C, residue 8; (B) chain C, residue 28; (C) chain D, residue 8; (D) chain D, residue 28. The maps depict σa weighted 2Fo−Fc electron density contoured at 1.0 σ.

Mutant	Rotamer	χ_1	χ_2	χ_3	χ_4	χ_5	C_α—S Dist. (Å)	Solvent Exposed	Structure
8R1$_A$	{**m,-120**}	-57	-120	-88	-109	61	4.2	No	β-sheet
8R1$_{B1}$	{**m,t**}	-60	168	85	68	12	3.7	No	
8R1$_{B2}$	{**t,p**}	-167	74	92	169	102	4.4	No	
8R1$_C$	{**m,m**}	-56	-74	-76	-93	77	3.6	Yes	
8R1$_{D1}$	{**m,t**}	-59	162	72	61	-94	4.6	Yes	
8R1$_{D2}$	{**m,m**}	-63	-58	-83	-86	74	3.4	Yes	
28R1$_A$	{**m, - **}	-61	-	-	-	-	-	No	α-helix
28R1$_B$	{**m,m**}	-61	-63	112	148	-	3.5	No	
28R1$_C$	{**m,m**}	-60	-61	96	140	90	3.3	Yes	
28R1$_D$	{**m,m**}	-58	-66	97	81	-73	3.5	Yes	

Table 2-2. Spin label side chain dihedral angles resolved for all rotamers in the N8R1/K28R1 crystal structure. The C_α—S distance is listed to show which rotamers contain the H_α—S hydrogen bond. The environment of each mutant was individually investigated to assess the validity of referring to each as a solvent-exposed site and the results are shown.

2.4.3 Spin – Label Rotamers

All R1 rotamers are defined by the five dihedral angles, χ_1 through χ_5 (Figure 2-1). In total, there are six β-sheet (8R1) and four α-helix (28R1) rotamers which were at least partially resolved in the crystal structure. The values of the dihedral angles χ_1 through χ_5 for all rotamers are displayed in Table 2-2. As a means of comparison with previous work,[11, 27] the **{χ_1,χ_2}** conformation for all rotamers are also presented in Table 2-2 in the m, p, and t notation (-60° ± 20°, +60° ± 20°, and 180° ± 20° respectively).[32]

At the four 28R1 α-helical sites in the asymmetric unit, only two of the four side chains, $28R1_C$ and $28R1_D$, are fully resolved. In this nomenclature, C and D refer to the GB1 chain on which the rotamer is located. Both rotamers have {χ_1,χ_2} conformations of {**m,m}** type. The C_α to S_δ distance for $28R1_C$ and $28R1_D$ is 3.3 Å and 3.5 Å respectively. These distances suggest the possible presence of a C_α—H_α• • •S_δ hydrogen bond.[28] The **{m,m}** conformation and the presence of this hydrogen bond have been reported previously for rotamers located within a solvent exposed, α-helix environment.[27] Thus, our results are reasonable given the location of 28R1 within GB1. In addition, the disulfide dihedral angle (χ_3) for both $28R1_C$ and $28R1_D$ is ~+90°. This agrees with previous computational results that χ_3 of either +90° or -90° are commonly populated.[36, 91] Previously published R1 residues in a solvent-exposed α-helix with a **{χ_1,χ_2}** conformation of **{m,m}** show a value of -90° for χ_3.[27]

Four unique 8R1 sites exist within the crystal asymmetric unit, and two of these sites (on chains B and D) contain two unique partially-occupied R1 rotamers. Therefore, there are a total of six different β-sheet rotamers (Table 2-2). In order to determine which positions were good models for a "solvent-exposed" β-sheet, the environment of each site was individually examined.

The $8R1_A$ and $8R1_B$ side chains appear to be perturbed by local contacts. For $8R1_A$ (Figure 2-6a), residue $Y33_A$, from a symmetry copy of chain A is in close proximity to the nitroxide (~3.5 Å). At site $8R1_B$ two resolved R1 rotamers were observed ($8R1_{B1}$ and $8R1_{B2}$ in Figure 2-6b). Both $8R1_B$ rotamers point into the same physical space as the $28R1_A$ rotamer. Although $28R1_A$ is not fully resolved, it is likely that $28R1_A$ and $8R1_B$ are interacting with each other especially given the proximity of the S_γ from $28R1_A$ to both of the $8R1_B$ rotamers (~5.1 Å). For these reasons, $8R1_A$, $8R1_{B1}$, and $8R1_{B2}$ were considered to be perturbed by the non-native environment surrounding each site and thus excluded from further analysis.

The environments for $8R1_C$, $8R1_{D1}$ and $8R1_{D2}$ show no local contacts arising from crystal lattice artifacts and can therefore be characterized as "solvent exposed". Residues from chains A and C form an open pocket around $8R1_C$ (Figure 2-7a). A similar pocket formed by the residues from chains B and D is observed for the rotamers $8R1_{D1}$ and $8R1_{D2}$ (Figure 2-7b). The environments of these rotamers can be further categorized as an interior strand on a twisted sheet (Figures 2-7a and 2-7b). The **{χ_1,χ_2}** conformations of these three solvent-exposed, interior strand, twisted β-sheet rotamers are **{m,m}** ($8R1_C$ and $8R1_{D2}$) and **{m,t}** ($8R1_{D1}$) (see Table 2-2) and an overlay of all three rotamers is shown Figure 2-7c.

2.4.4 8R1/28R1 GB1 DEER Data and Comparison with Crystal Structure

DEER spectroscopy was performed to examine the interspin distance in the context of side chain packing trends. DEER resolves the interspin distance distribution between two spin labeled sites. However, it is important to know whether the DEER results are obtained from dimerized GB1 (as in the crystal) or GB1 monomers.

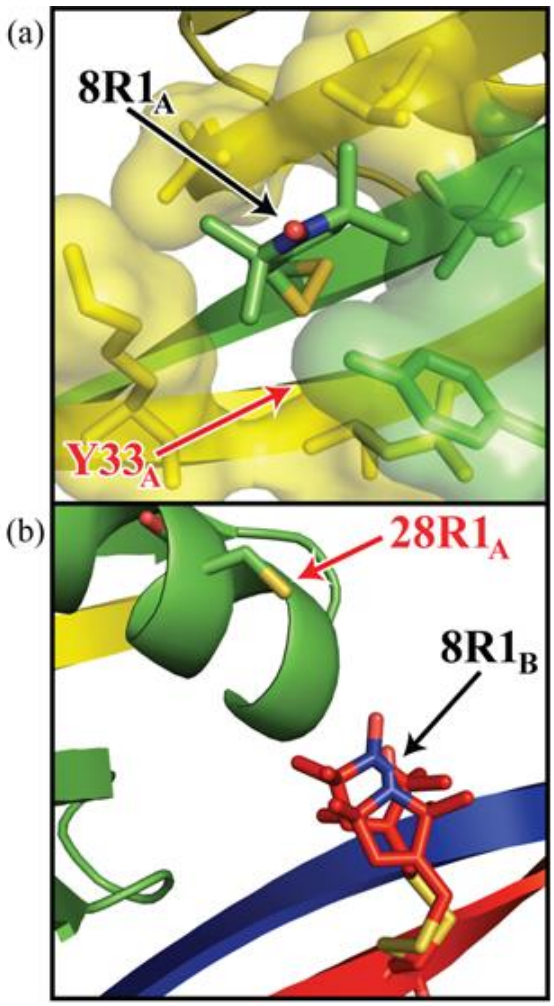

Figure 2-6. The local environments for $8R1_A$, $8R1_{B1}$ and $8R1_{B2}$. (a) For $8R1_A$, residue $Y33_A$ from a crystallographic symmetry mate is located within 3.5 Å of the nitroxide and is likely perturbing this rotamer. (b) For $8R1_B$, the $28R1_A$ rotamer is pointing directly into the same physical space as both $8R1_B$ rotamers. Although $28R1_A$ is not fully resolved, the 5.1 Å distance between the $28R1_A$ S_γ and $8R1_B$ makes it highly likely that the movement/conformation of both sites are influenced by each other.

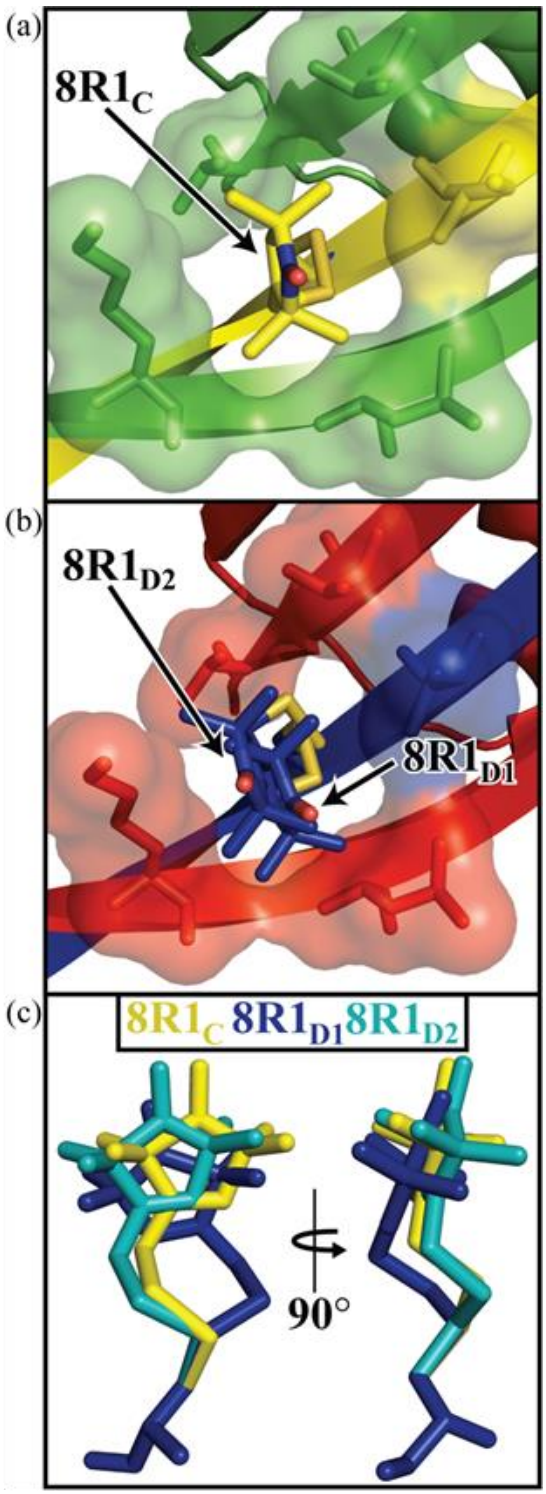

Figure 2-7. The local environments for (a) $8R1_C$, (b) $8R1_{D1}$ and $8R1_{D2}$. All three rotamers shown above are undisturbed by any protein contacts arising from the crystal lattice. (c) An overlay of the three 8R1 rotamers from (a) and (b).

In the crystal, the doubly labeled GB1 mutant formed a domain swapped dimer, but this is not the native structure of wild type GB1. This is not unexpected, since certain GB1 mutants have shown a propensity for formation of domain-swapped dimers; however, the specific mechanism here (interchange of the N-terminal strands between two chains) has not been previously observed for GB1.[117, 118] The closest literature precedent for the fold we observed is swapping of the C-terminal β-strand in the B1 domain of protein L.[119, 120] Figure 2-8a shows the structure of a single WT GB1 molecule, as determined by solid-state NMR [PDB ID 2LGI][105] and Figure 2-8c shows the domain-swapped dimer structure presented here. Figure 2-8b shows the structure of a single GB1 chain within this specific domain swapped dimer. The first strand of the β-sheet (upon which 8R1 is located) is swapped with the same domain from the second GB1 monomer resulting in this specific domain-swapped dimer structure. Size exclusion chromatography was used to investigate the oligomeric state of GB1 in solution under buffer conditions used for the DEER measurements. After three days in the same buffer conditions as the DEER measurements, the elution volume for the mutant was consistent with that of the GB1 monomer (Figure 2-9), indicating that the formation of the domain-swapped dimer is a result of the specific crystallization conditions employed in this study.

The results of the DEER experiment are shown in Figure 2-10. The main panel shows the baseline subtracted time domain data (black) with the best fit (red) from *DEER Analysis 2008.*[116] The inlay in Figure 2-10 shows the resulting distance distribution (red). These results further support the existence of only monomers and no dimers within the ESR sample. If dimers were present, multiple distances would be observed (as there would be four different labels within close proximity) but in this case, only one is observed. Also, the modulation depth of the time domain signal is consistent with only two spins per species:[121]

Figure 2-8. The primary sample within this study is N8R1/K28R1 GB1. (a) The native monomer fold of GB1 (PDB: 2LGI); the two labeled sites are shown in yellow (N8R1 on the β-sheet) and green (K28R1 on the α-helix). Exchange of the N-terminal β-strand between two chains (b) results in the domain-swapped dimer observed in the crystal structure (c).

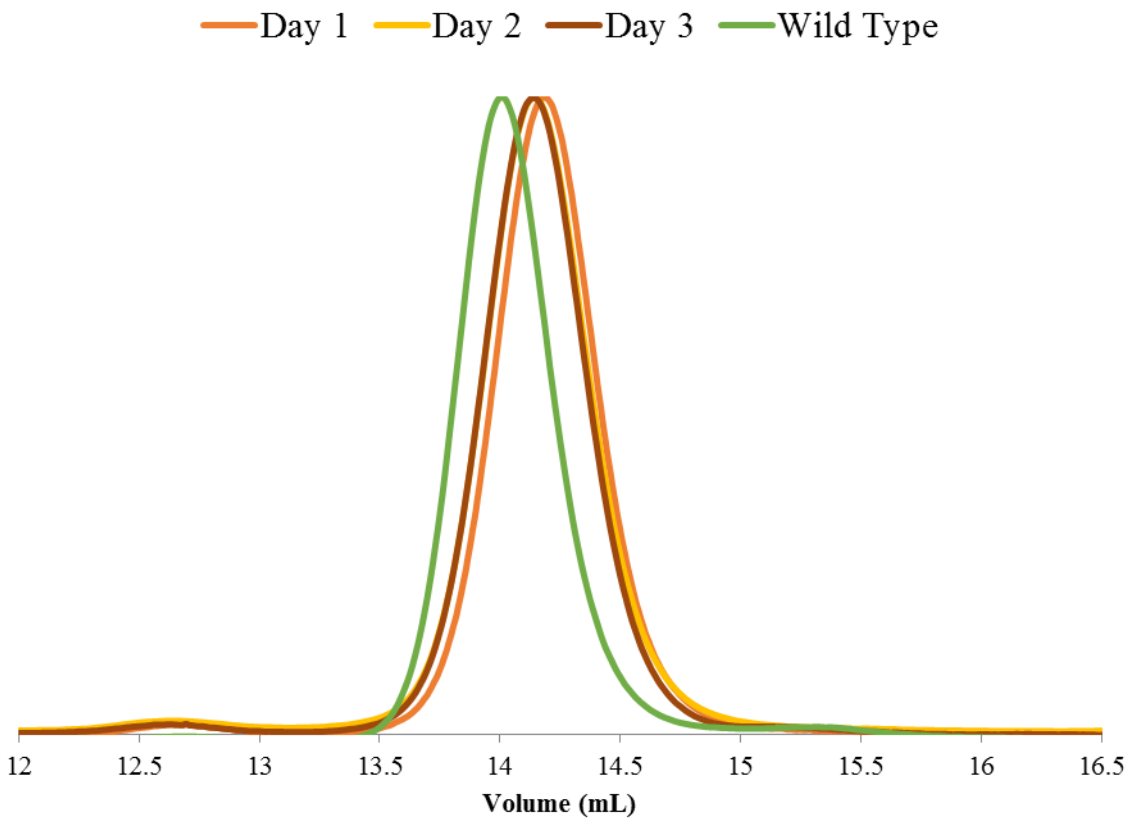

Figure 2-9. Overlay of chromatogram over the period of three days of the doubly labeled N8C/K28C GB1 mutant.

$$V_\lambda = (1 - \lambda_B)^{(n-1)} \qquad (2\text{-}1)$$

Where V_λ is the modulation depth from the DEER experiment (0.6 as seen in Figure 2-8), λ_B is the fraction of B spins inverted by the pump pulse (for a pump pulse of 16 ns, this value was determined to be 0.4), and n is spins per cluster. Rewriting equation 2-1 to solve for n:

$$n = \frac{\ln V_\lambda}{\ln(1-\lambda_B)} + 1 = \frac{\ln(0.6)}{\ln(0.6)} + 1 = 2 \qquad (2\text{-}2)$$

It is clear in equation 2-2 that the number of spins per cluster within the DEER sample is 2. If the mutant had formed a domain swapped dimer, this value would be 4 since each monomer contains 2 spins each.

DEER was also performed on the doubly labeled GB1 mutant in the crystal buffer conditions (150 mM sodium acetate pH 4.5, 18% w/v PEG 3350) as a means to further investigate the formation of the dimer. The result (data not shown) is very similar to the data shown in Figure 2-10 and thus the buffer did not play a role in the dimerization.

The distance distribution between the two spin labels in the monomeric GB1 shows a most probable distance of 25 Å. In the high-resolution structure of monomeric WT GB1 (PDB ID 2LGI), the N8 C_α to K28 C_α distance is 18 Å.[105] Given this distance and the approximate length of the R1 chain, a 25 Å interspin is reasonable for monomeric GB1. The DEER data was also compared to the modeling program by Polyhach and Jeschke,[122] and the model predicts the same most probable distance of 25 Å, but the overall distribution was narrower.

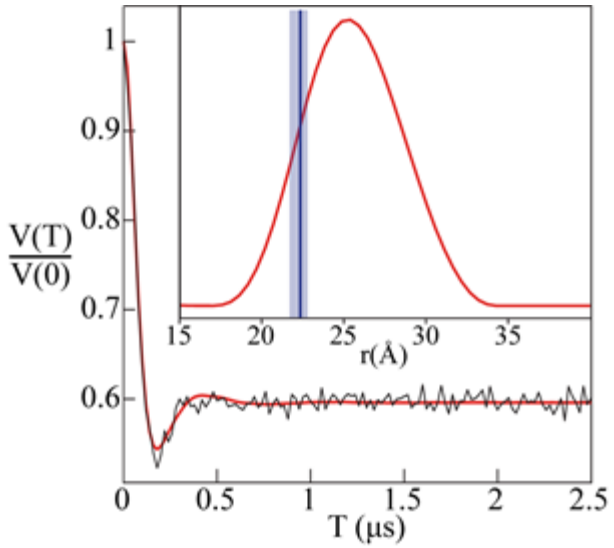

Figure 2-10. The baseline subtracted time-domain DEER data from the N8R1/K28R1 GB1 sample (black) with the best fit from *DEER Analysis 2008* (red). The inset figure is the resulting distance distribution (red) with an average distance of 26 Å and a most probable distance of 25 Å (suggesting a naturally folded GB1 sample). As a means of comparison between this distribution and the resolved rotamers, the 8R1 and 28R1 rotamers were modeled into a NMR WT GB1 structure (PDB ID 2LGI shown in Figure 2-8a) at the appropriate locations. The vertical line (blue) represents the average nitroxide N—N distance between the 8R1 β-sheet rotamers and the 28R1 α-helix rotamers (22.4 Å) and the light blue is the full range of generated distances.

2.4.5 8R1 Variable Temperature CW ESR

As a means to explore the dynamics that coincide with rotamers described above for 8R1, variable temperature CW spectra were collected for the 8R1 single mutant. Buffer conditions for the single mutant were identical to that of the DEER sample (buffer B) except for the addition of 30% w/v Ficoll 70. With 30% w/v Ficoll 70 the rotational correlation time of global tumbling of GB1 (≈ 80ns) does not substantially contribute to the CW ESR spectral lineshapes.[10] The spectra were collected over the range of 253 K to 293 K in increments of 10 K. The resultant spectra with best fits from the MOMD model are shown in Figure 2-11 and the parameters for the best fits are listed in Table 2-3.

At temperatures between 293 – 273, the 8R1 CW spectra contain only a single highly mobile component. An additional component, which exhibits lower mobility, is observed at lower temperatures. At 293 K the spectrum could be fit with a rotational correlation time, τ, of 2.1 ns, and an order parameter, S_{20}, of 0. The order parameter and rotational correlation times are, however, correlated and we could obtain similar quality fits with τ in the range of 2.0 – 2.1 ns by adjusting S_{20} in the range of 0 – 0.11. These parameters are indicative of highly mobile (low τ) and, at best, a very weakly-ordered motional behavior for R1 at this site. The spectra collected at 263 K and 253 K contained two components. The main component was consistent with the motion described above with $\tau \approx 4.1$ ns and 4.6 ns, respectively. At these values of τ the order parameter was 0. However a second component, with a much slower rotational rate, was necessary for a quality fit. Given the low population of this component (~15-19% at 263 K), the spectral features could be simulated with a range of correlation times and order parameters for the second component. At 263 K, for example, we could obtain similar quality fits with τ in the

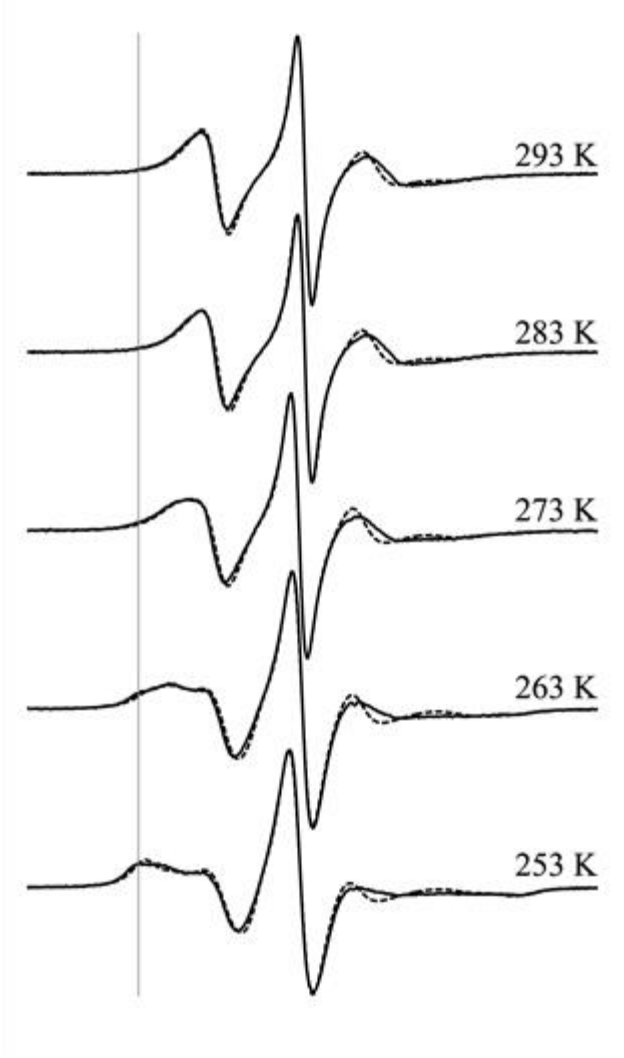

Figure 2-11. Variable temperature CW spectra for the N8R1 single mutant in buffer B with 30% w/v Ficoll 70 (solid) and the fits (dotted) using the MOMD model. For all spectra, β_D was held constant at 32°. For the spectra at 263 K and 253 K, a second component was necessary for fitting and the vertical gray line is a guide for visualizing the second component. All simulations shown in the figure are with $S_{20} = 0$. Adequate fits to the second component could, however, be achieved using a range of order parameters (see text).

Temperature (K)	N	τ (ns)	S_{20}	Occupancy (%)
293	25	2.1	0	100
283	25	2.2	0	100
273	25	2.9	0	100
263 comp 1	25	4.1	0	81
263 comp 2	20	14.2	0	19
253 comp 1	25	4.6	0	62
253 comp 2	20	19.6	0	38

Table 2-3. The parameters used to simulate the N8R1 single mutant CW spectra. The fits are shown in Figure 2-11. Note that a range of order parameters could be used for the second component (see text).

range of a 13.2 – 21 ns by adjusting S_{20} in the range of 0 – 0.7 as well as the relative populations of the two components.

2.5 DISCUSSION

The goals of this study were to resolve the R1 side chain in both an α-helix and β-sheet using X-ray crystallography and to compare these results with distances determined by DEER spectroscopy. Additionally, the β-sheet site was further investigated using variable temperature CW to assess the dynamics of R1. The first goal was accomplished by spin labeling the N8C/K28C double mutant of GB1 and solving the crystal structure. The structure that resulted was a domain-swapped dimer. While this result was unexpected, it offers multiple advantages. First, we were able to obtain twice as many resolved R1 residues as compared to the natively folded GB1 structure. In addition, the asymmetric unit of this crystal provided two dimers and therefore a total of four different GB1 chains with resolved R1 residues. This greatly increased the probability of finding fully resolved residues that were not perturbed by crystal lattice contacts, which in this case led to two α-helix rotamers ($28R1_C$ and $28R1_D$) and three β-sheet rotamers ($8R1_C$, $8R1_{D1}$ and $8R1_{D2}$). The α-helix rotamers allow the direct comparison with literature and the β-sheet rotamers offer new insight into the spin label conformations in a β-sheet context. An additional advantage that came about with the dimer was the β-sheet environment where 8R1 is located. In monomeric GB1 (Figure 2-8a), the 8R1 site (yellow) lies very close to a β-turn. Within the dimer (Figures 2-3a, 2-8b and 2-8c), the 8R1 site is located in

the middle of a lengthy β-sheet, far away from any turns. This removes the possibility that the 8R1 rotamers are biased by the proximity to a β-turn.

The β-sheet (8R1) rotamers found here are categorized as solvent-exposed rotamers on an interior strand of a twisted β-sheet and have either an **{m,m}** ($8R1_C$ and $8R1_{D2}$) or an **{m,t}** ($8R1_{D1}$) **{χ_1,χ_2}** conformation. An assessment of the environment is important due to the previous dynamics study by Lietzow et al in which a similar environment was investigated on cellular retinol binding protein (CRBP) using CW ESR methods.[11] Using the ESR data as a starting point, they modeled the R1 side chain into the respective locations on the WT CRBP crystal structure. They predicted that in the solvent-exposed, interior strand, twisted sheet environment, the two most likely conformations for R1 would be **{m,m}** and **{m,t}**. The unperturbed 8R1 rotamers presented here match these predictions. Another interesting correlation with Lietzow's work is the H_α—S_δ hydrogen-bond that is probable for many of the rotamers. For both **{m,m}** 8R1 rotamers, this bond is likely present (as indicated by the relatively short C_α—S_δ distance in Table 2-2) while in the **{m,t}** rotamer, the C_α—S_δ distance is too long and precludes the existence of a H_α—S_δ hydrogen-bond. This trend matches well with Lietzow's observations in that **{m,m}** conformations are able to form this hydrogen-bond leading to more restrained movement. However, **{m,t}** conformations are unable to form this hydrogen-bond and are therefore more mobile.[11]

The rotamers $8R1_C$, $8R1_{D1}$, and $8R1_{D2}$ were found to be viable solvent-exposed, interior strand, twisted β-sheet rotamers. This work is important due to the general lack of X-ray crystallography based, β-sheet R1 rotameric studies. The recent study by Freed et al[93] thoroughly investigated β-sheet rotameric states of R1 within membranes. The membrane-embedded environment of Freed's study differs from the solvent-exposed environment

investigated here. Such differences in environment are important given that the **{χ_1,χ_2}** conformations of the rotamers found in that study are different (**{t,m}** and **{p,p}** vs. **{m,m}** and **{m,t}**).

While the rotameric states of R1 observed in the crystal provide new structural information for R1 in β-sheets, the variable temperature CW spectra of 8R1 allow analysis of the R1 side chain dynamics at this site. The spectra observed over the measured temperature range show a main component of very fast, non or weakly ordered motion with a second less mobile component becoming more obvious at low temperatures (Figure 2-11).

