CHARACTERIZING PROTEIN CONFORMATIONS USING COPPER-BASED PULSED DIPOLAR ESR SPECTROSCOPY AND COMPLIMENTARY BIOPHYSICAL METHODS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Gregory Edward Merz

February 2016

CHARACTERIZING PROTEIN CONFORMATIONS USING COPPER-BASED PULSED DIPOLAR ESR SPECTROSCOPY AND COMPLIMENTARY BIOPHYSICAL METHODS

Gregory Edward Merz, Ph. D.

Cornell University 2016

Proteins play an important and diverse role in all living organisms. If proteins are unable to carry out their prescribed functions, the results can be problematic, or even fatal, for an organism. For example, in humans, Alzheimer's and Parkinson's are just 2 of the many diseases caused by proteins which either do not function, or function incorrectly. Proteins can also perform chemical transformations which are very difficult via synthetic methods, such as the oxidation of methane to methanol and the fixation of dinitrogen to ammonia. Understanding the mechanisms of these processes may lead to much more efficient catalysts, greatly reducing the large energy expenditures currently required.

In biochemistry, the link between structure and function has been well established, and so in order to understand the mechanisms and functions of proteins, we must understand their structures. In many cases, the structures of flexible proteins can be difficult to elucidate, especially if multiple conformations exist simultaneously. Here, we use copper-based pulsed dipolar ESR spectroscopy (PDS) and other, complimentary biophysical and biochemical methods to characterize protein

conformations in flexible proteins. These include mutants of Superoxide Dismutase 1 (SOD1) which cause familial ALS, and the *Drosophila melanogaster* circadian clock protein Period.

By using these techniques, we show that fALS mutants of SOD1 tend to aggregate in solution as opposed to the wild-type (WT) protein which does not. Furthermore, we propose a structural mechanism by which this aggregation occurs. In the Period protein, we have discerned small differences in the conformation of mutants that mimic phosphorylation vs. the WT. These subtle changes may cause differences in circadian behavior observed in fruit flies.

BIOGRAPHICAL SKETCH

Gregory Edward Merz was born on May 23, 1988 to Gary and Regina Merz in Rochester, New York. He attended St. John the Evangelist elementary school through the 6th grade, where his interest in the sciences was initially piqued. For middle and high school, he attended McQuaid Jesuit High School, a Catholic college preparatory school in the nearby town of Brighton. Here, he enjoyed many of his classes, especially Fr. Wroblewski's Sophomore Advanced English, and Brad Bovenzi's chemistry and AP chemistry classes. Upon graduating from McQuaid, Gregory moved on to Allegheny College in Meadville, Pennsylvania, where he obtained a Bachelor of Science Degree, with a double major in chemistry and economics. While in college, he participated in intercollegiate athletics, playing on the golf team, as well as student government and Greek life. He was also a researcher in the biochemistry lab or Dr. Marty Serra, and wrote his senior thesis on the thermodynamics and structure of single-nucleotide bulge loops in RNA hairpins. Dr. Serra was a wonderful research advisor and mentor, and encouraged his continued study in chemistry. Following graduation from Allegheny, Gregory came to Cornell University, where, for the last 5 years, he has been studying biochemistry under Dr. Brian Crane. His work in the Crane lab has specifically focused on using Copper-Based Pulsed Dipolar ESR Spectroscopy along with other biochemical methods to understand protein conformations.

To my parents

ACKNOWLEDGMENTS

I would like to first thank my parents and my family, as they have taught me a great deal about how to live my life. They showed me the importance of hard work, which has certainly been necessary during this journey. My family has always fully supported me and helped me in whatever ways they could, and for that I am eternally grateful.

I would also like to thank my advisor, Brian Crane. He has been a wonderful mentor to me, and has taught me more than I could have imagined about science and also about being a scientist and a person. I cannot imagine working in any other lab, as I have enjoyed my time in the Crane lab immensely. His hands-off approach has really suited the way I like to work (especially the hours I like to work) and I believe I have been more productive because of it. I have also been encouraged to form my own ideas for my projects and experiments, and have been allowed to explore those ideas, something which I think will really benefit me going forward. I cannot say enough "thank-yous" to Brian for taking me into his lab.

My committee members have also helped me greatly along the way, especially Professor Freed. My collaborators, especially Peter Borbat, have helped me tremendously during my time at Cornell, and for that I am very grateful.

I would like to thank all of my lab-mates, both current and former, and they have made the Crane lab a fantastic place to work. My original mentor in the lab, Xiaoxiao, taught me many of the techniques I still use today and passed on her generous wisdom in the lab. Anand was and still is my best friend from the Crane lab. I really enjoyed our late-night discussions on science, (repetitive) music, culture, and life in general. I still enjoy our chats on the phone from time to time, but it isn't quite the same as having him sitting 2 seats down from me in the office. The night shift is quite a bit lonelier now that no one is here, but as Anand predicted, I am also quite a bit more efficient. I would be remiss if I didn't thank the other member of the night-shift crew, Ken Gee. It was always amusing hearing his hilarious one-liners and thoughts on life. The days of playing HORSE in the lab with wet paper towels and garbage cans will not soon be forgotten. Ria was also a great companion in the lab, especially when I wanted to waste time chatting in her office. She always had a kind word or something funny to say, and also really understood my experiences in the lab. I miss having her around. My office mates Craig and Dipanjan have made for a great office experience. Although they generally keep to themselves, they are always there to answer my questions or engage in some friendly conversation. I've also had the privilege of training several students, which has been very rewarding. Alise and Mike have worked on the synthetic amino acid project with me, and Kritika on the Period project. All 3 of them are doing really great work in the lab (even if they don't believe it) and I'm happy that my projects are in such good hands as theirs. I have also

worked with several wonderful undergrads, Matt Byrne and Vasia Giannakakos and a high school student Connor McAllister. It was amazing to see their growth and development in their short time in the lab, and I'm really grateful for the work they did to help me out. Of course I cannot forget all of the other members of the Crane lab, both past and present: Karen, Sarah, Anna, Tom, Magali, Bee, Estella, Angela, and our undergrads Pete, Camille, Nancy, Erika, Sophie, Pete and Paul. Last but of course not least: Joanne. She has kept the lab on track and done all of the jobs that don't seem important or glamorous, but they keep the lab running smoothly. And of course it wouldn't be the same if we weren't all worried about getting the classic "Joanne finger-wag" for doing something improperly.

Besides my lab, there are others in the chemistry department (and beyond) whom I need to thank as well. David Bunck helped me with my synthesis project and it would not have been successful without him. The same goes for the Lancaster, Lewis, and Dichtel labs, for allowing me to use their equipment for that project. My collaborators Ashley Pratt, formerly at Scripps and Deniz Top at Rockefeller have also done some really nice work to compliment mine, and I thank them for that.

Finally I need to thank my friends. Kait, you hold a special place in my heart, and I have cherished all of the good times we have had together and am looking forward to even more adventures. Thank you for everything. Brian has constantly been there with me, from living together to watching sports and bad local

commercials, to our shenanigans at the bars, to stopping down and chatting in your office to "sir….", it's been a really fun time. I'm going to miss having you around. Michael has always been the quiet compliment to my non-stop talking, and my incessant Liverpool updates. I'm not sure how he's put up with it for 5 years, but thanks for being a good sport. I'd like to thank Jay for letting me live in his living room for the past few months, and also for all the good times before that. Like Michael, Dan has always been a quiet guy, but a really good friend. I also need to thanks several groups of friends. Team Charts, our trivia team at The Nines for the past 5 years, has been awesome. Dan, Jay, Kait, Corinne, and Christian have made my Tuesday nights something to look forward to. Also, I will always remember the Friday and Saturday nights in the winter at Lynah Rink, watching Cornell hockey games. Katie, Joerg, Kait, Seth, Michelle, Dan, and most especially Spencer have made those games much more enjoyable.

There are many, many more people with whom I have crossed paths during my 5-plus years at Cornell, who have improved my graduate career and experience in large and small ways, even if our interactions were brief. To those people, and anyone I may have forgotten to mention, my sincerest thanks.

TABLE OF CONTENTS

Appendix 1: Towards *in vivo* Pulsed Dipolar ESR Spectroscopic Measurements

Appendix 2: Towards Crystallization of the First 700 Residues of the *Drosophila melanogaster* Circadian Clock Protein Period

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3:

Chapter 4:

Appendix 2:

LIST OF TABLES

Chapter 2:

Chapter 4:

LIST OF ABBREVIATIONS

ALS – Amyotrophic Lateral Sclerosis

CFP – Cyan Fluorescent Protein

CK2 – Casein Kinase 2

CLC – Cycle

CLK – Clock

CRY – Cryptochrome

CW – Continuous Wave

DBT – DoubleTime

DEER – Double Electron Electron Resonance

DNA – Deoxyribonucleic acid

DQC – Double Quantum Coherence

DTT - Dithiothreitol

ENDOR – Electron Nuclear Double Resonance

EPR – Electron Paramagnetic Resonance

ESEEM – Electron Spin Echo Envelope Modulation

ESR –Electron Spin Resonance

ET – Electron Transfer

FALS – Familial Amyotrophic Lateral Sclerosis

FRET – Forster or Fluorescent Resonance Energy Transfer

FT – Fourier Transform

FWHM – Full Width at Half Maximum

GFP – Green Fluorescent Protein

GST – Glutathione S-Transferase

IMAC – Immobilized Metal ion Affinity Chromatography

JET – Jetlag

LB – Luria-Bertani

LOV – Light-Oxygen-Voltage

MALS – Multi Angle Light Scattering

MPD – 2-methyl-2,4-pentane-diol

MTSSL – 1-oxyl-2,2,5,5-tetramethylpyrroline-3-(methyl)-methanethiosulfate

NMR – Nuclear Magnetic Resonance

NO – Nitroxide

PAS – Period-Aryl hydrocarbon receptor nuclear translocator-Single minded

PER – Period

PDS – Pulsed Dipolar ESR Spectroscopy

PELDOR – Pulsed Electron Double Resonance

PYTYR – 3-pyrazolyltyrosine

RNA – Ribonucleic Acid

SAXS – Small Angle X-ray Scattering

SEC – Size-Exclusion Chromatography

SNR – Signal-to-Noise Ratio

SOD1 – Cu/Zn Superoxide Dismutase

TIM – Timeless

WT – Wild Type

Chapter 1

OVERVIEW OF PULSED DIPOLAR ESR SPECTROSCOPY AND OTHER PROTEIN CHARACTERIZATION METHODS

1.1 Introduction

It is estimated that there are 10^5 to 10^6 unique proteins in humans, and 10^7 to 10^8 unique proteins across all life forms [1, 2]. Each protein plays an important role in their respective organism. These roles are incredibly diverse, from transporting oxygen in humans to modulating the circadian rhythms of fruit flies to controlling the motion of single celled bacteria [3-5]. If proteins are unable to carry out their prescribed functions, the results can be problematic, or even fatal, for an organism. Issues may also arise from proteins that perform other tasks for which they are not designed, known as a gain of function. In humans, Alzheimer's, Parkinson's, and even inherited cataracts are just 3 of the many diseases caused by proteins which either do not function, or function incorrectly [6]. Gaining insight into the mechanisms by which proteins carry out their given tasks is essential in attempting to prevent and treat these diseases.

Direct human health benefits are not the only motivation for understanding protein function. The study of protein systems in model organisms has provided tremendous insight into complex cellular processes such as signal transduction, gene regulation and molecular trafficking [4, 5]. Together, these can lead to the

understanding of higher order functions such as an organism's behavior. Furthermore, certain proteins can perform chemical transformations which are very difficult via synthetic methods. The oxidation of methane to methanol and the fixation of dinitrogen to ammonia are two important transformations that nature carries out much more efficiently than does the industrial chemistries of man [7-9]. Elucidating how proteins and enzymes catalyze these reactions could lead to much more efficient catalysts, greatly reducing the large energy expenditures currently required.

In biochemistry, the link between structure and function has been well established, and so in order to understand the mechanisms and functions of proteins, we must understand their structures. Protein structure is broken down into 4 classes, each representing an increasing level of complexity. Primary structure is the sequence of amino acids which constitute a given protein. Secondary structure is the way in which these amino acids fold locally, into motifs such as beta sheets or alpha helices. Tertiary structure is the 3-dimensional folding of the entire protein, and quaternary structure is the way that fully-folded protein molecules interact with one another. Determination of primary and secondary structure is a fairly well solved problem. For primary structure, proteomics and genomics have become reliable methods for determining amino acid sequences and for secondary structure, prediction of local 3-dimensional folding based upon primary structure is possible through computation

and comparison to already solved protein structures. In fact, for secondary structure prediction there are over 50 such programs, which compute everything from alpha-helical sections to homology modeling to *ab initio* folding [10]. However determining the 3-dimensional structure of an entire folded protein based upon the sequence is not currently feasible, despite recent advances in theory and computation. Here, experimental methods are needed to solve this important problem.

1.2 Methods of Determining 3-Dimensional Protein Structures

1.2.1 X-Ray Crystallography

X-ray crystallography is by far the most common method of determining protein structures. In the Protein Data Bank, almost 90% of structures have been solved using this method [11]. As the name implies, crystallography involves growing crystals of the purified protein or protein complex of interest and exposing them to x-rays of wavelength commensurate with the atomic spacings in the molecule. The x-rays scatter off of electrons in the protein and surrounding solvent onto a detector to produce a diffraction pattern. Through application of a Fourier Synthesis, the diffraction patterns can be converted into an electron density map. From these maps, the structure of the protein can be elucidated (Fig. 1.1).

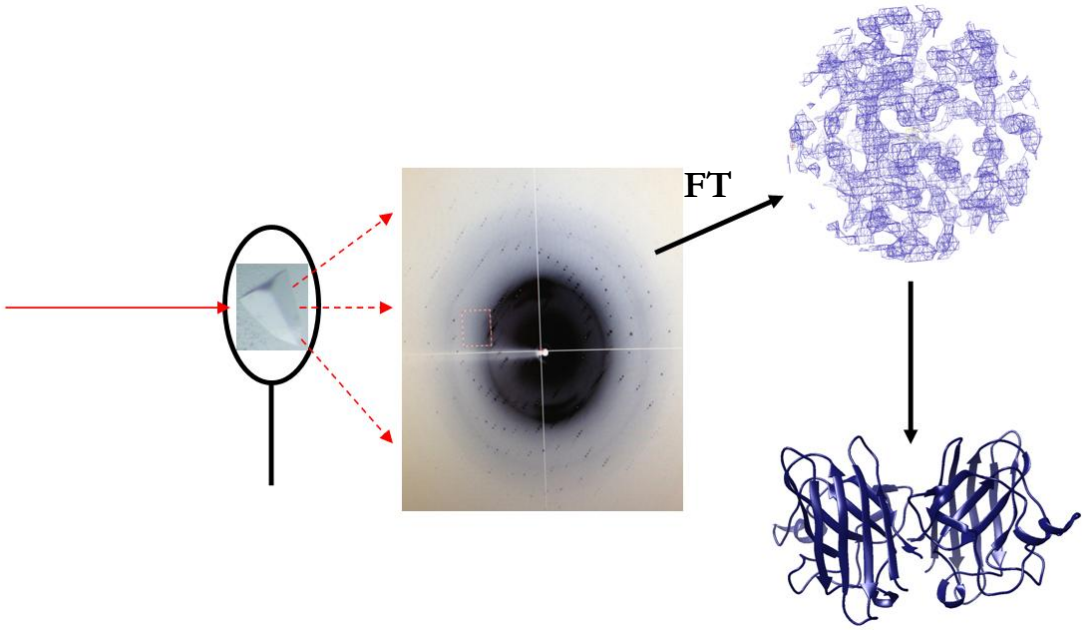

Figure 1.1: Schematic of an x-ray crystallography experimental set-up. A crystal is placed in front of an x-ray source (red), and the x-rays are diffracted onto the detector. The diffraction pattern can then be Fourier Transformed to generate an electron density map. The protein is then overlaid onto the electron density map to give a 3-dimensional structure.

The amount of information gained from x-ray crystallography depends on the diffraction resolution of the crystals. For low or moderate resolution data, the general shape of the protein backbone and the position of side chains can often be identified. However, in cases of very high resolution (< 1.2 Å), each individual atom in the protein can be discerned [12].

Despite the high resolution information potentially available, crystallography does have some limitations. As indicated previously, protein crystals are required for this

technique, and these are not always trivial to grow. Even if crystals are grown, the resolution must be high enough for useful information to be extracted. Other issues may arise during data processing, such as twinning and difficulty in phasing, making the conversion from a diffraction pattern into a structure challenging [12]. Furthermore, there is a certain loss of information involved in crystallizing a protein, as it is not in the natural solution state. In order for a protein to crystallize, it generally adopts a single, uniform conformation. This conformation may not be the sole or even active form of the protein, which may need to adopt multiple conformations in order to be functional. As crystallography generally produces a time-averaged structure, it can be difficult to characterize intermediates or transient states relevant to function. Furthermore, complexes representing functionally critical but weak protein-protein interactions are very difficult to crystallize. Dynamics are not easily inferred from crystal structures. To overcome these limitations, there are several methods of structurally characterizing proteins in solution, which can be used in conjunction with crystallography.

1.2.2 NMR Spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy offers a method of full structural characterization of proteins in solution. This is done from the scalar coupling, or J-coupling, between atoms [13]. This J-coupling arises from the nuclei in adjacent atoms changing the electronic environment of their neighbors through bonds

between these neighboring atoms [13]. Each amino acid side chain has a unique coupling spectrum and by looking at the coupling between side chains, as well as distance restraints between protons, a model for the structure of the protein is generated [13]. Multiple models are often generated to examine their reproducibility, as well as the number of conformations sampled by the protein.

Because protein NMR is a solution state technique, it can report on properties such as dynamics that crystallography cannot. Further, crystals are not needed, so in theory NMR spectroscopy can be performed on any protein in the solution state. Of course, this technique does have limitations. Large proteins pose a problem for NMR studies, due to the fact that T_2 relaxation times decrease as protein size increase, and this eliminates many proteins as candidates for this technique. Often isotopic labeling with NMR active nuclei (^{13}C and ^{15}N) is required to achieve necessary signal to noise ratios [13]. This can be expensive and also reduce protein yields significantly. Many metalloproteins are also not suitable for protein NMR, as their paramagnetic centers broaden linewidths and make NMR spectra very difficult to interpret. In some cases, however, the paramagnetic ions can be used in conjunction with NMR spectroscopy to decipher structural information of the protein of interest [14].

1.2.3 Small Angle X-Ray Scattering

Small angle X-ray scattering, or SAXS, has characteristics of both NMR spectroscopy and X-ray crystallography. Instead of a crystal, a solution sample is

exposed to an x-ray source. The observed result is an isotropic scattering profile where intensity decreases as the scattering angle increases [15] (Fig. 1.2). As opposed to crystallography, the scattering profile from SAXS cannot be used to gain high-resolution information. In SAXS, the high-resolution information is contained at larger angles, where the signal to noise is poor. This, along with orientational averaging of the protein molecules in solution, limits the resolution of SAXS [15]. Despite these limitations, SAXS is a useful tool, as it can measure properties such as the radius of gyration, disorder or unfolding, and gives a low-resolution particle shape, all with relatively simple sample preparation [15]. It is also a useful technique for observing the association/dissociation of protein complexes and stimuli-generated conformational changes [15, 16].

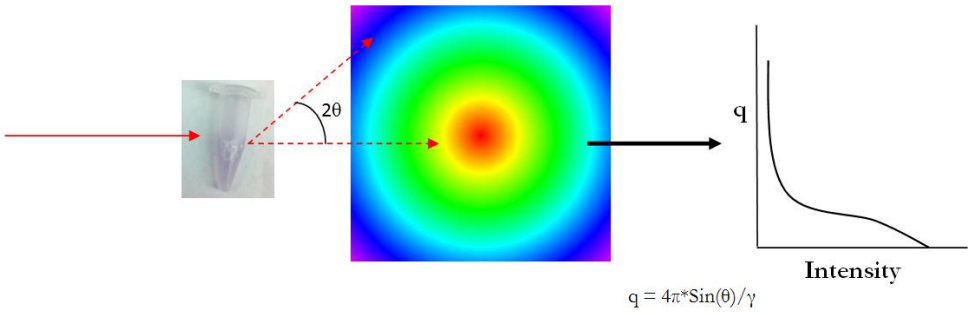

Figure 1.2: Schematic of a SAXS experimental setup. A solution state sample is placed in front of an x-ray source. The scattering profile is collected on a detector, and then the intensity is plotted against the angle of the scattering. From this, low-resolution information is extracted.

1.3 ESR Spectroscopy

1.3.1 Theory

Whereas NMR spectroscopy observes atomic interactions through the effect of an applied magnetic field on nuclei, electron spin resonance spectroscopy (ESR, also known as electron paramagnetic resonance, EPR) observes those interactions though effects on electrons, specifically unpaired electrons [17]. Generally, the unpaired electrons are in the form of organic or protein radicals, or transition metals. ESR spectroscopy has many applications, including elucidating the electronic structure around an unpaired spin and observing the coupling between unpaired spins and neighboring nuclei [17]. Pulsed ESR measurements can also be used to measure distances between unpaired spins [18], which is the main focus of this thesis.

The main theoretical consideration for ESR spectroscopy is that in an applied magnetic field, an unpaired electron can either adopt a spin-up or spin-down conformation. The spin-up conformation is lower in energy (i.e. more stable), and thus slightly favored. As the applied magnetic field strength increases, the difference in energy between the 2 states also increases, which is known as the Zeeman Splitting (Fig. 1.3). By delivering a quantized unit of energy equal to the Zeeman Splitting, the spin may be flipped from spin-up to spin-down. For a standard continuous wave (CW) ESR experiment, the energy, in the form of microwave radiation, is held constant and a field sweep is done by increasing the magnetic strength. When the

resonance energy is reached, an absorption peak is recorded. From the position of the absorption peaks, g-values, which represent the electron's local magnetic environment, (equation 1) can be easily calculated. These g-values, along with the lineshape of the spectrum, provide ample information on the electronic structure surrounding the unpaired spin(s) [17].

$$\Delta E = g\beta \mathbf{B} \quad (1)$$

Where ΔE is the Zeeman splitting energy, β is the Bohr magneton, and $\mathbf{B}$ is the applied magnetic field.

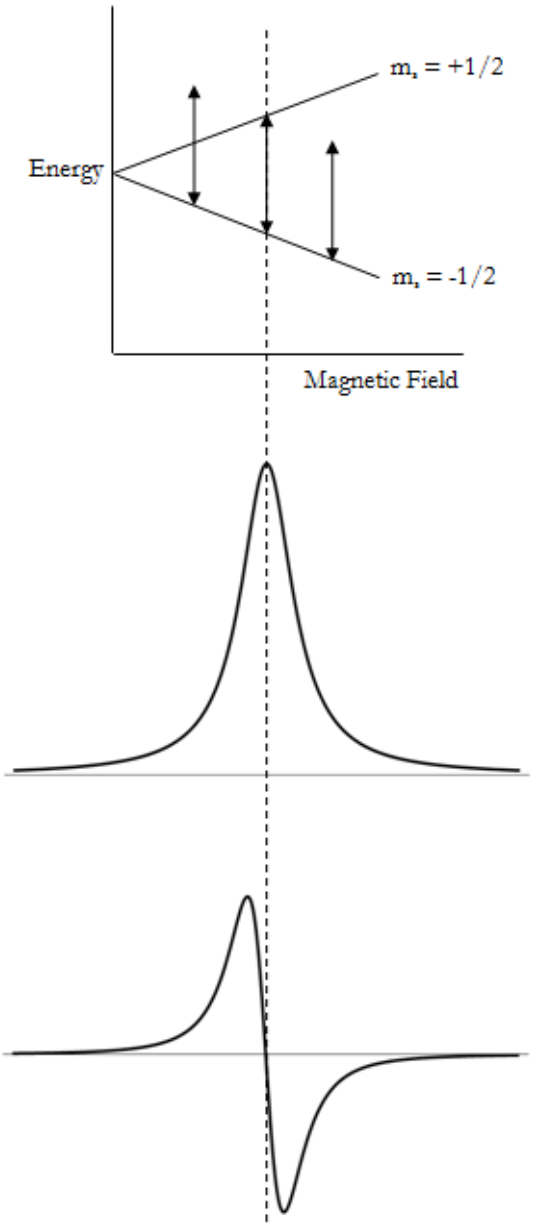

Figure 1.3: (a) Zeeman splitting diagram. The vertical arrows represent the microwave energy input into the system. As the magnetic field is swept, that energy becomes in resonance with the Zeeman splitting, causing a spin flip, which is recorded by the absorption event in (b). Due to electronic circuitry considerations, the absorption spectra are routinely converted to their 1st derivative (c).

There is another class of ESR experiments based on similar principles, pulsed ESR spectroscopy, which measures the interactions between 2 spins [19]. Techniques such as ESEEM (Electron Spin Echo Envelope Modulation) [20] and ENDOR (Electron Nuclear Double Resonance) [21] measure interactions between nuclear spins and unpaired electron spins. Another type of pulsed ESR experiment is pulsed dipolar ESR spectroscopy (PDS) [22] . PDS measures the dipolar coupling between 2 unpaired spins, usually denoted as A and B spins [23, 24] (Fig. 1.4). From the strength of the dipolar coupling, a distance between the 2 spins can be extracted. In general, distances of ~15-90 angstroms can be measured. Measuring distances in this way is a central component of this thesis.

In PDS, unlike CW ESR, the magnetic field is kept constant and instead of a constant application of microwave radiation, microwaves are applied to the sample in very short pulses, on the order of tens of nanoseconds. For this thesis, the focus is on DEER (Double Electron Electron Resonance) also known as PELDOR (Pulsed Electron Double Resonance). In a DEER experiment, the microwave pulses are given at 2 different frequencies, one which flips the A spins and the other which flips the B spins. The A and B spins can, but do not have to have the same chemical origin. Here, a 4-pulse DEER sequence (Fig. 1.4) is used. The 3 A-spins are subject to the detection pulses, whereas the B-spins are subject to the pump pulse. In the cases

where the A and B spins have the same chemical origin, they are distinguished in frequency by slightly shifting the detection and pump frequencies from each other.

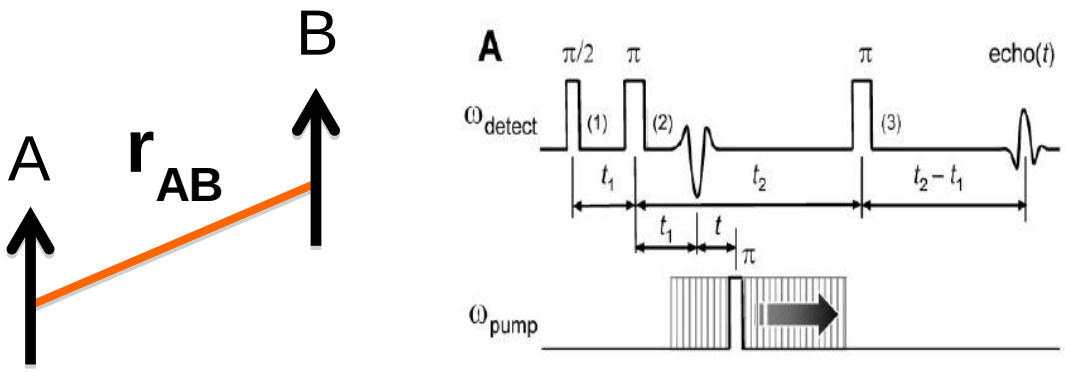

Figure 1.4: The 4-pulse DEER sequence at right is used to measure the distance r_{AB} between unpaired spins A and B. The 3 detection pulses at ω_{detect} flip the A-spins and the pump pulse at ω_{pump} flips the B-spins. The amplitude of the spin echo generated by the A-spins as a function of time T gives the time-domain signal (Fig. 1.5)

The 3 A-pulses cause 2 refocusings of the A spins in the XY-plane; these are known as "spin-echoes" and the second spin-echo is recorded. In the absence of the B pulse, the A spins will refocus at precisely the same time, generating a maximal spin-echo intensity. When the B-spins are flipped by the pump π pulse, it causes a change in the local magnetic field around the A-spins, due to the dipolar coupling between the A and B spins. This causes defocusing of the A-spins in the XY-plane, and in turn, causes the spin-echo to have a lower intensity than in the absence of the pump pulse. As the pump pulse is shifted in time, the effect on the intensity of the spin-echo

of the A spins also changes. The intensity of the spin-echo as a function of time T constitutes the time-domain signal in DEER experiments, which has the form

$$V(T) = V(0)[1-p_b(1-\cos(DT))] \quad (2)$$

Where V(0) is the echo amplitude in the absence of a pump pulse, p_b is the probability of a spin-flip, and

$$D = \omega_D(1-3\cos^2(\theta)) + J \quad (3)$$

Where θ is the angle between the external magnetic field and r_{AB} and J is the exchange interaction constant. In practice, over long distances, J is weak and the approximation J = 0 is valid. Since spin-labels are generally isotropic, θ is approximated as 90° and thus the $3\cos^2(\theta)$ term goes to 0 (see Section 2.5 for a discussion on the orientation of θ). So then $D = \omega_D$, where

$$\omega_D = (\gamma^2\hbar)/r^3_{AB} \quad (4)$$

The overall equation then becomes:

$$V(T) = V(0)[1-p_b(1-\cos((\gamma^2\hbar)/r^3_{AB})T)] \quad (5)$$

r_{AB} is the distance between the A and B spins. Solving for r_{AB} gives the distribution of distances between the 2 spins in the sample (Fig. 1.5).

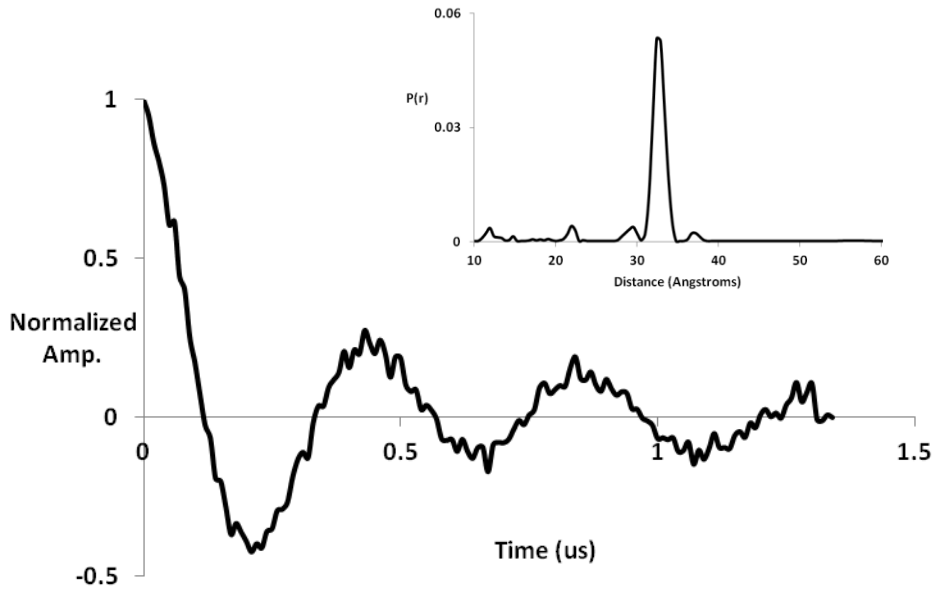

Figure 1.5: Time-domain signal and corresponding distance distribution from a DEER experiment.

A second method of measuring distances via pulsed dipolar ESR spectroscopy is Double Quantum Coherence (DQC). Unlike DEER, all of the (6) microwave pulses are applied at a single frequency. The DQC method can improve upon DEER in some ways, such as greater signal-to-noise and the ability to measure distances shorter than 15 angstroms [25]. However, DQC requires that the A and B spins have significant spectral overlap, as the pulses are applied at a single frequency [23]. Although many of the PDS measurements carried out in this study were between species with small spectral overlap, DQC generally produced similar results as DEER (see Chapter 2.3).

1.3.2 Applications

In practice, DEER is mostly used to measure distances in proteins or other biomolecules, such as nucleic acids. The range of DEER spectroscopy, ~15-90 Å, is well-suited to measuring relatively long distances in proteins. These measurements are done on frozen solutions, with the temperature varied between 15 and 70 K depending on the nature of the sample. Low temperatures are needed to reduce T2 relaxation and allow for sufficiently long spin echoes. Despite the solution being frozen, if freezing is rapid, solution characteristics such as sample heterogeneity and disorder can be trapped in the solid solution and observed through the modality and breadth of the distance distributions. The vast majority of proteins, however, do not contain paramagnetic species, and so they need to be added artificially. Most commonly, a nitroxyl radical (MTSSL [1-oxyl-2,2,5,5-tetramethylpyrroline-3-(methyl)-methanethiosulfate]) is covalently attached to an exposed cysteine residue [26] (Fig. 1.6). The target residues may be natural cysteines, or ones mutated into the protein at an opportune position. A nitroxyl radical side chain covalently attached to a cysteine residue is commonly referred to as "R1".