At solvent-exposed sites on α-helices, an order parameter, S_{20}, is commonly needed to fit the spectra in the context of the MOMD model.[6, 10, 28] The order parameter reflects the restriction in amplitude of the spin label motion due to the backbone. The need for S_{20} is commonly attributed to the ordering that results from the $H_\alpha - S_\delta$ interaction as described by the χ_4/χ_5 model. Therefore for solvent-exposed α-helical sites where the internal motion of R1 is well understood, decreased order is interpreted as an increase in backbone fluctuations. In the higher temperature spectra for 8R1 in GB1, S_{20} is not necessary for quality fits using the MOMD model. For the 8R1 site, the relatively low crystallographic B-factor C_α suggests very limited fluctuation of the backbone. Therefore the most likely cause for the lack of order is that the $H_\alpha - S_\delta$ hydrogen bond is not present at this site. Thus, this component is likely due to the **{m,t}** rotamer observed in the crystal structure for $8R1_{D1}$ in which R1 is geometrically unable to form the $H_\alpha - S_\delta$ interaction. At decreased temperatures, the appearance of a less mobile component suggests a weak interaction that restricts the motion of R1. The spectral features from this component were weakly sensitive to order parameters. Since **{m,m}** is the only other rotamer found at the 8R1 site within the crystal, this weak interaction may be the formation of the $H_\alpha - S_\delta$ hydrogen bond that is present

in the **{m,m}** rotameric state. The CW results suggests that the **{m,t}** conformer is dominant at room temperature, which indicates that the interpretation of order parameters in terms of backbone fluctuations should be undertaken with caution in the case of β-sheets, a point that was also made recently by Freed et. al.[93]

DEER ESR was performed on the doubly spin-labeled GB1 mutant. DEER spectroscopy is able to resolve distance distributions between two spin labeled sites. While this technique has been advantageously utilized in many previous studies, the added distance of the spin label side chain has been a source of complication as the experiment does not directly report on the C_α–C_α distance. Here, DEER was performed on the doubly labeled N8C/K28C GB1 mutant in order to compare the resulting distance distribution with distances predicted in the crystal structure. The DEER results were obtained on GB1 in its monomeric form and therefore are not directly comparable to the crystal structure. However this domain-swapping allowed the opportunity to apply the rotamers to a related system with very similar rotameric environments and compare the distances between rotamers with the DEER distribution. This was accomplished by modeling in the R1 side chain into the high resolution WT GB1 structure (PDB: 2LGI) at sites N8 and K28, then adjusting all the R1 dihedral angles (χ_1 through χ_5) to match the conformations of the R1 side chains found within the crystal structure.

The distance distribution obtained by DEER is compared to the distances obtained by modeling the rotamers in the inlay of Figure 2-10. The three 8R1 and two 28R1 rotamers were modeled at the respective locations within WT GB1 and nitroxide N to nitroxide N distances were generated. The solid blue line represents the average distance from this model (22.4 Å) and the light blue represents the full range of generated distances (21.7 Å to 22.8 Å). These distances are ~ 2 – 3 Å shorter than the most probable DEER distance of 25 Å. The results are

encouraging given that the error in the most probable distance is ~ ±1 Å. The slight difference between the DEER and modeling results can possibly be due to minor changes in R1 rotamer packing that may occur upon crystallization or due to the presence of glycerol in the DEER experiments. Interestingly, when the χ_3 value for the α-helix (28R1) rotamers is altered from the +90° orientation observed in the crystal to the other common χ_3 orientation of -90°,[27, 36, 91] the mean distance between rotamers increases to 25.5 Å, a distance that closely matches the most probable distance from the DEER distribution. The results are interesting as they indicate by direct experimental measurement that the rotameric states of R1 found in this crystal provide a very close match to the most probable distance measured by DEER. These results foreshadow future work that explores the full conformational preferences of the spin label in solvent accessible β-sheets by modeling. An understanding of the range of orientations of the spin label can lead to methods by which DEER distributions is analyzed to extract constraints on C_α–C_α distance distributions. One such approach was presented recently by Sarver et al.[25] In this work, a series of DEER measurements between multiple sites on a protein-DNA complex were coupled with molecular dynamics (MD) simulations of R1 at the labeled sites. This combination of techniques was able to report directly on the C_α–C_α distribution between the labeled sites.

The distance comparison performed in this work demonstrates the utility of coupling X-ray crystallography resolved rotamers with DEER spectroscopy. Simply modeling resolved rotamers of similar environment (solvent accessibility, secondary structure, etc.) into appropriate locations within a wild type monomer structure resulted in the generation of a mean distance within ~ 2 Å of the most probable DEER distance. In other words, this combination of techniques could possibly be used to elucidate the C_α – C_α distance simply from the most

probable distance between two labeled sites and an appropriate selection of rotamers. In addition to the characterization of R1 motional modes using X-ray crystallography and variable temperature CW, these results provide important first steps in a greater understanding of R1 side-chain packing trends in solvent-exposed β-sheets.

2.6 ACKNOWLEDGEMENTS

We thank Prof. Angela M. Gronenborn (Department of Structural Biology, University of Pittsburgh) for her generous gift of wild type GB1 plasmid. We thank Prof. Michael Trakselis (Department of Chemistry, University of Pittsburgh) for use of his facilities for purifying GB1, as well as his lab member, Robert Bauer, for his helpful discussions. We thank Jessica Sarver for her helpful comments on the manuscript. This work was supported by the National Science Foundation (MCB 1157712 grant) to S.S. and the University of Pittsburgh (W.S.H.). The protein mutation work was supported by the National Science Foundation CAREER Award (MCB 0745754) to CPJ. Reprinted (adapted) with permission from Biochemistry, 2012, V. 51, pages 6350-6359.

3.0 R1 SPIN LABEL ROTAMERIC BEHAVIOR IN H-BONDING AND NON H-BONDING EDGE BETA-STRAND SITES AND IMPLICATIONS FOR INTERPRETATION OF DISTANCE DISTRIBUTIONS

This work was written in collaboration with W. Seth Horne and Sunil Saxena. The thesis author set up crystal trays, collected ESR data, analyzed data, and prepared the manuscript.

3.1 CHAPTER SYNOPSIS

The R1 spin label has successfully been used to explore structure and dynamics of many proteins. However R1 measurements are complicated by the inherent flexibility of the side chain. While the solvent-exposed α-helical environment has been thoroughly explored as a means to address this, similar work in β-sheets is lacking. The goal of this study is to explore R1 specifically within two edge β-strands using X-ray crystallography and double electron electron resonance (DEER) distance measurements. Crystal structures yielded seven distinct rotamers for the non-hydrogen bonded site and three rotamers for the hydrogen bonded site. The observed rotamers indicate contextual differences in rotameric preferences as compared to other solvent-exposed environments. However one rotamer from the hydrogen bonded site is similar to what is commonly observed for R1 in α-helices. For the DEER measurements, each strand site was

paired with the same α-helical site that has been characterized previously. The most probable distance for DEER can be adequately rationalized based on the crystal rotamers. Additionally, the appropriateness of common DEER analysis methods are assessed for this environment including molecular dynamics (MD), rotamer libraries with MMM, and MtsslWizard. The MD results display an inability to model both sites correctly, likely because of undersampling issues. MMM and MtsslWizard exhibit some quality fits but no clear trend was observed for fitting DEER distances featuring edge β-strand R1 sites. Rotameric preferences observed in crystals here are not well modeled by current analysis methods and thus further work is needed to correctly relate DEER distances to protein structure.

3.2 INTRODUCTION

The incorporation of unpaired electron spins into proteins is an increasingly useful and powerful tool for elucidation of protein structure and dynamics using magnetic resonance techniques.[123-125] This process, known as site-directed spin labeling (SDSL), involves the site specific placement of either stable organic radicals or paramagnetic metals within a protein. While metals such as Mn^{2+}, Cu^{2+}, and Gd^{3+} are becoming increasingly advantageous as site specific probes of late,[61, 126-131] the most commonly utilized spin probe is still the nitroxide based side chain R1. R1 is generated by reacting the thiol specific methanothiosulfonate spin label (MTSSL) with cysteine residues. The resulting R1 side chain, shown in Figure 3-1, is a useful structural probe that has been shown to provide site specific dynamics information in some protein environments[13] as well as long range distance constraints about the protein or protein complex of interest.[1, 123]

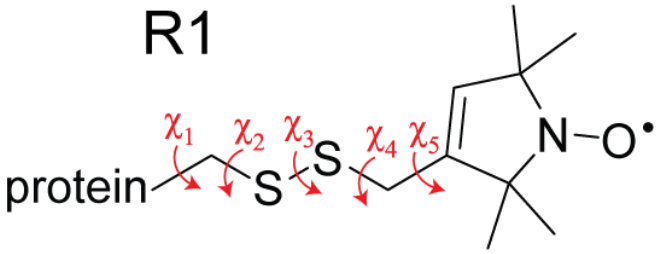

Figure 3-1. The common R1 spin label and the five associated dihedral angles that define R1 rotameric states.

Using current spin labelling methods, nanoscale distances can be measured between multiple R1 sites providing long range (15 – 100 Å) distance constraints.[51, 132, 133] However the measured distance is between the N – O bonds of the nitroxide rings. R1 flexibility leads to a range of nitroxide ring locations relative to its protein attachment site. In turn any distance measurement made between an R1 pair is dominated by R1 rotameric motions.

The flexibility of the linker connecting the nitroxide to the backbone is a critical limitation of the R1 label. The breadth of R1 rotameric states makes it challenging to extract C_α – C_α distributions between R1 labeled sites. For example, we previously generated quantitative estimates of the C_α – C_α distribution within a protein-DNA complex by comparing double electron electron resonance (DEER) with molecular dynamics (MD) and the C_α – C_α distribution from MD had a width approximately six times narrower than the DEER measurement.[25] Therefore to account for R1 flexibility in distance measurements, various computational tools have been developed.[37, 39] While very useful, these tools cannot consistently account for R1 flexibility in all cases. Consequently, a number of DEER measurements are typically required before quantitative estimates of assembly structure or protein conformational changes can be made.[20-25] Additionally, most of these tools were constructed only considering R1 free in solution[37] or R1 in α-helices[36, 134] without considering other protein secondary structure environments.

One important aspect of the current work is the focus on solvent-exposed β-sheet sites. Given the prevalence of β-sheets in protein structure, such knowledge is essential. A previous X-band continuous wave (CW) ESR dynamics study of R1 in a variety of solvent-exposed β-sheet locations indicated that R1 rotameric behavior is dominated by local interactions at many internal strand sites and some external strand sites.[11] Recently, we used a β-sheet site within a

protein-DNA complex and gained insight into rotameric preferences by comparing DEER and MD simulations.[25] The comparison allowed for extraction of the $C_\alpha - C_\alpha$ distribution as well as the identification of MD rotamers possibly sampled at the β-sheet site.[25] To gain further insight into the β-sheet environment, we later explored an internal strand site with a variety of techniques and resolved R1 rotamers not previously observed.[135] Also we indicated that methods currently used to interpret site specific dynamics information using R1 in α-helices are likely unreliable in the interior-strand β-sheet environment. These data suggest that the β-sheets exhibit complex effects on R1 and thus prediction of R1 rotameric preferences to aid in DEER interpretation is challenging.

While interior strands are likely influenced by local residues, logic suggests that edge strand locations have a higher likelihood of being free of interactions with neighboring residues as a neighboring β-strand only exists on one side of the R1 site. However, edge strand sites must be separated into two categories: either the backbone of the R1 site is hydrogen bonded or not hydrogen bonded to the sole neighboring strand. This particular distinction is important because it dictates the orientation of the $C_\alpha - C_\beta$ bond relative to the rest of the β-sheet. When the backbone carbonyl and amide bonds are not hydrogen bonded to the neighboring strand (Figure 3-2), the $C_\alpha - C_\beta$ bond points into the rest of the sheet, increasing the possibility of residues on the neighboring chain affecting the rotameric preferences of R1. All edge sites explored in the X-band CW study fall into this non-hydrogen-bonding category.[11] The alternative edge strand category is where the backbone is hydrogen bonded to the neighboring strand (Figure 3-2). At these sites, the $C_\alpha - C_\beta$ bond is pointing away from the rest of the sheet into open space which may allow for R1 to sample full conformational space. In turn, this may allow for internal,

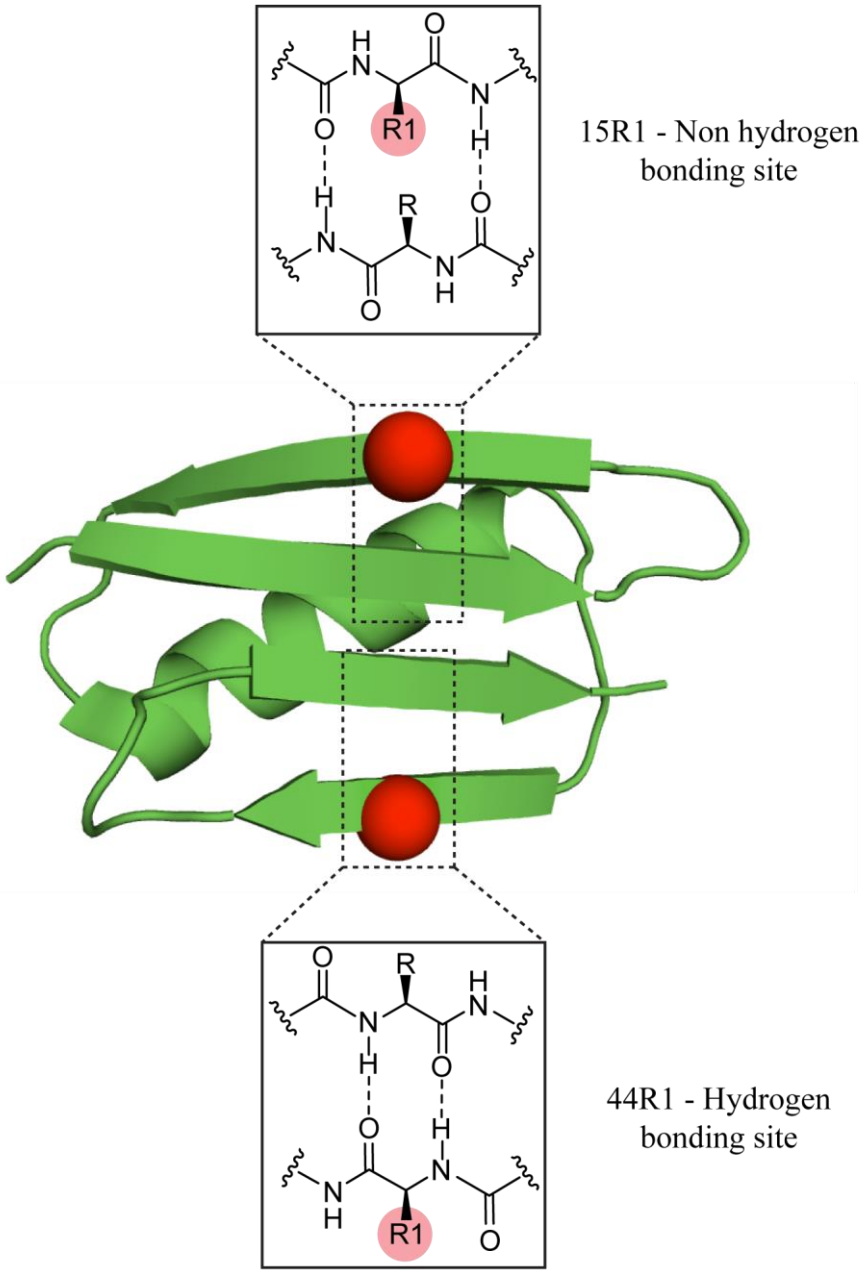

Figure 3-2. Cartoon displaying the contextual differences between edge β-strand sites. Either the backbone of the site is not hydrogen bonded to the neighboring strand (15R1), or it is hydrogen bonded to the neighboring strand (44R1). Non-hydrogen bonded sites are expected to experience clashes with the neighboring chain because of the direction of the $C_\alpha - C_\beta$ vector.

restricting interactions, similar to α- helix sites. Thus far, no such environment has been systematically investigated.

To this end, we explore the rotameric behavior of the spin-labeled side chain R1 in the two solvent-exposed edge strand environments described above. Using the B1 immunoglobulin binding domain of protein G (GB1),[135] we explore differences in rotameric behavior of R1 located at a non-hydrogen-bonding site (15R1) and a hydrogen-bonding site (44R1) through a variety of methods. DEER distance measurements were performed between each β-sheet site and a consistent α-helix site. Multiple crystal structures show rotamers of R1 at each location with χ_1/χ_2 dihedral angle combinations that are unique amongst the two sites and unlike any previously observed rotamers.[12, 27, 28, 135] Modeling of the most probable distance using the crystal resolved rotamers was performed to allow comparison of the rotameric preferences in solution as compared to the rotmaers observed in solution. Additionally, the appropriateness of common computational tools used for predicting DEER distances was also assessed in these edge strand environments. Taken together, the results presented here demonstrate that R1 rotameric behavior in edge strands of solvent exposed β-sheets depend strongly on whether the edge strand R1 site is hydrogen bonded to the neighboring strand.

3.3 MATERIALS AND METHODS

3.3.1 Generation, Expression, Purification, and Labeling of GB1 Mutants

The mutant plasmids coding for E15C, E15C/K28C, T44C, and K28C/T44C GB1 were generated using previously described methods.[130] The resulting plasmids were then transformed into BL21(DE3) *E. coli* cells and subsequently utilized for overexpression, purification, and eventual spin labeling via methods described previously.[135]

3.3.2 Crystallization, Data Collection, and Structure Determination

Crystals of the nitroxide-labeled E15C and T44C mutants of GB1 were grown by hanging drop vapor diffusion. Solutions of protein in water varying in concentration between 20 and 30 mg/mL were mixed (1 μL + 1 μL) with the well solutions. A single crystal was solved for 15R1 (Figure 2a) and the well solution was 0.1 M magnesium chloride, 0.1 M Tris pH 4.5, and 20% w/v PEG 4000. Two distinct crystals forms were obtained for the 44R1 mutant, each from a different buffer, and are denoted crystal form A ($44R1_A$) and crystal form B ($44R1_B$) and are shown in Figures 3a and 3b respectively. The well solution for the $44R1_A$ crystal was 0.1 M HEPES pH 7.5 and 1.4 M sodium citrate. For the $44R1_B$ crystal, the well solution was 0.2 M potassium sodium tartrate tetrahydrate, 0.2 M sodium citrate pH 6.0, and 2 M ammonium sulfate. The drops were allowed to equilibrate at room temperature over wells containing the crystallization buffers. A single crystal was flash frozen in liquid nitrogen after cryoprotection in well buffer supplemented with 25% v/v glycerol.

Diffraction data were collected using CuKα radiation on a Rigaku/MSC diffractometer with FR-E generator. Data for 15R1 and $44R1_A$ crystals were obtained on a Saturn 944 CCD detector, while data for the $44R1_B$ crystal were obtained on an R-AXIS HTC image plate detector. All diffraction experiments were carried out at a temperature of 100 K. Raw images were indexed, integrated, and scaled with d*TREK. Structure determination and refinement were carried out using the Phenix software suite. Structures were solved by molecular replacement using a published structure of wild-type GB1 (PDB 2QMT) as the search model. In the case of 15R1, the data indexed as tetragonal but the actual space group was found to be P21 with pseudomerohedral twinning (twin law L, -K, H). Refinement for 15R1 made use of the twin refinement algorithm in Phenix. All data collection and refinement details are displayed in Table 3-1.

3.3.3 ESR Measurements and Analysis

All X-band ESR experiments were performed on a Bruker Elexsys 580 spectrometer with a Bruker ER 4118X-MD5 resonator or a Bruker ElexSys E680 X-band FT/CW spectrometer with a Bruker EN4118X-MD4 resonator. For the DEER experiment, the pulse sequence used was $(\pi/2)\nu_1$-τ_1-$(\pi)\nu_1$-T-$(\pi)\nu_2$-τ_2-$(\pi)\nu_1$-τ_2-echo.[90] The pump frequency (ν_2) was placed at the maximum of the nitroxide spectrum and the observer frequency (ν_1) was offset at the maximum of the low field component (~67 MHz downfield). The $(\pi/2)\nu_1$ and $(\pi)\nu_1$ pulses were set to lengths of 16 ns and 32 ns respectively while the $(\pi)\nu_2$ pulse was set to 16 ns. τ_1 and T were set to 200 ns and 160 ns respectively, and T was incremented by a stepsize of 10ns for 128 points; τ_2 was adjusted such that $T + \tau_2 = 1300$ ns. The raw time domain DEER spectrum was analyzed using *DEER Analysis*

Protein	15R1	$44R1_A$	$44R1_B$
Data Collection			
Unit cell dimensions (Å, °)	$a = 52.3$, $b = 79.5$, $c = 52.4$	$a = 25.0$, $b = 37.5$, $c = 26.4$	$a = 25.0$, $b = 37.2$, $c = 49.2$
	$\alpha = 90$, $\beta = 90.1$, $\gamma = 90$	$\alpha = \gamma = 90$, $\beta = 108$	$\alpha = \beta = \gamma = 90$
Space group	$P2_1$	$P2_1$	$P2_12_12_1$
Resolution (Å)	27.11–2.20 (2.28–2.20)	25.03–2.50 (2.59–2.50)	25.05–1.60 (1.66–1.60)
Total observations	66,931	5,475	38,883
Unique observations	21,244	1,575	6,458
Redundancy	3.2 (2.5)	3.5 (3.6)	6.0 (4.4)
Completeness (%)	97.1 (90.9)	96.1 (95.9)	100 (97.2)
I/σ	9.1 (3.7)	13.6 (2.3)	24.0 (3.1)
R_{merge} (%)	8.4 (22.0)	6.4 (22.8)	5.2 (31.6)
Refinement			
Resolution (Å)	27.08–2.20	25.03–2.50	24.6–1.60
R (%)	18.9	23.6	15.6
R_{free} (%)	21.8	25.1	17.5
Avg. B factor ($Å^2$)	33.6	49.9	19.3
RMSD			
Bonds (Å)	0.007	0.003	0.014
Angles (°)	1.3	1.03	1.4

Table 3-1. Data collection and refinement statistics for the three crystal structures.

2013.[116] The temperature for all experiments was controlled with an Oxford ITC503 temperature controller and an Oxford ER 4118CF gas flow cryostat with liquid nitrogen as the primary cryogen.

3.3.4 Molecular Dynamics Simulations

The R1 spin labeled side was constructed at sites 15, 28, and 44 of a wild type (WT) structure of GB1 (PDB: 2LGI)[105] in PyMol.[136] R1 at each site were initially constructed to match rotamers resolved in the crystal structures from each site. The resultant structures were solvated in an explicit water box and counter ions were added to the box to neutralize the system. Molecular dynamics simulations were performed using NAMD[137] with the CHARMM22[138] force field for the entire protein with the exception of R1, which was parameterized using force fields developed by Sezer et al.[36] Energy of the system was initially minimized using a conjugate gradient method. The system was then heated to 300 K and equilibrated in an NPT ensemble for 1 ns at 1 atm via a Langevin piston. The backbone of the protein was restrained for all minimization and equilibration steps. Three 10 ns MD runs were performed for each site label site at steps of 1 fs. Convergance for all runs occurred by 2 ns and the final 8 ns of each run were analyzed to obtain the rotameric states of R1. The resulting dihedral angles were extracted with VEGA ZZ[139] and visualized using PyMol. The MD simulations were performed separately and thus distance distributions were generated by modeling the MD frames at the appropriate locations on a static WT GB1 structure (PDB: 2LGI) and measuring 25,000 distances between randomly selected frames.

3.4 RESULTS

3.4.1 Design of the Mutants

The primary focus of this work is to explore R1 rotameric behavior in edge β-strand environments and assess how this behavior affects measured distance distributions. This goal was accomplished through use of the protein GB1, which contains a single, four stranded β-sheet. Two sites in GB1 were selected to explore edge strand environments. One site is E15 where the backbone atoms of the residue are non-hydrogen-bonding to the neighboring strand. This site is expected to project the $C_\alpha - C_\beta$ vector of the side chain into the neighboring strand such that sterics may dictate rotameric preferences. The other site is T44 where the backbone atoms are hydrogen-bonding to the neighboring strand. This site is expected to project the $C_\alpha - C_\beta$ vector of the side chain into open space. R1 labeling is dependent upon site specific reactions with cysteine residues. Thus cysteine mutations were incorporated at these desired sites (i.e. E15C and T44C) then subsequently labeled (15R1 and 44R1) for crystallization experiments. As further means to explore the rotameric preferences of R1 at these sites, DEER distance measurements were also performed. This was accomplished by pairing each strand label with a second label at residue K28C. We have shown previously through crystallography that 28R1 displays rotameric preferences representative of a solvent-exposed α-helix site.[135] Therefore the double mutants E15C/K28C and K28C/T44C were generated and in turn labeled to yield 15R1/28R1 and 28R1/44R1.

3.4.2 Crystal Structures of Spin Labeled GB1

Solving the structure of spin-labeled protein crystals allows for the direct investigation of R1 rotameric preferences. The structure of the 15R1-GB1 crystal was refined to 2.2 Å yielding eight monomers in the asymmetric unit (Table 3-1). All 15R1-GB1 monomers exhibit a WT-like fold. Given eight monomers within the asymmetric unit, eight different 15R1 sites are present within the structure. Seven of the sites contain fully resolved R1 rotamers (Figure 3-3 and Table 3-2). The individual rotamers are referred to as $15R1_{1\text{-}8}$, with the only rotamer not fully resolved being $15R1_3$. For each of the monomers, the R1 side chains are in close proximity to other monomers, possibly influencing the R1 rotameric structure.

The structures of two distinct crystal forms of 44R1-GB1 were solved. The first form, designated $44R1_A$, was refined to 2.5 Å. The asymmetric unit of $44R1_A$ contains a single monomer of GB1 with a single fully resolved rotamer, $44R1_{A1}$ (Figure 3-4 and Table 3-2). The second crystal form, designated $44R1_B$, was refined to 1.6 Å. $44R1_B$ also contains a single GB1 monomer, however the labeled site contains two partially occupied rotamers and the rotamers found at this site are designated as $44R1_{B1}$ and $44R1_{B2}$ (Figure 3-4 and Table 3-2). Similar to all monomers of 15R1-GB1, both 44R1-GB1 crystals exhibit WT GB1-like fold. The spin labeled sites in both crystals exhibit potential interactions with symmetry-related chains in the crystal lattice.

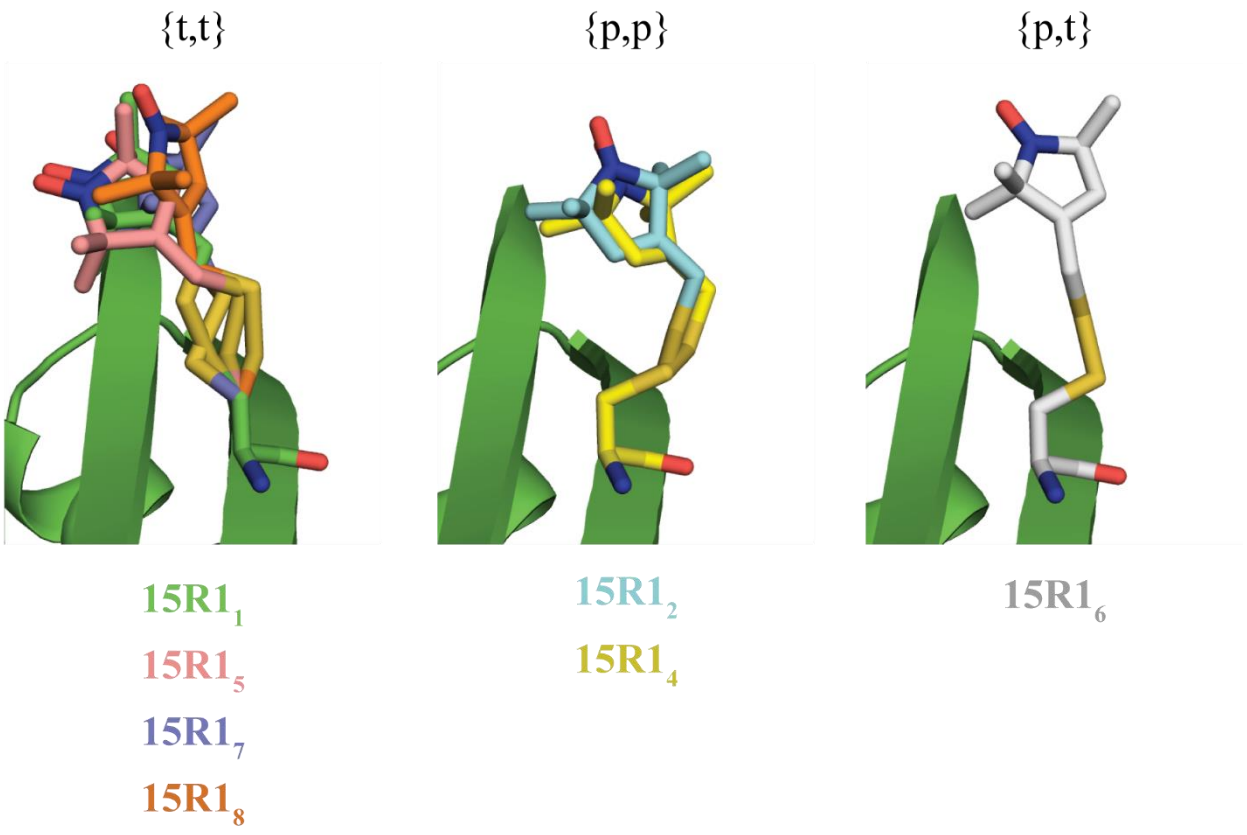

Figure 3-3. Fully resolved R1 rotamers from the 15R1 crystal structure categorized by the χ_1/χ_2 values.