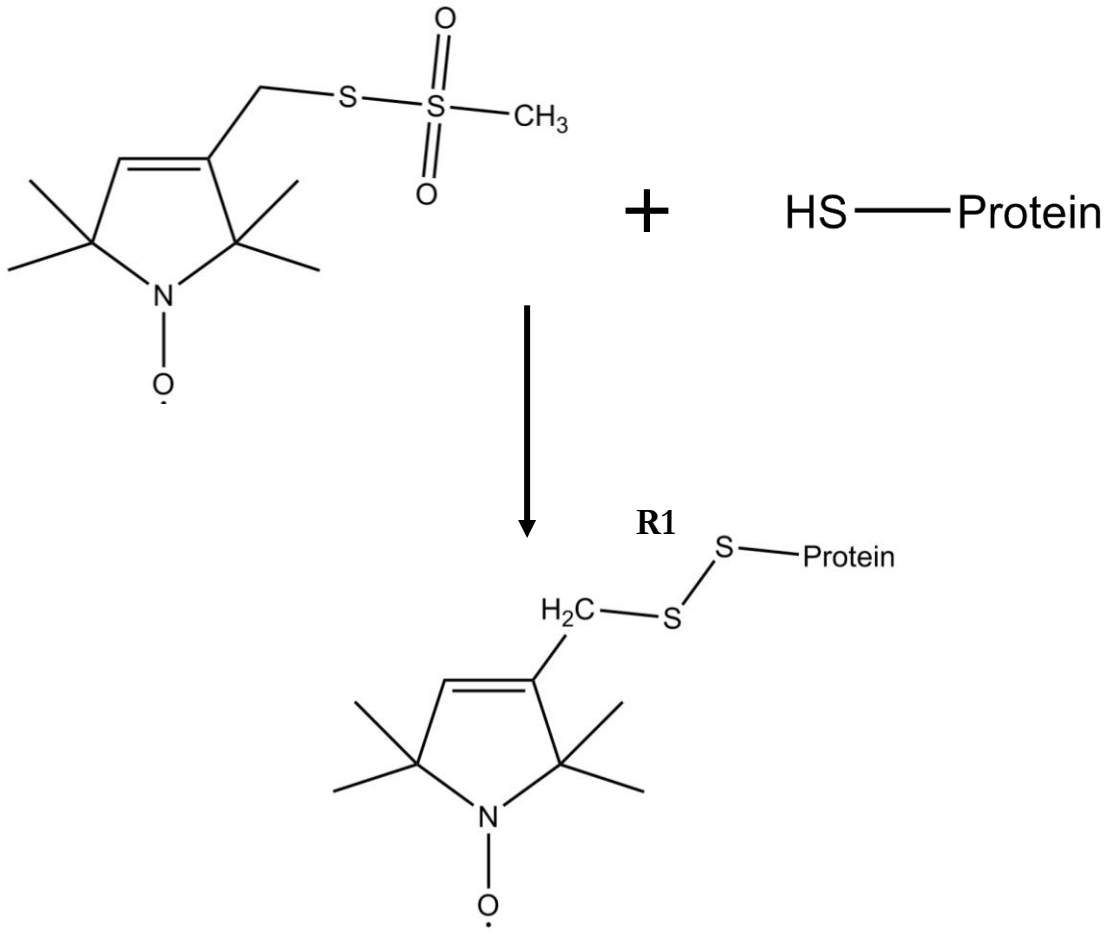

Figure 1.6: Schematic of an MTSSL covalently attached to the –SH group of an exposed cysteine residue. This sidechain is commonly referred to as "R1".

Despite the many successful applications of spin-labeling used in concert with PDS [23, 27-29] (including several in this thesis), nitroxyl-based labels do have several limitations. The most obvious, from a structural standpoint, is the issue of multiple cysteines. In a monomeric sample, two cysteine residues are needed to provide two unpaired spins in order to measure a distance. If there are less than 2 cysteines, this

problem is easily corrected through site-directed mutagenesis. However, in many proteins, especially large eukaryotic systems, there are many cysteines in the sample, and these cysteines cannot always be replaced while maintaining the structural integrity (and activity) of the protein. Moreover, nitroxide labels are spectroscopically indistinguishable from each other in an ESR experiment, meaning that distance distributions resulting from more than 2 labels per sample can be challenging to unambiguously deconvolute. Finally, the reducing nature of a cell makes *in vivo* work with nitroxides very difficult, although some advancements have recently been made [10, 30]. In order to circumvent these and other limitations, metal ions have increasingly been used as paramagnetic labels for PDS measurements.

Metal ions as paramagnetic centers are advantageous in that they are distinguishable from nitroxides (and from other metals) through their different g-values. Metal ions used in PDS include Fe (III) [31, 32], Mo (V) [31], Gd (III) [33-36], and, most commonly, Cu (II) [37-43]. Cu^{2+} has a d^9 electron configuration, meaning it only has 1 unpaired electron, as is the case with nitroxide label. However, the magnetic spectrum of Cu^{2+} is approximately an order of magnitude broader, leading to much less spectral coverage by pulse and pump sequences as compared to nitroxides. Furthermore, Cu^{2+} has a much faster intrinsic relaxation time, which also contributes to the challenge of using Cu^{2+} in PDS experiments. Despite this, there have been many examples of Cu^{2+} used as a spin probe, including double labeling

experiments where Cu-Cu, Cu-NO and NO-NO distances have all been measured from one sample [42].

Copper has several advantages over other metals in its use as a spin-label for PDS. The only paramagnetic state of copper, Cu^{2+}, has a single unpaired electron. This is in contrast to other biologically relevant metals, such as iron or manganese, which can have multiple unpaired electrons at the metal center, making PDS measurements complicated. Spin states also do not matter, as there is no difference between "low-spin" and "high-spin" Cu^{2+}, as there is with other metals. For example, low-spin Fe^{3+} only has 1 unpaired electron, whereas high-spin Fe^{3+} has 5 unpaired electrons, meaning that only the low-spin species can be measured by PDS, and any contamination from high-spin Fe^{3+} will create difficult complications. Furthermore, Cu^{1+} has a d^{10} electron configuration, making it diamagnetic and ESR silent. Therefore, if there is a mixture of oxidation states in the sample, only those in the Cu^{2+} state (the desired spins) will be measured.

Despite this, there have been several instances of other metals being used for PDS measurements. The distance between Mo^{5+} and Fe^{3+} in sulfite oxidase has been measured [31], as have numerous distances involving gadolinium [33-36]. Gd^{3+} has recently become more popular as a spin-label for DEER measurements. Spin properties of Gd^{3+} allow for high-field (W-band) measurements, meaning greater sensitivity [44]. Initially, this would seem counterintuitive, as Gd^{3+} is high-spin with an

f^7 electron configuration, meaning that there are 7 potential transitions (±7/2 to ±5/2, ±5/2 to ±3/2, ±3/2 to ±1/2, and -1/2 to +1/2). This is in contrast to the cases of Cu^{2+}, or nitroxides, s = 1/2 systems, where there is only the -1/2 to 1/2 transition to consider. However, in the Gd^{3+} spectrum, the -1/2 to 1/2 transition is very narrow and very large compared to all other transitions, which makes specific selection feasible, especially at high field, where it narrows even further [11]. Because the transition is so narrow, and the background is relatively featureless, orientational selectivity is not a concern, as it would be for nitroxides at high field [45]. The higher field strength also leads to a greater transition probability for the -1/2 to 1/2 transition, which leads to a better signal to noise ratio (SNR) and the ability to probe shorter distances as compared to nitroxides [45]. Furthermore, since there is no naturally occurring gadolinium in cells, the background noise that occurs in other attempts at *in vivo* PDS measurements should be negligible. A PDS measurement has been made on Gd^{3+} labels in cells at high-frequency, but those measurements were not strictly *in vivo*, and required covalent attachment of a Gd^{3+} label *in vitro* [36]. See Appendix 1 for efforts to make *in vivo* PDS measurements.

1.3.3 Comparison with FRET

FRET (Forster or Fluorescent Resonance Energy Transfer) is another method of measuring the distance between 2 spectral probes within a macromolecule. In this case, as the name implies, fluorescent probes are utilized, with one being a donor

fluorophore and the other an acceptor fluorophore. When the donor probe is stimulated, it transfers energy to the acceptor probe. The FRET efficiency, or the quantum yield of the excitation, is based on the distance separating the two probes. By measuring the relative quantum yield of fluorescence from both probes, usually through fluorescence microscopy, a distance can be calculated [46]. Because both techniques are able to measure distances between probes, there is a natural comparison between PDS and FRET. However, just as the techniques have much different theoretical backgrounds, their applications are much different as well.

The applications of PDS have been discussed above, but to summarize: PDS very accurately measures distance distributions on frozen solution state samples at very low temperatures. Thus, a PDS measurement is essentially a snapshot of all of the distances between spin probes in a sample. The precise nature of conformational changes can be elucidated by taking multiple snapshots and observing small changes in distance distributions. In FRET, the exact distance between probes cannot be measured as accurately, as there are some simplifying approximations made in the calculations [47]. Moreover, the spin probes for FRET are generally much larger than the nitroxide labels or metal centers found in PDS [46]. Thus, the reporter is generally far away from the area of interest and insensitive to very minor conformational changes which can be detected by PDS.

While FRET cannot measure distance distributions with the sensitivity of PDS, it can look at real-time changes, which PDS cannot do. Because FRET can be done at ambient temperatures, the interaction of the 2 probes can be monitored in the time-domain [48]. For example, the association of 2 subunits can be monitored in real-time, or the frequency of some association or dissociation event can be measured on long time scales [49]. FRET allows kinetics data to be generated in a way that is not possible with PDS [50]. Furthermore, because of the ambient experimental conditions, it is possible to collect FRET data *in vivo* [51], which, due to the challenges outlined above, has not currently been achieved with PDS. However, while FRET is an excellent tool for observing dynamic processes, PDS has much better sensitivity for elucidating detailed structural features.

REFERENCES

1. Levitt, M. 2009. Nature of the protein universe, *P Natl Acad Sci USA*. 106: 11079-11084.

2. Available from: http://proteomics.cancer.gov/whatisproteomics.

3. Hill, R.J., Konigsberg, W., Guidotti, G., and Craig, L.C. 1962. Structure of Human Hemoglobin .1. Separation of Alpha and Beta Chains and Their Amino Acid Composition. *J Biol Chem*. 237: 1549-1554.

4. Crane, B.R. and Young, M.W. 2014. Interactive Features of Proteins Composing Eukaryotic Circadian Clocks. *Ann Rev Biochem*. 83: 191-219.

5. Wadhams, G.H. and Armitage, J.P. 2004. Making sense of it all: Bacterial chemotaxis. *Nat Rev Molec Cell Bio*. 5: 1024-1037.

6. Reynaud, E. 2010. Protein Misfolding and Degenerative Diseases. *Nat Edu*. 3: 28-29.

7. Merkx, M., Kopp, D.A., Sazinsky, M.H., Blazyk, J.L., Muller, J., and Lippard, S.J. 2001. Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: A tale of two irons and three proteins. *Angew Chem Int Edit*. 40: 2782-2807.

8. Seefeldt, L.C., Hoffman, B.M., and Dean, D.R. 2009. Mechanism of Mo-Dependent Nitrogenase. *Ann Rev Biochem*. 78: 701-722.

9. Lancaster, K.M., Roemelt, M., Ettenhuber, P., Hu, Y., Ribbe, M.W., Neese, F., Bergmann, U., and DeBeer, S. 2011. X-ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor. *Science*. 334: 974-977.

10. Dunkel, S., Pulagam, L.P., Steinhoff, H.J., and Klare, J.P. 2015. In vivo EPR on spin labeled colicin A reveals an oligomeric assembly of the pore-forming domain in E. coli membranes. *Phys Chem Chem Phys*. 17: 4875-4878.

11. Raitsimring, A.M., Atashkin, A.V., Poluektov, O.G., and Caravan, P. 2005. High-field pulsed EPR and ENDOR of Gd3+ complexes in glassy solutions. *App Mag Res*. 28: 281-295.

12. Rupp, B. 2010. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. New York: Garland Science.

13. Cavanagh, F.J., Palmer, W.J., Rance, A.G. III, Skelton, M., and Nicholas, J. 2007. *Protein NMR Spectroscopy: Principles and Practice*. 2 ed. Burlington, MA: Elsevier.

14. Otting, G. 2010. Protein NMR Using Paramagnetic Ions. *Ann Rev Biophys*. 39: 387-405.

15. Mertens, H.D.T. and Svergun, D.I. 2010. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J Struc Biol*. 172: 128-141.

16. Vaidya, A.T., Top, D., Manahan, C.C., Tokuda, J.M., Zhang, S., Pollack, L., Young, M.W., and Crane, B.R. 2013. Flavin reduction activates Drosophila cryptochrome. *P Natl Acad Sci USA*. 110: 20455-20460.

17. *Physical Methods in Bioinorganic Chemistry: Spectroscopy and Magnetism*. 2010, Sausalito, California: University Science Books.

18. Tsvetkov, Y.D., Milov, A.D., and Maryasov, A.G. 2008. Pulse electron-electron double resonance (PELDOR) as nanometre range EPR spectroscopy. *Uspekhi Khimii*. 77: 515-550.

19. Prisner, T., Rohrer, M., and MacMillan, F. 2001. *P*ulsed EPR spectroscopy: Biological applications. *Ann Rev Phys Chem*. 52: 279-313.

20. Dikanov, S.A. and Tsvetkov, Y.D. 1992. Electron Spin Echo Envelope Modulation (ESEEM) Spectroscopy. Boca Raton, Florida: CRC Press.

21. Kevan, L. and Kispert, L.D. 1976. Electron Spin Double Resonance Spectroscopy. Wiley.

22. Borbat, P.P. and Freed, J.H. 2007. Measuring distances by pulsed dipolar ESR spectroscopy: spin-labeled histidine kinases. *Methods Enzymol*. 423:52-116.

23. Borbat, P.P. and Freed, J.H. 2007. Measuring distances by pulsed dipolar ESR spectroscopy: Spin-labeled histidine kinases. *Two-Component Signaling Systems, Pt B*. 423: 52-116.

24. Jeschke, G. 2012. DEER Distance Measurements on Proteins. *Ann Rev Phys Chem*. 63: 419-446.

25. Borbat, P.P., Mchaourab, H.S., and Freed, J.H. 2002. Protein structure determination using long-distance constraints from double-quantum coherence ESR: Study of T4 lysozyme. *J Am Chem Soc*. 124: 5304-5314.

26. Hubbell, W.L., Gross, A., Langen, R., and Lietzow, M.A. 1998. Recent advances in site-directed spin labeling of proteins. *Curr Op Struc Biol*. 8:649-656.

27. Bhatnagar, J., Borbat, P.P., Pollard, A.M., Bilwes, A.M., Freed, J.H., and Crane, B.R. 2010. Structure of the Ternary Complex Formed by a Chemotaxis Receptor Signaling Domain, the CheA Histidine Kinase, and the Coupling Protein CheW As Determined by Pulsed Dipolar ESR Spectroscopy. *Biochemistry*. 49: 3824-3841.

28. Sircar, R., Borbat, P.P., Lynch, M.J., Bhatnagar, J. Beyersdorf, M.S., Halkides, C.J., Freed, J.H., and Crane, B.R. 2015. Assembly States of FliM and FliG within the Flagellar Switch Complex. *J Mol Bio*. 427: 867-886.

29. Samanta, D., Borbat, P.P., Dzikovski, B., Freed, J.H., and Crane, B.R. 2015. Bacterial chemoreceptor dynamics correlate with activity state and are coupled over long distances. *P Natl Acad Sci USA*. 112: 2455-2460.

30. Schmidt, M.J., Borbas, J., Drescher, M., and Summerer, D. 2014. A Genetically Encoded Spin Label for Electron Paramagnetic Resonance Distance Measurements. *J Am Chem Soc*. 136: 1238-1241.

31. Astashkin, A.V., Rajapakshi, A., Cornelison, M.J., Johnson-Winters, K., and Enemark, J.H. 2012. Determination of the Distance between the Mo(V) and Fe(III) Heme Centers of Wild Type Human Sulfite Oxidase by Pulsed EPR Spectroscopy. *J Phys Chem B*. 116: 1942-1950.

32. Ezhevskaya, M., Bordignon, E., Polyhach, Y., Moens, L., Dewilde, S., Jeschke, G., and Van Doorslaer, S. 2013. Distance determination between low-spin ferric haem and nitroxide spin label using DEER: the neuroglobin case. *Mol Phys*. 111: 2855-2864.

33. Yagi, H., Banerjee, D., Graham, B., Hubert, 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.

34. Song, Y., Meade, T.J., Astashkin, A.V., Klein, E.L., Enemark, J.H., and Raitsimring, A., 2011. Pulsed dipolar spectroscopy distance measurements in biomacromolecules labeled with Gd(III) markers. *J Mag Res.* 210: 59-68.

35. Garbuio, L., Bordignon, E., Brooks, E.K., Hubbell, W.L., Jeschke, G., and Yulikov, M. 2013. Orthogonal Spin Labeling and Gd(III)-Nitroxide Distance Measurements on Bacteriophage T4-Lysozyme. *J Phys Chem B.* 117: 3145-3153.

36. 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.

37. Jun, S., Becker, J.S., Yonkunas, M., Coalson, R., and Saxena, S. 2006. Unfolding of alanine-based peptides using electron spin resonance distance measurements. *Biochemistry.* 45: 11666-11673.

38. Ruthstein, S., Ji, M., Mehta, P., Jen-Jacobson, L., and Saxena, S. 2013. Sensitive Cu2+-Cu2+ Distance Measurements in a Protein-DNA Complex by Double-Quantum Coherence ESR. *J Phys Chem B.* 117: 6227-6230.

39. Yang, Z.Y., 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.

40. 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 Edit.* 52: 1990-1993.

41. Ji, M., Tan, L., Jen-Jacobson, L., and Saxena, S. 2014. Insights into copper coordination in the EcoRI-DNA complex by ESR spectroscopy. *Mol Phys.* 112: 3173-3182.

42. 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.

43. Pratt, A.J., Shin, D.S., Merz, G.E., Rambo, R.P., Lancaster, W.A., Dyer, K.N., Borbat, P.P., Poole II, F.L., Adams, M.W.W., Freed, J.H., Crane, B.R., Tainer, J.A., and Getzoff, E.D. 2014. Aggregation propensities of superoxide

dismutase G93 hotspot mutants mirror ALS clinical phenotypes. *P Natl Acad Sci USA*. 111: E4568-E4576.

44. Gordon-Grossman, M., Kaminker, I., Gofman, Y., Shai, Y., and Goldfarb, D. 2011. W-Band pulse EPR distance measurements in peptides using Gd3+-dipicolinic acid derivatives as spin labels. *Phys Chem Chem Phys*. 13: 10771-10780.

45. Potapov, A., Song, Y., Meade, T.J., Goldfarb, D., Astashkin, A.V., and Raitsimring, A. 2010. Distance measurements in model bis-Gd(III) complexes with flexible "bridge". Emulation of biological molecules having flexible structure with Gd(III) labels attached. *J Mag Res*. 205: 38-49.