Rotamer	Conformation	χ_1	χ_2	χ_3	χ_4	χ_5	C_α—S Dist. (Å)
15R1$_1$	{t,t?}	-171	155	131	120	179	4.4
15R1$_2$	{p,p?}	69	104	133	-79	26	4.0
15R1$_3$	{t?,-}	135	-	-	-	-	-
15R1$_4$	{p,p?}	66	89	132	-67	47	3.8
15R1$_5$	{t,t}	171	-167	-68	-109	-51	4.5
15R1$_6$	{p,t?}	81	135	93	175	-63	4.3
15R1$_7$	{t?,t?}	-151	-145	135	129	127	4.3
15R1$_8$	{t?,t}	144	-177	-90	-157	70	4.5
44R1$_{A1}$	{t,p?}	-163	101	72	106	28	3.9
44R1$_{B1}$	{t,t?}	-166	152	69	45	-104	4.4
44R1$_{B2}$	{m,t?}	-61	-153	-75	96	-100	4.4

Table 3-2. Summary of all rotamers observed in the crystal structures.

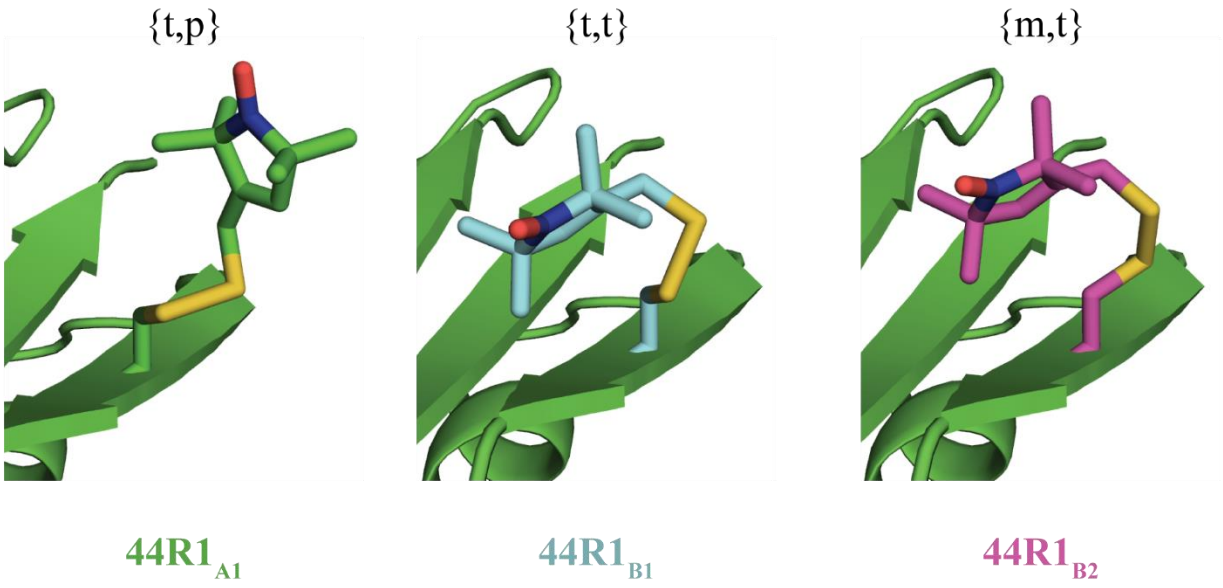

Figure 3-4. Fully resolved R1 rotamers from the two different crystal forms $44R1_A$ and $44R1_B$.

An important consideration in applying crystallography to investigate R1 conformational preferences is whether the crystallization process is artificially influencing the observed rotamers. Contacts with other protein chains in the asymmetric unit or symmetry mates in the crystal have the potential to bias the resulting rotamers. In our previous work using crystallography to investigate R1, the immediate environment surrounding each R1 residue was investigated for crystal contacts. Certain rotamers were dismissed when unnatural contacts were observed.[135] Given the β-sheet environment, the dihedral angles χ_1 to χ_3 are more dependent on native local residues due to packing and thus unnatural crystal contacts possibly impact only χ_4 and χ_5. We note that commonly at solvent-exposed α-helical sites, only χ_1 to χ_3 are resolved. This was rationalized as a highly dynamic rotameric structure about χ_4 and χ_5 leading to insufficient electron density to resolve the remaining atoms. Nevertheless, χ_1 through χ_3 contain essential information about R1 rotameric preferences at α-helical sites. In the β-sheet environment, the dihedral values observed for χ_1 through χ_3 are likely to be reflective of R1 preferences due to local packing. This is indeed the case for the $8R1_{B1}$, a rotamer we previously dismissed for unnatural contacts, but matches well with rotamers we considered "solvent exposed."[135] Therefore despite the unnatural crystal contacts for the 15R1 and 44R1 rotamers presented here, the $\chi_1 - \chi_3$ rotameric preferences are likely representative of R1 at these sites.

3.4.3 Resolved Rotamers of R1

A rotamer of R1 is defined by the five dihedral angles about the five rotatable bonds between the protein backbone and the nitroxide ring (Figure 3-1). A summary of all rotamers for 15R1 and 44R1 are displayed in Table 3-2. The table also contains the χ_1/χ_2 values in the m (-60° ± 20°), p

(60° ± 20°), t (180° ± 20°) notation as commonly utilized.[32] The preferred χ_1/χ_2 conformations of {m,m} and {t,p} have been resolved previously for solvent-exposed α-helix sites.[12, 27, 28] These observations support the possible existence of a C_α—H_α• • •S_δ hydrogen bond that restricts R1 flexibility at α-helical locations. Accordingly, the distance between the H_α and S_δ atoms are also provided to further investigate the possible existence of this C_α—H_α• • •S_δ hydrogen bond.

15R1 is a non-hydrogen-bonding β-strand site and as such, sterics with the neighboring chain likely preclude the possibility of a restricting C_α—H_α• • •S_δ H-bond.[11] The χ_1/χ_2 conformations of four rotamers ($15R1_1$, $15R1_5$, $15R1_7$, and $15R1_8$) match most closely with {t,t}. The {t,t} conformation, to the best of our knowledge, is unique among R1 rotamers in solvent-exposed environments. The remaining three rotamers most closely resemble either {p,p} ($15R1_2$ and $15R1_4$) or {p,t} ($15R1_6$) conformations. All three conformations for 15R1 have previously not been observed in any other solvent exposed context and thus are unique to this site though {p,p} has been observed previously at a membrane embedded β-sheet site.[31] As such none of these observed 15R1 rotamers match the {m,m} or {t,p} conformations that potentially allow for the C_α—H_α• • •S_δ H-bond to form

44R1 is a hydrogen-bonding edge strand site which, potentially, could allow for the C_α—H_α• • •S_δ H-bond to form. One of the observed rotamers at this site, $44R1_{A1}$, is very similar to the {t,p} conformation expected for such an interaction. Additionally, the χ_3 value for $44R1_{A1}$ is 72°, most similar to an expected value of χ_3 of +90°. For the {t,p} conformation, a χ_3 of +90° results in an R1 conformation that extends furthest away from the protein. The two remaining rotamers do not match the same trend and are similar to {t,t} ($44R1_{B1}$) or {m,t} ($44R1_{B2}$). {t,t} matches the most commonly observed state for 15R1 while {m,t} matches a previously observed rotamer for a solvent-exposed internal β-strand site.[135]

3.4.4 DEER Distance Measurements

The DEER experiment yields distance distributions between two R1 sites. The distance is measured between the unpaired electrons located within the nitroxide ring. However, the nitroxide ring is tethered to the protein through the five rotatable bonds. These bonds are capable of placing the unpaired electron up to 7 Å away from the C_α atom of the backbone. Also the flexibility of the side chain starkly increases the number of possible distances between the two spins. Thus the spatial sampling of R1 at each site highly influences both the most probable distance and the distribution width.[124] Here, two such DEER distance measurements were performed to explore the rotameric differences between the two edge strand sites 15R1 and 44R1. This was accomplished through pairing each β-sheet site with the same α-helix site, 28R1. Solvent-exposed α-helical sites are well characterized and we have previously observed that 28R1 rotameric behavior exhibits expected behavior.[135]

The resulting distance distributions for each measurement are displayed in Figure 3-5a and 3-5b respectively. For 15R1-28R1, the most probable DEER distance is ~ 25 Å. The C_α-C_α distance for 15R1-28R1 in a WT structure of GB1 (PDB: 2LGI) is 16 Å and thus the labels add 9 Å to the measurement. Additionally the distribution width (between 16% and 84% probability) of this DEER distribution is 7 Å. In the case of 28R1-44R1, the most probable distance is ~ 24.5 Å and the WT C_α-C_α distance is 12 Å. This displays a label contribution of ~ 12.5 Å, or more than half the most probable distance. In addition, the distribution width is 4 Å.

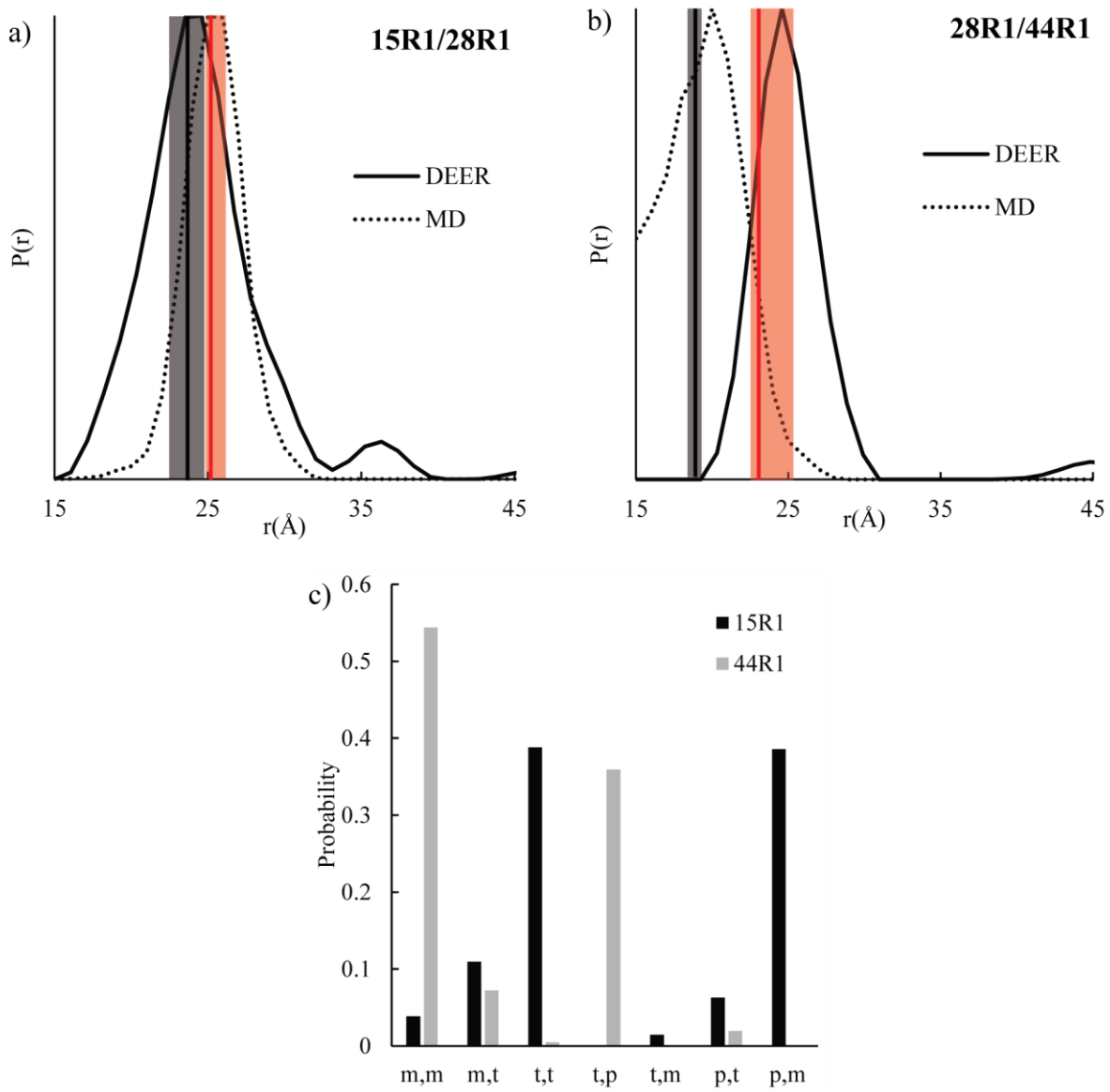

Figure 3-5. DEER results compared with MD and crystal (vertical black lines) for 15R1/28R1 (a) and 28R1/44R1 (b). For the 28R1 α-helical rotamers, χ_3 was adjusted to -90° and the distances again calculated (red). (c) The most commonly populated MD rotamers for each site; {m,p} and {p,p} were not sampled at either site.

3.4.5 Molecular Dynamics Simulations

Molecular dynamics simulations allow for an investigation of R1 spin label preferences in various protein environments. Currently, the only available MD force fields for R1 were generated with R1 at a solvent - exposed α-helix site.[36] The appropriateness of this force field for modeling the two edge β-strand sites investigated here was assessed. Rotameric preferences predicted by MD could be directly compared to the crystal rotamers. As a means to provide comparison with the DEER measurements, simulations were also performed at the solvent exposed α- helix site 28R1 such that comparable distance distributions could be generated with MD. In the case of 28R1, R1 populated the {m,m} conformation in 84 % of the frames and χ_3 was consistently populated close to -90°.

The distance distribution comparison for 15R1/28R1 is displayed in Figure 3-5a. The most probable distance predicted by MD is remarkably similar (~ 1 Å longer) to that of the DEER distribution. The MD generated width is ~ 50% more narrow compared to DEER, however there is substantial overlap between the experimental and MD distributions (Figure 3-5a). The similar distance distributions indicate that the rotameric sampling in the MD simulations likely reflect the rotameric preferences at site 15. Interestingly, the most probable $\{\chi_1/\chi_2\}$ conformations from MD is {t,t} (Figure 3-5c). The second most populated conformation observed in MD is {p,m}. This {p,m} conformation was predicted previously using spatial modeling of R1 at a different non-hydrogen-bonding edge strand site.[11]

The MD generated distribution for 28R1/44R1 is overlaid with the DEER distribution in Figure 3-5b. As compared to the case of 15R1/28R1, the 28R1/44R1 distributions do not match well. The MD predicted most probable distance is ~ 5 Å shorter than DEER and the majority of

the distribution predicts shorter distances. Interestingly, 58 % of the MD frames exhibit either {m,m} or {t,p} conformations (Figure 3-5c), the two expected conformations that spatially allow for the C_α—H_α• • •S_δ H-bond. Additionally, despite three separate 10 ns MD runs using these force fields, 44R1 only sampled χ_3 values close to -90°.

3.4.6 Alternate Distance Comparisons

Various tools have been created to aid in predicting DEER distributions requiring only a protein structure on which to simulate R1 motion. These include MtsslWizard,[39] a PyMol plugin that predicts R1 behavior through simple spatial sampling that searches for conformations free from steric clashes, and MMM,[37] a Matlab plugin that uses rotamer libraries generated through MD simulations of a single R1 in solution to predict rotameric preferences. Unlike the MD simulations described, these models use a static structure of the protein such that fluctuations in both backbone motion and rotameric sampling of local residues is not considered. Regardless, both are very useful and have been used in conjunction then compared with experimental DEER measurements to validate results and draw conclusions.[40-42] Here, both tools were used to assess their validity in the solvent-exposed edge strand environment of β-sheets. The same WT GB1 structure (PDB: 2LGI) was used with both tools. Additionally MMM was used at both 175 K and 298 K.

The alternate distance comparisons for 15R1/28R1 is overlaid with the DEER distance distribution in Figure 3-6a. MtsslWizard predicts a most probable distance ~ 5 Å longer than DEER. However the most probable distance from MMM is ~3 Å longer for the 298 K run and ~1 Å longer for the 175 K run as compared to the DEER measurement. Similar comparisons are

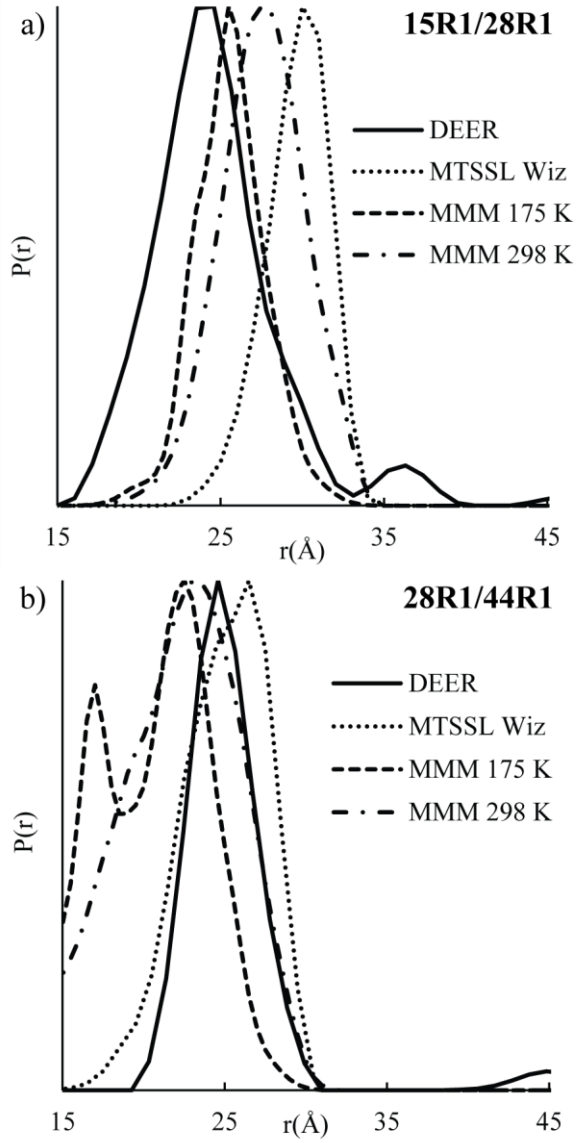

Figure 3-6. The DEER distance distributions compared with common, simple modeling techniques for both 15R1/28R1 (a) and 28R1/44R1 (b).

shown for 28R1/44R1 in Figure 3-6b. The 28R1/44R1 distance generated by MtsslWizard matches well with a most probable distance ~ 2 Å longer and a similar distribution width. On the other hand, MMM also predicts most probable distances similar to DEER with the 175 K run predicting ~ 2 Å shorter and the 298K run predicting ~ 1 Å shorter. Additionally, the MMM distributions display a significant population of shorter distance, including a peak ~ 8 Å shorter than DEER.

3.5 DISCUSSION

The goal of this work is to investigate rotameric preferences of the R1 spin label in two distinct edge β-strand environments and to assess the utility of common distance modeling programs in these contexts. We used mutants of the model protein GB1 to explore a site where the backbone atoms are non-hydrogen-bonding to the neighboring strand (15R1) and a site where the backbone atoms are hydrogen-bonding to the neighboring strand (44R1). Rotameric preferences of R1 dictate the location of the nitroxide relative to the protein and thus understanding these preferences aid in directly relating DEER distributions to protein structure. The DEER distributions themselves establish a basis of comparison for how R1 is behaving experimentally. X-ray crystallography allows for direct observation of R1 rotameric preferences. Through comparison of the crystal generated distances with DEER, crystal results can be validated and, in turn, allow for a direct interpretation of the rotameric preferences. Additionally, the appropriateness of common computational tools used for modeling of R1 rotameric behavior are also explored for each site.

In order to use DEER as a means of understanding these sites, each strand site was paired with the same solvent-exposed α-helical site, 28R1. The behavior of R1 in a solvent-exposed α-helix environment is well understood and thus this site was used as a control. Previously observed 28R1 {m,m} rotamers are used here for generating distances with the edge strand rotamers.[135] An important consideration for these rotamers is the observed value for χ_3. Both rotamers exhibit values close to +90° but other crystals display values of -90°.[27] Furthermore calculations suggest that χ_3 values of ±90° are equally populated at α-helical sites. [36, 134] Therefore distances were also calculated after altering χ_3 to -90°. Additionally, current MD force fields are based on R1 within a solvent-exposed α-helical environment. Thus it is fair to assume that the solvent-exposed helix site, 28R1, is modelled correctly. This assumption is reasonable given that the {χ_1/χ_2} conformations from MD predict the {m,m} state as the most populated state. The {m,m} conformation has been directly observed in crystal rotamers of R1 in solvent-exposed α-helices by us[135] and others.[12, 27, 28] Therefore differences between the MD distance distributions and the DEER distributions are attributed to differences of R1 at the β-sheet locations.

Previously we have shown that comparing DEER distributions with crystal rotamer generated distances is a valuable means of exploring rotameric preferences.[135] All crystal structures solved here display a WT-like fold of GB1 so all distance modeling was performed using the various resolved monomers that contain each β-sheet rotamer. As none of the crystals included the helix site, 28R1 rotamers were modeled into the GB1 monomers at the appropriate location and the distances calculated. The comparisons are shown in Figures 3-5a and 3-5b. The black bar represents the average modeled distance and the grey bar represents the range of calculated distances for the different rotamers. In the case of 15R1/28R1, the distances match up

very well. The most probable distance predicted by the rotamers is 24 Å, within 1 Å of the DEER most probable distance. This indicates that the rotamers likely represent the rotameric preferences at this non-hydrogen-bonding site. Interestingly, when χ_3 of the 28R1 rotamers is adjusted to -90°, the most probable distance shifts only slightly to 26 Å (Figure 3-5a).

Due to the difference in context between the MD force fields for R1 and the β-sheet sites investigated here, the validity of the MD runs were assessed through comparison with the DEER distance distributions. In the case of 15R1/28R1, MD compares well with the DEER most probable distance (Figure 3-5a) though the MD distribution is narrower. Additionally, the predicted MD rotamers were compared with the observed crystal rotamers. The most commonly populated rotamer from MD is {t,t} which matches closely with four of the seven observed crystal rotamers. Taken together, the MD force fields generated for solvent-exposed α-helical sites seem reasonable for modeling rotameric preferences for 15R1.

MtsslWizard and MMM are increasingly used as a simple and quick means of gaining insight into protein structure from R1 generated distance distribution. MtsslWizard uses a purely sterics approach while MMM uses the rotamer library approach. 15R1 is a non-hydrogen-bonding edge site where interactions with the neighboring chain are expected. Figure 3-6a displays the poor fit MtsslWizard provides for 15R1. This poor fit may be due to MtsslWizard relying only on local sterics. Since the local protein environment is static, rotamers that fail to meet the selected criterion of the model will be rejected. In solution, bulky residues also experience rotameric sampling and which may allow for different R1 rotamers that are unavailable using this tool. Therefore this model may create an artificial steric bias depending on the selected protein structure. Figure 3-6a also displays the results from MMM for 15R1/28R1. 298 K exhibits a broader selection of representative rotamers and this results in a

distribution with a most probable distance that is ~ 3 Å too long. Interestingly, MMM at 175 K provides a very similar distribution as compared to the MD distribution.

Similar comparisons were also performed to explore R1 at the hydrogen-bonding site, 44R1. Similar to 15R1, all resolved crystal structures of 44R1-GB1 were used for calculating the rotamer distances through modeling the 28R1 sites into the structures. As compared to the DEER distribution, the initial rotamer comparison does not match well. The rotamers predict a most probable distance of ~ 19 Å, 6 Å shorter than the DEER most probable distance. However, the χ_3 value of the 28R1 rotamers was again adjusted to -90° and the change is drastic as the most probable distance shifts to ~ 23 Å (Figure 3-5b). In particular, the case of $44R1_{A1}$ is noteworthy because this rotamer leads to the two longest modeled distances of 24 Å and 25 Å. These measurements match very well with the DEER most probable distance of 25 Å. Interestingly, $44R1_{A1}$ is the rotamer that exhibits a conformation similar to {t,p} that spatially allows for the stabilizing C_α—H_α• • •S_δ H-bond to form.

The appropriateness of the MD force fields for use at this hydrogen-bonding site was also investigated. This was accomplished through comparison with the 28R1/44R1 DEER distance distribution. As Figure 3-5b displays, MD does not match well with the DEER distribution. Most distances are too short as compared to DEER. Exploration of the predicted rotameric preferences by MD display the majority of frames exhibit either {m,m} or {t,p} conformation with a consistent χ_3 value of -90°. Although {m,m} and {t,p} are expected given the hydrogen-bonding location, it is likely that these predicted conformations are not representative of what is occurring in solution. Likely the MD force fields used here are oversampling these rotameric states.

Despite the poor results from MD, it is noteworthy that the crystal rotamer which provides the best distance comparison is the {t,p}-like rotamer, $44R1_{A1}$. The essential difference between this rotamer and the MD {t,p} rotamers is χ_3. The $44R1_{A1}$ rotamer exhibits a χ_3 of +90° while all of the MD {t,p} rotamers sample a χ_3 close to -90°. Accordingly, in an attempt to recreate the DEER distribution with these MD force fields, an additional 10 ns MD run was performed with χ_1 to χ_3 of 44R1 restricted to a {t,p,+90°} conformation. The results of this comparison are shown in Figure 3-7a. Restriction of these dihedral angles was also unable to recreate the DEER distribution correctly. However, a feature in the distribution is observed that matches well with the most probable DEER distance. As the first three dihedral angles are restricted to specific values, this feature that matches DEER must result from specific values of χ_4 and χ_5. Thus the 44R1 MD rotamers that resulted in this longer distance feature were investigated. Figures 3-7b and 3-7c display the χ_4 and χ_5 probabilities for the entire restricted MD run as compared to the rotamers that resulted in the distances matching DEER. While the χ_5 sampling is approximately the same for both, interestingly χ_4 sampling displays a clear trend. The MD rotamers that result in the longer calculated distance feature exhibit a distribution of χ_4 values centered at 100° which matches well with the 106° χ_4 value observed in the $44R1_{A1}$ crystal rotamer. Despite these efforts of constraining 44R1, current MD force fields clearly undersample R1 rotameric states that are able to recreate the experimental DEER distribution.

MtsslWizard and MMM were also used to predict DEER distances for 28R1/44R1. In the case of MtsslWizard, the purely sterics approach matches well with the DEER distribution (Figure 3-6b). Although the most probable distance differs slightly, the distribution matches well. This solvent-exposed hydrogen-bonding site likely lacks the potential for steric bias and MtsslWizard matches DEER better in this context as compared to the non-hydrogen-bonding site

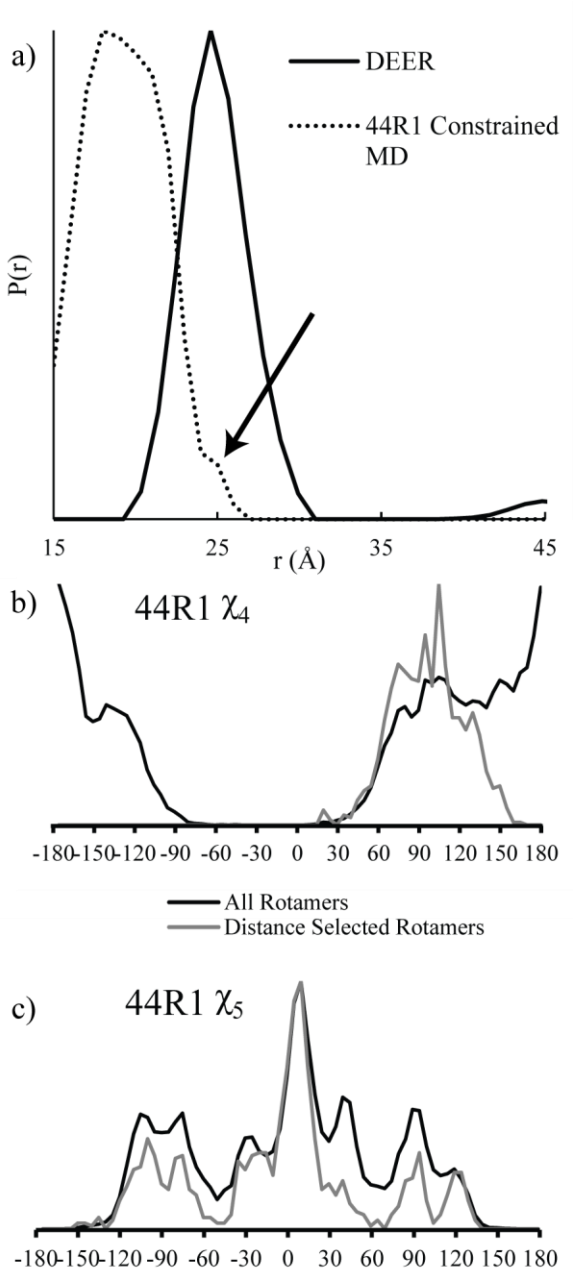

Figure 3-7. (a) Comparison of distance distributions from DEER and the MD run in which χ_1 to χ_3 of 44R1 was constrained to a {t,p,+90°} conformation. The χ_4 (b) and χ_5 (c) values of rotamers that led to calculated distances most similar to DEER (black arrow in (a)).

15R1. On the other hand, MMM at both 298 K and 175 K matches fairly well with regards to most probable distance, but a substantial amount of the distribution predicts distances much shorter compared to DEER.