46. Piston, D.W. and Kremers, G.J. 2007. Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem Sci*. 32: 407-414.

47. Vekshin, N.L. 1997. Energy Transfer in Macromolecules. Bellingham, Washington: *SPIE*.

48. Padilla-Parra, S., Auduge, M., Coppey-Moisan, M. and Tramier, M. 2008. Quantitative FRET analysis by fast acquisition time domain FLIM at high spatial resolution in living cells. *Biophys J*. 95: 2976-2988.

49. Bova, M.P., Ding, L.L., Horwitz, J., and Fung, B.K. 1997. Subunit exchange of alpha A-crystallin. *J Biol Chem*. 272: 29511-29517.

50. Xing, J., Jayasundar, J.J., Ouyang, Y., and Dong, W.J. 2009. Forster Resonance Energy Transfer Structural Kinetic Studies of Cardiac Thin Filament Deactivation. *J Biol Chem*. 284: 16432-16441.

51. Jares-Erijman, E.A. and Jovin, T.M. 2006. Imaging molecular interactions in living cells by FRET microscopy. *Curr Opin Chem Biol*. 10: 409-416.

Chapter 2

COPPER BASED PULSED DIPOLAR ESR SPECTROSCOPY AS A PROBE OF PROTEIN CONFORMATION LINKED TO DISEASE STATES*

2.1 Abstract

We demonstrate the ability of Pulsed Dipolar Electron Spin Resonance (ESR) Spectroscopy (PDS) to report on the conformation of Cu-Zn superoxide dismutase (SOD1) through the sensitive measurement of dipolar interactions between inherent Cu^{2+} ions. Although the anisotropy of the Cu ESR spectrum provides challenges for PDS, Ku-band (17.3 GHz) double electron-electron resonance (DEER) and double-quantum coherence (DQC) variants of PDS coupled with distance reconstruction methods recover Cu-Cu distances in good agreement with crystal structures. Moreover, Cu-PDS measurements expose distinct differences between the conformational properties of wild-type (WT) SOD1 and a single-residue variant (I149T) that leads to the disease amyotrophic lateral sclerosis (ALS). The I149T protein displays a broader Cu-Cu distance distribution within the SOD1 dimer compared to WT. Distance distributions obtained from Cu-Cu, Cu-NO and NO-NO separations in an NO-labeled sample reveal increased structural heterogeneity within

*The material in this chapter is reprinted with permission from Merz, G.E., Borbat, P.P., Pratt, A.J., Getzoff, E.D., Freed, J.H., and Crane, B.R. Copper-Based Pulsed Dipolar ESR Spectroscopy as a Probe of Protein Conformation Linked to Disease States. Biophysical Journal 107. (2014). Copyright (2014).

the protein and a tendency for mutant dimers to associate. In contrast, perturbations caused by the ALS mutation are completely masked in the crystal structure of I149T. Thus, PDS readily detects alterations in metalloenzyme solution properties not easily deciphered by other methods and in doing so supports the notion that increased dynamics and associations of SOD1 ALS variants contribute to disease progression.

2.2 Introduction

Only a few biophysical methods provide high-resolution long-range distance restraints for proteins in solution [1-4]. One such technique is pulsed dipolar ESR spectroscopy (PDS) [5-7], which when coupled with nitroxide radical spin labeling [8], can measure distances to within one Å accuracy and yield distance distributions within and between proteins that range up to 90 Å. However, despite the many advantages of PDS, nitroxide labels have certain limitations: their flexibility reduces effective distance resolution due to uncertainty in the positions of backbone atoms; their placement may perturb structure and/or function; multiple labels cannot be distinguished from each other (based for example on a difference in g-values); high labeling efficiency is necessary; and use is limited to environments where labels are stable; this usually necessitates biochemical reconstitution. Cu^{2+} ions offer an alternative spin-probe [9] that for Cu-metalloproteins requires no perturbation of the sample [10]. Moreover, when Cu^{2+} and nitroxide labels are combined in a single

sample, triangulation of distance restraints [11, 12] can provide a more global view of structure and dynamics [13]. Here we apply Cu-PDS to the enzyme Cu-Zn superoxide dismutase (SOD1, Figure 2.3) and demonstrate its ability to reveal aberrant conformational properties of SOD1 mutants linked to the disease amyotrophic lateral sclerosis (ALS). These differences are not well discriminated by other methods and their consequences speak to the mechanism of ALS.

Whereas mainstream use of PDS is based on nitroxide spin-labels, recent work has demonstrated the ability of PDS to provide distance distributions for separations between paramagnetic metal centers in peptides and proteins [13-18]. PDS is mainly represented by double electron-electron resonance (DEER, aka PELDOR) and double-quantum coherence (DQC), with DEER currently implemented over a wide range of microwave frequencies (from cm to mm waves). Saxena et al. have developed techniques to measure Cu-Cu and Cu-nitroxide (NO) distances with DEER [17] and variants [13, 15-17] of the generic DQC method [7, 19, 20] in both synthetic peptides [13, 15, 17], and proteins [17]. In other studies, DEER was also used to determine distances between the Type 1 and 2 Cu centers of nitrite reductase [10]. Similarly, DEER has been applied to Fe^{3+} and Mo^{5+} systems [14], and Gd^{3+} labels are increasingly employed in DEER measurements conducted at mm-wave frequencies [21-23]. Here, we apply both DEER and DQC PDS experiments (Section 2.5, Methods) to attribute such differences to increased structural heterogeneity in the

protein subunits themselves. DEER and DQC provide similar distance data when implemented in their most common 4-pulse and 6-pulse sequences, respectively; although other variants have been developed [24-27]. DEER provides an advantage for spin-probes with widely-separated ESR spectra and good isolation from nuclear coherence effects, whereas DQC has a wider distance range and often yields greater sensitivity. Our use of intense pulses with proportionally wide spectral excitation of ~40 G or more significantly increases signal-to-noise ratio (SNR) for these Cu^{2+} studies (cf. Section 2.5, Methods)

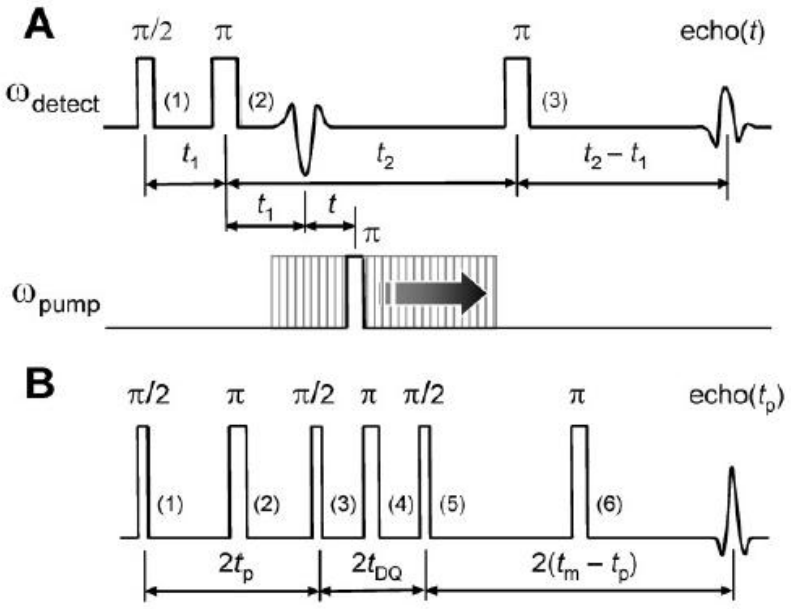

Figure 2.1: (A) The 4-pulse DEER and (B) 6-pulse DQC PDS pulse sequences. Both pulse sequences are well described in the literature. Briefly in the (4-pulse) DEER method, a refocused primary echo is created by the sequence of mw pulses 1, 2, and 3, $\pi/2$-t_1-π-t_2-π-(t_2 - t_1)-echo applied at the frequency ω_{detect}, thus selecting for detection of the electron spins at the desired field position, while the "pump" π-pulse is applied at a sufficiently different frequency, ω_{pump}. Hence it flips the electron spins corresponding to a different part of the ESR spectrum. The time position of the pump pulse is advanced in small steps from the 2nd to the 3rd detection pulse, producing an amplitude modulation of the echo. The modulation pattern thus mainly represents the "dipolar" oscillations from the dipole-dipole interaction between the electron spins. From the oscillation frequency, ω_{dip} is proportional to $1/r^3$, the distance, r, can be accurately determined. The position of the echo does not change in this version of DEER, therefore variable relaxation and nuclear ESEEM contributions to the echo amplitude are relatively small. In the 6-pulse DQC sequence, based on intense mw pulses applied at a single frequency, pulses 2 to 6 move in such a manner that the position of the echo also does not change. Pulses 3, 4, and 5 move as a group, while refocusing π-pulses 2 and 6 are always in the middle of respective intervals "$2t_p$" and "$2t_m - 2t_p$". Dipolar oscillations recorded as a function of $t_\zeta \equiv t_m - 2t_p$ are separated from other contributions to the echo by phase cycling.

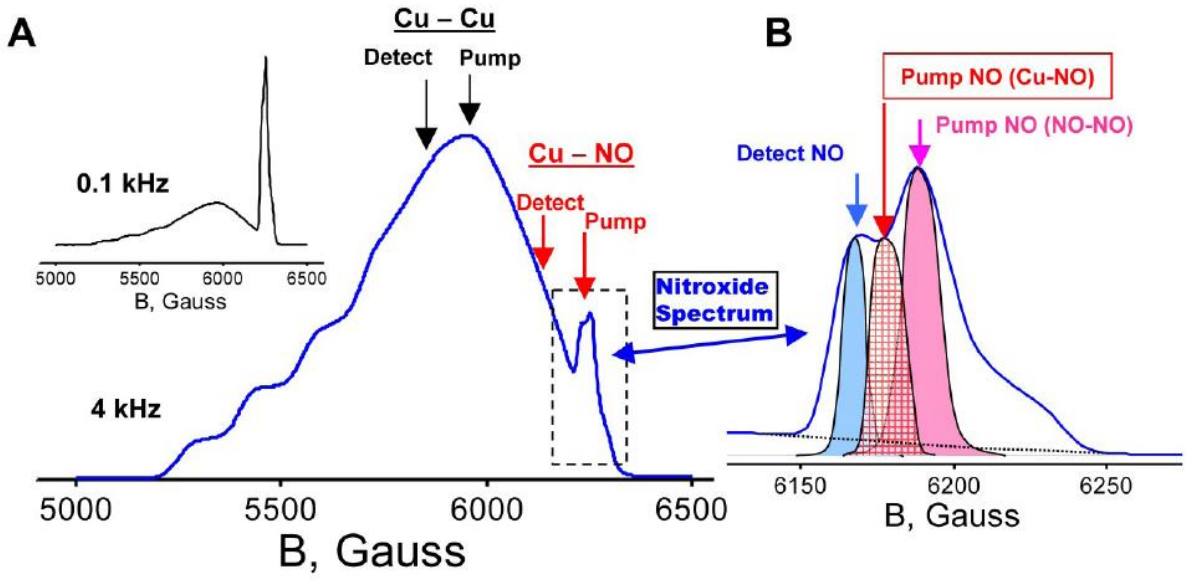

Figure 2.2: (A) Field-sweep echo-detected SOD1 Cu^{2+} ESR spectrum was recorded at 15 K and 4 kHz repetition rate. On the high-field side the spectrum from nitroxide labels is visible. The arrows indicate spectral positions of detection and pumping used in these sets of measurements. The inset on the left was recorded at low repetition rate and is dominated by the narrow nitroxide spectrum. (B) The nitroxide field-sweep echo-detected spectrum is plotted on an expanded field scale to show more clearly the positions of spectral excitations. The spectrum was recorded at 30 K and 0.5 kHz repetition rate. NO-NO DEER data were obtained in the standard manner, as shown. The temperature range used was 20-30 K, with the optimum being at 30 K.

SOD1 is a 32 kDa homodimer, with one Cu^{2+} ion tightly bound in each monomer (Figure 2.3, 2.5). The Cu^{2+} centers are 32 Å apart, which is well within the 10-90 Å range of PDS (Figure 2.3, 2.5). Over 150 mutants of SOD1 have been implicated in the familial form of ALS (fALS), a late-onset terminal

neurodegenerative disease [28]. The connection between the SOD1 mutations and fALS is unclear because the mutant proteins have very similar x-ray crystal structures [29, 30] and enzymatic activities [31] when compared to wild type (WT) SOD1. Decreases in thermal stability versus unfolding have been observed in ALS mutants [32], and recently, dimer destabilization leading to aggregation has been proposed as a potential mechanism of disease progression [33-36]. Here we apply Cu-PDS to probe the structural properties of the fALS SOD1 variant I149T, which we find has one of the most aberrant Cu-Cu interactions of several fALS variants tested.

2.3 Results

2.3.1 Cu-Cu ESR on WT SOD1

Human SOD1 and fALS variants in the background of the stabilizing mutations C6A and C111S [37, 38] were expressed in *E. coli* cells and purified as described below. The C6A and C111S substitutions aid expression, but do not effect SOD1 structure, activity [37, 38] or PDS signals, and are henceforth referred to as WT (Sec. 2.5, Figure 2.3). The resulting time-domain DEER traces of SOD1 revealed sufficiently deep DEER modulation (6-10%) with pronounced oscillations, characteristic of a single dominant dipolar interaction well within the optimal DEER distance range (Figures 2.3, 2.4). Tikhonov regularization [39] and the maximum entropy method [40] produced a distance distribution function that peaks sharply at

32 Å, the Cu-Cu separation in the SOD1 dimer (Figure 2.4A). The standard orientation-free kernel was effective in reconstruction, despite substantial anisotropy of the SOD1 ESR spectrum [41]. DQC produced similar results compared to DEER, but after just an hour of data averaging.

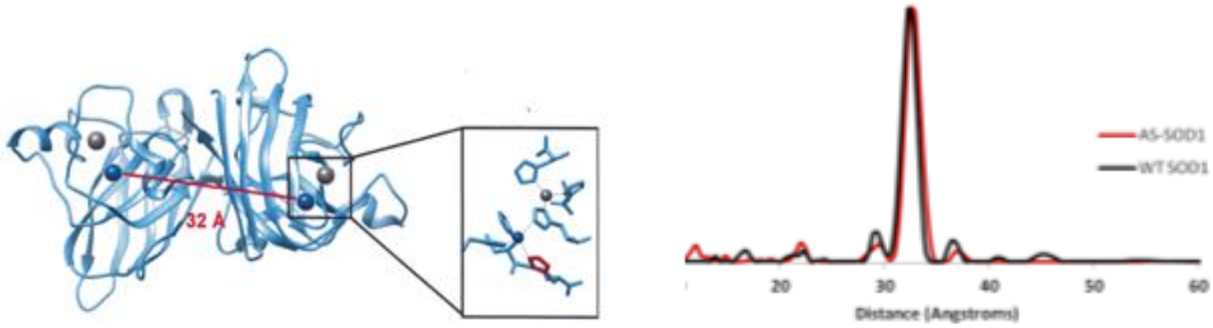

Figure 2.3: (Left) Crystal structure of human SOD1 (PDB ID: 2V0A). 32 Å separate the copper ions across the dimer interface (blue). Zinc ions (grey) reside beside the copper ions. Inset shows the active site metal coordination environment, with the histidine 48 side chain shown in red. Variant H48Q, for which Cu-Cu distance is 34 Å (Figure 2.4), substitutes a ligand to Cu^{2+}. (Right) Cu-Cu distance distributions for WT SOD1 vs. its C6A/C111S mutant, (abbreviated as AS-SOD1) reconstructed with the L-curve Tikhonov regularization method and refined using the Maximum Entropy Method.

2.3.2 Cu-Cu ESR on SOD1 fALS Mutants

PDS measurements on fALS mutants H48Q and I149T produced Cu-Cu distance distributions decidedly different from WT (Figure 2.4). H48Q substitutes a copper ligand in the active center and hence was expected to perturb the Cu-Cu separation (Figure 2.3). Indeed the Cu-Cu distribution from H48Q is widened and peaks at ~34 Å (Figure 2.4B). Remarkably, I149T, a substitution remote from the

active center, shows a distribution that is severely broadened compared to WT (Figure 2.4). Such behavior could arise from increased conformational variation within the protein itself, relocating the copper coordination site by ~1-2 Å.

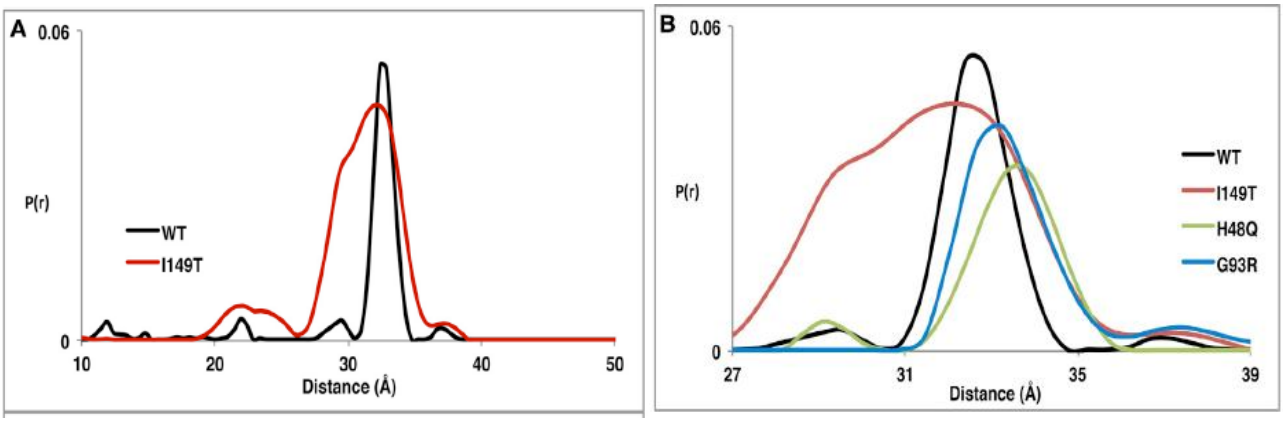

Figure 2.4: (A) Distance distributions for separation between Cu^{2+} centers of Cu/Zn SOD WT and I149T variant. (B) Distance distributions for WT Cu/Zn SOD and 3 ALS mutants at 350μM. The I149T P(r) is multiplied by 4 to aid comparison.

2.3.3 Crystal Structure of SOD1 fALS Mutant I149T

To further investigate the conformational properties of the I149T variant, we crystallized the protein and determined its structure to 2.4 Å resolution (Figure 2.5, Table 2.1; PDB accession code 4OH2). The crystallographic electron density clearly reflects loss of the Ile149 Cδ1 atom (Figure 2.5 inset), but otherwise, the I149T structure shows very little change compared to that of the WT. The 149 residue lies at the periphery of the dimer interface and the side chain faces into the hydrophobic core of the subunit β-barrel. The Ile→Thr substitution produces no change in β-

sheet conformation or subunit association that can be observed among the five unique dimers of the crystal structure. Comparison of the Cu-Cu distances in the I149T and WT structures (determined from refinement at equivalent resolutions) shows a slightly broader distribution for the mutant dimers (σ = 0.07 vs. 0.02 Å), but the overall widening is extremely small (and likely insignificant) compared to that observed in the Cu-PDS experiment (FWHM = 5.3 vs. 1.8 Å). Note that the increased width in the distance distribution cannot be reasonably attributed to orientational correlation effects that differ in I149T (cf. Section 2.5, Methods).

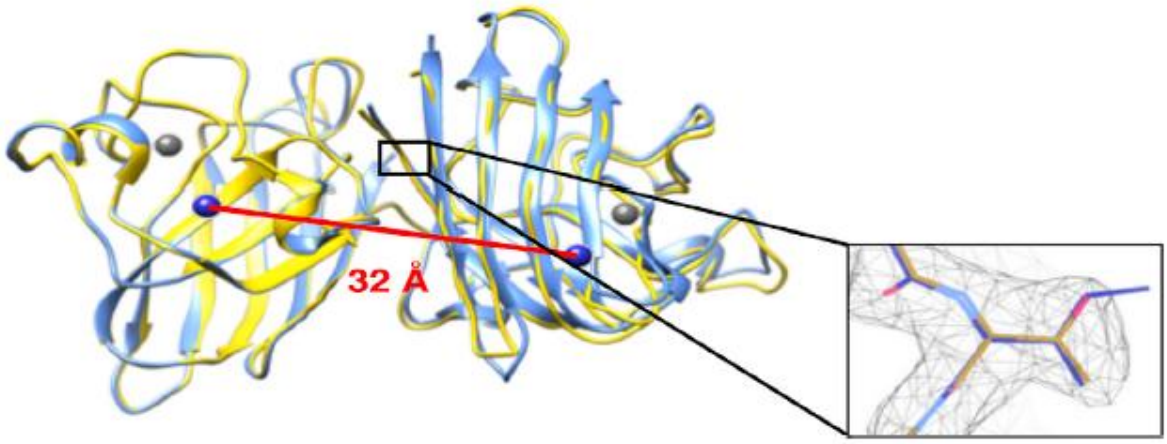

Figure 2.5: Cu-Cu separation in the SOD dimer detected by PDS. The I149T structure (yellow) is nearly identical to that of WT (blue). Inset: superposition of residue 149 with 2.38 Å resolution F_o-F_c omit electron density (3σ for Thr149).

Table 2.1: Crystallographic data collection and refinement statistics.

Data Collection Metric	
Space Group	$C222_1$
Unit Cell Dimensions, (Å)	a = 165.10
	b = 203.81
	c = 144.41
Unique Reflections	95,238
Resolution, (Å)	36.10-2.38
Last Shell, (Å)	2.44-2.38
Redundancy, (%/last shell)	6.4/2.9
Completeness, (%/last shell)	98.35/79.40
I/σ, (overall/last shell)	10.9/1.52
Linear R-fac (R_{merge}), (overall/last shell)	.209/.714
CC ½ (overall/last shell)	.973/.543
Refinement Metric	
R-Work, (%/last shell)	0.166/.225
R-Free, (%/last shell)	0.215/.290
No. atoms	22,063
No. water molecules	891
Mean B value	21.9
RMSD from ideal bond lengths, (Å)	0.014
RMSD from ideal bond angles, (Å)	1.4

2.3.4 Double Labeling of I149T

To further investigate the source of the broad Cu-Cu distribution in the non-crystalline state of the I149T variant, we labeled the SOD1 dimer with a nitroxide at position V94 by mutating the residue to Cys and reacting it with methanethiosulfonate spin label to yield the spin-bearing side-chain, commonly known as R1. The 94 position, which resides at an exposed β-turn, was chosen as a site that would generate unique Cu-NO, Cu-Cu and NO-NO distances, but whose labeling would not interfere with dimerization or metal binding. The combined set of three measurements for Cu and NO sites obtained from the WT protein show four distinct distances that match very well with the expectations from the crystal structure (Figure 2.6). However, the distance distributions derived from the I149T DEER data are much broader with all three types of spin interactions than those derived from the WT (Figure 2.6). Thus, the I149T dimer is structurally heterogeneous and this likely reflects increased protein dynamics in the subunits, perhaps further superimposed on subunit rearrangements about the dimer interface. Indeed, the 24 Å intra-subunit Cu-NO separation in I149T is greatly diminished and merged with the longer Cu-NO inter- subunit separation (Figure 2.6c). The small chemical change caused by the mutation (Figure 2.5) thus increases the amplitude (and/or nature) of structural fluctuations within the entire protein, which nonetheless still center on the WT conformation. Crystallization of I149T must mask this behavior by stabilizing the WT state and damping any

deviations from it. A similar pattern of behavior was recently reported for a membrane protein [42, 43], for which a wide range of structural states were sampled by DEER but only two distinct states were captured by x-ray crystallography.

Importantly, new short distances appear for I149T in both the Cu-Cu and NO-NO data (Figure 2.6), a fact that can only be explained by higher-order associations of SOD dimers (Figure 2.6, 2.9). Particularly, the strong ~25 Å NO-NO signal requires an association of SOD subunits through an interface distinct from that of the WT SOD dimer. Differences in the high SNR time-domain data between WT and I149T are also in full support of this interpretation (Figure 2.7). Although short distance components indicative of aggregation clearly manifest in both I149T and I149T:V94R1, they are more pronounced when the spin-label is present. Thus, V94R1 further destabilizes the I149T variant, even though V94R1 alone produces no conformational destabilization or aggregation (Figure 2.3, 2.6).

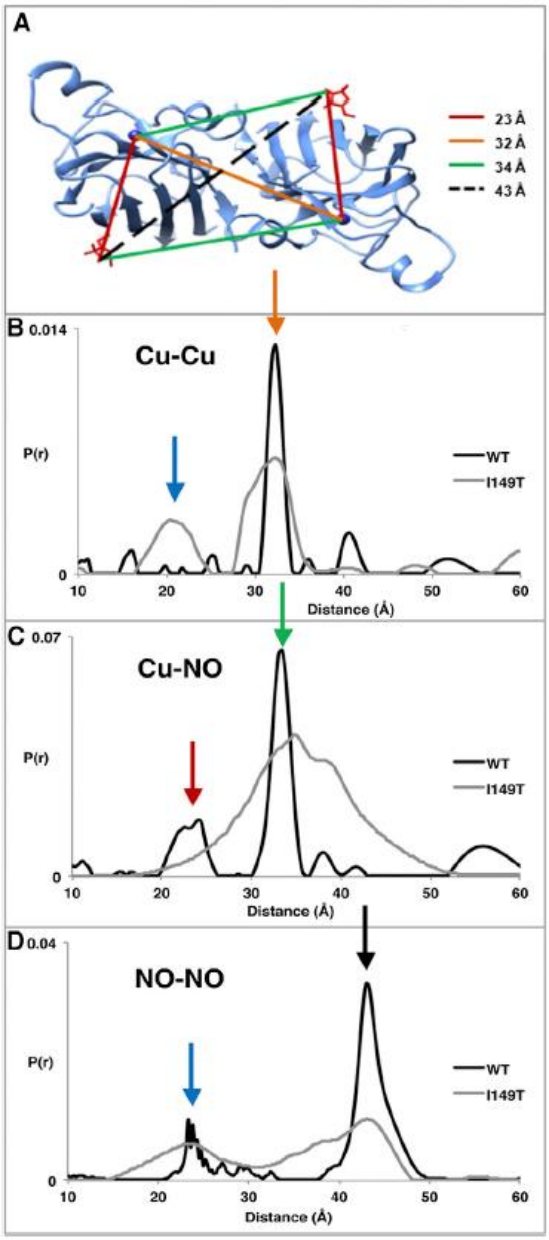

Figure 2.6: Cu-NO PDS experiment reveals conformational differences of the I149T variant. (A) Distances based on the I149T crystal structure (modeled nitroxides shown in red). (B) Cu-Cu, (C) Cu-NO and (D) NO-NO distance distributions from WT (C6A:V94C:C111S) and with addition of I149T shown with colored arrows designating separations in (A). Short distances can be explained by dimer association (blue). The I149T trace in (C) has been multiplied by 3 to aid comparison. Sharp 23 Å peak in NO-NO WT due to D_2O modulation is not present in protiated buffer. WT measurements were made at 150μM, I149T at 350 μM.

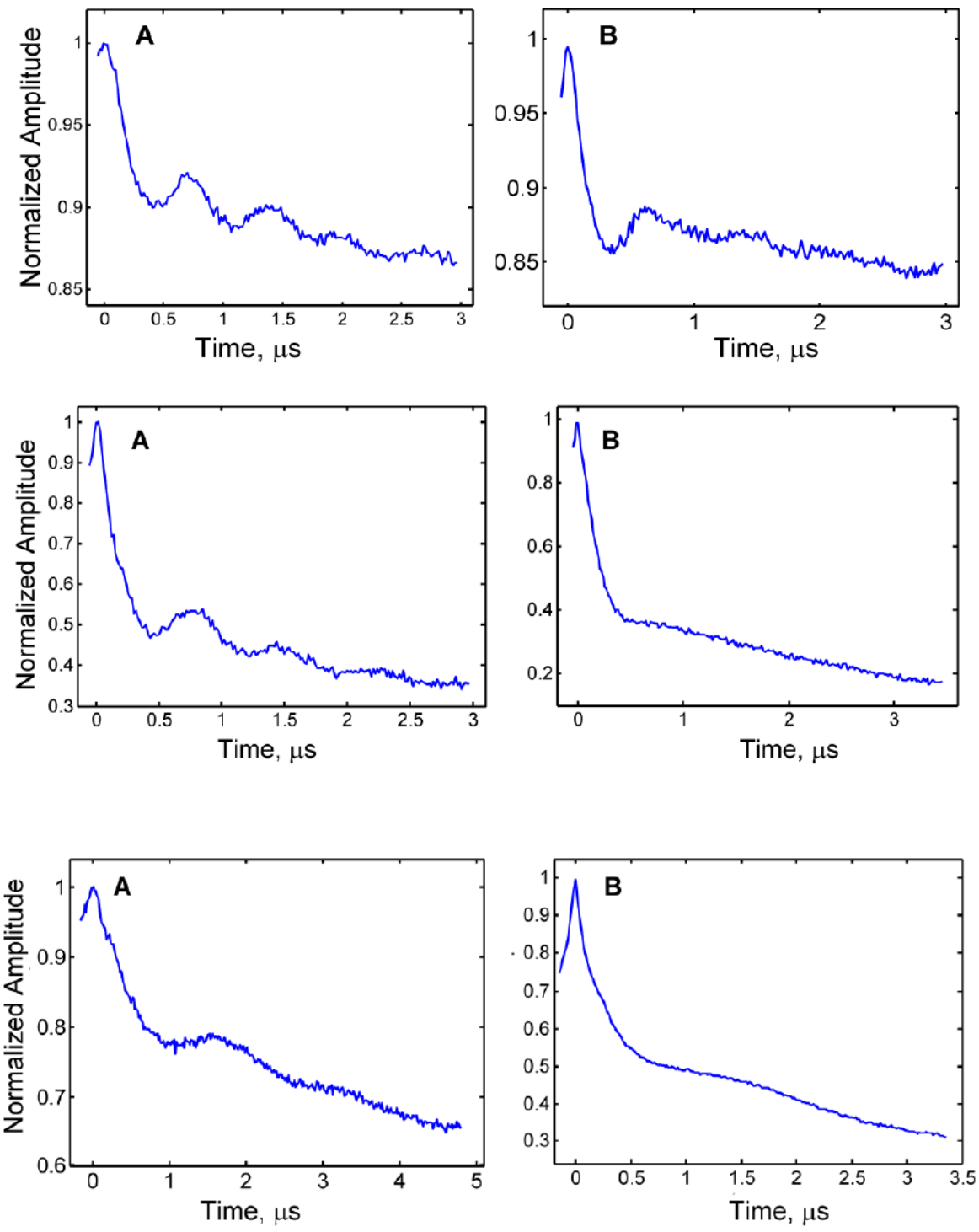

Figure 2.7: DEER raw time domain data for Cu-Cu (top), Cu-NO (middle) and NO-NO (bottom) for SOD1 V94R1 (A) vs. V94R1:I149T (B).

Aggregation of I149T is also supported by the increase in DEER modulation depth in all cases (Figure 2.7) and by multi-angle light scattering (MALS) experiments that show a concentration-dependent increase in the average molecular weight of the I149T variant, in contrast to the WT, which remains at the dimeric MW over the same concentration range (Figure 2.8). Hence, at least this particular ALS mutation appears to promote increased oligomerization of the protein.

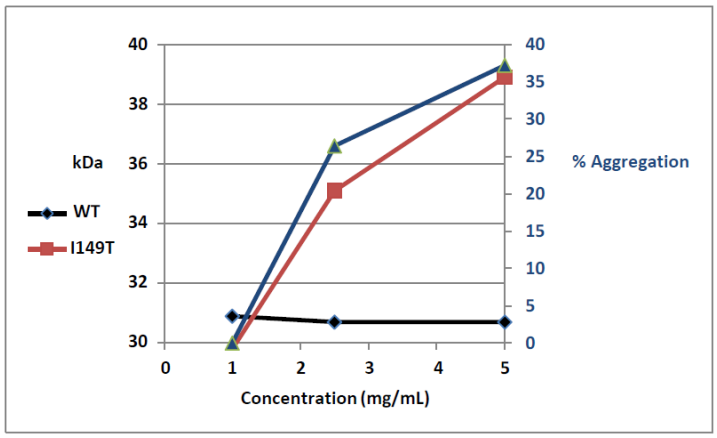

Figure 2.8: MALS data shows that the I149T mutant undergoes concentration dependent aggregation, which is completely absent for the WT. The black and red traces show average molecular weights for WT and I49T respectively. The blue trace shows the percentage of aggregation in the I149T samples.

The plausibility of a tetrameric state containing short spin separations is supported by a dimer-to-dimer contact in the I149T crystal structure that produces close Cu-Cu and NO-NO separations consistent with those observed by PDS (Figures 2.9 and 2.10). Due to the moderate intensity of the component derived from

a dimer-of-dimers, even less probable, higher-order aggregates are not expected to significantly change the signal. Indeed, a long-distance component in the range of the SOD1 dimer size (ca. 60 Å), which would result from abundant multimeric aggregates, was not observed. However, characterization of large protein aggregates currently presents a difficult task for PDS and is often only evident by an increase in the baseline slope [7, 44]. This issue could be explored further in the future.

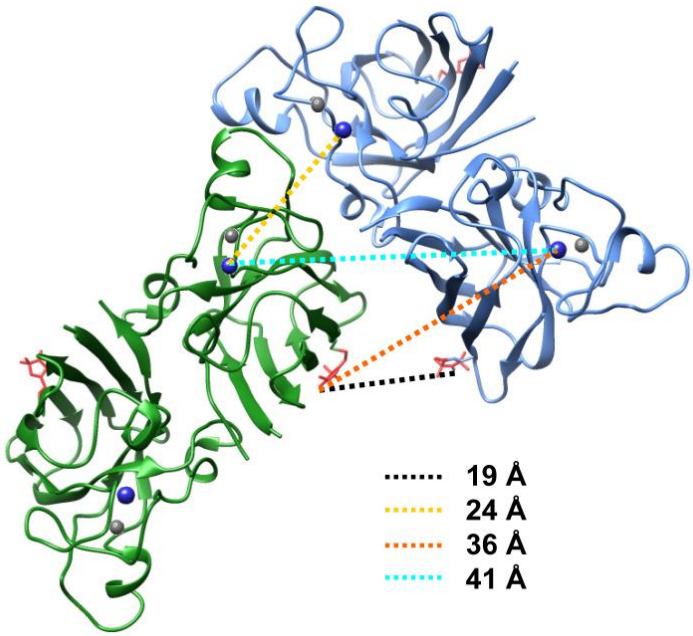

Figure 2.9: One possible model of the close contact between two I149T dimers (green and blue) based on one of several lattice contacts in the I149T crystal. Interdimer distances are shown with dashed lines. The inter-dimer Cu-Cu distance agrees well with the new short distance found in I149T by PDS, but rotation about this Cu-Cu distance is required to bring the nitroxide moieties into reasonable agreement with the PDS data (See Fig. 2.10.)