3.6 CONCLUSIONS

The main goal of this study was to assess R1 rotameric preferences according to contextual differences between a non-hydrogen-bonding edge strand site and a hydrogen-bonding edge strand site using a variety of methods. DEER provides the experimental basis for comparison with the various techniques. Crystallography was used to resolve R1 rotamers and interestingly, the rotamers resolved at both sites are distinct when compared with previously observed rotamers from other sites. This indicates that context clearly matters with regards to rotameric preferences at these solvent-exposed edge β-strand sites. The DEER distribution is adequately rationalized by the distances generated with modeling crystal rotamers indicating that though crystal contacts were observed for all rotamers, these rotamers are relevant. One rotamer for the hydrogen bonding site displays a {t,p} rotamer possibly indicative of the C_α—H_α• • •S_δ H-bond though further analysis is required to explore this potential. When attempting to model these distances with common analysis techniques, no consistent trend was observed. While some good fits were obtained, rotameric preferences of R1 at edge strand sites is not consistently well modeled using common techniques. Accordingly, further work is necessary to fully understand R1 in these environments for edge strand sites to be used for gaining protein structural information from DEER results.

4.0 CYSTEINE SPECIFIC CU^{2+} CHELATING TAGS USED AS PARAMAGNETIC PROBES IN DOUBLE ELECTRON ELECTRON RESONANCE

This work, written in collaboration with Matthew D. Shannon, Miriam R. Putterman, Rajith J. Arachchige, Ishita Sengupta, Min Gao, Christopher P. Jaroniec, and Sunil Saxena, has been published in the Journal Physical Chemistry B, 2015, V. 119, pages 2839-2843. The thesis author prepared protein mutants, labeled samples, performed ESR experiments, analyzed data, and prepared the manuscript.

4.1 CHAPTER SYNOPOSIS

Double electron electron resonance (DEER) is an attractive technique that is utilized for gaining insight into protein structure and dynamics via nanometer scale distance measurements. The most commonly used paramagnetic tag in these measurements is a nitroxide spin label, R1. Here, we present the application of two high-affinity Cu^{2+} chelating tags, based on the EDTA and cyclen metal-binding motifs as alternative X-band DEER probes, using the B1 immunoglobulin-binding domain of protein G (GB1) as a model system. Both types of tags have been incorporated into a variety of protein secondary structure environments and exhibit high spectral sensitivity. In particular, the cyclen-based tag displays distance distributions with

comparable distribution widths and most probable distances within 1 Å to 3 Å when compared to homologous R1 distributions. The results display the viability of the cyclen tag as an alternative to the R1 side-chain for X-band DEER distance measurements in proteins.

4.2 INTRODUCTION

Double electron electron resonance (DEER) spectroscopy is an attractive electron spin resonance (ESR) technique that has allowed for the experimental measurement of distance distributions between multiple paramagnetic species in a variety of biological systems.[140-142] Paramagnetic species are typically not native to many protein systems and thus are introduced using a technique known as site directed spin labeling (SDSL).[143-145] In SDSL, paramagnetic tags are commonly incorporated through direct attachment to cysteine residues which have been engineered into the protein at sites of interest via mutagenesis. By far, the most common paramagnetic tag is the methanothiosulfonate spin label, or MTSSL. MTSSL reacts specifically with the free thiol group of cysteine residues and the result is the nitroxide side chain known as R1, as shown in Figure 4-1a. The use of R1 in DEER distance measurements as well as its various other applications have been reviewed extensively.[143-148]

In addition to stable organic radicals, an alternate source of ESR active species within proteins is paramagnetic metal ions. The simplest cases are those proteins that bind these paramagnetic metals naturally and if a protein contains multiple metal centers, DEER can be utilized to elucidate structural and dynamical information. Additionally, SDSL can be used in conjunction with these native metal binding sites and DEER can be performed between the metal

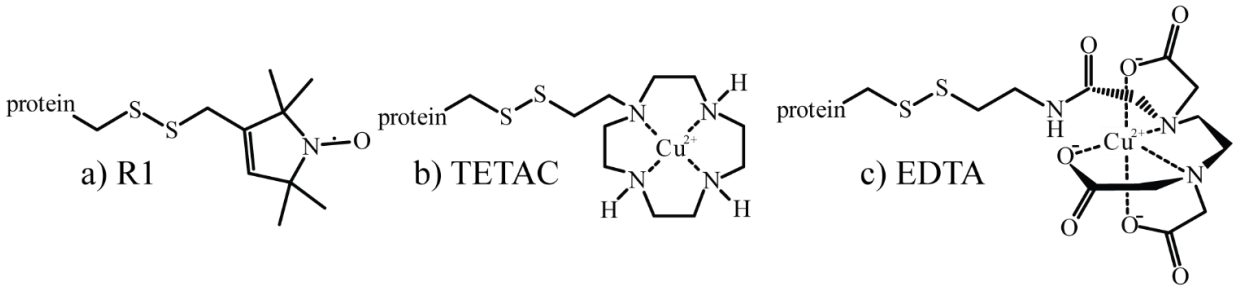

Figure 4-1. The three paramagnetic side chains used as DEER probes after attachment to a cysteine residue: (a) R1 is the common nitroxide side chain, while (b) TETAC and (c) EDTA are both Cu^{2+} chelating tags utilized here.

center and the spin labeled site(s). Applications of DEER measurements using natural metal binding sites has been reviewed recently.[126]

An alternate means of utilizing paramagnetic metal ions as DEER probes is through the site-specific incorporation of tags that strongly chelate paramagnetic metals such as Gd^{3+}.[149] In addition to Gd^{3+} tags being able to take advantage of the increased sensitivity at high field,[150] these metal chelating tags have displayed distinct advantages over R1 in highly relevant biological environments. Within lipid membranes, certain Gd^{3+} tags have displayed less conformational bias due to the hydrophobic environment as compared to R1 and thus may provide a more representative distance measurement within the membrane.[151] Metal based DEER measurements also appear to less affected by multi-spin effects in proteins containing more than two spins.[152] Additionally, Gd^{3+} tags have displayed much greater stability to the reducing conditions of living cells as compared to R1 for *in-cell* ESR distance measurements.[128] Taken together, metal chelating tags are advantageous for measuring distances in some biological environments.

The high field Gd^{3+} DEER measurements have been performed primarily at W-band (~95 GHz), or in some instances at Ka-band (~32 GHz). While Gd^{3+}-R1 DEER measurements have been performed at X-band (~9.5 GHz) on a model system, the measurement suffered from low signal to noise[153] due to the broadening of the central adsorption in the Gd^{3+} spectrum. Given the prevalence of X-band instruments and the advantages these tags can offer, it is important to develop alternative metal chelating tags for use at X-band frequencies.

An additional group of metal chelating tags being developed for protein structural studies are those that strongly chelate Cu^{2+}. Such tags have been successfully utilized recently for the

measurement of electron-nucleus distance dependent paramagnetic relaxation enhancements by solid state nuclear magnetic resonance (NMR) spectroscopy.[154-157] Similar to MTSSL and all the Gd^{3+} labels, the tags utilized thus far react specifically with cysteine residues. While intrinsically bound Cu^{2+} ions have been used extensively for X-band DEER measurements in model systems and several proteins,[126] the use of Cu^{2+} chelating tags has not been explored in this context. Here, we utilize two such Cu^{2+} tags as X-band DEER probes and compare the results with the common R1 spin label. The tags selected are the 1-(2-(pyridin-2-yldisulfanyl)ethyl)-1,4,7,10-tetraazacyclododecane (TETAC) tag[157] and the commercially available ethylenediaminetriacetic acid (EDTA) tag,[158, 159] the latter of which has been also utilized to chelate Mn^{2+} for use in DEER distance measurements.[149] The resultant side chains for the TETAC and EDTA tags are shown in Figures 4-1b and 4-1c. Note that these tags are ~20 % and ~40 % larger than the R1 side chain. All tags were incorporated into various double cysteine mutants of the 56-residue B1 immunoglobulin binding-domain of protein G (GB1) and the results presented here display the utility of these tags as X-band, paramagnetic metal probes for use in protein distance measurements.

4.3 METHODS

4.3.1 Construction of GB1 Mutants

The plasmid encoding for wild type GB1 was kindly provided by Prof. Angela Gronenborn (University of Pittsburgh). The cysteine mutation locations were chosen to represent a variety of

solvent-exposed β-sheet locations (I6, N8, and E15) as well as a solvent-exposed α-helix location (K28). Mutations were performed one at a time and mutant plasmid was used to create the desired double mutants (6/28, 8/28, and 15/28). Each mutant was created using the appropriate plasmid DNA, the primers encoding for the desired mutation (Invitrogen, Carlsbad, CA), and the KAPA Hifi Hotstart Ready Mix (Kapa Biosystems, Cape Town, South Africa). Resultant PCR reaction mixtures were treated with DpnI (New England Biolabs, Boston, MA), transformed into XL1-Blue Supercompetent cells (New England Biolabs), and grown overnight on culture plates containing Luria-Bertani (LB) broth with 100 mg/mL ampicillin. Colonies were picked and grown overnight in 50 mL of LB with 100 mg/mL ampicillin and the plasmid DNA was purified using the PureYield Plasmid Midiprep System (Promega, Madison, WI). All mutations were confirmed by DNA sequencing (Genomic and Proteomics Core Laboratories, University of Pittsburgh, Pittsburgh, PA) and subsequently transformed into BL21(DE3) competent cells (New England Biolabs) for expression.

4.3.2 Protein Expression, Purification, and Labeling

Expression and purification of all GB1 mutants was performed as was previously described.[135] The labels used were the TETAC tag (synthesis described previously),[80] the methanothiosulfonate spin label (MTSSL), or the [S-Methanethiosulfonylcysteaminyl] ethylenediamine-N,N,N',N'-Tetraacetic Acid chelating tag (MTS-EDTA); the latter two tags were both purchased from Toronto Research Chemicals (Toronto, Canada). Labeling with MTSSL was performed as previously described.[135] Labeling with MTS-EDTA and TETAC were performed similarly except since both tags are water soluble, no DMSO was utilized in the

labeling process. Removal of the excess chelating tags was performed similar to the procedure for MTSSL.[135]

4.3.3 ESR Measurements

DEER distance measurements were performed on MTSSL labeled mutants as previously described.[135] For the MTS-EDTA and TETAC labeled samples, the samples were at a concentration of 0.5mM in 50mM N-ethyl morpholine (NEM) at pH 7.4, 25% glycerol, and 0.95mM isotopic $^{65}CuCl_2$ (Cambridge Isotopes, Tewksbury, MA). The Cu^{2+} chelating samples were slightly under loaded with Cu^{2+} due to Cu^{2+} interacting with GB1 elsewhere and leading to unwanted signals.[160]

All DEER distance measurements were performed on a Bruker Elexsys 580 spectrometer equipped with a Bruker ER4118X-MD5 resonator or a Bruker ElexSys E680 X-band FT/CW spectrometer equipped with a Bruker EN4118X-MD4 resonator. The temperature for all experiments was controlled using an Oxford ITC503 temperature controller with an Oxford ER 4118CF gas flow cryostat. Distance measurements for both Cu^{2+} samples were performed at 20K. The four pulse sequence utilized for the measurements was $(\pi/2)\nu1\text{-}\tau1\text{-}(\pi)\nu1\text{-}T\text{-}(\pi)\nu2\text{-}\tau2\text{-}(\pi)\nu1\text{-}\tau2\text{-echo}$. For all distance measurements, the pump pulse ($\nu2$) was placed at the maximum of the Cu^{2+} spectrum and the observer pulses ($\nu1$) are placed 150 MHz downfield from the maximum. The observer pulse ($\nu1$) lengths were 16 ns and 32 ns for the $\pi/2$ and π pulses, and the pump pulse ($\nu2$) length was 16ns. The parameter τ_1 was set to 200 ns and T to 160 ns initially with T being increased by 10ns for 128 steps. τ_2 was adjusted such that $\tau_2 + T = 1300$

ns. DEER data collection times for all spectra are comparable (~ 20 hrs for all R1, EDTA-Cu^{2+} and TETAC-Cu^{2+} labeled samples).

4.4 RESULTS AND DISCUSSION

The Cu^{2+} binding tags and the common R1 spin label were attached to several double cysteine mutants of GB1 as described below. The three-dimensional structure of GB1 is well known[100, 103, 161, 162] consisting of a single α-helix and a four-stranded β-sheet (Figure 4-2). For this comparison, three different double cysteine mutants were utilized, each containing the same solvent-exposed α-helical site (K28C) and a single solvent-exposed β-sheet site. The three β-sheet sites were selected to represent a variety of β-sheet locations including an internal strand site (I6C), a previously investigated internal strand site neighboring a β-turn (N8C),[135] and an external strand site (E15C). The relative location of all the labeled sites can be seen in Figure 4-2. For all three GB1 double mutants (6/28, 8/28, and 15/28), each was labeled with the three different tags (R1, TETAC-Cu^{2+}, and EDTA-Cu^{2+}), and X-band DEER measurements were performed.

As a basis for comparison with the Cu^{2+}-chelating tags, DEER distance measurements were performed on the three corresponding double R1 labeled cysteine mutants of GB1. The baseline subtracted time domain data for 6R1/28R1, 8R1/28R1, and 15R1/28R1 are displayed in the top panels of Figure 4-3. Overlaid on the data is the best simulated fits using Tikhonov regularization (black) and single Gaussian model fits (dotted). The model free Tikhonov

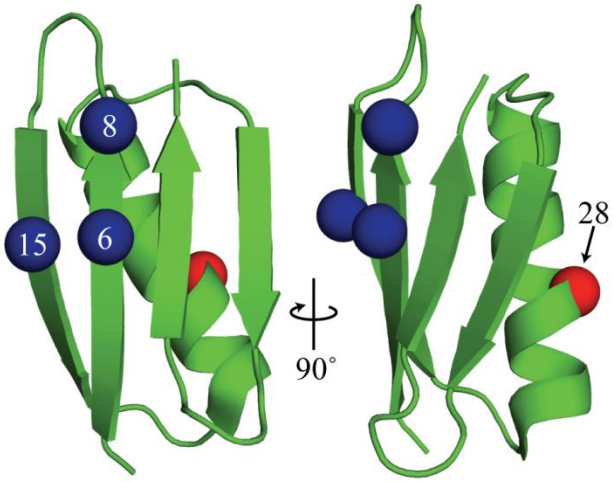

Figure 4-2. An NMR structure of GB1 (PDB ID: 2LGI)[162] showing all the labeled sites. Each DEER pair included the α-helix site 28, and one of the β-sheet sites, either 6, 8, or 15.

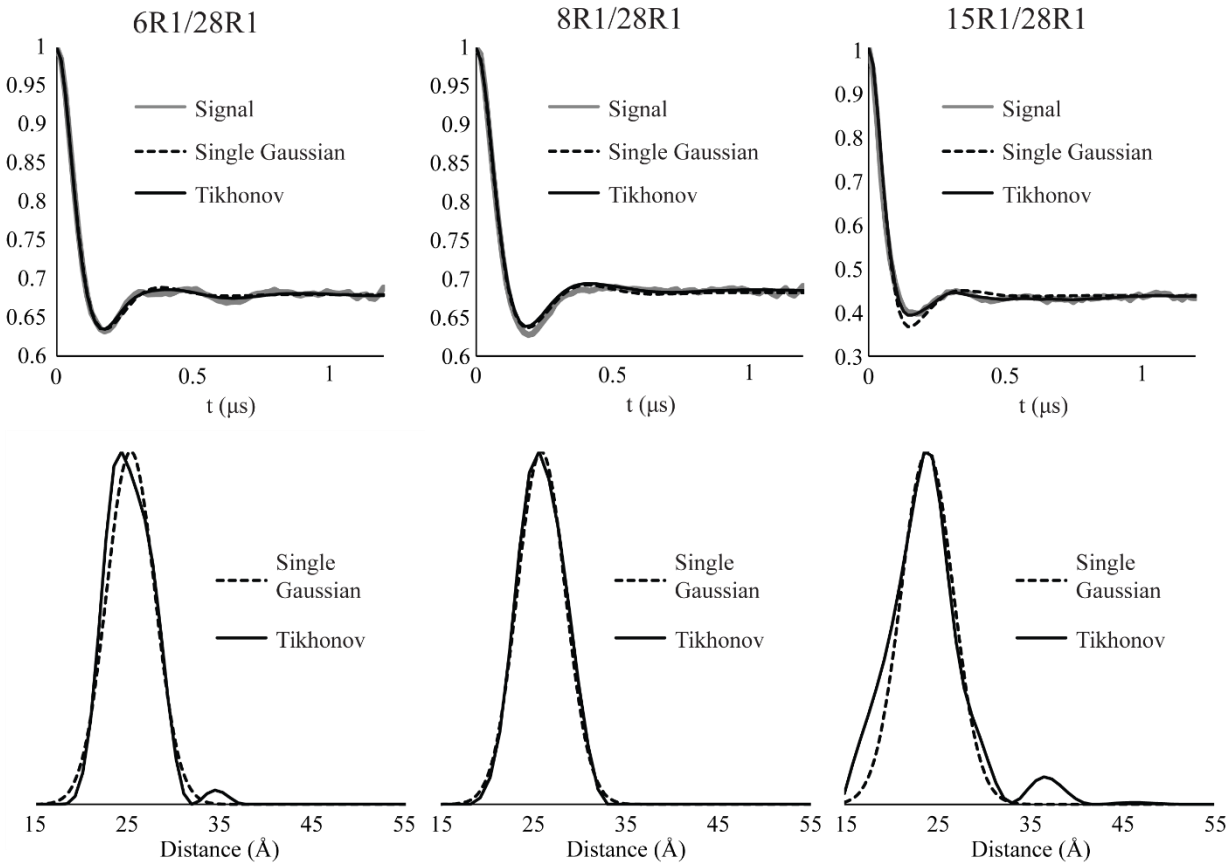

Figure 4-3. The single Gaussian and Tikhonov regularization fits for the DEER data of the 6/28, 8/28, and 15/28 mutants labeled with R1. The bottom panels displayed the similarities of the two different fits.

regularization fits can provide better visual fits of the time domain data, however single Gaussian model fits can improve the reliability of the fit. The bottom panels of Figure 4-3 show the resulting distance distributions for all of the time domain fits. While the Tikhonov fits are better in the time domain, the most probable distance and distribution widths for all three data sets are comparable. For all R1 and Cu^{2+} measurements, single distances were expected and therefore single Gaussian fits were generated in Deer Analysis for all of the time domain data fits.[116] Additionally, the modulation depth of R1 measurements (~40%) is inherently larger than comparable Cu^{2+} measurements (~10%) and thus the time domain data for R1 vs Cu^{2+} measurements are displayed separately.

Cu^{2+}-Cu^{2+} DEER measurements can be problematic due to the presence of orientational selectivity. Though the flexibility of the tags was expected to wash out this effect, nevertheless both Cu^{2+} tags were probed by performing the DEER experiment at different magnetic fields as to excite different Cu^{2+} orientations. The results for the EDTA-Cu^{2+} tag attached to 8/28 are shown in Figure 4-4. Due to the unchanging time domain data and Pake pattern despite the pulse locations, orientational effects are not a concern for this EDTA tag. Therefore, the pulse positions resulting in the best signal to noise (position A in Figure 4-4) was used for the overall tag comparison. Similar data was collected for the TETAC-Cu^{2+} attached to 8/28 and this data is displayed in Figure 4-5. While the time domain signals for the different magnetic fields appear similar, slight differences are observed in the Pake patterns indicating very weak orientational effects. As a means to further explore these effects, the baseline subtracted time domain data was fit with both Tikhonov regularization and single Gaussian fits. The fits and subsequent distance distributions are shown in Figure 4-6. Fits to the data obtained at different magnetic

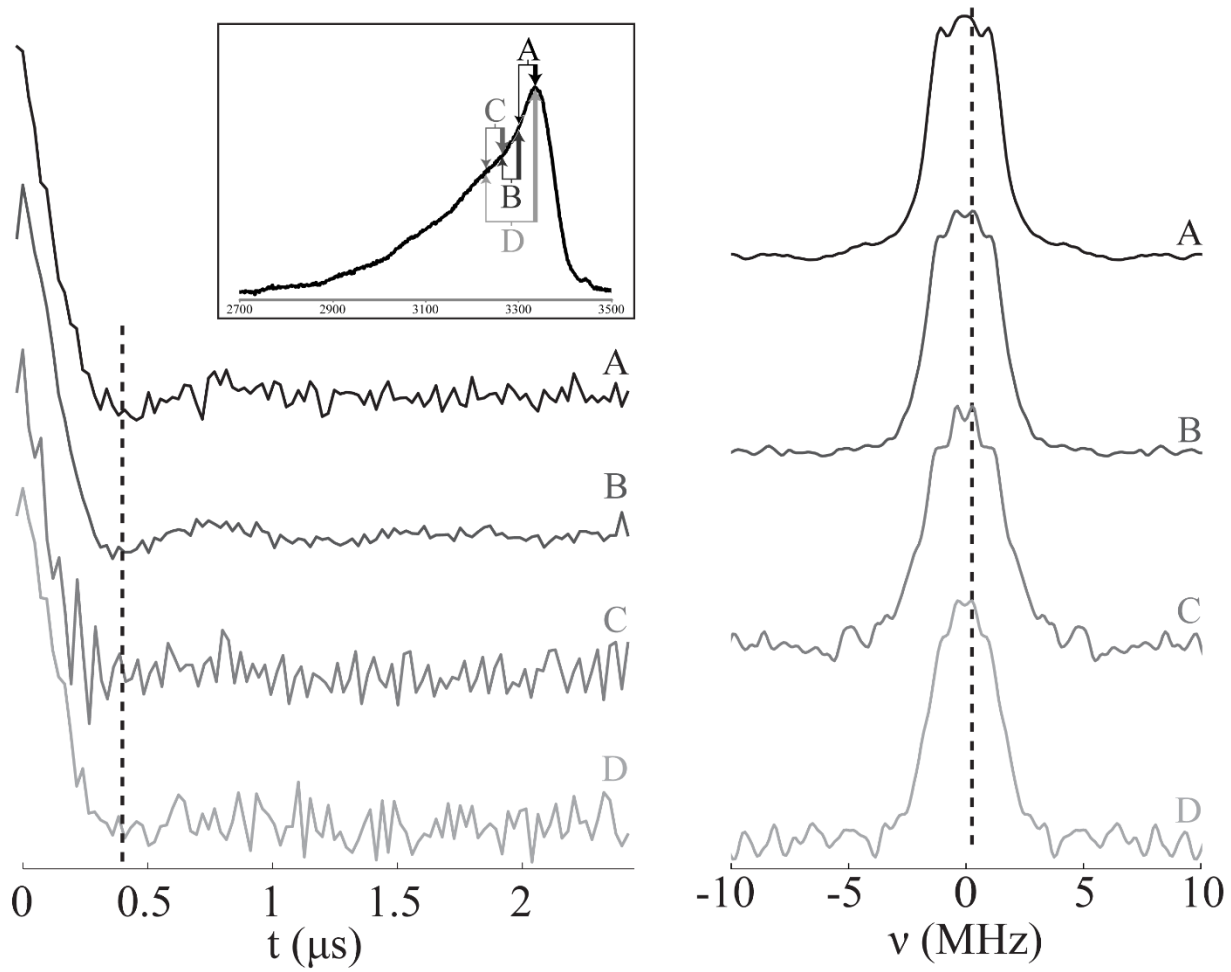

Figure 4-4. The baseline subtracted time domain and subsequent Pake patterns for 8/28 GB1 labeled with EDTA-Cu^{2+} with the observer (thin arrow) and pump (thick arrows) placed at various field strengths. The similar modulation period and dipolar frequency indicate minimal orientation selectivity.

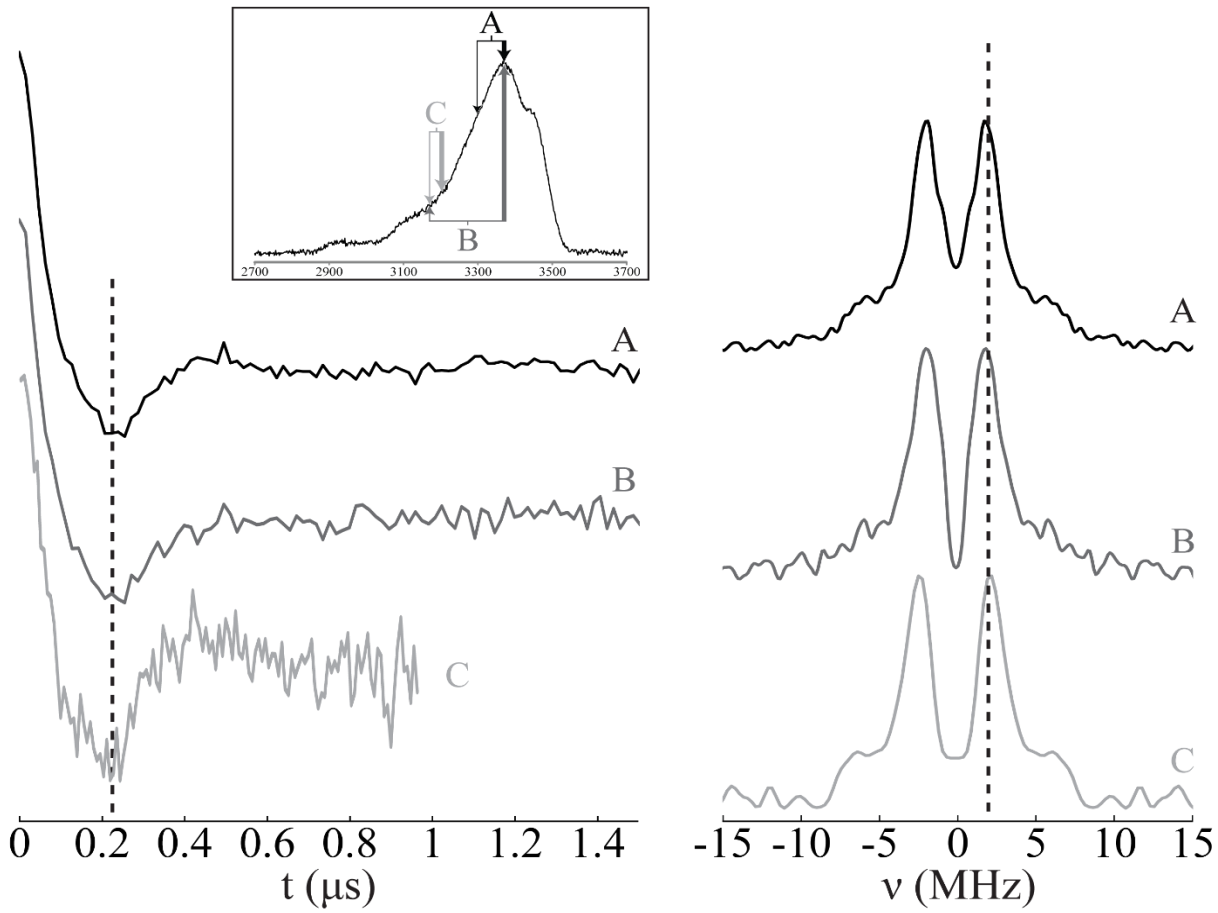

Figure 4-5. The baseline subtracted time domain and subsequent Pake patterns for 8/28 GB1 labeled with TETAC-Cu^{2+} with the observer (thin arrow) and pump (thick arrows) placed at various field strengths. Slight changes were observed in the Pake pattern and thus these data were further investigated in Figure 4-6.

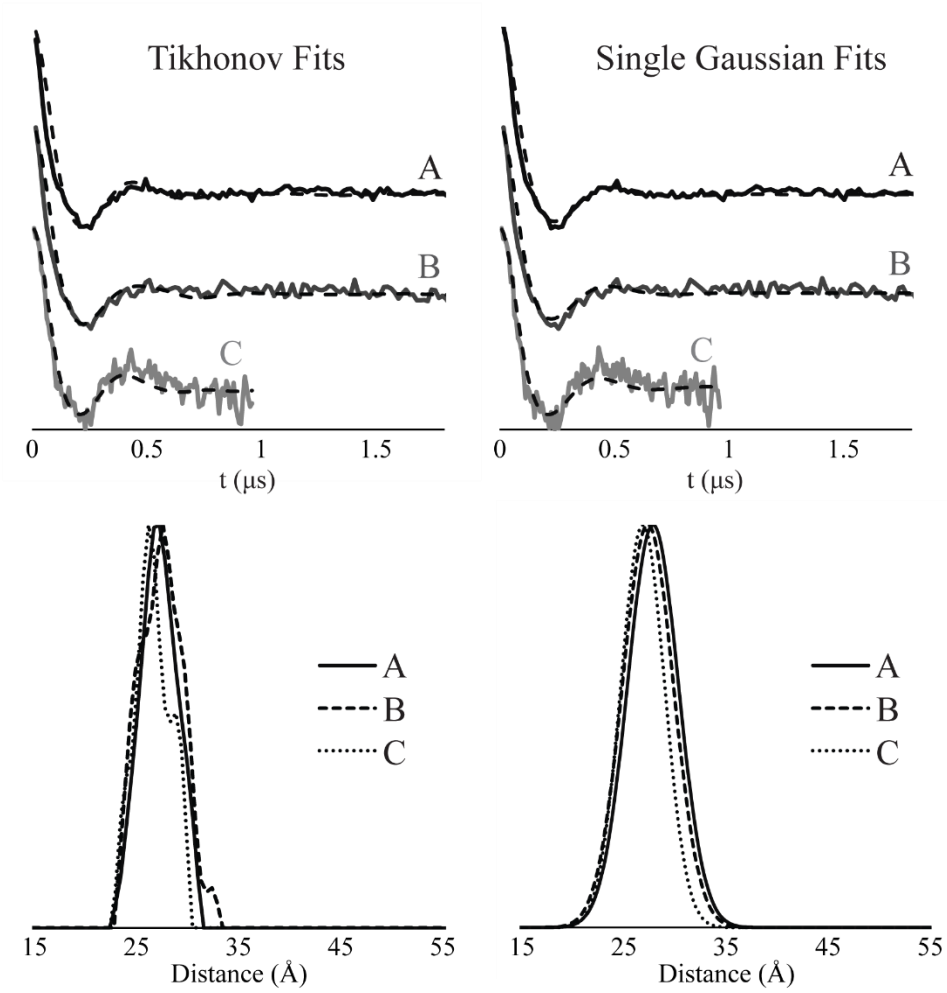

Figure 4-6. The multiple DEER spectra collected for 8/28-TETAC-Cu^{2+} fit using Tikhonov regularization or model single Gaussian fits as well as the resultant distance distributions for all fits.

fields indicate similar the distribution widths and the most probable distance varies only by ~ 1Å**.** Thus orientational effects are largely reduced possibly due to an orientational distribution between the two Cu^{2+} centers.[66, 163] Similar to the EDTA-Cu2+ data, the pulse positions yielding the best signal to noise (position A in Figure 4-5) is used for the remainder of the comparison.

Figure 4-7 shows the baseline subtracted Cu^{2+} signal for the 6/28, 8/28, and 15/28 mutants respectively. The time domain signals for the Cu^{2+} tagged mutants mostly display modulation depths of up to 10%., which is sufficient to achieve a quality S/N at X-band. A lower modulation depth was achieved in 8/28 TETAC-Cu^{2+}, which is likely due to incomplete labeling (estimated to be ~ 74% from the modulation depth).[121] Nevertheless, reasonable signal to noise was achieved for this mutant. For Gd^{3+} DEER measurements, the commonly low modulation depths of ~1% have been attributed to the presence of free Gd^{3+} ions masking the DEER effect.[59] Cu^{2+}-DEER offers a significant advantage in that free Cu^{2+} is EPR silent[164] in the N-ethylmorpholine (NEM) buffer at pH 7.4, which was used in these measurements. Thus with NEM buffer, the presence of free ions is not a concern allowing for maximum modulation depth, and in turn maximum sensitivity to be achieved. As discussed above, since single distances are expected for these Cu^{2+} measurements, single Gaussian model fits were performed and these fits are overlaid with the data in Figure 4-7.

The resultant distance distributions from all measurements are shown in the Figures 4-8. For each of the three mutants, the EDTA-Cu^{2+} tag shows the longest most probable distance when compared to the other two tags. These longer distances are expected given the relative structures of the tags, as shown in Figure 4-1. R1 has five bonds between the protein backbone and the nitroxide ring, TETAC has six between the backbone and the chelating motif, while

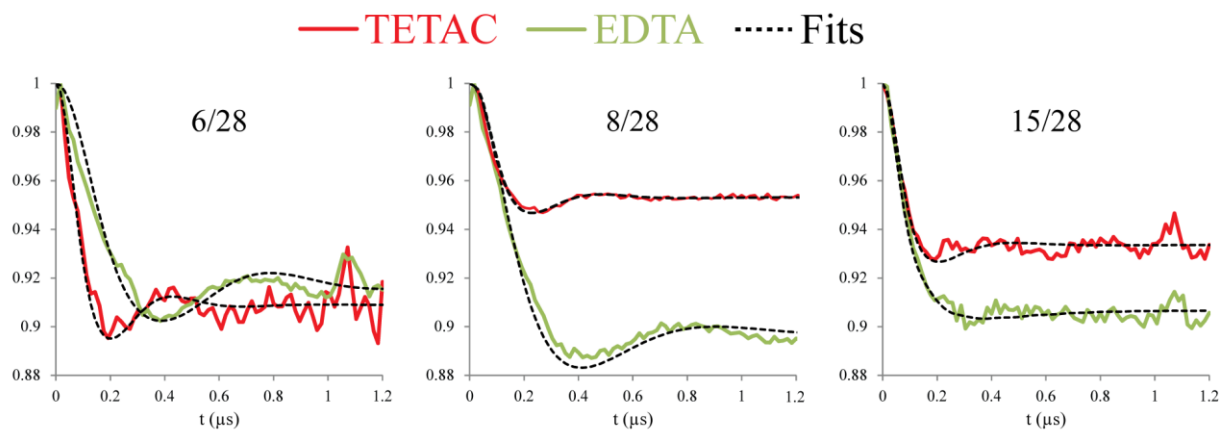

Figure 4-7. The baseline subtracted time domain DEER data for the double cysteine GB1 mutants 6/28, 8/28, and 15/28 tagged with TETAC-Cu^{2+} (red) or EDTA-Cu^{2+} (green) and the best single Gaussian fits (dotted black) using DEER Analysis.

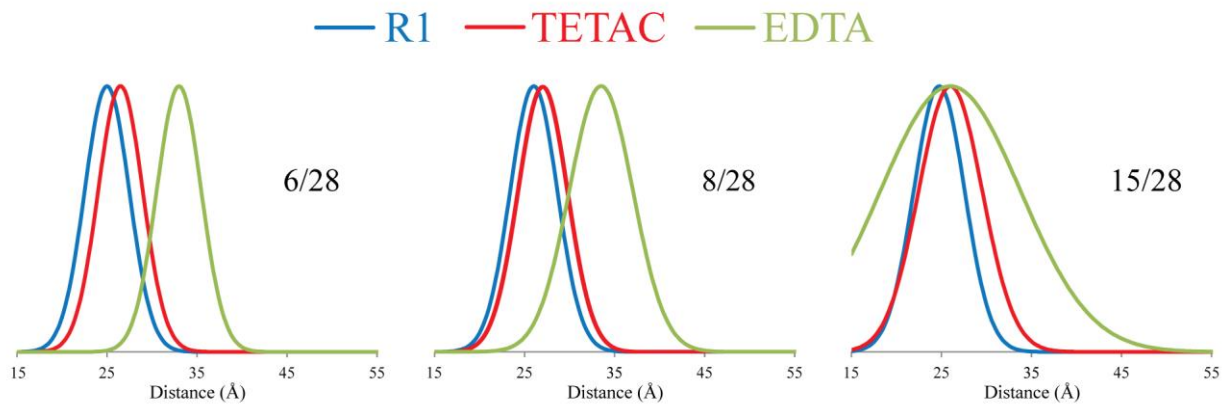

Figure 4-8 The resultant distance distributions from the single Gaussian fits for the double cysteine mutants 6/28, 8/28, and 15/28 labeled with R1 (blue), TETAC-Cu^{2+} (red), and EDTA-Cu^{2+} (green). The R1 time domain fits are shown in Figure 4-3 and the Cu^{2+} time domain fits are in Figure 4-7.

EDTA has nine. Consequently, the resultant TETAC-Cu^{2+} DEER distance distributions show remarkable similarity with each of the R1 distributions (Figure 4-8). For the various mutants, the most probable distance for the TETAC-Cu^{2+} distributions, as compared to the R1 distributions, only differ by 3 Å for the 6/28 mutant, 1 Å for the 8/28 mutant, and 1 Å for the 15/28 mutant. Additionally, the breadth of the distributions are comparable to the R1 distributions for all cases. Despite the variety of solvent-exposed protein environments probed here, the corresponding DEER distributions match well suggesting that the TETAC-Cu^{2+} side chain is a reasonable alternative to the widely utilized R1 spin label for use in DEER distance measurements.

In summary, this work displays the utility of Cu^{2+}-chelating tags as X-band DEER spin probes in various solvent-exposed protein environments. Both Cu^{2+}-chelating tags used here exhibit minimal orientational selectivity at X-band frequencies. Additionally the tags display sufficient sensitivity, partly due to the use of NEM buffer which eliminates the potential negative effects caused by the presence of excess Cu^{2+} ions. The TETAC-Cu^{2+} tag is closest to R1 in terms of the measured distance distribution as compared to other transition metal chelating tags utilized thus far, making it a viable alternative spin label for use as an X-band DEER probe. While the environments probed here are all solvent exposed sites, it will be interesting to determine if these Cu^{2+} tags offer distinct advantages in the membrane environment given the structural similarity of the Gd^{3+} and Cu^{2+} chelating tags. Additionally the possible use of these Cu^{2+} tags as *in vivo* DEER probes is an exciting prospect that may allow for *in-cell* measurements to be made with the more commonly utilized X-band instruments.

4.5 ACKNOWLEDGEMENT

This work was supported by grants from the National Science Foundation (MCB-1157712 to S.S., MCB-1243461 to C.P.J.) and the Camille & Henry Dreyfus Foundation (Camille Dreyfus Teacher-Scholar Award to C.P.J.). The Bruker E680 was purchased with funds from the National Institutes of Health Grant 1S10RR028701. Reprinted (adapted) with permission from Journal Physical Chemistry B, 2015, V. 119, pages 2839-2843.

5.0 THE DOUBLE HISTIDINE CU^{2+}-BINDING MOTIF: A HIGHLY RIGID, SITE-SPECIFIC SPIN PROBE FOR ELECTRON SPIN RESONANCE DISTANCE MEASUREMENTS

This work, written in collaboration with Miriam R. Putterman, Astha Desai, W. Seth Horne, and Sunil Saxena, has been published in Angewandte Chemie, 2015, V. 54, pages 6330–6334. The thesis author collected ESR and CD data, analyzed data, and prepared the manuscript.

5.1 CHAPTER SYNOPSIS

The development of ESR methods that measure long-range distance distributions has advanced biophysical research. However, the spin labels commonly employed are highly flexible, which leads to ambiguity in relating ESR measurements to protein backbone structure. Here we present the double histidine (dHis) Cu^{2+}-binding motif as a rigid spin probe for use in double electron electron resonance (DEER) distance measurements. The spin label is assembled in situ from natural amino acid residues and a metal salt, requires no post-expression synthetic modification, and provides distance distributions dramatically narrower than a common nitoxide protein spin label. Simple molecular modeling based on an X-ray crystal structure of unlabeled protein led to a predicted most probable distance within 0.5 Å of experiment. The use of Cu^{2+} DEER based on

the dHis motif shows great promise for resolving precise, unambiguous distance constraints that report directly on protein backbone structure and flexibility.

5.2 INTRODUCTION

In recent years, the measurement of 1-10 nm distances between paramagnetic species in proteins via pulsed electron spin resonance (ESR) methods have greatly advanced biophysical research.[123, 165] Most current applications of these methods rely on covalent modification of cysteine residues with thiol-reactive spin labels.[123] The most commonly utilized spin-labeling reagent reacts to form the nitroxide-functionalized residue **R1** (Figure 5-1b), and the nitroxide is the reporter for ESR distance measurements. As such, site-directed spin labeling (SDSL), and the Cys-derived R1 residue in particular, has found widespread use in protein ESR as well as NMR.[124] Despite its promise in these applications, SDSL suffers from a significant limitation: R1 measurements are dominated by the conformational dynamics of its flexible side chain rather than local protein backbone fluctuations.

Significant efforts to investigate R1 conformational preferences through X-ray crystallography[12, 29, 31, 135, 166] and computational techniques[25, 35, 37, 167] have helped to address this limitation in part, though the accurate prediction of interspin distances in proteins remains challenging. Much work has also been devoted to develop alternate protein spin labels that are more rigid than R1. Slight changes to the basic R1 residue through modification of the nitroxide ring[6, 168] or its replacement with alternate heterocycles[45] can rigidify its structure. To date the most promising label for proteins appears to be the bifunctional RX side chain in which the

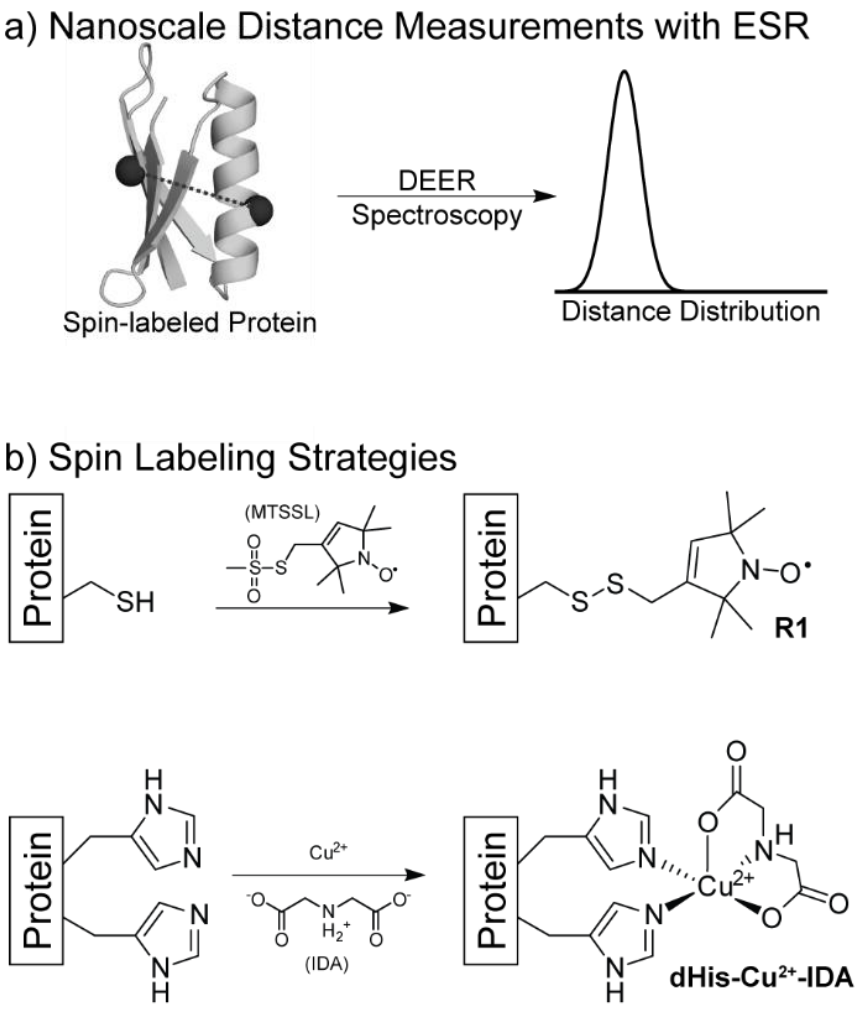

Figure 5-1. (a) Overview of the DEER experiment. (b) The mutations and resulting side chains of the most common spin label (R1) and the paramagnetic metal label reported here (dHis-Cu^{2+}-IDA). Note that the widths of the resultant distance distributions are heavily influenced by the side chain flexibility.

nitroxide ring is covalently attached to two neighboring cysteines.[46] This added rigidity comes at the price of a more complex synthetic scheme necessary to introduce RX into expressed proteins. Thus, despite significant progress, there remains need for a protein spin labeling method that provides narrow, easily interpretable distances and minimizes post-expression synthetic manipulations.

An alternate approach to nitroxides for spin labeling biological systems is paramagnetic metal centers.[64, 65, 67, 169, 170] These probes have been introduced into proteins utilizing both native metal-binding sites[68, 71, 72] as well as unnatural chelating side chains introduced by SDSL-like modification of cysteine.[163, 171, 172] Natural metal-binding sites can provide excellent rigidity,[71, 72, 76, 163, 173] but can only be used when the system of interest natively binds metals. The best unnatural metal-chelating tag shows distance distribution widths comparable to R1, suggesting similar flexibility.[129] Paramagnetic metals can alternatively be incorporated into a protein through the creation of artificial metal binding sites.[73, 74, 76] One such structure that has been utilized for a variety of applications is the double histidine (dHis) motif in which two strategically placed His sides chains are used to chelate Cu^{2+} (Figure 5-1b).[73, 174-178] The dHis motif has been used to chelate Cu^{2+} in the context of an α-helix in T4-lysozyme to measure short-range, relaxation-based, average distances at room temperature.[176]

Presented here is the first use of the dHis-Cu^{2+} motif as a spin probe for double electron electron resonance (DEER) distance measurements. We show that this readily introduced motif is ideal as a highly rigid spin label in both α-helix and β-sheet environments. Individual histidines contain only two rotatable bonds (compared to five in R1), and the simultaneous coordination of Cu^{2+} by two histidines highly restricts the movement of the metal center relative

to the protein backbone. As a means to investigate the use of the dHis-Cu^{2+} motif, two unique sites were incorporated into the B1 immunoglobulin-binding domain of protein G (GB1), one in an α-helix, the other in a β-sheet, and the resultant DEER distance measurement is compared to an analgous measurement utilizing the R1 spin label.

5.3 METHODS

5.3.1 Construction of GB1 Mutants

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

using the PureYield Plasmid Midiprep System (Promega, Madison, WI). All mutations were confirmed by DNA sequencing (Genomic and Proteomics Core Laboratories, University of Pittsburgh, Pittsburgh, PA) and subsequently transformed into BL21(DE3) competent cells (New England Biolabs) for expression.

5.3.2 Protein Expression, Purification, and Labeling

Expression/purification of all GB1 mutants and the labeling of the 6C/28C mutant with MTSSL was performed as was described previously.[135]

5.3.3 Crystallization, Data Collection, and Structure Determination

Crystals of the 6H/8H/28H/32H mutant of GB1 were grown by hanging drop vapor diffusion method. 1 μL of an 18 mg/mL solution of protein in water was mixed with 1 μL of a crystallization buffer composed of 0.2 M NaCl, 0.1 M HEPES pH 7.0, and 1.75 M $(NH_4)_2SO_4$. The drop was allowed to equilibrate at room temperature over a well containing the crystallization buffer. A single crystal was flash frozen in liquid nitrogen after cryoprotection in crystallization buffer supplemented with 30% glycerol by volume. Diffraction data were collected using CuKα radiation on a Rigaku/MSC diffractometer (FR-E generator, Saturn 944 CCD detector) equipped with an X-Stream 2000 low temperature system operated at 100 K. Raw diffraction data were processed with d*TREK. Structure determination and refinement were carried out using the Phenix software suite. The structure was solved by molecular replacement using a published structure of wild-type GB1 (PDB 2QMT) as the search model.

Coordinates and structure factors have been deposited in the Protein Data Bank (PDB: 4WH4). Data collection and refinement statistics are in Table 5-1.

5.3.4 Circular Dichroism

CD data were collected using an Olis DSM17 Circular Dichroism Spectrometer. All samples were 40 μM protein in 20 mM sodium phosphate buffer (pH 7.0) and measured in 2 mm quartz cells. 120 μM of both Cu^{2+} and IDA were included for the appropriate samples. The scans were collected at a temperature of 25 °C from 200 nm to 260 nm with 1 nm increments and a 2 nm bandwidth. The melts were collected by monitoring the signal at 220 nm from 4 °C to 98 °C in 2 °C increments with a 0.5 °C dead band and 2 min equilibration time at each temperature. All measurements were baseline corrected with buffer alone, or buffer including the Cu^{2+} and IDA for the appropriate samples. The melts were fit to a two-state unfolding model to calculate the melting temperatures.

5.3.5 ESR Measurements

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

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

Table 5-1. X-ray crystal data collection and refinement statistics for 6H/8H/28H/32H-GB1.

(IDA) if the IDA was specified as part of the sample. Protein, Cu^{2+}, and IDA concentrations are dependent on the experiment performed.

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

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

The DEER distance measurement was performed on the MTSSL labeled mutant as described previously.[135] The Cu^{2+} distance measurement was performed at 20K. The four pulse sequence utilized for the measurements was $(\pi/2)\nu_1$-τ_1-$(\pi)\nu_1$-T-$(\pi)\nu_2$-τ_2-$(\pi)\nu_1$-τ_2-echo. For the distance measurements displayed in Figure 5-11, the pump pulse (ν_2) was placed at the maximum of the Cu^{2+} spectrum and the observer pulses (ν_1) are placed 150 MHz downfield from the maximum. The observer pulse (ν_1) lengths were 16 ns and 32 ns for the $\pi/2$ and π pulses, and the pump pulse (ν_2) length was 16 ns. The parameter τ_1 was set to 200 ns and T to 160 ns initially with T being increased by 10 ns for 128 steps. τ_2 was adjusted such that τ_2 + T = 1300 ns. Shot repetition time was set to 2000 μs and collection time was approximately 10 hours for

the R1 measurement, and for the Cu^{2+} measurement, shot repetition time was set to 300 μs and the collection time was approximately 16 hours.

5.4 RESULTS AND DISCUSSION

Two Cu^{2+}-binding dHis sites were introduced into GB1 – one in the α-helix and one in the β-sheet. For the helix site, K28H and Q32H mutations were performed to yield **28H/32H-GB1**. The *i* and *i+4* placement of the histidines accounts for the turn of the helix and produces a known metal-binding environment.[174] In the sheet, I6H and N8H were performed to yield **6H/8H-GB1**. The *i* and *i+2* residue placement creates a dHis site by placement of both histidines on the solvent-exposed face of the sheet. Unlike the helix site, to the best of our knowledge, the β-sheet design has yet to be demonstrated experimentally. For Cu^{2+}-DEER measurements, the helix and sheet mutations described above were combined into a single **6H/8H/28H/32H-GB1** variant. Finally, a 6C/28C double mutant of GB1 was created and spin-labeled to yield **6R1/28R1-GB1**.

5.4.1 dHis Mutations Do Not Affect GB1 Folding

Due to the number of mutations involved in dHis labeling, the stability and structure of the mutants was assessed with circular dichroism (CD) and X-ray crystallography. In the CD scans (Figure 5-2), the signatures of all mutants are similar in shape and magnitude to WT GB1 indicating that the fold of all mutants is maintained, even in the presence of Cu^{2+} and IDA. Analysis of temperature-induced unfolding (Figure 5-3) reveals cooperative two-state unfolding

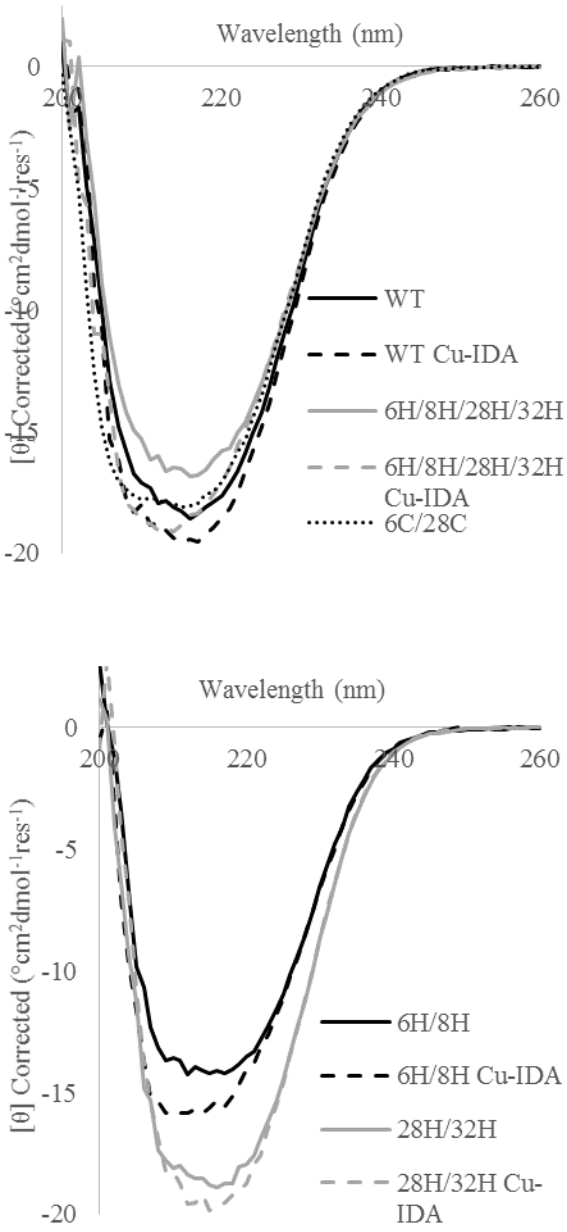

Figure 5-2. The CD scans of WT, 6H/8H, 28H/32H, and 6H/8H/28H/32H GB1 with and without Cu^{2+}-IDA as well as 6C/28C.

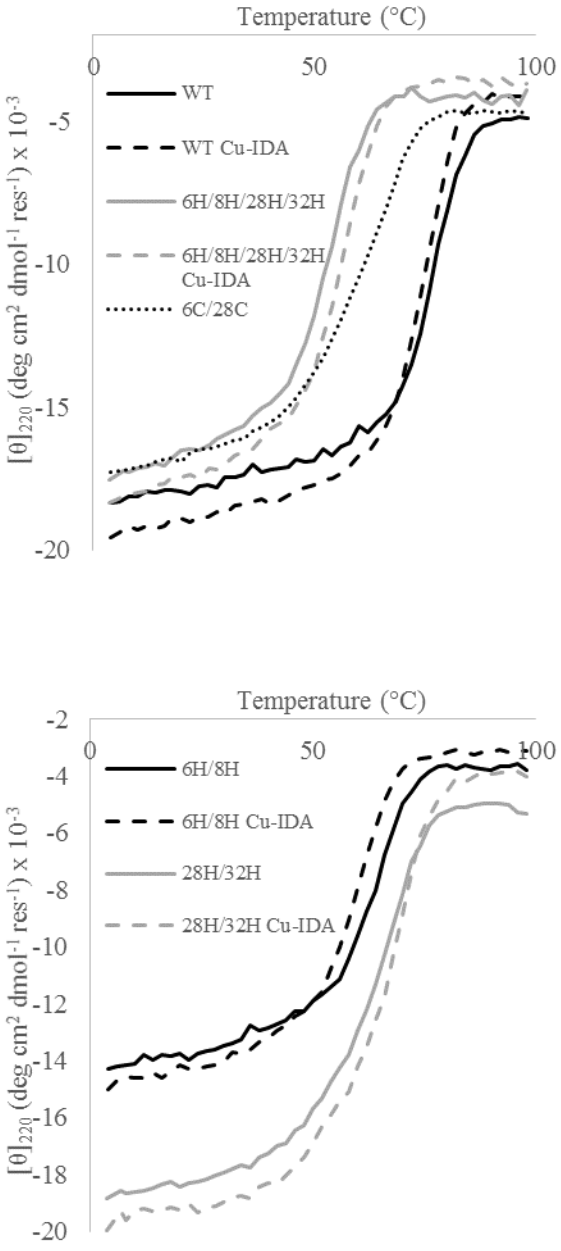

Figure 5-3. The CD thermal melts of WT, 6H/8H, 28H/32H, and 6H/8H/28H/32H GB1 with and without 3:1 Cu^{2+}-IDA, in addition to 6C/28C, monitored at 220 nm. The calculated melting temperatures are shown in Table 5-2.

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

Table 5-2. The calculated melting temperatures for all of the thermal melts displayed in Figure 5-3.

behavior in all samples, consistent with WT-GB1. The melting temperatures, T_m, are somewhat decreased in the mutants indicating a loss in stability from either cysteine or histidine modifications (Figure 5-3 and Table 5-2). Interestingly, the 28H/32H and 6H/8H/28H/32H mutants show increased T_m in the presence of Cu^{2+}-IDA. This observation is consistent with similar stabilization observed previously for a dHis site in an α-helix.[178] The destabilization of the protein upon His incorporation in the β-sheet is consistent with the low sheet folding propensity of the residue.[179]

The effect of the mutations on the global structure of GB1 was also assessed through X-ray crystallography. The crystal structure of **6H/8H/28H/32H-GB1** is shown in Figure 5-4a (PDB 4WH4). As can be seen, the wild type fold of GB1 is maintained despite the mutations. Numerous attempts of soaking the crystal in Cu^{2+} or co-crystallizing with Cu^{2+} proved to be unsuccessful.