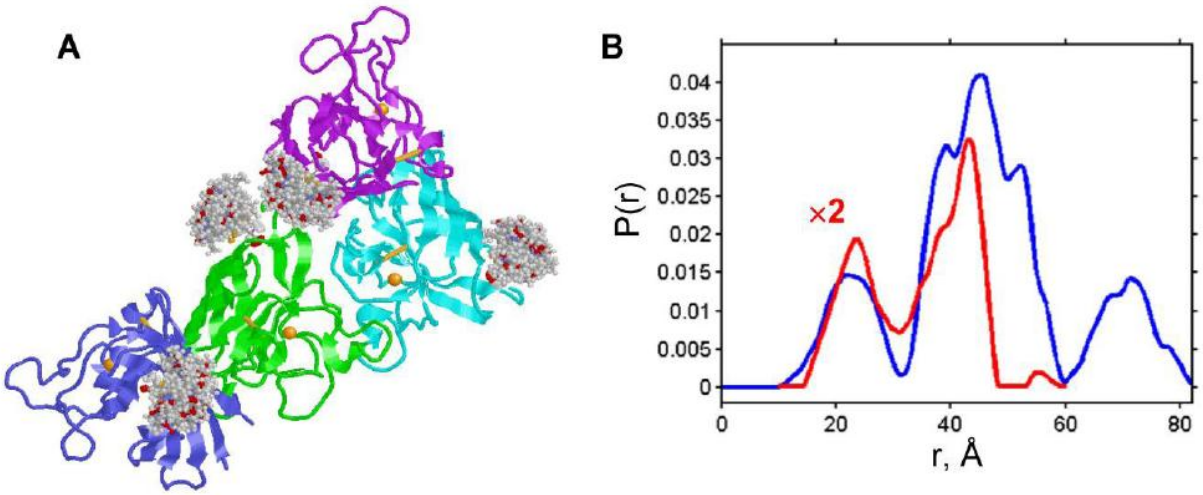

Figure 2.10: (A) R1 side-chain modeling was carried out on the crystal structure of Fig. 2.5. R1 conformers were simulated using rotamers implemented in MMM2013 (Molecular Multiscale Modeling, software package). (B) Distance distributions obtained by rotamer modeling (blue) are compared to the NO-NO experiment on R1:I149T (red). As expected, the experimental solution DEER data indicate that the structure must deviate from that suggested by contacts found in the crystal lattice. However, a relatively small restructuring, as noted under Fig. 2.9, can easily bring all distances into the proper range.

2.4 Discussion

The Cu-PDS experiments on SOD I149T indicate that the disease-causing residue substitution increases conformational heterogeneity in the overall protein fold. These data along with MALS experiments further show that the increase in heterogeneity leads to aggregation. Previous studies on the solution properties of SOD1 fALS mutants demonstrated that the residue substitutions caused local

destabilization [45] or produced metal loss, but that such changes were only detected in denaturing conditions [46, 47]. An exception is an ALS mutational site in contact with I149 (I113T). The I113T variant was also found not to affect the SOD1 crystal structure, yet the solution structure was considerably altered, as judged by small-angle x-ray scattering (SAXS) [30]. Our method is the first to reveal structurally defined conformational differences between WT SOD and an ALS mutant that promotes aggregation under native solution conditions. We have found that substitutions at the G93 position of SOD1 (a hot spot for ALS mutations) also give perturbed Cu-Cu PDS spectra, albeit to a lesser extent than I149T (Figure 2.4). Thus, an altered Cu-Cu PDS distribution, which appears as a consistent feature of ALS variants tested thus far, may prove a useful diagnostic tool for identifying mutations in *sod1* that lead to ALS. In conclusion, Cu-PDS, when combined with orthogonal (e.g. nitroxide) labeling, enables a straightforward assessment of metalloprotein conformation and association state relative to the active site metal ions. This information fills an important gap between low-resolution methods such as EM, SAXS, FRET and high-resolution techniques such as NMR and x-ray crystallography.

2.5 Methods

2.5.1 Protein Expression and Purification

Human SOD1 DNA in the pET28 expression vector was provided by the Getzoff lab at the Scripps Research Institute. ALS mutations were made using the

QuikChange method. Proteins were expressed in *E. coli* BL21-DE3 cells. Cells were grown at 37 °C until they reached an OD_{600} of 0.75, at which point they were induced with 0.3 mM IPTG and 0.25 mM $CuSO_4$. Additional growth for 18 hours at 25 ˚C was followed by harvesting by centrifugation. Cells were lysed by sonication, insoluble debris was pelleted by centrifugation, and the supernatant was run over a nickel immobilized metal ion affinity chromatography (IMAC) column. After elution with excess imidazole (200 mM), the His-tag was cleaved with an overnight thrombin digest of 400 μg at room temperature. SOD1 was further purified on a Superdex 75 (Pharmacia) size-exclusion column. Exogenous metal was removed using EDTA at low pH as previously described [48] and copper and zinc were reconstituted by direct addition of 1-20 μM $ZnSO_4$ (equi-molar with protein) at 4°C overnight followed by the same concentration of $CuSO_4$ at 4°C for 4 hours. The protein was then concentrated to 350 μM. Spin-labeling of Cys variants with MTSSL was carried out on these purified, metallated samples, as previously described [44].

Upon purification, the metal content of the fALS variants (I149T, H48Q) was shown by ICP-MS (University of Georgia, Department of Biochemistry [49]) not to vary significantly compared to that of the WT (all samples had 71-83% Cu^{2+} loading). Uniform metal content among the samples was also supported by only modest variations in ESR CW spectra and DEER echo modulation depth, the latter of which may in part be due to minor orientational effects (cf. Section 2.5, Methods).

2.5.2 Crystallization, X-ray Diffraction Data Collection, and Structure Solution

Crystals were grown by vapor diffusion using the hanging drop method. The well solution consisted of 100 mM NaCl, 100 mM Tris-HCl pH 7.6, 2.8 M $(NH_4)_2SO_4$ and a protein concentration of 10 mg/mL in 2.25 mM potassium phosphate buffer, pH 7.0 and 160 mM NaCl. Crystals were soaked in a solution of 25% glycerol in mother liquor before flash-cooling in the cryostream. Data was collected at 100 K using 0.979 Å x-ray radiation at beamline A1 at the Cornell High Energy Synchrotron Source (CHESS). Data was reduced and scaled using HKL2000 [50]. Initial phases were determined by molecular replacement in AutoMR (Phenix) [51] using WT SOD1 (PBD ID 2V0A [52]) as a model. Further refinement was carried out with XFIT [53], CNS [54], and Phenix [55] (Table 2.1). For comparison of Cu ion positions in the WT and I149T SOD, diffraction data for was downloaded from the PDB and then limited to the same resolution as of I149T. Cu ions were than randomly moved 0.5-1.0 Å and the structures refined in Phenix by conjugate gradient minimization. No restraints were placed on metal ion ligands. Variances of the Cu-Cu interdimer separations were calculated from the five unique SOD dimers in each asymmetric unit.

2.5.3 Pulsed Dipolar ESR Measurements

The 4-pulse DEER and the 6-pulse DQC PDS pulse sequences (Figure 2.1) were used to measure distances between unpaired electron spins of catalytic Cu^{2+} ions in SOD1 and fALS mutants. DEER was also applied to measure distances between

Cu^{2+} ions and MTSSL spin labels (R1) introduced at site V94, mutated to cysteine, as well as between the NO groups of R1 nitroxide side-chains. Cu-Cu DEER experiments were conducted by flipping Cu-spins and observing echo modulation at the field positions well outside of the NO spectrum (Figure 2.2). The field positions of pump and detection pulses, located at the Cu^{2+} g_y maximum (~5900-6000 G), were typically separated by 60 G. Experiments where pulse and pump positions were swapped and shifted by ~200 G toward the NO spectrum in Cu-NO measurements produced no detectable difference in DEER modulation. NO-NO interactions were determined by confining both frequencies to the nitroxide spectrum (at ~6200 G, Figure 2.2) where Cu^{2+} absorption is low and will not contribute more than ~1% to the deep (~50%) modulation of DEER signal when NO is pumped. Cu-NO interactions were measured by pumping the nitroxide spins between the pump and detection frequencies of the NO-NO case and detecting the Cu well outside of the NO envelope (Figure 2.2). DQC experiments were carried out at the Cu spectrum maximum in the g_y region. In general, the DQC method can easily circumvent the minimal 20 Å limit of standard DEER, as well as provide increased SNR.

The Cu^{2+} magnetic tensor principal axis system of bovine SOD [56] has g_z moderately tilted away from the inter-copper vector (**u**), which could introduce orientation correlation effects with respect to g_z but not for the $g_{x,y}$ spectral regions. Moreover, inhomogeneous hyperfine broadening comparable to the difference

between the pump and detection frequencies further randomizes orientational effects over the $g_{x,y}$, region. Thus, due to the stronger orientational correlations at g_z, the g_z spectral region was avoided. Fortunately, reconstruction artifacts caused by reduced contributions at g_z are expected to be minimal because the orientations selected at g_z have **u** aligned parallel to the static magnetic field and orientations of **u** less than the magic angle do not contribute substantially in the Tikhonov treatment. As expected, the DEER data recorded at g_x and g_y were identical and distance reconstruction exhibited few spurs. It is important to note that this simplification may not be appropriate for other copper systems: for example, one with an axial g-tensor that has z perpendicular to **u.** For such a case, orientational effects will likely necessitate a more complicated distance analysis, with related examples known for nitroxides [57].

2.5.4 Multi-angle Light Scattering Coupled to Size-exclusion Chromatography

Size-exclusion chromatography coupled with multi-angle light scattering (MALS) was used to study the molar mass of WT SOD1 and the I149T mutant. Proteins (1-5 mg/mL) were run at room temperature on an SEC column (WTC050N5 - Wyatt) pre-equilibrated with 50mM Tris-HCl pH 7.5, 150mM NaCl buffer. Analysis and molecular weight determination was carried out with Wyatt technologies ASTRA. Bovine serum albumin (Sigma) was used as a control for data quality.

REFERENCES

1. Jain, D. and Lamour, V. 2010. Computational tools in protein crystallography. *Methods Mol Biol.* 673: 129-156.

2. Clore, G.M. and Gronenborn, A.M. 1998. NMR structure determination of proteins and protein complexes larger than 20 kDa. *Curr Opin Chem Biol.* 2: 564-570.

3. Moerner, W.E. 2002. Single molecule spectroscopy of autofluorescent proteins. *J Chem Phys.* 117: 10925-10937.

4. Klare J.P. and Steinhoff, H.J. 2009. Spin labeling EPR. *Photosynth Res.* 102: 377-390.

5. Borbat, P.P. and Freed, J.H. 2007. Measuring distances by pulsed dipolar ESR spectroscopy: spin-labeled histidine kinases. *Methods Enzymol.* 423: 52-116.

6. Jeschke, G. 2012. DEER distance measurements on proteins. *Annu Rev Phys Chem.* 63: 419-446.

7. Borbat, P.P. and Freed, J.H. 2014. Pulse Dipolar ESR: Distance Measurements, In *Structure and Bonding.*152: 1-82.

8. 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.

9. 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.

10. 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.

11. Gaffney, B.J., Bradshaw, M.D., Frausto, S.D., Wu, F.Y., Borbat, P.P., and Freed, J.H. 2012. Locating a Lipid at the Portal to the Lipoxygenase Active Site. *Biophys. J.* 103: 2134-2144.

12. Yang, Z.Y., 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. *P Natl Acad Sci USA*. 109: E993-E1000.

13. Sarver, J., Silva, K.I., and Saxena, S. 2013. Measuring Cu2+-Nitroxide Distances Using Double Electron-Electron Resonance and Saturation Recovery. *App Mag Res*. 44: 583-594.

14. Astashkin, A.V., Rajapakshe, A., Cornelison, M.J., Johnson-Winters, K., and Enemark, J.H. 2012. Determination of the distance between the Mo(V) and Fe(III) heme centers of wild type human sulfite oxidase by pulsed EPR spectroscopy. *J Phys Chem B*. 116: 1942-1950.

15. Jun, S., Becker, J.S., Yonkunas, M., Coalson, R., and Saxena, S. 2006. Unfolding of alanine-based peptides using electron spin resonance distance measurements. *Biochemistry*. 45: 11666-11673.

16. Ruthstein, S., Ji, M., Mehta, P., Jen-Jacobson, L., and Saxena, S. 2013. Sensitive Cu2+-Cu2+ distance measurements in a protein-DNA complex by double-quantum coherence ESR. *J Phys Chem B*. 117: 6227-6230.

17. 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.

18. Ezhevskaya, M., Bordignon, E., Polyhach, Y., Moens, L., Dewilde, S., Jeschke, G., and van Doorslaer, S. 2013. Distance determination between low-spin ferric haem and nitroxide spin label using DEER: the neuroglobin case. *Mol Phys*. 111: 2855-2864.

19. Borbat, P.P., and Freed, J.H. 2000. Distance Measurements in Biological Systems by EPR. In *Biological Magnetic Resonance*, ed. LJ Berliner, GR Eaton, SS Eaton, 19: 385-459. New York: Academic/Plenum Publishers.

20. Borbat, P.P. and Freed, J.H. 1999. Multiple-quantum ESR and distance measurements. *Chem Phys Lett*. 313 :145-154.

21. 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.

22. Song, Y., Meade, T.J., Astashkin, A.V., Klein, E.L., Enemark, J.H., and Raitsimring, A. 2011. Pulsed dipolar spectroscopy distance measurements in biomacromolecules labeled with Gd(III) markers. *J Magn Reson.* 210: 59-68.

23. Garbuio, L., Bordignon, E., Brooks, E.K., Hubbell, W.L., Jeschke, G. and Yulikov, M. 2013. Orthogonal Spin Labeling and Gd(III)-Nitroxide Distance Measurements on Bacteriophage T4-Lysozyme. *J Phys Chem B.* 117: 3145-3153.

24. Lovett, J.E., Lovett, B.W., and Harmer, J. 2012. DEER-Stitch: Combining three- and four-pulse DEER measurements for high sensitivity, deadtime free data. *J Magn Reson.* 223: 98-106.

25. Borbat, P.P., Georgieva, E.R., and Freed, J.H. 2013. Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance. *J Phys Chem Lett.* 4: 170-175.

26. Borbat, P.P., Davis, J.H., Butcher, S.E., and Freed, J.H. 2004. Measurement of large distances in biomolecules using double-quantum filtered refocused electron spin-echoes. *J Am Chem Soc.* 126: 7746-7747.

27. Jeschke, G., Bender, A., Paulsen, H., Zimmermann, H., and Godt, A. 2004. Sensitivity enhancement in pulse EPR distance measurements. *J Magn Reson.* 169: 1-12.

28. Pratt, A.J., Getzoff, E.D., and Perry, J.J. 2012. Amyotrophic lateral sclerosis: update and new developments. *Degener Neurol Neuromuscul Dis.* 2012: 1-14.

29. Galaleldeen, A., Strange, R.W., Whitson, L.J., Antonyuk, S.V., Narayana, N., Taylor, A.B., Schuermann, J.P., Holloway, S.P., Hasnain, S.S., and Hart, P.J. 2009. Structural and biophysical properties of metal-free pathogenic SOD1 mutants A4V and G93A. *Arch Biochem Biophys.* 492: 40-47.

30. Hough, M.A., Grossmann, J.G., Antonyuk, S.V., Strange, R.W., Doucette, P.A., Rodriguez, J.A., Whitson, L.J., Hart, P.J., Hayward, L.J., Valentine, J.S., and Hasnain, S.S. 2004. Dimer destabilization in superoxide dismutase may result in disease-causing properties: structures of motor neuron disease mutants. *Proc Natl Acad Sci U S A.* 101: 5976-5981.

31. Borchelt, D.R., Lee, M.K., Slunt, H.S., Guarnieri, M., Xu, Z.S., Wong, P.C., Brown Jr., R.H., Price, D.L., Sisodia, S.S., and Cleveland, D.W. 1994.

Superoxide dismutase 1 with mutations linked to familial amyotrophic lateral sclerosis possesses significant activity. *Proc Natl Acad Sci U S A*. 91: 8292-8296.

32. Rodriguez, J.A., Valentine, J.S., Eggers, D.K., Roe, J.A., Tiwari, A., Brown Jr., R.H., and Hayward, L.J. 2002. Familial amyotrophic lateral sclerosis-associated mutations decrease the thermal stability of distinctly metallated species of human copper/zinc superoxide dismutase. *J Biol Chem*. 277: 15932-15937.

33. DiDonato, M., Craig, L., Huff, M.E., Thayer, M.M., Cardoso, R.M., Cassmann, C.J., Lo, T.P., Bruns, C.K., Powers, E.T., Kelly, J.W., Getzoff, E.D., and Tainer, J.A. 2003. ALS mutants of human superoxide dismutase form fibrous aggregates via framework destabilization. *J Mol Biol*. 332: 601-615.

34. Bruijn, L.I., Houseweart, M.K., Kato, S., Anderson, K.L., Anderson, S.D., Ohama, E., Reaume, A.G., Scott, R.W., and Cleveland, D.W. 1998. Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. *Science*. 281: 1851-1854.