5.4.2 IDA Prevents Nonspecific Cu^{2+} Binding to WT-GB1

One potential difficulty of the dHis motif is relatively low Cu^{2+} binding affinity. Apparent dissociation constants ranging from 200 μM to 2 μM have been observed for various α-helix sites.[74] The Cu^{2+} concentrations necessary to fully populate these sites can be accompanied by nonspecific binding of the metal elsewhere in the protein. Previous NMR work has suggested the existence of a nonspecific Cu^{2+} binding site in GB1.[160] In order to assess this binding, continuous wave (CW) spectra were collected for WT-GB1 in the presence of Cu^{2+}. The CW spectrum is sensitive to the number of of equatorially coordinating oxygens or nitrogens (Figure 5-5b) and thus is indicative of the immediate coordination environment. Figure 5-5a displays the

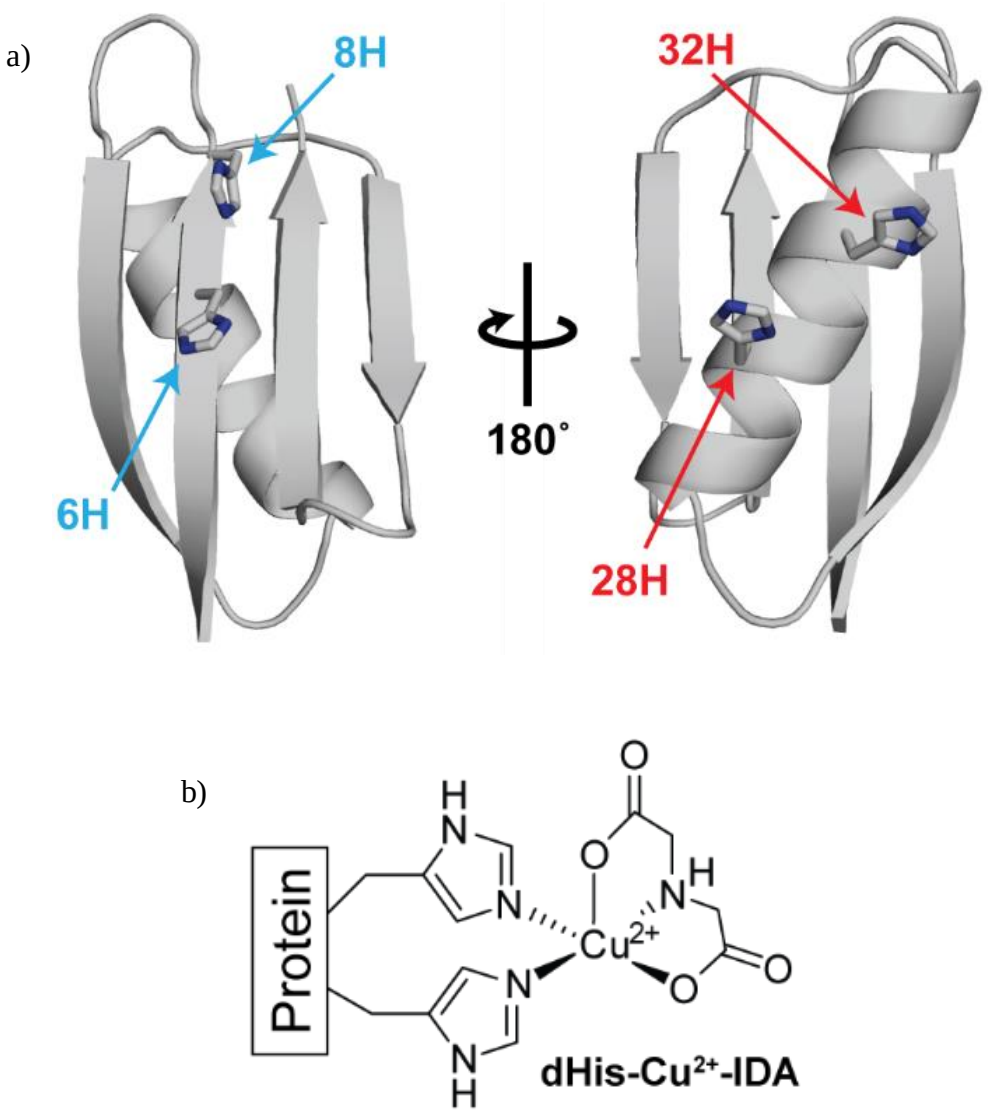

Figure 5-4. (a) Crystal structure of 6H/8H/28H/32H-GB1 (PDB: 4WH4). (b) The expected Cu^{2+} binding environment when simultaneously coordinated to two histidines and iminodiacetate.

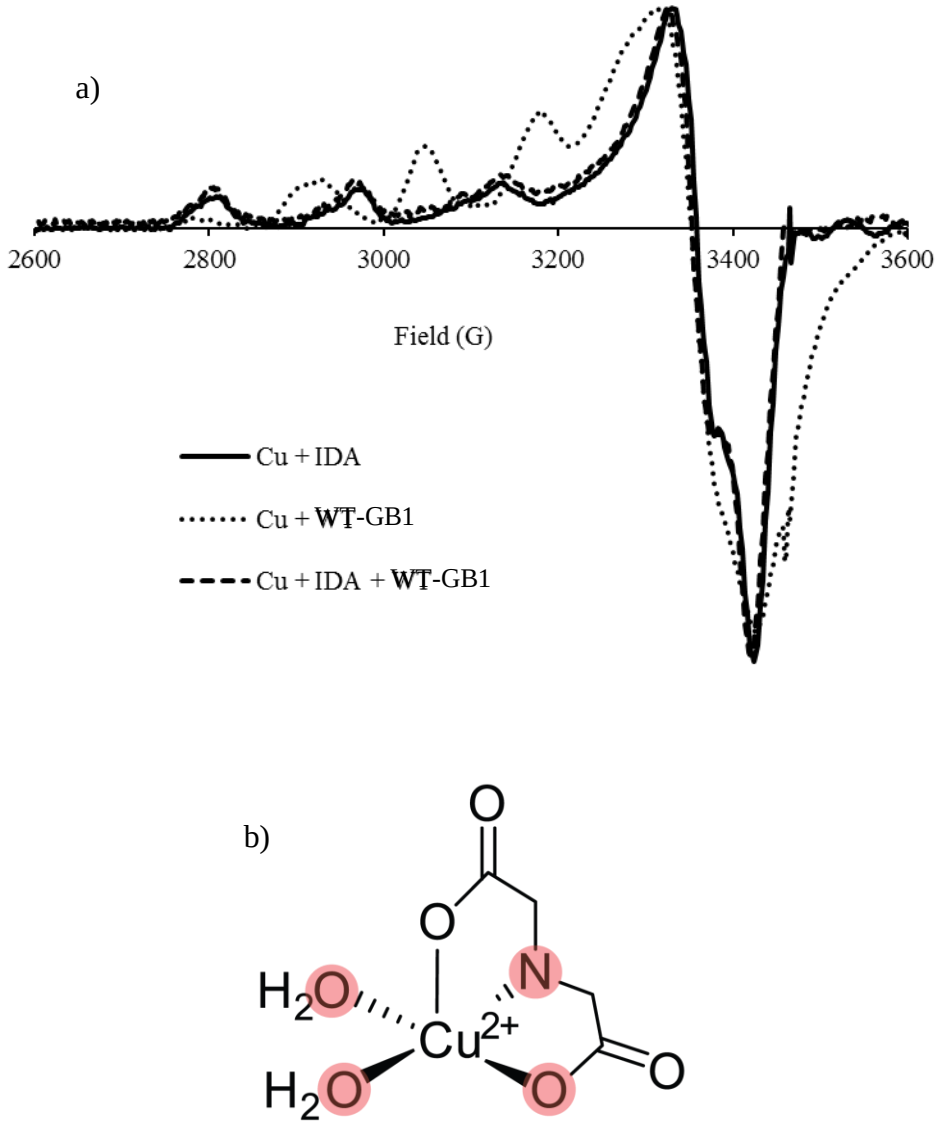

Figure 5-5. (a) Overlay of the continuous wave spectra for Cu^{2+}-IDA, Cu^{2+}-WT GB1, Cu^{2+}-IDA-WT-GB1. (b) The expected Cu^{2+} coordination environment in the Cu^{2+}-IDA complex. The g_I and A_I values are indicative of atoms highlighted in red.

CW spectrum for Cu^{2+} mixed with WT-GB1. The spectrum was collected in NEM buffer. In this buffer, unbound Cu^{2+} forms the insoluble and ESR silent $Cu(OH)_2$. Accordingly, the presence of a Cu^{2+} spectrum indicates that Cu^{2+} is indeed binding to WT-GB1.

As a means to prevent nonspecific binding, Cu^{2+} was introduced as a complex with iminodiacetate (IDA) (Figure 5-4b).[73] Based on a previous crystal structure of Cu^{2+} coordinated to two imidazoles and an IDA derivative,[180] IDA is expected to provide an one equatorial nitrogen, one equatorial oxygen, and one axial oxygen. Accordingly, Cu^{2+} CW spectra were collected for WT-GB1 and IDA and these spectra are displayed in Figure 5-5a. The Cu^{2+} CW spectrum for IDA is distinctly different from the Cu^{2+} spectrum for WT-GB1 which is indicative of a different coordination environment. When WT-GB1 was added to the Cu^{2+} – IDA solution, the resultant CW spectrum clearly matches that of Cu^{2+} – IDA. This result indicates that Cu^{2+} is preferentially binding to IDA and IDA is preventing Cu^{2+} from occupying the natural WT-GB1 site. Adding IDA to WT-GB1 already coordinating Cu^{2+} resulted in a CW spectrum identical to Cu^{2+} – IDA, again indicating the removal of nonspecific binding (data not shown).

ESEEM was also used to assess Cu^{2+} binding to WT-GB1 and prevention of this binding with IDA. ESEEM is capable of detecting remote (~ 5 Å and 8 Å away) nuclear spins and thus ESEEM can further characterize the binding environment. Figure 5-6 display the Cu^{2+} ESEEM spectrum for WT-GB1 and the peak observed at 2.8 MHz is characteristic of a remote backbone nitrogen atom (see Figure 5-6 inset).[181] ESEEM was also used to explore prevention of nonspecific Cu^{2+} binding to WT-GB1 with IDA. Figure 5-6 displays Cu^{2+} ESEEM spectra of IDA and WT-GB1. For Cu^{2+}-IDA, no features indicative of remotely coordinated nitrogens are expected (Figure 5-5b). Accordingly the Cu^{2+} ESEEM spectrum of IDA displays no features in

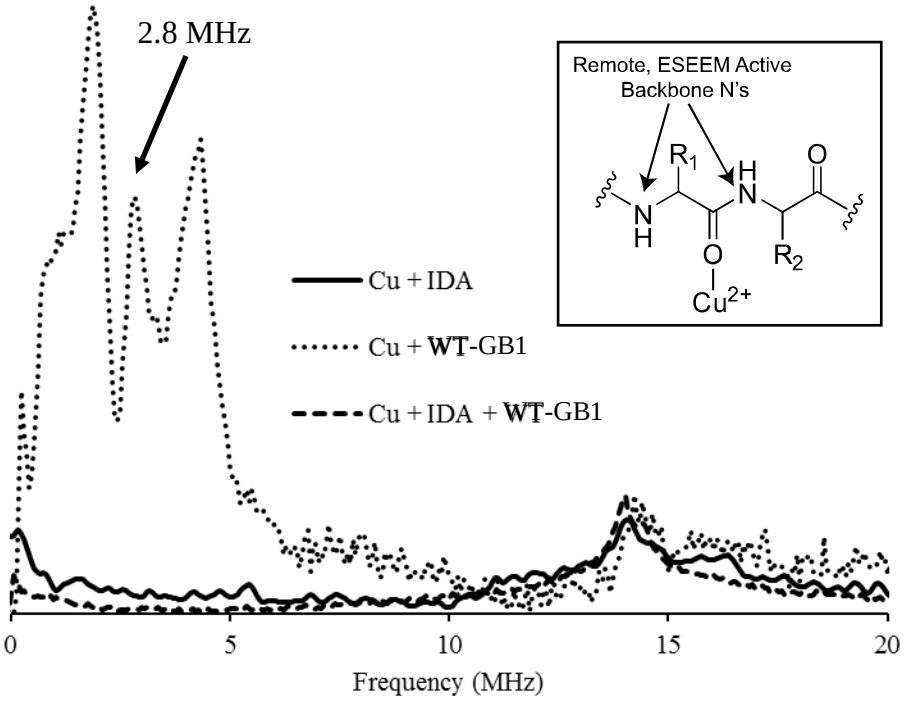

Figure 5-6. Overlay of the ESEEM spectra for Cu^{2+}-IDA, Cu^{2+}-WT GB1, Cu^{2+}-IDA-WT GB1. The spectra have been normalized according to ^{1}H ESEEM intensity. The peak at 2.8 MHz is indicative of a remotely coordinating backbone nitrogen atom (see inset).

the 0 MHz to 4 MHz region of the spectrum indicating the absence of remote nitrogens. When Cu^{2+} - IDA is introduced to WT-GB1, the backbone nitrogen peak is not present and the spectrum is similar to the spectrum observed for Cu^{2+}-IDA alone (Figure 5-6). Taken together, the CW and ESEEM results indicate that IDA prevents the Cu^{2+} from binding to the nonspecific site within WT-GB1.

5.4.3 The dHis-Cu^{2+}-IDA Binding Environment

The Cu^{2+}-IDA complex retains two open equatorial binding sites, potentially allowing for *cis*-coordination by two histidines. In order to investigate the immediate binding environment of the dHis-Cu^{2+} sites, we carried out CW experiments on **6H/8H-GB1** (β-sheet site) and **28H/32H-GB1** (α-helix site) in the presence of Cu^{2+}-IDA (Figure 5-7). Both helix and sheet dHis sites show similar CW spectra, with $g_{\parallel}$ and $A_{\parallel}$ values of 2.263 ± 0.002 and 172 ± 1 G respectively. These values are consistent with three nitrogens and one oxygen directly coordinating to Cu^{2+} in the equatorial plane (see Figure 5-7 inset).[182] The remarkable similarity between the spectra of each site suggests that the dHis motif provides a relatively consistent binding environment in both α-helix and β-sheet contexts.

In the inset of Figure 5-7, five of the expected six ligands chelating to the Cu^{2+} ion in a dHis-Cu^{2+}-IDA complex are shown. As a means to investigate the final ligand in the axial position, an ESEEM spectrum was collected for **28H/32H-GB1** with 1 equivalent of Cu^{2+}-IDA in D_2O with all experimental conditions kept identical. The resulting spectrum is displayed in Figure 5-8. As can be seen, the ^1H ESEEM peak at ~14 MHz has essentially disappeared indicating that solvent H_2O was the primary cause of this signal. The ^2H peak appears at ~2.1

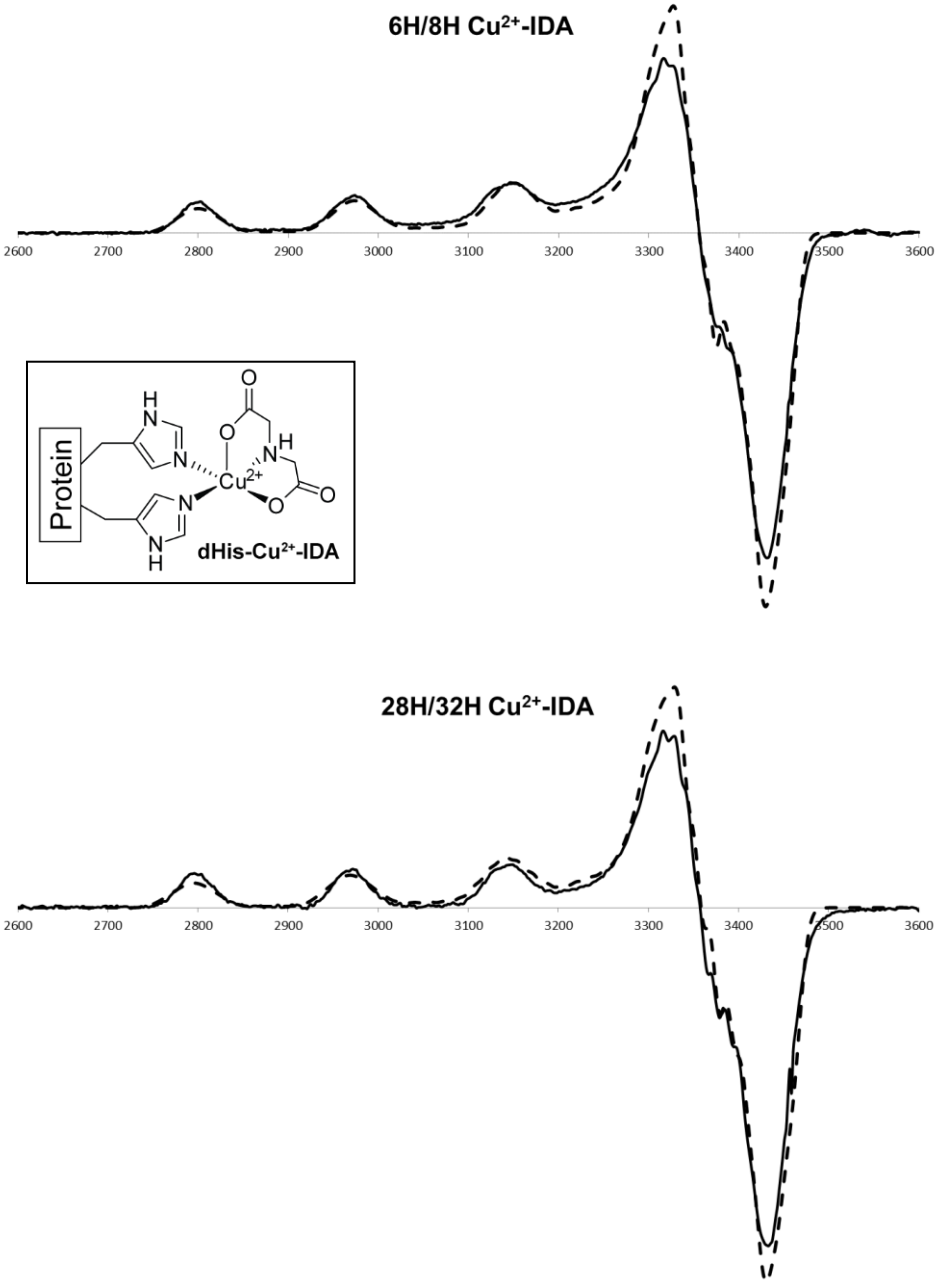

Figure 5-7. CW spectra of 6H/8H and 28H/32H GB1 with Cu^{2+}-IDA (solid) and the corresponding fits (dashed) from Bruker Simfonia software. Inset shows the expected binding environment assembling at each of the dHis sites.

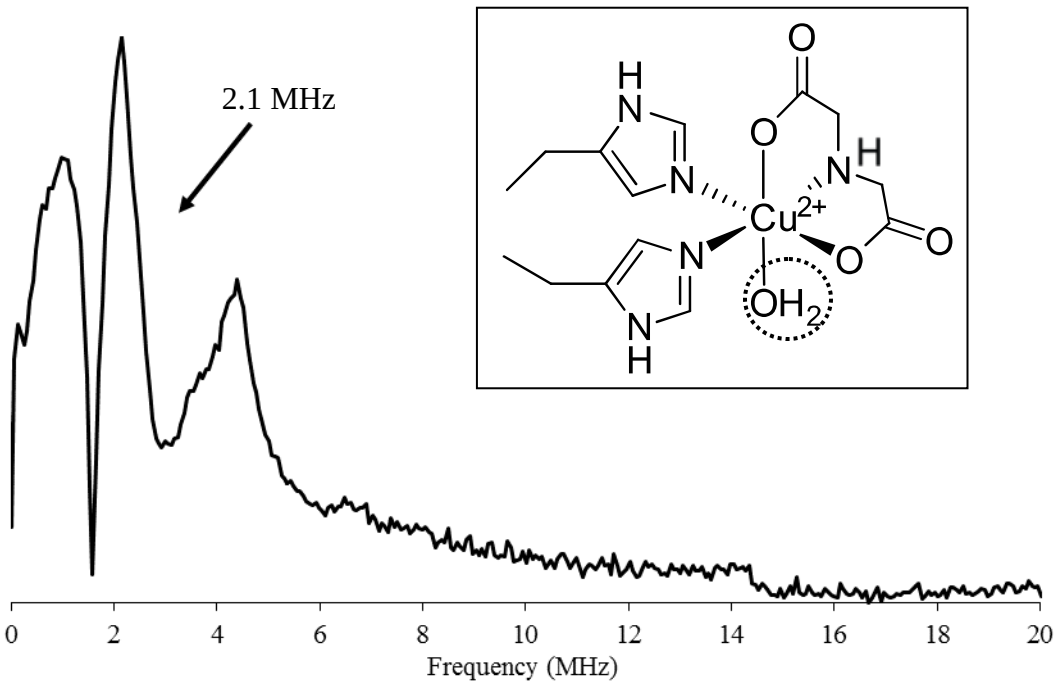

Figure 5-8. The ESEEM spectra of 1:1 Cu^{2+}-IDA:28H/32H-GB1 in D_2O. The stark decrease of the ^1H peak at ~14 MHz and the breadth of the ^2H at ~2.1 MHz both suggest axial H_2O coordination (see inset).

MHz and the breadth of this peak can distinguish between equatorially coordinated, axially coordinated, and ambient D_2O.[183] The peak breadth observed in Figure 5-8 matches best with axial coordination (Figure 5-8 inset) and thus completes the established coordination sphere of the Cu^{2+} centers.

The normalized Cu^{2+}-ESEEM spectrum for each dHis mutant (Figure 5-9) display features common for histidine coordination[184] and the similar spectra suggest similar binding environments for both sites. In particular, the appearance of the slight feature at ~ 8 MHz suggests multiple histidine coordination.[185, 186] Additionally, the ratio of the nuclear quadrapole interaction (~2 MHz) feature and the double quantum (~4 MHz) feature is similar in both complexes, indicating that the number of histidines that coordinate to Cu^{2+} is the same in both mutants.[185, 187, 188] The ratio also matches well with comparable data for a two histidine model complex.[188] Taken together, the ESEEM data and the CW data support the assembly of the binding environment as shown in the inset of Figure 5-8 for both sites.

5.4.4 dHis vs R1 DEER Measurements

Cu^{2+}-Cu^{2+} DEER measurements were performed on the **6H/8H/28H/32H-GB1** sample. The DEER data was acquired at 2 positions in the $g_{\perp}$ and one position in the $g_{\parallel}$ region of the Cu^{2+} spectrum to explore the potential for orientational effects. The locations of the observer and pump pulses are shown in Figure 5-10a. The baseline subtracted time domain data (Figure 5-10b) and the resultant Pake patterns (Figure 5-10c) for each spectrum display no appreciable differences. Additionally, the time domain data is overlaid with the best single Gaussian fits (dotted red in Figure 5-10b) resulting in the distance distributions shown in Figure 5-10d. For

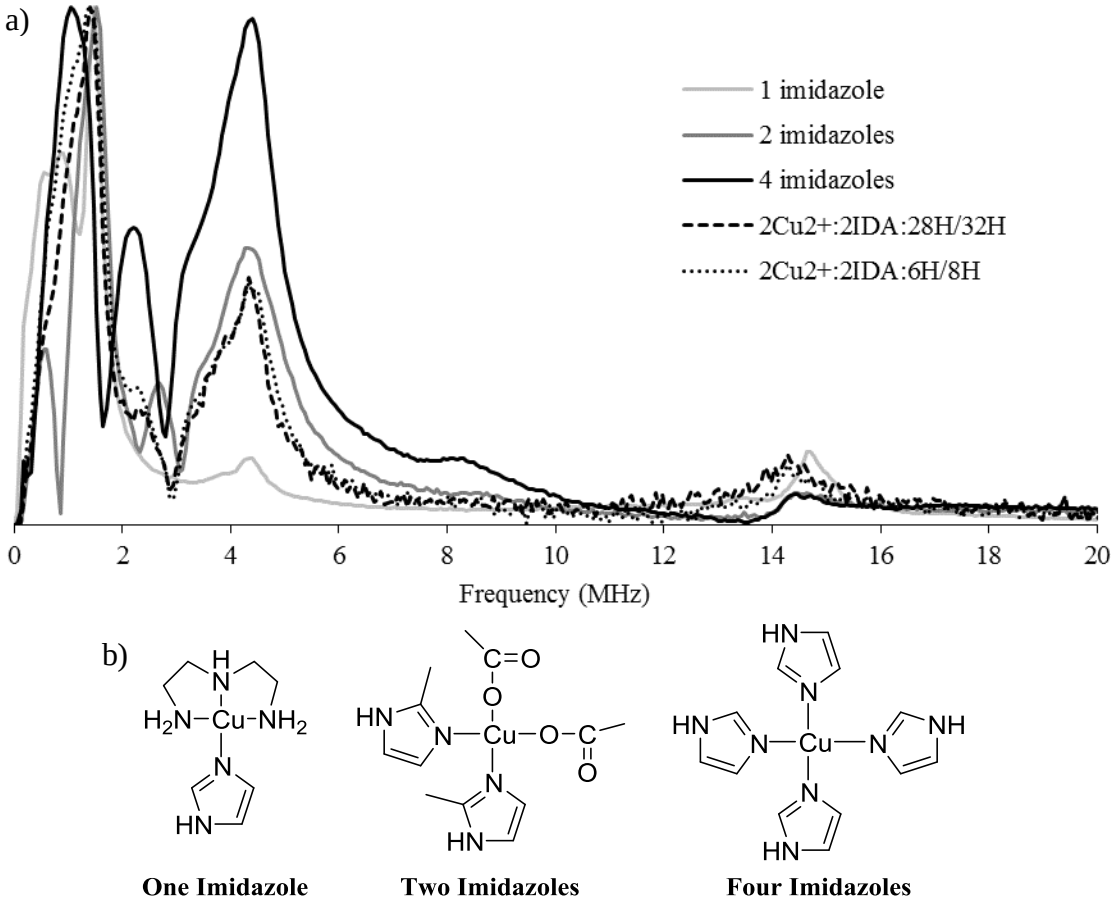

Figure 5-9. (a) Comparison of the ESEEM spectra of the 6H/8H and 28H/32H mutants with Cu^{2+}-IDA compared to ESEEM spectra for the three model imidazole complexes displayed in b. The spectra are normalized to the maximum NQI intensity.

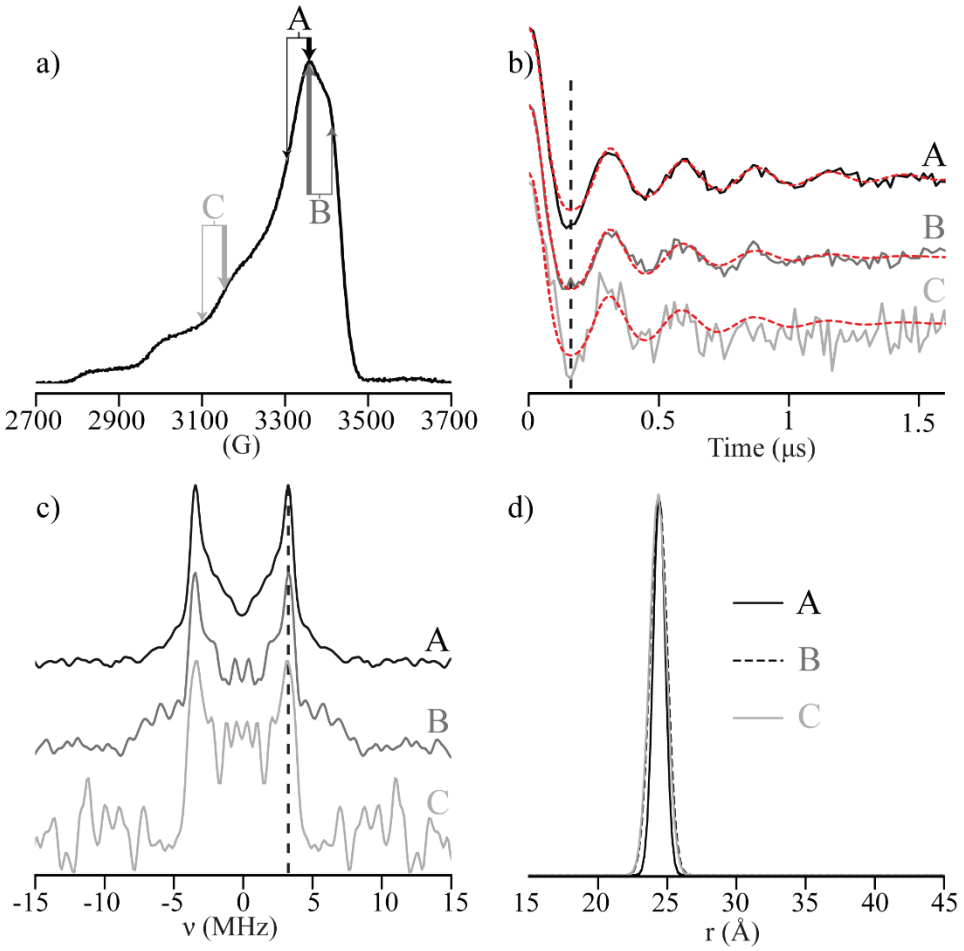

Figure 5-10. (a) Pulse positions (thin arrow indicate observer pulses and thick arrows indicated pump pulses) probed for orientational effects in the Cu^{2+}-Cu^{2+} DEER measurement. (b) The baseline subtracted time domain data and the best single Gaussian fits (dotted red lines), and (c) the Pake patterns for each of these measurements. (d) The single Gaussian distributions for each of the fits in (b).

these distributions, the most probable distance only varies by approximately 0.1 Å with only the breadth of the distributions changing, which is likely due to the lower signal to noise at positions B and C. Taken together, there are negligible orientational effects and as such, only the data from position A is presented in Figure 5-11.