35. Stathopulos, P.B., Rumfeldt, J.A., Scholz, G.A., Irani, R.A., Frey, H.E., Hallewell, R.A., Lepock, J.R., and Meiering, E.M. 2003. Cu/Zn superoxide dismutase mutants associated with amyotrophic lateral sclerosis show enhanced formation of aggregates in vitro. *Proc Natl Acad Sci U S A*. 100: 7021-7026.

36. Elam, J.S., Taylor, A.B., Strange, R., Antonyuk, S., Doucette, P.A., Rodriguez, J.A., Hasnain, S.S., Hayward, L.J., Valentine, J.S., Yeates, T.O., and Hart, P.J. 2003. Amyloid-like filaments and water-filled nanotubes formed by SOD1 mutant proteins linked to familial ALS. *Nat Struct Biol*. 10: 461-467.

37. Hallewell, R.A., Imlay, K.C., Lee, P., Fong, N.M., Gallegos, C., Getzoff, E.D., Tainer, J.A., Cabelli, D.E., Tekamp-Olson, P., Mullenbach, G.T., and Cousens, L.S. 1991. Thermostabilization of recombinant human and bovine CuZn superoxide dismutases by replacement of free cysteines. *Biochem Biophys Res Commun*. 181: 474-480.

38. Lepock, J.R., Frey, H.E., and Hallewell, R.A. 1990. Contribution of conformational stability and reversibility of unfolding to the increased thermostability of human and bovine superoxide dismutase mutated at free cysteines. *J Biol Chem*. 265: 21612-21618.

39. Chiang, Y.W., Borbat, P.P., and Freed, J.H. 2005. The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J Magn Reson*. 172: 279-295.

40. Chiang, Y.W., Borbat, P.P., and Freed, J.H. 2005. Maximum entropy: a complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *J Magn Reson*. 177: 184-196.

41. Lieberman, R.A., Sands, R.H., and Fee, J.A. 1982. A study of the electron-paramagnetic resonance properties of single monoclinic crystals of bovine superoxide-dismutase. *J Biol Chem*. 257: 336-344.

42. Georgieva, E.R., Borbat, P.P., Ginter, C., Freed, J.H., and Boudker, O. 2013. Conformational ensemble of the sodium-coupled aspartate transporter. *Nat Struct Mol Biol.* 20: 215-221.

43. Hanelt, I., Wunnicke, D., Bordignon, E., Steinhoff, H.J., and Slotboom, D.J. 2013. Conformational heterogeneity of the aspartate transporter Glt(Ph). *Nat Struct Mol Biol.* 20: 210-214.

44. Bhatnagar, J., Borbat, P.P., Pollard, A.M., Bilwes, A.M., Freed, J.H., and Crane, B.R. 2010. Structure of the ternary complex formed by a chemotaxis receptor signaling domain, the CheA histidine kinase, and the coupling protein CheW as determined by pulsed dipolar ESR spectroscopy. *Biochemistry*. 49: 3824-3841.

45. Museth, A.K., Brorsson, A.C., Lundqvist, M., Tibell, L.A., and Jonsson, B.H. 2009. The ALS-associated mutation G93A in human copper-zinc superoxide dismutase selectively destabilizes the remote metal binding region. *Biochemistry*. 48: 8817-8829.

46. Kitamura, F., Fujimaki, N., Okita, W., Hiramatsu, H., and Takeuchi, H. 2011. Structural instability and Cu-dependent pro-oxidant activity acquired by the apo form of mutant SOD1 associated with amyotrophic lateral sclerosis. *Biochemistry*. 50: 4242-4250.

47. Rumfeldt, J.A., Lepock, J.R., and Meiering, E.M. 2009. Unfolding and folding kinetics of amyotrophic lateral sclerosis-associated mutant Cu,Zn superoxide dismutases. *J Mol Biol.* 385: 278-298

48. Lyons, T.J., Nersissian, A., Goto, J.J., Zhu, H., Gralla, E.D., and Valentine, J.S. 1998. Metal ion reconstitution studies of yeast copper-zinc superoxide

dismutase: the "phantom" subunit and the possible role of Lys7p. *JBIC*. 3: 650-662.

49. Vaccaro, B., Menon, A.L., Lancaster, W.A., and Adams, M.W.W. 2009. Metallomics Using Inductively Coupled Plasma Mass Spectrometry. In *Current Protocols in Chemical Biology*: John Wiley & Sons, Inc.

50. Otwinowski, Z., and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, ed. CW Carter, Jr.: 307-326. San Diego: Academic Press.

51. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., and Zwart, P.H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 66: 213-221.

52. Strange, R.W., Yong, C.W., Smith, W., and Hasnain, S.S. 2007. Molecular dynamics using atomic-resolution structure reveal structural fluctuations that may lead to polymerization of human Cu-Zn superoxide dismutase. *Proc Natl Acad Sci U S A*. 104: 10040-10044.

53. McRee, D.E. 1992. XtalView: A visual protein crystallographic software system for X11/Xview. *J Mol Graph*. 10: 44-47.

54. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*. 54: 905-921.

55. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*. 68: 352-367.

56. Lieberman, R.A., Sands, R.H., and Fee, J.A. 1982. A study of the electron-paramagnetic resonance properties of single monoclinic crystals of bovine superoxide-dismutase. *J Biol Chem*. 257: 336-344.

57. Marko, A., Margraf, D., Yu, H., Mu, Y., Stock, G., and Prisner, T. 2009. Molecular orientation studies by pulsed electron-electron double resonance experiments. *J Chem Phys*. 130: 64102-64111.

Chapter 3

SITE-SPECIFIC INCORPORATION OF A Cu^{2+} SPIN-LABEL INTO PROTEINS FOR MEASURING DISTANCES BY PULSED DIPOLAR ESR SPECTROSCOPY

3.1 Abstract

Pulsed dipolar ESR spectroscopy (PDS) is a powerful tool for measuring solution-state distances in macromolecules. Recently, paramagnetic metal ions, specifically Cu^{2+}, have been utilized as spin labels on proteins as they can be spectroscopically distinguished from traditional nitroxide (NO)-based labels. However, the application of metal ions for PDS has been largely limited to natural metalloproteins. Here we demonstrate site-specific incorporation of Cu^{2+} into non-metalloproteins through use of a genetically encodable non-natural amino acid, 3-pyrazolyltyrosine (PyTyr). We first establish proof-of-principle for PyTyr-based PDS by measuring Cu-to-nitroxide distances within an engineered cyan-fluorescent protein (CFP) and then apply the method to track interactions and domain positions in the complex formed by the large multi-domain histidine kinase CheA and its target response regulator CheY. These measurements reveal new mechanistic insights into the dynamics of the histidine substrate domain during the phosphotransfer cycle.

3.2 Introduction

Pulsed dipolar ESR spectroscopy (PDS) in concert with site-specific spin labeling provides long distance restraints (up to 90 Å) to characterize protein conformation and assembly states in solution [1-3]. In the most common usage, nitroxyl-based spin-probes (NO) are covalently attached to engineered cysteine residues [4]. However, large eukaryotic proteins often contain many cysteine residues, which cannot always be replaced to ensure specific labeling. Furthermore, nitroxide radicals are spectroscopically indistinguishable from each other and thus most experiments are limited to two interacting sites. In contrast, paramagnetic metal ions, such as Cu^{2+}, have anisotropic g-tensors that allow their specific excitation and detection. Moreover, the combination of nitroxide and Cu^{2+}-based spin labels in the same sample allows for Cu-Cu, NO-NO and Cu-NO separations to be uniquely determined [5,6]. Nonetheless, specificity of binding is a major limitation in the use of metal ions as spin probes. Natural metal-binding centers in active sites and metal-binding loops of metalloproteins have been exploited to this end [7-10], but these complex centers have drawbacks in control over positioning and transfer to non-cognate systems. To overcome these limitations, a small, specific metal-binding probe that can be incorporated at any position in a protein is desirable. To this end, we have employed a genetically-encodable non-natural derivative of tyrosine, 3-pyrazolyltyrosine (PyTyr, Figure 3.1), to site-specifically incorporate Cu^{2+} ions for

PDS measurements. PyTyr was previously used as an electron acceptor in protein electron transfer (ET) experiments [11]. However, its ability to bind and stabilize Cu^{2+} with reasonably high affinity makes it a suitable candidate for use as a spin label. Here, we incorporate PyTyr into two different proteins and demonstrate its utility for PDS. For proof of principle we measure Cu-NO distances in cyan fluorescent protein (CFP, Figure 3.1), and then we characterize the interaction between the multi-domain histidine kinase CheA and its phosphorylation target, the response regulator protein CheY (Figure 3.2). In the latter case, Cu-PDS reveals domain juxtapositions within the complex and tracks conformational changes upon binding protein effectors and nucleotide modulators. Assessment of the spin-echo modulation depth further reveals the stoichiometry of interaction [12].

3.3 Results

In order to incorporate PyTyr into a protein of interest, a mutant tRNA/tRNA synthetase pair (PyTyrRS) was evolved to deliver PyTyr to the ribosome [11] through suppression of the amber (TAG) stop codon [13]. PyTyr was synthesized from tyrosine and pyrazole by a CuI catalyzed cross-coupling reaction, with some modification to the published protocol (see Methods).

3.3.1 Incorporation of PyTyr into CFP

We first chose to incorporate PyTyr into cyan-fluorescent protein (CFP, Figure 3.1), a derivative of green fluorescent protein (GFP), that offers a colored, structurally characterized system to readily evaluate yields and test feasibility [14]. PyTyr was placed at position 151 (CFP151PyTyr) and a Ser208Cys substitution was introduced by mutagenesis to allow additional labeling with a nitroxyl radical. The double mutant was cloned into pET28 and co-expressed with PyTyrRS in *E. coli* BL21(DE3) cells. The His_6-tagged protein was purified by Ni-NTA affinity chromatography, during which Cys208 was labeled on-column with the nitroxide spin label (MTSSL [1-oxyl-2,2,5,5-tetramethylpyrroline-3-(methyl)-methanethiosulfate]). After removal of the His_6-tag by thrombin digest and subsequent gel filtration chromatography, Cu^{2+} was added as $CuSO_4$ (1.5 equivalents) at 4˚C overnight. The protein was then buffer exchanged to remove excess Cu^{2+} and concentrated to 150 μM for PDS measurements. .

Cu-NO PDS experiments were carried out on a home-built Ku-band ESR spectrometer operating at 17.0-17.6 GHz with a four-pulse double electron electron resonance (DEER) pulse sequence [15]. Protein samples were 150 uM, and the data was collected at 15-20 K for 8-12 hours. For Cu-NO measurements, nitroxide spins were pumped at frequencies corresponding to 6180 G and Cu^{2+} spins were detected at those corresponding to 6100 G. Time-domain traces were converted to distance distributions using Tikhonov regularization [16] followed by the maximum entropy

method [17]. The bimodal distance distribution (Figure 3.1) obtained agrees well with a model based on the crystallographic structure of CFP [18]. The two peaks in the distribution likely reflect two distinct conformations of the spin label.

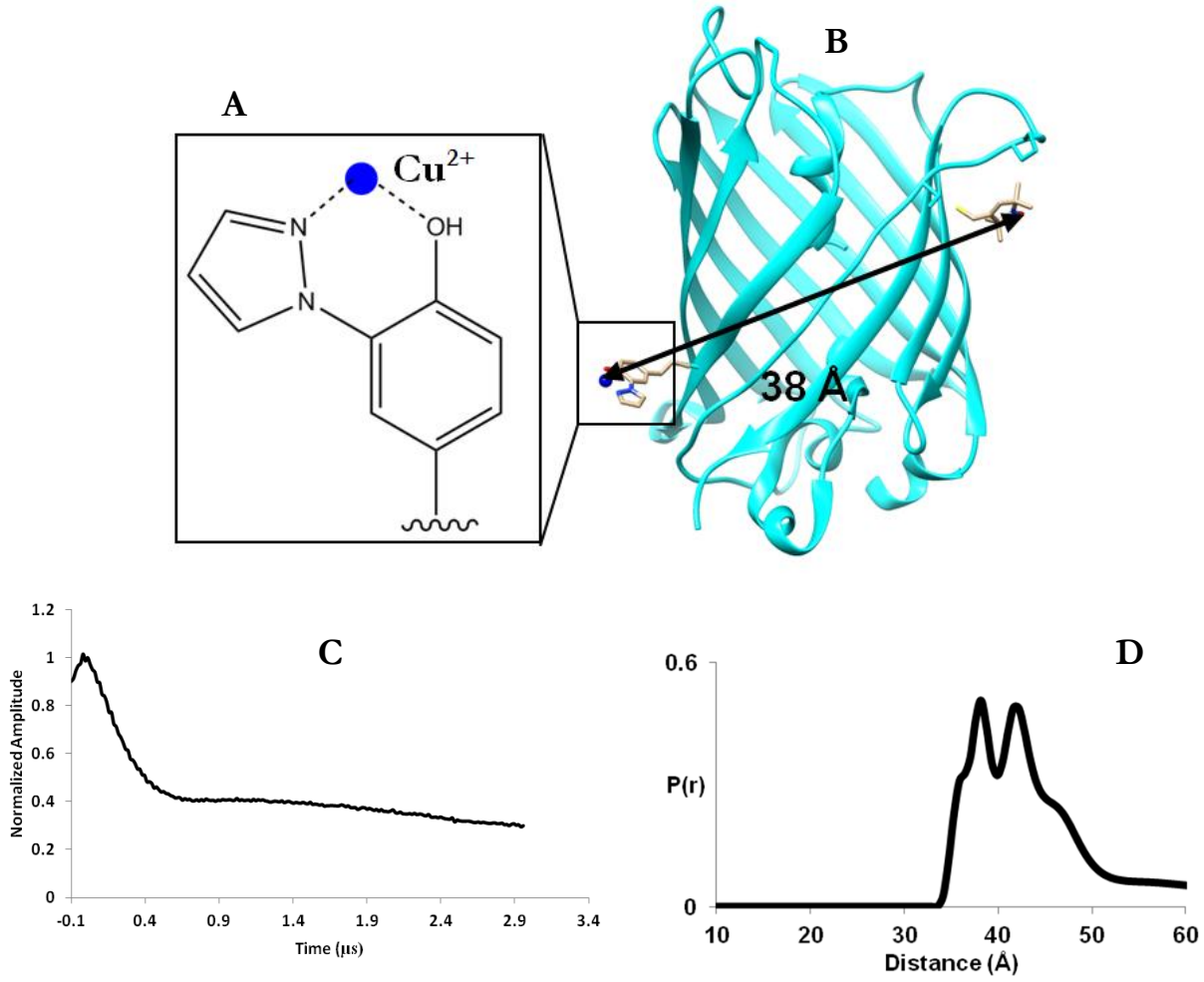

Figure 3.1: (A) Cu^{2+} binding amino acid PyTyr. (B) Expected Cu-NO separation in CFP from the crystal structure CFPpyTyr (PDB:3ZTF [18]) modeled with a nitroxide on residue 208. (C) Cu-NO baseline corrected time domain data. (D) Cu-NO distance distribution obtained by DEER spectroscopy.

3.3.2 Incorporation of PyTyr into T. maritima CheY

Next, we applied PyTyr Cu^{2+}-labeling to characterize domain juxtaposition in the complex formed by the histidine kinase CheA and its target response-regulator protein CheY. CheA and CheY form a two-component system [19] that regulates bacterial motility in response to chemoattractants. *T. maritima* CheA is a dimeric kinase, each subunit composed of five domains (named P1 through P5). CheY docks at the CheA P2 domain to receive phosphate from the P1 domain, which undergoes autophosphorylaion on a specific histidine residue by the ATP-binding P4 domain. P1 and P2 are connected to each other and the CheA core (P3P4P5) by long flexible linkers [20, 21]. Crystal and NMR structures are known for all CheA domains [22-25], but not as an intact protein. The modulation of interactions among P1, P2 and P4 underlie kinase regulation and thereby biological function. There is currently little structural information for P1 and P2 when attached to the kinase core.

Following procedures analogous to those described above, we labeled the response regulator CheY with PyTyr-Cu^{2+} and the P1 domain of CheA with a nitroxide spin-label. CheA dimerization then generates two nitroxide labels per kinase unit; likewise we expect two CheY proteins to bind a dimeric kinase.

PyTyr was introduced at position 41 of *T. maritima* CheY and a Cys residue was introduced at position 76 of P1 in an otherwise cys-less version of CheA (C63S C208S). The CheY variant was expressed and purified in the same manner as

CFP151PyTyr, except that there was no need for nitroxide labeling. Expression and purification of CheY41PyTyr yielded 510 ug of protein, which was concentrated down to 600 uM. Complexes were formed by direct addition of each component and incubation overnight at 4˚C, after which they were flash frozen for PDS measurements. CheA and CheW were expressed and purified, and CheA was spin labeled, as previously described [26, 27].

Cu-NO DEER time-domain traces (Figure 3.5) were collected in the same manner as for CFP151PyTyr. NO-NO DEER data was collected on the same spectrometer at 20-30 K for 8-12 hours, with standard pulse sequences [5]. Tikhonov regularization followed by the maximum entropy optimization produced Cu-NO distances. Weak Cu-Cu and NO-NO signals were observed, but these were not strong enough to yield usable distance distributions (cf. Figure 3.4).