Cu^{2+} and nitroxide DEER experiments were performed on **6H/8H/28H/32H-GB1** and **6R1/28R1-GB1**, respectively. Inspection of the baseline subtracted DEER signals and the resultant distance distributions (Figure 5-11a,b) allows for a direct comparison of the behavior of the two different spin labels. The modulation depths (Figure 5-11a) of the Cu^{2+}-DEER was 2.2% and that of R1 was 33%. Nevertheless, quality signal to noise was achieved in a reasonable collection time (~16 hrs) for the Cu^{2+} measurement. A dramatic increase in the number of modulation periods was observed for dHis-Cu^{2+} relative to R1 (Figure 5-11a), suggesting a much narrower distance distribution for the former. The data were analyzed with Tikhonov regularization[116] (Figure 5-11b), revealing a distance distribution width (between 16% and 84% probability) for dHis remarkably narrower (1.0 Å) than the analogous R1 sample (5.2 Å). This narrow distribution suggests the bound Cu^{2+} centers are highly localized in space providing a very precise distance measurement.

We rationalized the inter-spin distances observed in the DEER experiments by a combination of X-ray crystallography and molecular modeling. For the nitroxide sample, we modeled R1-modified Cys residues into a published crystal structure of WT-GB1; rotamers were based on a previous crystal structure (Figure 5-11c).[135] The observed Cu^{2+}–Cu^{2+} distance was interpreted based on the X-ray crystal structure of **6H/8H/28H/32H-GB1** crystalized in the absence of Cu^{2+} (Figure 5-4). Efforts to derivative these crystals by soaking in Cu^{2+} were

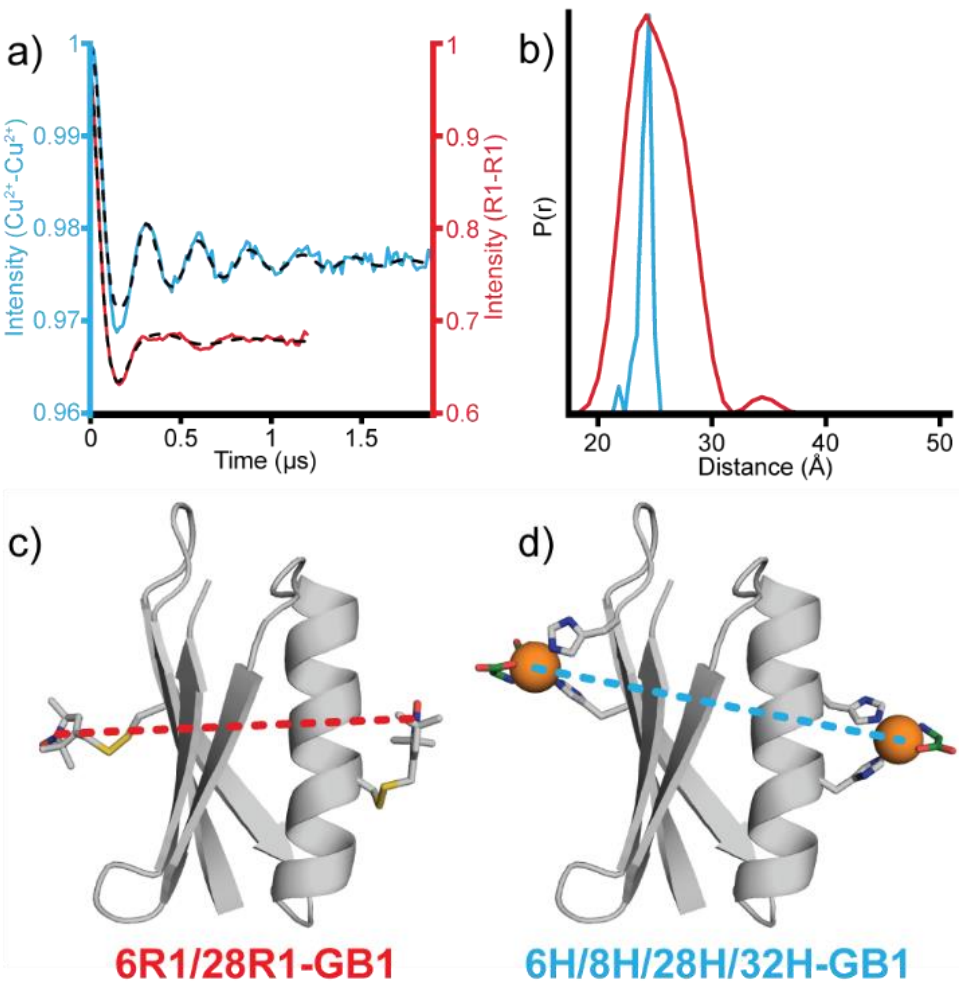

Figure 5-11. (a) Baseline subtracted time domain DEER data for both the R1 (red) and the Cu^{2+} (blue) samples as well as the best fit (dotted black) from Tikhonov regularization; modulation depths have been scaled for comparison. (b) The distance distributions from each sample. (c, d) Structural models of **6R1/28R1-GB1** and **6H/8H/28H/32H-GB1** complexed with Cu^{2+}-IDA with measured inter-spin distances.

unsuccessful, however force field minimization at both the helix and sheet dHis sites provided similar histidine rotamers and coordination geometries as a published crystal structure of Cu^{2+} bound to an IDA derivative and two imidazoles.[180] The predicted Cu^{2+}–Cu^{2+} distance in this model (Figure 5-11d) was 25 Å, in excellent agreement with the measured most probable distance by DEER of 24.5 Å. The ease with which this distance was precisely predicated is remarkable. Despite significant efforts, it is still difficult to predict R1 based distances by modelling techniques.[20, 23-25, 38, 39, 135, 189, 190]

The results reported here on dHis-Cu^{2+} as a protein spin label for DEER are noteworthy in light of the many successful applications of the bifunctional RX label.[47-51] While RX shows greatly increased rigidity compared to R1, our findings suggest dHis-Cu^{2+} is superior in terms of the width of distance distributions obtained by DEER and the simplicity of structural interpretation. In the first use of RX, distribution widths of 2.8 Å and 2.0 Å were reported in two different model proteins where analogous R1 measurements displayed widths of 6.5 Å and 4.7 Å respectively.[46] Additionally one of the distances was modelled using a crystal structure of the RX-modified protein and the predicted distance was 1.4 Å longer than the DEER distance.[46] The DEER distance distribution between the two dHis-Cu^{2+} sites here is considerably narrower (1.0 Å) and the predicted inter-spin distance was within 0.5 Å of experiment. In total, the dHis-Cu^{2+}-IDA motif displays distinct advantages over RX as a highly precise, site-specific spin probe. These advantages are multiplied by the operational simplicity of introducing multiple dHis sites in a single chain and modifying those sites by addition of a metal salt and ligand to solution prior to ESR measurement.

5.5 CONCLUSIONS

In summary, we have shown here that the dHis-Cu^{2+}-IDA motif is a highly rigid, site-specific spin label useful for ESR distance measurements in proteins. The label can be incorporated in both α-helix and β-sheet contexts using natural amino acids and requires no post-expression synthetic manipulations. The distance distribution between dHis sites in a small protein observed by DEER is significantly narrower than that observed for an analogous mutant with R1 labels. The most probable distances are readily interpreted through simple modeling. Together, these results display that Cu^{2+}-DEER using dHis motifs is a simple yet highly effective means of overcoming the inherent limitations of the commonly used R1. In addition, this histidine dependent motif is an ideal alternative in proteins in which incorporation of non-native cysteines for R1 labeling is too perturbative to the native structure. Given that dHis coordination greatly restricts motion of the Cu^{2+} ion, resultant distance measurements can provide structural constraints that can be more easily related to backbone flexibility. With further development, it may be possible to elucidate C_α to C_α distance distributions from these measurements. Such Cu^{2+}-DEER measurements will additionally be important for paramagnetic relaxation enhancement based structure determinations in NMR,[79] and could possibly be combined with crystallography and/or NMR for accurate structure determinations of very large, multicomponent biological systems.[21, 191, 192] Finally, this labeling may be also found suitable for in-cell metal-metal ESR measurements[55] or triangulation of paramagnetic metal binding locations.[193]

5.6 ACKNOWLEDGEMENTS

This work was supported by grants from the National Science Foundation (MCB-1157712 to S.S.) and the Bruker E680 was purchased with funds from the National Institutes of Health Grant 1S10RR028701. Reprinted (adapted) with permission from Angewandte Chemie, 2015, V. 54, pages 6330-6334. Copyright 2015 John Wiley and Sons, License Number: 3674400672859.

6.0 SUMMARY OF MAJOR ACHIEVMENTS

The work presented here is important progress within the field of protein spin labeling. R1 has thus far been the primary spin label used for better understanding protein structure, dynamics, and subsequently function. While most previous efforts have focused on R1 specifically in α-helix or random coil environments, the work presented here provides the necessary data indicating the complexity of R1 in β-sheets. The presented crystallography results of R1 at interior strand sites, non-hydrogen bonded edge strand sites, and hydrogen bonded edge strand sites indicate that the environment surrounding R1 in each of these β-sheet locations dictate different rotameric preferences of R1. As compared with DEER distance measurements, clearly the crystallographic observations are likely representative of R1 rotameric preferences in solution. Additionally, it is clear that current modelling approaches are unable to adequately describe R1 in the varyingly complex environments of β-sheets.

In order to establish alternate spin labels that address the limitations of R1, site specific incorporation of Cu^{2+} into proteins for ESR distance measurements was developed here. First are site specific chelating tags. Similar to R1, these tags take advantage of the specificity associated with cysteines to specifically place tags within a protein structure. One of the tags, the cyclen tag, exhibits a flexibility comparable to R1. While this result indicates specifically this cyclen tag as an adequate alternate tag for R1, potentially in the cellular environment, further improvement upon labeling was necessary to address the flexibility concerns. As such, Cu^{2+} was

incorporated through an alternate means; an artificial metal binding motif. The motif, specifically the double histidine binding motif, is capable of site specifically directing the binding location of Cu^{2+} within various protein secondary structure contexts. In addition, specific introduction of Cu^{2+} complexed with IDA provides the specificity to select for the double histidine sites and prevent Cu^{2+} binding to alternate, nonspecific locations. At these double histidine sites, the Cu^{2+} centers are highly localized and thus provide highly precise distance measurements exhibiting distribution widths potentially on par with backbone fluctuations. As compared to every other spin labeling method, this procedure provides a simple protocol that yields the most narrow and thus most precise protein distance measurements.

These results, specifically the double histidine results, exhibit a substantial step forward in protein spin labeling methodologies. Spin labels which generate distances that are reflective of relative backbone fluctuations have always been the goal and this motif seems like the best candidate for providing such results. Given the simplicity of its implementation, we foresee this spin labeling method becoming useful to investigate a number of different protein systems. The general procedure may be further developed such that different chelators besides IDA could be used, or potentially different paramagnetic metals could be used, to adapt this protocol for various applications. Additionally, the method may be useful in the cellular environment.

My efforts at the University of Pittsburgh have yielded the following publications so far:

Ruthstein, S., Stone, K., Cunningham, T. F., Ji, M., Cascio, M., and Saxena, S. (2010) Pulsed electron spin resonance resolves the coordination site of Cu^{2+} ions in α1-glycine receptor. *Biophys. J. 99*, 2497–2506.

Mamonov, A. B., Lettieri, S., Ding, Y., Sarver, J. L., Palli, R., Cunningham, T. F., Saxena, S., and Zuckerman, D. M. (2012) Tunable, mixed-resolution modeling using library-based Monte Carlo and graphics processing units. *J. Chem. Theory Comput. 8* (8), 2921-2929.

Cunningham, T. F., McGoff, M., Sengupta, I., Jaroniec, C., Horne, W., and Saxena, S. (2012) High-resolution structure of a protein spin-label in a solvent-exposed β-sheet and comparison with DEER spectroscopy. *Biochemistry 51*, 6350–9.

Sengupta, I., Gao, M., Arachchige, R., Nadaud, P. S., Cunningham, T. F., Saxena, S., Schwieters, C., and Jaroniec, C. P. (2015) Protein Structural Studies by Paramagnetic Solid-State NMR Spectroscopy Aided by a Compact Cyclen-Type Cu(II) Binding Tag, *J. Biomol. NMR. 61*, 1-6.

Cunningham, T. F., Shannon, M. D., Putterman, M. R., Arachchige, R., Sengupta, I., Gao, M., Jaroniec, C. P., and Saxena, S. (2015) Cysteine Specific Cu^{2+} Chelating Tags used as Paramagnetic Probes in Double Electron Electron Resonance, *J. Phys. Chem. B 119*, 2839-2843.

Cunningham, T. F., Putterman, M. R., Desai, A., Horne, W., and Saxena, S. (2015) The double histidine Cu^{2+}-binding motif – A highly rigid, site-specific spin probe for electron spin resonance distance measurements, *Angew. Chem. Int. Ed., 54*, 6330-6334.

Yang, Z., Ji, M., Cunningham, T.F., and Saxena S. (2015) Cu^{2+} as an ESR Probe of Protein Structure and Function, *Methods in Enzymology*, DOI: 10.1016/bs.mie.2015.05.026

Cunningham, T. F., Horne, W., and Saxena, S. (2015) R1 spin label rotameric behavior in H-bonding and non H-bonding edge Beta-strand sites and Implications for Interpretation of Distance Distributions. In preparation

7.0 BIBLIOGRAPHY

[1] Schiemann, O., and Prisner, T. F. (2007) Long-range distance determinations in biomacromolecules by EPR spectroscopy, *Q Rev Biophys 40*, 1-53.

[2] Cafiso, D. S. (2014) Identifying and quantitating conformational exchange in membrane proteins using site-directed spin labeling, *Acc Chem Res 47*, 3102-3109.

[3] Hubbell, W. L., and Altenbach, C. (1994) Investigation of structure and dynamics in membrane proteins using site-directed spin labeling, *Current Opinion in Structural Biology 4*.

[4] Berliner, L. J., Grunwald, J., Hankovszky, H. O., and Hideg, K. (1982) A novel reversible thiol-specific spin label: papain active site labeling and inhibition, *Anal Biochem 119*, 450-455.

[5] Nadaud, P. S., Helmus, J. J., Hofer, N., and Jaroniec, C. P. (2007) Long-range structural restraints in spin-labeled proteins probed by solid-state nuclear magnetic resonance spectroscopy, *J Am Chem Soc 129*, 7502-7503.

[6] Columbus, L., Kalai, T., Jeko, J., Hideg, K., and Hubbell, W. L. (2001) Molecular motion of spin labeled side chains in alpha-helices: analysis by variation of side chain structure, *Biochemistry 40*, 3828-3846.

[7] McHaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics, *Biochemistry 35*, 7692-7704.

[8] McHaourab, H. S., Oh, K. J., Fang, C. J., and Hubbell, W. L. (1997) Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition studied by site-directed spin labeling, *Biochemistry 36*, 307-316.

[9] McHaourab, H. S., Kalai, T., Hideg, K., and Hubbell, W. L. (1999) Motion of spin-labeled side chains in T4 lysozyme: effect of side chain structure, *Biochemistry 38*, 2947-2955.

[10] Columbus, L., and Hubbell, W. L. (2004) Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA, *Biochemistry 43*, 7273-7287.

[11] Lietzow, M. A., and Hubbell, W. L. (2004) Motion of spin label side chains in cellular retinol-binding protein: correlation with structure and nearest-neighbor interactions in an antiparallel beta-sheet, *Biochemistry 43*, 3137-3151.

[12] Langen, R., Oh, K. J., Cascio, D., and Hubbell, W. L. (2000) Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure, *Biochemistry 39*, 8396-8405.

[13] Columbus, L., and Hubbell, W. L. (2002) A new spin on protein dynamics, *Trends Biochem Sci 27*, 288-295.
[14] Meirovitch, E., Nayeem, A., and Freed, J. H. (1984) Analysis of protein-lipid interactions based on model simulations of electron spin resonance spectra, *The Journal of Physical Chemistry 88*, 3454-3465.
[15] Bracken, C., Carr, P. A., Cavanagh, J., and Palmer, A. G., 3rd. (1999) Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: implications for the entropy of association with DNA, *Journal of molecular biology 285*, 2133-2146.
[16] Milov, A. D., Ponomarev, A. B., and Tsvetkov, Y. D. (1984) Electron-electron double resonance in electron spin echo: Model biradical systems and the sensitized photolysis of decalin, *Chemical Physics Letters 110*, 67-72.
[17] Martin, R. E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M., and Spiess, H. W. (1998) Determination of End-to-End Distances in a Series of TEMPO Diradicals of up to 2.8 nm Length with a New Four-Pulse Double Electron Electron Resonance Experiment, *Angewandte Chemie International Edition 37*, 2833-2837.
[18] Ward, R., Bowman, A., Sozudogru, E., El-Mkami, H., Owen-Hughes, T., and Norman, D. G. (2010) EPR distance measurements in deuterated proteins, *J Magn Reson 207*, 164-167.
[19] Bordignon, E. (2012) Site-directed spin labeling of membrane proteins, *Top Curr Chem 321*, 121-157.
[20] Jeschke, G., Bender, A., Schweikardt, T., Panek, G., Decker, H., and Paulsen, H. (2005) Localization of the N-terminal domain in light-harvesting chlorophyll a/b protein by EPR measurements, *J Biol Chem 280*, 18623-18630.
[21] Park, S. Y., Borbat, P. P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A. M., Freed, J. H., Bilwes, A. M., and Crane, B. R. (2006) Reconstruction of the chemotaxis receptor-kinase assembly, *Nat Struct Mol Biol 13*, 400-407.
[22] Gaffney, B. J., Bradshaw, M. D., Frausto, S. D., Wu, F., Freed, J. H., and Borbat, P. (2012) Locating a lipid at the portal to the lipoxygenase active site, *Biophys J 103*, 2134-2144.
[23] Finiguerra, M. G., Prudencio, M., Ubbink, M., and Huber, M. (2008) Accurate long-range distance measurements in a doubly spin-labeled protein by a four-pulse, double electron-electron resonance method, *Magn Reson Chem 46*, 1096-1101.
[24] Swanson, M. A., Kathirvelu, V., Majtan, T., Frerman, F. E., Eaton, G. R., and Eaton, S. S. (2011) Electron transfer flavoprotein domain II orientation monitored using double electron-electron resonance between an enzymatically reduced, native FAD cofactor, and spin labels, *Protein Sci 20*, 610-620.
[25] Sarver, J. L., Townsend, J. E., Rajapakse, G., Jen-Jacobson, L., and Saxena, S. (2012) Simulating the dynamics and orientations of spin-labeled side chains in a protein-DNA complex, *J Phys Chem B 116*, 4024-4033.
[26] Guo, Z., Cascio, D., Hideg, K., Kalai, T., and Hubbell, W. L. (2007) Structural determinants of nitroxide motion in spin-labeled proteins: tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme, *Protein Sci 16*, 1069-1086.
[27] Guo, Z., Cascio, D., Hideg, K., and Hubbell, W. L. (2008) Structural determinants of nitroxide motion in spin-labeled proteins: solvent-exposed sites in helix B of T4 lysozyme, *Protein Sci 17*, 228-239.

[28] Fleissner, M. R., Cascio, D., and Hubbell, W. L. (2009) Structural origin of weakly ordered nitroxide motion in spin-labeled proteins, *Protein Sci 18*, 893-908.
[29] Kroncke, B. M., Horanyi, P. S., and Columbus, L. (2010) Structural origins of nitroxide side chain dynamics on membrane protein alpha-helical sites, *Biochemistry 49*, 10045-10060.
[30] Freed, D. M., Horanyi, P. S., Wiener, M. C., and Cafiso, D. S. (2010) Conformational exchange in a membrane transport protein is altered in protein crystals, *Biophys J 99*, 1604-1610.
[31] Freed, D. M., Khan, A. K., Horanyi, P. S., and Cafiso, D. S. (2011) Molecular origin of electron paramagnetic resonance line shapes on beta-barrel membrane proteins: the local solvation environment modulates spin-label configuration, *Biochemistry 50*, 8792-8803.
[32] Lovell, S. C., Word, J. M., Richardson, J. S., and Richardson, D. C. (2000) The penultimate rotamer library, *Proteins 40*, 389-408.
[33] Tombolato, F., Ferrarini, A., and Freed, J. H. (2006) Dynamics of the nitroxide side chain in spin-labeled proteins, *J Phys Chem B 110*, 26248-26259.
[34] Warshaviak, D. T., Serbulea, L., Houk, K. N., and Hubbell, W. L. (2011) Conformational analysis of a nitroxide side chain in an alpha-helix with density functional theory, *J Phys Chem B 115*, 397-405.
[35] Fajer, M. I., Li, H., Yang, W., and Fajer, P. G. (2007) Mapping electron paramagnetic resonance spin label conformations by the simulated scaling method, *J Am Chem Soc 129*, 13840-13846.
[36] Sezer, D., Freed, J. H., and Roux, B. (2008) Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine alpha-helix, *J Phys Chem B 112*, 5755-5767.
[37] Polyhach, Y., Bordignon, E., and Jeschke, G. (2011) Rotamer libraries of spin labelled cysteines for protein studies, *Phys Chem Chem Phys 13*, 2356-2366.
[38] Klose, D., Klare, J. P., Grohmann, D., Kay, C. W., Werner, F., and Steinhoff, H. J. (2012) Simulation vs. reality: a comparison of in silico distance predictions with DEER and FRET measurements, *PLoS One 7*, e39492.
[39] Hagelueken, G., Ward, R., Naismith, J. H., and Schiemann, O. (2012) MtsslWizard: In Silico Spin-Labeling and Generation of Distance Distributions in PyMOL, *Appl Magn Reson 42*, 377-391.
[40] Engelhard, C., Raffelberg, S., Tang, Y., Diensthuber, R. P., Moglich, A., Losi, A., Gartner, W., and Bittl, R. (2013) A structural model for the full-length blue light-sensing protein YtvA from Bacillus subtilis, based on EPR spectroscopy, *Photochem Photobiol Sci 12*, 1855-1863.
[41] Lemmin, T., Dimitrov, M., Fraering, P. C., and Dal Peraro, M. (2014) Perturbations of the straight transmembrane alpha-helical structure of the amyloid precursor protein affect its processing by gamma-secretase, *J Biol Chem 289*, 6763-6774.
[42] Kerry, P. S., Turkington, H. L., Ackermann, K., Jameison, S. A., and Bode, B. E. (2014) Analysis of influenza A virus NS1 dimer interfaces in solution by pulse EPR distance measurements, *J Phys Chem B 118*, 10882-10888.
[43] Marchetto, R., Schreier, S., and Nakaie, C. R. (1993) A Novel Spin-Labeled Amino Acid Derivative for Use in Peptide Synthesis: (9-Fluorenylmethyloxycarbonyl) -2,2,6,6-tetramethylpiperidine-N-oxyl-4-amino-4-carboxylic *Journal of the American Chemical Society*.

[44] Fawzi, N. L., Fleissner, M. R., Anthis, N. J., Kalai, T., Hideg, K., Hubbell, W. L., and Clore, G. M. (2011) A rigid disulfide-linked nitroxide side chain simplifies the quantitative analysis of PRE data, *J Biomol NMR 51*, 105-114.
[45] Warshaviak, D., Khramtsov, V. V., Cascio, D., Altenbach, C., and Hubbell, W. L. (2013) Structure and dynamics of an imidazoline nitroxide side chain with strongly hindered internal motion in proteins, *J Magn Reson 232*, 53-61.
[46] Fleissner, M. R., Bridges, M. D., Brooks, E. K., Cascio, D., Kalai, T., Hideg, K., and Hubbell, W. L. (2011) Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy, *Proc Natl Acad Sci U S A 108*, 16241-16246.
[47] Sahu, I. D., McCarrick, R. M., Troxel, K. R., Zhang, R., Smith, H. J., Dunagan, M. M., Swartz, M. S., Rajan, P. V., Kroncke, B. M., Sanders, C. R., and Lorigan, G. A. (2013) DEER EPR measurements for membrane protein structures via bifunctional spin labels and lipodisq nanoparticles, *Biochemistry 52*, 6627-6632.
[48] Vanea, E., Gruian, C., Rickert, C., Steinhoff, H. J., and Simon, V. (2013) Structure and dynamics of spin-labeled insulin entrapped in a silica matrix by the sol-gel method, *Biomacromolecules 14*, 2582-2592.
[49] Sahu, I. D., Kroncke, B. M., Zhang, R., Dunagan, M. M., Smith, H. J., Craig, A., McCarrick, R. M., Sanders, C. R., and Lorigan, G. A. (2014) Structural investigation of the transmembrane domain of KCNE1 in proteoliposomes, *Biochemistry 53*, 6392-6401.
[50] Vendome, J., Felsovalyi, K., Song, H., Yang, Z., Jin, X., Brasch, J., Harrison, O. J., Ahlsen, G., Bahna, F., Kaczynska, A., Katsamba, P. S., Edmond, D., Hubbell, W. L., Shapiro, L., and Honig, B. (2014) Structural and energetic determinants of adhesive binding specificity in type I cadherins, *Proceedings of the National Academy of Sciences of the United States of America 111*, E4175-4184.
[51] Bowman, A., Hammond, C. M., Stirling, A., Ward, R., Shang, W., El-Mkami, H., Robinson, D. A., Svergun, D. I., Norman, D. G., and Owen-Hughes, T. (2014) The histone chaperones Vps75 and Nap1 form ring-like, tetrameric structures in solution, *Nucleic Acids Res 42*, 6038-6051.
[52] Igarashi, R., Sakai, T., Hara, H., Tenno, T., Tanaka, T., Tochio, H., and Shirakawa, M. (2010) Distance determination in proteins inside Xenopus laevis oocytes by double electron-electron resonance experiments, *J Am Chem Soc 132*, 8228-8229.
[53] Azarkh, M., Okle, O., Eyring, P., Dietrich, D. R., and Drescher, M. (2011) Evaluation of spin labels for in-cell EPR by analysis of nitroxide reduction in cell extract of Xenopus laevis oocytes, *J Magn Reson 212*, 450-454.
[54] Jagtap, A. P., Krstic, I., Kunjir, N. C., Hansel, R., Prisner, T. F., and Sigurdsson, S. T. (2015) Sterically shielded spin labels for in-cell EPR spectroscopy: analysis of stability in reducing environment, *Free Radic Res 49*, 78-85.
[55] Qi, M., Gross, A., Jeschke, G., Godt, A., and Drescher, M. (2014) Gd(III)-PyMTA label is suitable for in-cell EPR, *J Am Chem Soc 136*, 15366-15378.
[56] Carr, D. H., Brown, J., Bydder, G. M., Steiner, R. E., Weinmann, H. J., Speck, U., Hall, A. S., and Young, I. R. (1984) Gadolinium-DTPA as a contrast agent in MRI: initial clinical experience in 20 patients, *AJR Am J Roentgenol 143*, 215-224.

[57] Matalon, E., Huber, T., Hagelueken, G., Graham, B., Frydman, V., Feintuch, A., Otting, G., and Goldfarb, D. (2013) Gadolinium(III) spin labels for high-sensitivity distance measurements in transmembrane helices, *Angew Chem Int Ed Engl 52*, 11831-11834.
[58] Raitsimring, A. M., Gunanathan, C., Potapov, A., Efremenko, I., Martin, J. M., Milstein, D., and Goldfarb, D. (2007) Gd3+ complexes as potential spin labels for high field pulsed EPR distance measurements, *J Am Chem Soc 129*, 14138-14139.
[59] Potapov, A., Yagi, H., Huber, T., Jergic, S., Dixon, N. E., Otting, G., and Goldfarb, D. (2010) Nanometer-scale distance measurements in proteins using Gd3+ spin labeling, *J Am Chem Soc 132*, 9040-9048.
[60] Yagi, H., Banerjee, D., Graham, B., Huber, T., Goldfarb, D., and Otting, G. (2011) Gadolinium tagging for high-precision measurements of 6 nm distances in protein assemblies by EPR, *J Am Chem Soc 133*, 10418-10421.
[61] Edwards, D. T., Huber, T., Hussain, S., Stone, K. M., Kinnebrew, M., Kaminker, I., Matalon, E., Sherwin, M. S., Goldfarb, D., and Han, S. (2014) Determining the oligomeric structure of proteorhodopsin by Gd3+ -based pulsed dipolar spectroscopy of multiple distances, *Structure 22*, 1677-1686.
[62] Lueders, P., Jager, H., Hemminga, M. A., Jeschke, G., and Yulikov, M. (2013) Distance measurements on orthogonally spin-labeled membrane spanning WALP23 polypeptides, *J Phys Chem B 117*, 2061-2068.
[63] Martorana, A., Bellapadrona, G., Feintuch, A., Di Gregorio, E., Aime, S., and Goldfarb, D. (2014) Probing protein conformation in cells by EPR distance measurements using Gd3+ spin labeling, *J Am Chem Soc 136*, 13458-13465.
[64] Yang, Z., Becker, J., and Saxena, S. (2007) On Cu(II)-Cu(II) distance measurements using pulsed electron electron double resonance, *J Magn Reson 188*, 337-343.
[65] Yang, Z., Ji, M., and Saxena, S. (2010) Practical Aspects of Copper Ion-Based Double Electron Electron Resonance Distance Measurements, *Applied Magnetic Resonance 39*, 487-500.
[66] Yang, Z., Kise, D., and Saxena, S. (2010) An approach towards the measurement of nanometer range distances based on Cu2+ ions and ESR, *J Phys Chem B 114*, 6165-6174.
[67] Bode, B. E., Plackmeyer, J., Prisner, T. F., and Schiemann, O. (2008) PELDOR measurements on a nitroxide-labeled Cu(II) porphyrin: orientation selection, spin-density distribution, and conformational flexibility, *J Phys Chem A 112*, 5064-5073.
[68] van Amsterdam, I. M., Ubbink, M., Canters, G. W., and Huber, M. (2003) Measurement of a Cu-Cu distance of 26 A by a pulsed EPR method, *Angew Chem Int Ed Engl 42*, 62-64.
[69] Ruthstein, S., Stone, K. M., Cunningham, T. F., Ji, M., Cascio, M., and Saxena, S. (2010) Pulsed electron spin resonance resolves the coordination site of Cu(2)(+) ions in alpha1-glycine receptor, *Biophys J 99*, 2497-2506.
[70] Yang, Z., Kurpiewski, M. R., Ji, M., Townsend, J. E., Mehta, P., Jen-Jacobson, L., and Saxena, S. (2012) ESR spectroscopy identifies inhibitory Cu2+ sites in a DNA-modifying enzyme to reveal determinants of catalytic specificity, *Proceedings of the National Academy of Sciences of the United States of America 109*, E993-1000.
[71] van Wonderen, J. H., Kostrz, D. N., Dennison, C., and MacMillan, F. (2013) Refined distances between paramagnetic centers of a multi-copper nitrite reductase determined by pulsed EPR (iDEER) spectroscopy, *Angew Chem Int Ed Engl 52*, 1990-1993.

[72] Merz, G. E., Borbat, P. P., Pratt, A. J., Getzoff, E. D., Freed, J. H., and Crane, B. R. (2014) Copper-based pulsed dipolar ESR spectroscopy as a probe of protein conformation linked to disease states, *Biophys J 107*, 1669-1674.
[73] Arnold, F. H., and Haymore, B. L. (1991) Engineered metal-binding proteins: purification to protein folding, *Science 252*, 1796-1797.
[74] Regan, L. (1993) The design of metal-binding sites in proteins, *Annu Rev Biophys Biomol Struct 22*, 257-287.
[75] Voss, J., Hubbell, W. L., and Kaback, H. R. (1995) Distance determination in proteins using designed metal ion binding sites and site-directed spin labeling: application to the lactose permease of Escherichia coli, *Proceedings of the National Academy of Sciences of the United States of America 92*, 12300-12303.
[76] Yang, Z., Jimenez-Oses, G., Lopez, C. J., Bridges, M. D., Houk, K. N., and Hubbell, W. L. (2014) Long-range distance measurements in proteins at physiological temperatures using saturation recovery EPR spectroscopy, *J Am Chem Soc 136*, 15356-15365.
[77] Nadaud, P. S., Helmus, J. J., Kall, S. L., and Jaroniec, C. P. (2009) Paramagnetic ions enable tuning of nuclear relaxation rates and provide long-range structural restraints in solid-state NMR of proteins, *J Am Chem Soc 131*, 8108-8120.
[78] Nadaud, P. S., Helmus, J. J., Sengupta, I., and Jaroniec, C. P. (2010) Rapid acquisition of multidimensional solid-state NMR spectra of proteins facilitated by covalently bound paramagnetic tags, *J Am Chem Soc 132*, 9561-9563.
[79] Sengupta, I., Nadaud, P. S., Helmus, J. J., Schwieters, C. D., and Jaroniec, C. P. (2012) Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy, *Nat Chem 4*, 410-417.
[80] Sengupta, I., Gao, M., Arachchige, R. J., Nadaud, P. S., Cunningham, T. F., Saxena, S., Schwieters, C. D., and Jaroniec, C. P. (2015) Protein structural studies by paramagnetic solid-state NMR spectroscopy aided by a compact cyclen-type Cu(II) binding tag, *J Biomol NMR 61*, 1-6.
[81] Hubbell, W. L., and Altenbach, C. (1994) Investigation of structure and dynamics in membrane proteins using site-directed spin labeling, *Current Opinion in Structural Biology 4*, 566-573.
[82] Hubbell, W. L., Gross, A., Langen, R., and Lietzow, M. A. (1998) Recent advances in site-directed spin labeling of proteins, *Curr Opin Struct Biol 8*, 649-656.
[83] Hubbell, W. L., Cafiso, D. S., and Altenbach, C. (2000) Identifying conformational changes with site-directed spin labeling, *Nat Struct Biol 7*, 735-739.
[84] Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data, *Biochemistry 39*, 5355-5365.
[85] Liang, B., Bushweller, J. H., and Tamm, L. K. (2006) Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy, *J Am Chem Soc 128*, 4389-4397.
[86] Barnes, J. P., Liang, Z., McHaourab, H. S., Freed, J. H., and Hubbell, W. L. (1999) A multifrequency electron spin resonance study of T4 lysozyme dynamics, *Biophys J 76*, 3298-3306.

[87] Liang, Z., Lou, Y., Freed, J. H., Columbus, L., and Hubbell, W. L. (2004) A Multifrequency Electron Spin Resonance Study of T4 Lysozyme Dynamics Using the Slowly Relaxing Local Structure Model, *The Journal of Physical Chemistry B 108*, 17649-17659.
[88] Zhang, Z., Fleissner, M. R., Tipikin, D. S., Liang, Z., Moscicki, J. K., Earle, K. A., Hubbell, W. L., and Freed, J. H. (2010) Multifrequency electron spin resonance study of the dynamics of spin labeled T4 lysozyme, *J Phys Chem B 114*, 5503-5521.
[89] Milov, A. D., Ponomarev, A. B., and Tsvetkov, Y. D. (1984) Electron-electron double resonance in electron spin echo: model biradical systems and the sensitized photolysis of declain, *Chemical Physics Letters 110*, 67-72.
[90] Martin, R. E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M., and Spiess, H. W. (1998) Determination of end-to-end distances in a series of TEMPO diradicals of up to 2.8 nm length with a new four-pulse double electron electron resonance experiment, *Angewandte Chemie, International Edition in English 37*, 2833-2837.
[91] Tombolato, F., Ferrarini, A., and Freed, J. H. (2006) Dynamics of the nitroxide side chain in spin-labeled proteins, *Journal of Physical Chemistry B 110*, 26248-26259.
[92] Mchaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics, *Biochemistry 35*, 7692-7704.
[93] Freed, D. M., Khan, A. K., Horanyi, P. S., and Cafiso, D. S. (2011) Molecular Origin of Electron Paramagnetic Resonance Line Shapes on beta-Barrel Membrane Proteins: The Local Solvation Environment Modulates Spin-Label Configuration, *Biochemistry*.
[94] Munoz, V., Thompson, P. A., Hofrichter, J., and Eaton, W. A. (1997) Folding dynamics and mechanism of beta-hairpin formation, *Nature 390*, 196-199.
[95] Bauer, M. C., Xue, W. F., and Linse, S. (2009) Protein GB1 folding and assembly from structural elements, *Int J Mol Sci 10*, 1552-1566.
[96] Morrone, A., Giri, R., Toofanny, R. D., Travaglini-Allocatelli, C., Brunori, M., Daggett, V., and Gianni, S. (2011) GB1 is not a two-state folder: identification and characterization of an on-pathway intermediate, *Biophys J 101*, 2053-2060.
[97] Thoms, S., Max, K. E., Wunderlich, M., Jacso, T., Lilie, H., Reif, B., Heinemann, U., and Schmid, F. X. (2009) Dimer formation of a stabilized Gbeta1 variant: a structural and energetic analysis, *J Mol Biol 391*, 918-932.
[98] Walsh, J. D., Meier, K., Ishima, R., and Gronenborn, A. M. (2010) NMR studies on domain diffusion and alignment in modular GB1 repeats, *Biophys J 99*, 2636-2646.
[99] Sengupta, I., Nadaud, P. S., Helmus, J. J., Schwieters, C. D., and Jaroniec, C. P. (2012) Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy, *Nature Chem 4*, 410-417.
[100] Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. M. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G, *Science 253*, 657-661.
[101] Alexander, P., Fahnestock, S., Lee, T., Orban, J., and Bryan, P. (1992) Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures, *Biochemistry 31*, 3597-3603.

[102] Lian, L. Y., Yang, J. C., Derrick, J. P., Sutcliffe, M. J., Roberts, G. C., Murphy, J. P., Goward, C. R., and Atkinson, T. (1991) Sequential 1H NMR assignments and secondary structure of an IgG-binding domain from protein G, *Biochemistry 30*, 5335-5340.
[103] Gallagher, T., Alexander, P., Bryan, P., and Gilliland, G. L. (1994) Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR, *Biochemistry 33*, 4721-4729.
[104] Franks, W. T., Wylie, B. J., Stellfox, S. A., and Rienstra, C. M. (2006) Backbone conformational constraints in a microcrystalline U-15N-labeled protein by 3D dipolar-shift solid-state NMR spectroscopy, *J Am Chem Soc 128*, 3154-3155.
[105] Wylie, B. J., Sperling, L. J., Nieuwkoop, A. J., Franks, W. T., Oldfield, E., and Rienstra, C. M. (2011) Ultrahigh resolution protein structures using NMR chemical shift tensors, *Proc Natl Acad Sci U S A 108*, 16974-16979.
[106] (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr D Biol Crystallogr 50*, 760-763.
[107] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *Journal of Applied Crystallography 40*, 658-674.
[108] Frericks Schmidt, H. L., Sperling, L. J., Gao, Y. G., Wylie, B. J., Boettcher, J. M., Wilson, S. R., and Rienstra, C. M. (2007) Crystal Polymorphism of Protein GB1 Examined by Solid-State NMR Spectroscopy and X-ray Diffraction, *The Journal of Physical Chemistry B 111*, 14362-14369.
[109] Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallographica Section D-Biological Crystallography 53*, 240-255.
[110] Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallographica Section D-Biological Crystallography 60*, 2126-2132.
[111] Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7, *Nat. Protocols 3*, 1171-1179.
[112] Zielke, V., Eickmeier, H., Hideg, K., Reuter, H., and Steinhoff, H.-J. (2008) A commonly used spin label: S-(2,2,5,5-tetramethyl-1-oxyl-[Delta]3-pyrrolin-3-ylmethyl) methanethiosulfonate, *Acta Crystallographica Section C 64*, o586-o589.
[113] Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallographica Section D 66*, 12-21.
[114] Stoll, S., and Schweiger, A. (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR, *Journal of Magnetic Resonance 178*, 42-55.
[115] Budil, D. E., Lee, S., Saxena, S., and Freed, J. H. (1996) Nonlinear-Least-Squares Analysis of Slow-Motion EPR Spectra in One and Two Dimensions Using a Modified Levenberg–Marquardt Algorithm, *Journal of Magnetic Resonance, Series A 120*, 155-189.
[116] Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006 – a comprehensive software package for analyzing pulsed ELDOR data, *Appl Magn Reson*, 473-498.

[117] Louis, J. M., Byeon, I. J., Baxa, U., and Gronenborn, A. M. (2005) The GB1 amyloid fibril: recruitment of the peripheral beta-strands of the domain swapped dimer into the polymeric interface, *J Mol Biol 348*, 687-698.
[118] Jee, J., Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2008) The point mutation A34F causes dimerization of GB1, *Proteins 71*, 1420-1431.
[119] Kuhlman, B., O'Neill, J. W., Kim, D. E., Zhang, K. Y., and Baker, D. (2001) Conversion of monomeric protein L to an obligate dimer by computational protein design, *Proc Natl Acad Sci U S A 98*, 10687-10691.
[120] O'Neill, J. W., Kim, D. E., Johnsen, K., Baker, D., and Zhang, K. Y. (2001) Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus, *Structure 9*, 1017-1027.
[121] Bode, B. E., Margraf, D., Plackmeyer, J., Durner, G., Prisner, T. F., and Schiemann, O. (2007) Counting the monomers in nanometer-sized oligomers by pulsed electron-electron double resonance, *J Am Chem Soc 129*, 6736-6745.
[122] Polyhach, Y., and Jeschke, G. (2010) Prediction of favourable sites for spin labelling of proteins, *Spectroscopy 24*, 651-659.
[123] Hubbell, W. L., López, C. J., Altenbach, C., and Yang, Z. (2013) Technological advances in site-directed spin labeling of proteins, *Current Opinion in Structural Biology 23*, 725-733.
[124] Jeschke, G. (2013) Conformational dynamics and distribution of nitroxide spin labels, *Progress in nuclear magnetic resonance spectroscopy 72*, 42-60.
[125] Cafiso, D. S. (2014) Identifying and Quantitating Conformational Exchange in Membrane Proteins Using Site-Directed Spin Labeling, *Accounts of chemical research*.
[126] Ji, M., Ruthstein, S., and Saxena, S. (2014) Paramagnetic Metal Ions in Pulsed ESR Distance Distribution Measurements, *Accounts of Chemical Research 47*, 688-695.
[127] Goldfarb, D. (2014) Gd3+ spin labeling for distance measurements by pulse EPR spectroscopy, *Physical Chemistry Chemical Physics 16*, 9685-9699.
[128] Martorana, A., Bellapadrona, G., Feintuch, A., Di Gregorio, E., Aime, S., and Goldfarb, D. (2014) Probing Protein Conformation in Cells by EPR Distance Measurements using Gd3+ Spin Labeling, *Journal of the American Chemical Society 136*, 13458-13465.
[129] Cunningham, T. F., Shannon, M. D., Putterman, M. R., Arachchige, R. J., Sengupta, I., Gao, M., Jaroniec, C. P., and Saxena, S. (2015) Cysteine-Specific Cu2+ Chelating Tags Used as Paramagnetic Probes in Double Electron Electron Resonance, *The Journal of Physical Chemistry B 119*, 2839-2843.
[130] Cunningham, T. F., Putterman, M. R., Desai, A., Horne, W. S., and Saxena, S. (2015) The Double Histidine Cu^{2+}-Binding Motif: A Highly Rigid, Site-Specific Spin Probe for Electron Spin Resonance Distance Measurements, *Angewandte Chemie International Edition*, DOI: 10.1002/anie.201501968.
[131] Banerjee, D., Yagi, H., Huber, T., Otting, G., and Goldfarb, D. (2012) Nanometer-Range Distance Measurement in a Protein Using Mn2+ Tags, *The Journal of Physical Chemistry Letters 3*, 157-160.
[132] Martin, R. E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M., and Spiess, H. W. (1998) Determination of End to End Distances in a Series of TEMPO Diradicals of up to 2.8 nm Length with a New Four Pulse Double Electron Electron Resonance Experiment, *Angewandte Chemie International Edition 37*, 2833-2837.

[133] Milov, A. D., Ponomarev, A. B., and Yu, D. T. (1984) Electron-electron double resonance in electron spin echo: Model biradical systems and the sensitized photolysis of decalin, *Chemical Physics Letters 110*, 67-72.
[134] Tombolato, F., Ferrarini, A., and Freed, J. H. (2006) Modeling the effects of structure and dynamics of the nitroxide side chain on the ESR spectra of spin-labeled proteins, *J Phys Chem B 110*, 26260-26271.
[135] Cunningham, T. F., McGoff, M. S., Sengupta, I., Jaroniec, C. P., Horne, W. S., and Saxena, S. (2012) High-resolution structure of a protein spin-label in a solvent-exposed beta-sheet and comparison with DEER spectroscopy, *Biochemistry 51*, 6350-6359.
[136] Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
[137] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *Journal of computational chemistry 26*, 1781-1802.
[138] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J Phys Chem B 102*, 3586-3616.
[139] Pedretti, A., Villa, L., and Vistoli, G. (2004) VEGA--an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming, *Journal of computer-aided molecular design 18*, 167-173.
[140] Jeschke, G. (2002) Distance Measurements in the Nanometer Range by Pulse EPR, *ChemPhysChem 3*, 927-932.
[141] Steinhoff, H.-J. (2004) Inter- and Intra-Molecular Distances Determined by EPR Spectroscopy and Site-Directed Spin Labeling Reveal Protein-Protein and Protein-Oligonucleotide Interaction, *Biological Chemistry 385*, 913-920.
[142] Schiemann, O., and Prisner, T. (2007) Long-Range Distance Determinations in Biomacromolecules by EPR Spectroscopy, *Quarterly Reviews of Biophysics 40*, 1-53.
[143] Hubbell, W. L., Gross, A., Langen, R., and Lietzow, M. A. (1998) Recent Advances in Site-Directed Spin Labeling of Proteins, *Curr. Opin. Struct. Biol. 8*, 649-656.
[144] Hubbell, W. L., Cafiso, D. S., and Altenbach, C. (2000) Identifying Conformational Changes with Site-Directed Spin Labeling, *Nat. Struct. Mol. Biol. 7*, 735-739.
[145] Columbus, L., and Hubbell, W. L. (2002) A New Spin on Protein Dynamics, *Trends Biochem. Sci 27*, 288-295.
[146] Fanucci, G., and Cafiso, D. (2006) Recent Advances and Applications of Site-Directed Spin Labeling, *Current Opinion in Structural Biology 16*, 644-653.
[147] Jeschke, G. (2013) Conformational Dynamics and Distribution of Nitroxide Spin Labels, *Prog. Nucl. Magn. Reson. Spectrosc. 72*, 42-60.
[148] Hubbell, W., López, C., Altenbach, C., and Yang, Z. (2013) Technological Advances in Site-Directed Spin Labeling of Proteins, *Current Opinion in Structural Biology 23*, 725-733.
[149] Banerjee, D., Yagi, H., Huber, T., Otting, G., and Goldfarb, D. (2012) Nanometer-Range Distance Measurement in a Protein using Mn^{2+} Tags, *J. Phys. Chem. Lett. 3*, 157-160.

[150] Potapov, A., Yagi, H., Huber, T., Jergic, S., Dixon, N., Otting, G., and Goldfarb, D. (2010) Nanometer-Scale Distance Measurements in Proteins using Gd^{3+} Spin Labeling, *Journal of the American Chemical Society 132*, 9040-9048.

[151] Matalon, E., Huber, T., Hagelueken, G., Graham, B., Frydman, V., Feintuch, A., Otting, G., and Goldfarb, D. (2013) Gadolinium(III) Spin Labels for High-Sensitivity Distance Measurements in Transmembrane Helices, *Angew. Chem. 52*, 11831-11834.

[152] Edwards, D. T., Huber, T., Hussain, S., Stone, K. M., Kinnebrew, M., Kaminker, I., Matalon, E., Sherwin, M. S., Goldfarb, D., and Han, S. (2014) Determining the Oligomeric Structure of Proteorhodopsin by Gd^{3+}-Based Pulsed Dipolar Spectroscopy of Multiple Distances, *Structure*.

[153] Lueders, P., Jeschke, G., and Yulikov, M. (2011) Double Electron−Electron Resonance Measured Between Gd^{3+} Ions and Nitroxide Radicals, *Journal of Physical Chemistry Letters 2*, 604-609.

[154] Nadaud, P. S., Helmus, J. J., Kall, S. L., and Jaroniec, C. P. (2009) Paramagnetic Ions Enable Tuning of Nuclear Relaxation Rates and Provide Long-Range Structural Restraints in Solid-State NMR of Proteins, *J. Am. Chem. Soc. 131*, 8108-8120.

[155] Nadaud, P. S., Helmus, J. J., Sengupta, I., and Jaroniec, C. P. (2010) Rapid Acquisition of Multidimensional Solid-State NMR Spectra of Proteins Facilitated by Covalently Bound Paramagnetic Tags, *J. Am. Chem. Soc. 132*, 9561-9563.

[156] Sengupta, I., Nadaud, P., Helmus, J., Schwieters, C., and Jaroniec, C. (2012) Protein Fold Determined by Paramagnetic Magic-Angle Spinning Solid-State NMR Spectroscopy, *Nature Chemistry 4*, 410-417.

[157] Sengupta, I., Gao, M., Arachchige, R., Nadaud, P. S., Cunningham, T. F., Saxena, S., Schwieters, C., and Jaroniec, C. P. (2014) Protein Structural Studies by Paramagnetic Solid-State NMR Spectroscopy Aided by a Compact Cyclen-Type Cu(II) Binding Tag, *Journal of Biomolecular NMR*.

[158] Ebright, Y. W., Chen, Y., Pendergrast, P. S., and Ebright, R. H. (1992) Incorporation of an EDTA-Metal Complex at a Rationally Selected Site within a Protein: Application to EDTA-Iron DNA Affinity Cleaving with Catabolite Gene Activator Protein (CAP) and Cro, *Biochemistry 31*, 10664-10670.

[159] Ermacora, M. R., Delfino, J. M., Cuenoud, B., Schepartz, A., and Fox, R. O. (1992) Conformation-Dependent Cleavage of Staphylococcal Nuclease with a Disulfide-Linked Iron Chelate, *Proceedings of the National Academy of Sciences of the United States of America 89*, 6383-6387.

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

[161] Bouvignies, G., Meier, S., Grzesiek, S., and Blackledge, M. (2006) Ultrahigh-Resolution Backbone Structure of Perdeuterated Protein GB1 Using Residual Dipolar Couplings from Two Alignment Media, *Angewandte Chemie 118*, 8346-8349.

[162] Wylie, B. J., Sperling, L. J., Nieuwkoop, A. J., Franks, W. T., Oldfield, E., and Rienstra, C. M. (2011) Ultrahigh Resolution Protein Structures using NMR Chemical Shift Tensors, *Proc. Natl. Acad. Sci. U.S.A. 108*, 16974-16979.

[163] Ji, M., Ruthstein, S., and Saxena, S. (2014) Paramagnetic metal ions in pulsed ESR distance distribution measurements, *Acc Chem Res 47*, 688-695.
[164] Aronoff-Spencer, E., Burns, C., Avdievich, N., Gerfen, G., Peisach, J., Antholine, W., Ball, H., Cohen, F., Prusiner, S., and Millhauser, G. (2000) Identification of the Cu^{2+} Binding Sites in the N-Terminal Domain of the Prion Protein by EPR and CD Spectroscopy, *Biochemistry 39*, 13760-13771.
[165] Jeschke, G. (2012) DEER distance measurements on proteins, *Annual review of physical chemistry 63*, 419-446.
[166] Fleissner, M. R., Cascio, D., and Hubbell, W. L. (2009) Structural origin of weakly ordered nitroxide motion in spin-labeled proteins, *Protein Sci. 18*, 893-908.
[167] Hatmal, M. M., Li, Y., Hegde, B. G., Hegde, P. B., Jao, C. C., Langen, R., and Haworth, I. S. (2012) Computer modeling of nitroxide spin labels on proteins, *Biopolymers 97*, 35-44.
[168] Fawzi, N. L., Fleissner, M. R., Anthis, N. J., Kalai, T., Hideg, K., Hubbell, W. L., and Clore, G. M. (2011) A rigid disulfide-linked nitroxide side chain simplifies the quantitative analysis of PRE data, *J. Biomol. NMR 51*, 105-114.
[169] Narr, E., Godt, A., and Jeschke, G. (2002) Selective measurements of a nitroxide-nitroxide separation of 5 nm and a nitroxide-copper separation of 2.5 nm in a terpyridine-based copper(II) complex by pulse EPR spectroscopy, *Angewandte Chemie (English Edition) 41*, 3907-3910.
[170] Becker, J. S., and Saxena, S. (2005) Double quantum coherence electron spin resonance on coupled Cu(II)-Cu(II) electron spins, *Chemical Physics Letters 414*, 248-252.
[171] Goldfarb, D. (2012) Metal-Based Spin Labeling for Distance Determination, *Springer*.
[172] Goldfarb, D. (2014) Gd3+ spin labeling for distance measurements by pulse EPR spectroscopy, *PCCP 16*, 9685-9699.
[173] Yang, Z., Kurpiewski, M. R., Ji, M., Townsend, J. E., Mehta, P., Jen-Jacobson, L., and Saxena, S. (2012) ESR spectroscopy identifies inhibitory Cu2+ sites in a DNA-modifying enzyme to reveal determinants of catalytic specificity, *Proc. Natl. Acad. Sci. U. S. A. 109*, E993-1000.
[174] Todd, R. J., Van Dam, M. E., Casimiro, D., Haymore, B. L., and Arnold, F. H. (1991) Cu(II)-binding properties of a cytochrome c with a synthetic metal-binding site: His-X3-His in an alpha-helix, *Proteins 10*, 156-161.
[175] Higaki, J. N., Fletterick, R. J., and Craik, C. S. (1992) Engineered metalloregulation in enzymes, *Trends in biochemical sciences 17*, 100-104.
[176] Voss, J., Salwinski, L., Kaback, H. R., and Hubbell, W. L. (1995) A method for distance determination in proteins using a designed metal ion binding site and site-directed spin labeling: evaluation with T4 lysozyme, *Proceedings of the National Academy of Sciences of the United States of America 92*, 12295-12299.
[177] Jung, K., Voss, J., He, M., Hubbell, W. L., and Kaback, H. R. (1995) Engineering a metal binding site within a polytopic membrane protein, the lactose permease of Escherichia coli, *Biochemistry 34*, 6272-6277.
[178] Nicoll, A. J., Miller, D. J., Futterer, K., Ravelli, R., and Allemann, R. K. (2006) Designed high affinity Cu2+-binding alpha-helical foldamer, *Journal of the American Chemical Society 128*, 9187-9193.

[179] Minor, D. L., Jr., and Kim, P. S. (1994) Measurement of the beta-sheet-forming propensities of amino acids, *Nature 367*, 660-663.
[180] Brandi-Blanco, M. P., de Benavides-Giménez, M. M., González-Pérez, J. M., and Choquesillo-Lazarte, D. (2007) cis-[N-(4-Chlorobenzyl)iminodiacetato-[kappa]3N,O,O']bis(1H-imidazole-[kappa]N3)copper(II), *Acta Crystallographica Section E: Structure Reports Online 63*, m1678-m1679.
[181] Burns, C. S., Aronoff-Spencer, E., Dunham, C. M., Lario, P., Avdievich, N. I., Antholine, W. E., Olmstead, M. M., Vrielink, A., Gerfen, G. J., Peisach, J., Scott, W. G., and Millhauser, G. L. (2002) Molecular features of the copper binding sites in the octarepeat domain of the prion protein, *Biochemistry 41*, 3991-4001.
[182] Peisach, J., and Blumberg, W. E. (1974) Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins, *Archives of biochemistry and biophysics 165*, 691-708.
[183] Lu, J., Bender, C. J., McCracken, J., Peisach, J., Severns, J. C., and McMillin, D. R. (1992) Pulsed EPR studies of the type 2 copper binding site in the mercury derivative of laccase, *Biochemistry 31*, 6265-6272.
[184] Mims, W. B., and Peisach, J. (1978) The nuclear modulation effect in electron spin echoes for complexes of Cu2+ and imidazole with 14N and 15N, *Journal of Chemical Physics 69*, 4921-4930.
[185] Goldfarb, D., Fauth, J., Tor, Y., and Shanzer, A. (1991) Study of copper (II) binding to chiral tripodal ligands by electron spin echo spectroscopy, *Journal of the American Chemical Society 113*, 1941-1948.
[186] Hernandez-Guzman, J., Sun, L., Mehta, A. K., Dong, J., Lynn, D. G., and Warncke, K. (2013) Copper(II)-bis-histidine coordination structure in a fibrillar amyloid beta-peptide fragment and model complexes revealed by electron spin echo envelope modulation spectroscopy, *Chembiochem : a European journal of chemical biology 14*, 1762-1771.
[187] McCracken, J., Pember, S., and Benkovic, S. J. (1988) Electron spin-echo studies of the copper binding site in phenylalanine hydroxylase from Chromobacterium violaceum, *Journal of the American Chemical Society*.
[188] Silva, K. I., Michael, B. C., Geib, S. J., and Saxena, S. (2014) ESEEM analysis of multi-histidine Cu(II)-coordination in model complexes, peptides, and amyloid-beta, *Journal of Physical Chemistry B 118*, 8935-8944.
[189] Polyhach, Y., Bordignon, E., and Jeschke, G. (2011) Rotamer libraries of spin labelled cysteines for protein studies, *PCCP 13*, 2356-2366.
[190] Roux, B., and Islam, S. M. (2013) Restrained-ensemble molecular dynamics simulations based on distance histograms from double electron-electron resonance spectroscopy, *Journal of Physical Chemistry B 117*, 4733-4739.
[191] Duss, O., Yulikov, M., Jeschke, G., and Allain, F. H. (2014) EPR-aided approach for solution structure determination of large RNAs or protein-RNA complexes, *Nature communications 5*, 3669.
[192] Duss, O., Michel, E., Yulikov, M., Schubert, M., Jeschke, G., and Allain, F. H. (2014) Structural basis of the non-coding RNA RsmZ acting as a protein sponge, *Nature 509*, 588-592.

[193] Abdullin, D., Florin, N., Hagelueken, G., and Schiemann, O. (2015) EPR-Based Approach for the Localization of Paramagnetic Metal Ions in Biomolecules, *Angewandte Chemie (English Edition) 54*, 1827-1831.