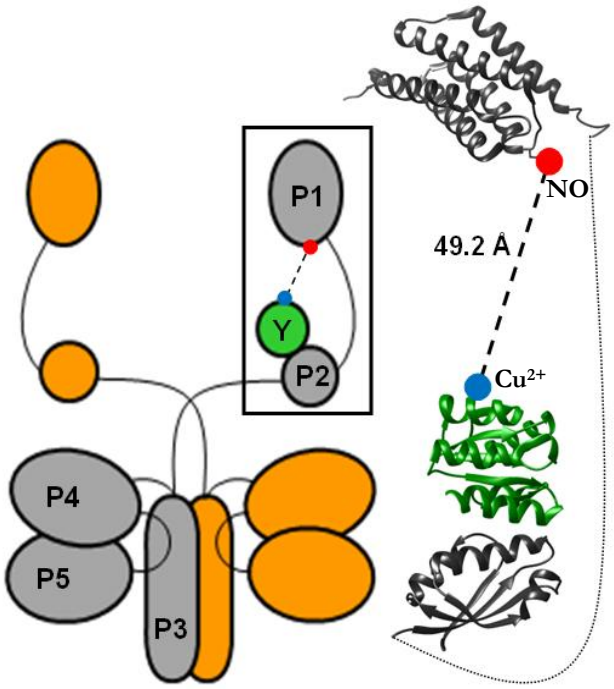

Figure 3.2: Schematic of the domain architecture of the CheA-CheY complex. A structural model of the P1-P2-CheY complex (P1, P2 – black ribbons, CheY-green ribbons), with the observed Cu-NO ESR distance indicated by a dotted line. (PDB: 1U0S [28] (P2-CheY) and 2LP4 [25] (P1))

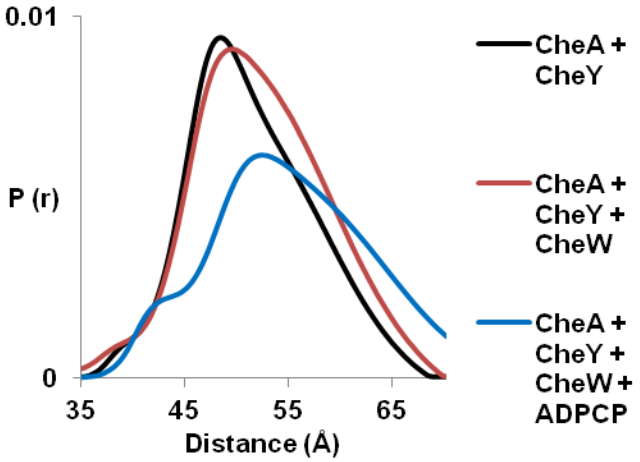

Figure 3.3: Cu-NO distance distributions for CheA-CheY complexed on their own (black), with unlabeled CheW (red) and with unlabeled CheW and ADPCP (blue).

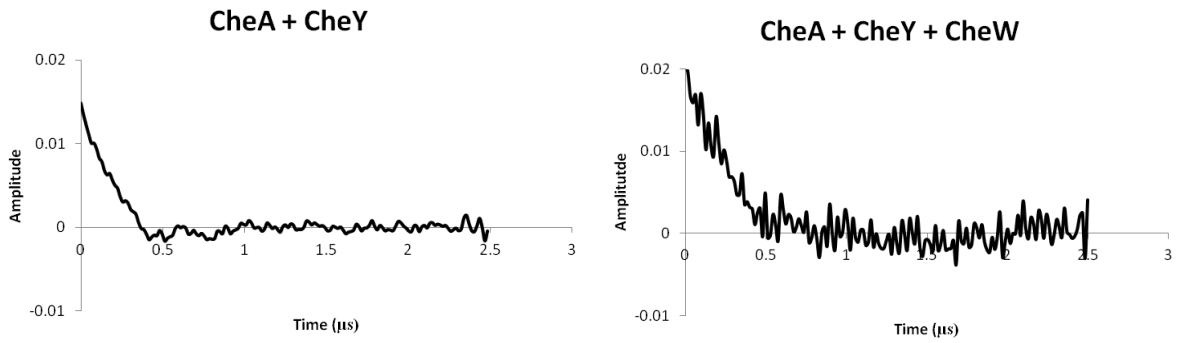

Figure 3.4: Cu-Cu (background corrected) time domain signals for CheA + CheY (left) and CheA + CheY + CheW samples (right).

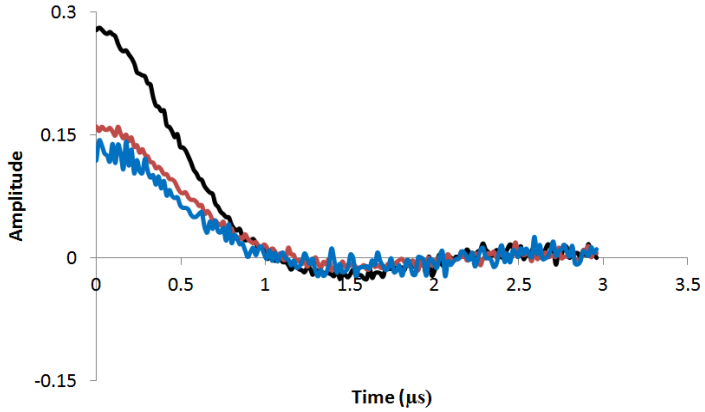

Figure 3.5: Cu-NO (background corrected) time domain signals for CheA + CheY (black), CheA + CheY + CheW (red) and CheA + CheY + CheW + ADPCP (blue).

3.4 Discussion

The obtained Cu-NO distance distributions (Figure 3.3) give insight into the structure of the CheA-CheY complex, in particular the relationship between the phosphate donating P1 domain and the acceptor CheY when bound to P2. In the resting state of CheA, with only CheY present, the Cu-NO maximal distance of 48 Å

suggests that the P1 domain resides some distance away from P2 and CheY. The weak Cu-Cu modulations indicate that the CheY molecules themselves are also far dispersed in the CheA dimer (> 60 Å). These signals collectively imply that in the absence of ATP, the structure of CheA is extended, with the P1 domain not docked on the P2-CheY complex. The protein CheW binds to the CheA regulatory domain P5 and is essential for coupling CheA to the receptor arrays [29]. The addition of CheW to CheA:CheY causes the Cu-NO distance distributions to broaden, and shift to slightly longer distances. These changes indicate that CheW can influence the position and dynamics of the P1 domain. Addition of the non-hydrolyzable ATP analog ADPCP to CheA + CheY + CheW, further shifts the Cu-NO distributions to longer distances. ADPCP-mediated release of P1 is also supported by SAXS data, which reveals a much more extended structure of CheA in the presence of the nucleotide [30].

The PDS measurements also enable the assessment of binding stoichiometry. The weak Cu-Cu signals from the complex could derive from long distances of separation or substoichiometric binding (i.e. negative cooperativity within the dimer producing a 1:2 CheY:CheA subunit complex). To confirm the binding stoichiometry, an NO-Cu DEER experiment was performed by detecting the NO spin echo while pumping the Cu^{2+} spins. The modulation depths of NO_{detect}-Cu_{pump} and Cu_{detect}-Cu_{pump} experiments are about the same (~.024), which strongly indicates that on average, 1

CheY proteins binds to one CheA subunit. Importantly, this measurement indicates that the Cu^{2+} binds specifically to PyTyr; if otherwise, there would not be close correspondence between the Cu-Cu and Cu-NO modulation depths.

This study has demonstrated the site-specific incorporation of Cu^{2+} for use as a spin-label in PDS experiments. In principle, any position on any protein expressed in *E. coli* can be targeted by this method. Furthermore, the labeling is quite specific, with little background from non-specific binding on a relatively large protein complex (~ 160 kDa) Thus, PyTyr incorporation extends PDS as a useful tool for exploring large dynamic protein complexes, the conformations of which are otherwise difficult to characterize. Moreover, the genetically encodable nature of the PyTyr label bypasses the need for covalent attachment. This advance overcomes limitations such as protein degradation during labeling and the lack of site-specificity when targeting Cys residues in large proteins. The spectroscopic challenges of Cu^{2+} such as an anisotropic g tensor and fast T2 relaxation times can be overcome [5, 9] and in some cases turned to advantage. For example, the broadness of the Cu^{2+} ESR spectra allows Cu signals to be distinguished from those of nitroxide; in CheA this allows for the measurement of 3 distinct distance distributions [5, 6]. It should also be noted that PyTyr is a close derivative of Tyr, and where such a substitution is made, the perturbation to protein structure on labeling will be minimal. Saxena et al have recently described an alternative strategy for incorporating Cu^{2+} specifically into proteins through

introduction of judiciously positioned His residues on known secondary structure elements [10]. The combination of that strategy, along with application of PyTyr could potentially allow high fidelity dual labeling by Cu. Finally, the ability to genetically encode Cu^{2+} labels, which should generally be more stable to reduction than nitroxides, also raises the possibility of their use as *in vivo* spin probes.

3.5 Methods

3.5.1 Synthesis of PyTyr

Production of Boc-L-3-iodotyrosine: 3-Iodo-L-Tyrosine was purchased from Aapptec. A 500 mL round bottom flask was charged with 3-Iodo-L-Tyrosine (2.0 g, 6.51 mmol), di-tert-butyl dicarbonate (1.42 g, 6.51 mmol), 40 mL dioxane, 60 mL 500 mM NaOH. A solution of 6 M NaOH was then added dropwise until the pH reached 9.5. The reaction stirred at room temperature for 4 hours. NaOH was added dropwise as needed to maintain pH at 9.5. The reaction was quenched with 6 M HCl until the pH reached 1, at which time the solution was extracted with ethyl acetate, and washed with 1 M HCl followed by brine. The organic layer was dried with $MgSO_4$ and filtered. Solvent was removed by rotary evaporation to afford a yellow oil. This oil was purified by column chromatography, using Silica Gel 60 Å with a gradient of dichloromethane/methanol (100/0 to 50/50 v/v). Drying under vacuum afforded an off-white solid (Boc-L-3-iodotyrosine) at 60% yield.

Production of PyTyr: Under a nitrogen atmosphere, Boc-L-3-iodotyrosine (2.0 g, 4.93 mmol) was added to a 100 mL schlenk flask, along with Cs_2CO_3 (3.21 g, 9.86 mmol), CuI (0.47g, 2.46 mmol), pyrazole (0.34 g, 4.93 mmol) and dry DMF (25 mL). The reaction was heated to 180˚C and stirred for 18 hours. After cooling, the reaction mixture was extracted with ethyl acetate and washed with 2 M HCl. The aqueous phase was concentrated by rotary evaporation to yield a brown solid. This solid was then purified by HPLC (Atlantis T3 C18 column 186003708, 40 mL/min flow rate, 0-60% acetonitrile in water over the course of 40 minutes) to a white solid at 15% yield.

3.5.2 Protein Expression and Purification

DNA coding for CFP was cloned into pET28 and co-transformed into *E. coli* BL21(DE3) cells with pEV01 containing PyTyrRS under control of an arabinose promoter [11]. A single colony was picked and placed in a 5 mL starter culture of Luria-Bertani (LB) broth with kanamycin (250 ug/mL) and chloroamphenicol (170 ug/mL). After overnight growth at 37˚C, 1 mL of the starter culture was aliquoted into 200mL of LB supplemented with the same antibiotics as above, as well as 30 mg of PyTyr. Cells were grown at 37˚C until A_{600} = 0.6, at which time they were induced with IPTG (1 mM final concentration) and arabinose (400 mg). The cells were grown at 37˚C overnight before being pelleted. The cell pellet was resuspended in 20 mL of lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, and 5 mM imidazole), followed by sonication to lyse the cells. The lysate was then centrifuged, and the supernatant

applied to Ni^{2+}-NTA affinity resin. Non-specific binding was removed by washing with a buffer of 50 mM HEPES pH 7.5, 500 mM NaCl, and 20 mM imidazole (wash buffer). A solution of 5 mg MTSSL [1-oxyl-2,2,5,5-tetramethylpyrroline-3-(methyl)-methanethiosulfate], 150 uL acetonitrile and 5 mL wash buffer was added to the resin for labeling C208 with the nitroxide side chain known as "R1". This solution was incubated with the resin for 5 hours at room temperature and then 7 hours at 4˚C. The labeled protein was then eluted with a buffer of 50 mM HEPES pH 7.5, 500 mM NaCl and 200 mM imidazole. The His_6-tag was removed by overnight incubation at room temperature with thrombin (200 ug) and the protein was further purified by size-exclusion chromatography (Superdex 200 10/300 analytical column). Cu^{2+} was added by direct addition of $CuSO_4$ (1.5 equivalents) and incubation at 4˚C overnight. The protein was then buffer exchanged into 50 mM Tris-HCl pH 7.5, 150 mM NaCl, and 30% glycerol. This process yielded 830 ug of CFP151PyTyr, which was concentrated to 150 uM and flash-frozen for PDS measurements. *T. maritima* CheY41PyTyr was produced identically to that described above, except that Glu41 was mutated to TAG with QuikChange (Stratagene, La Jolla, CA) mutagenesis and spin labeling did not take place. This process yielded 510 ug of CheY41PyTyr which was concentrated to 600 uM, incubated with other chemotaxis proteins as previously described, and then flash frozen for PDS measurements. *T. maritima* S76C and E387C CheA was expressed and purified as previously described [24]. Spin labeling followed as above.

3.5.3 Pulsed-Dipolar ESR Spectroscopy Measurements

For CFP151PyTyr, Cu-NO PDS experiments were carried out on a home-built Ku-band ESR spectrometer operating at 17.0-17.6 GHz [15] with a four-pulse double electron electron resonance (DEER) pulse sequence. Samples were 150 uM, and the data was collected at 15-20 K for 8-12 hours. Nitroxide spins were pumped at a frequency corresponding to 6180 G and Cu^{2+} spins were detected at that corresponding to 6100 G. For CheY41PyTyr and CheA76R1, Cu-NO DEER time-domain traces were collected in the same manner as for CFP151PyTyr. NO-NO DEER data was collected on the same spectrometer at 20-30 K for 8-12 hours, with standard pulse sequences [5]. Very weak Cu-Cu signals were observed, but these were not strong enough to yield usable distance distributions. Cf. section 3.3.2.

REFERENCES

[1] Borbat, P. P. and Freed, J. H. 2007. Measuring distances by pulsed dipolar ESR spectroscopy: Spin-labeled histidine kinases. *Method Enzymol.* 423: 52-116.

[2] Jeschke, G. 2012. DEER Distance Measurements on Proteins. *Annu Rev Phys Chem.* 63: 419-446.

[3] Crane, B. R., Borbat, P. P., and Freed, J. H. 2014. Defining Protein Complexes that Mediate Bacterial Chemotaxis by Pulsed Dipolar ESR Spectroscopy. *Biophys J.* 106: 685A-685A.

[4] Hubbell, W. L., Gross, A., Langen, R., and Lietzow, M. A. 1998. Recent advances in site-directed spin labeling of proteins. *Curr Opin Struc Biol.* 8: 649-656.

[5] 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.

[6] Jun, S., Becker, J. S., Yonkunas, M., Coalson, R., and Saxena, S. 2006. Unfolding of alanine-based peptides using electron spin resonance distance measurements. *Biochemistry.* 45: 11666-11673.

[7] Astashkin, A. V., Rajapakshe, A., Cornelison, M. J., Johnson-Winters, K., and Enemark, J. H. 2012. Determination of the Distance between the Mo(V) and Fe(III) Heme Centers of Wild Type Human Sulfite Oxidase by Pulsed EPR Spectroscopy. *J Phys Chem B.* 116: 1942-1950.

[8] 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 Edit.* 52: 1990-1993.

[9] Yang, Z. Y., Kise, D., and Saxena, S. 2010. An Approach towards the Measurement of Nanometer Range Distances Based on Cu^{2+} Ions and ESR. *J Phys Chem B.* 114: 6165-6174.

[10] 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. *Angew Chem Int Edit*. 54: 6330-6334.

[11] Liu, X. H., Li, J. S., Dong, J. S., Hu, C., Gong, W. M., and Wang, J. Y. 2012. Genetic Incorporation of a Metal-Chelating Amino Acid as a Probe for Protein Electron Transfer. *Angew Chem Int Edit*. 51: 10261-10265.

[12] Georgieva, E.R., Borbat, P.P., Norman, H.D., and Freed, J.H. 2015. Mechanism of influenza A M2 transmembrane domain assembly in lipid membranes. *Scientific Reports*. 5: 11757-11769.

[13] Wang, L., Brock, A., Herberich, B., and Schultz, P. G. 2001. Expanding the genetic code of Escherichia coli. *Science*. 292: 498-500.

[14] Hein, R. and Tsien, R. Y. 1996. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. 6: 178-182.

[15] Borbat, P. P., Crepeau, R. H., and Freed, J. H. 1997. Multifrequency two-dimensional Fourier transform ESR: An X/Ku-band spectrometer. *J Magn Reson*. 127: 155-167.

[16] Chiang, Y. W., Borbat, P. P., and Freed, J. H. 2005. The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J Magn Reson*. 172: 279-295.

[17] Chiang, Y. W., Borbat, P. P., and Freed, J. H. 2005. Maximum entropy: A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *J Magn Reson*. 177: 184-196.

[18] Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimousin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. J., and Royant, A. 2012. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat Commun*. 3: 751-760.

[19] Falke, J.J., Bass, R.B., Butler, S.L., Chervitz, S.A., and Danielson, M.A. 1997. The two-component signaling pathway of bacterial chemotaxis: A molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Ann Rev Cell Dev Biol*. 13: 457-512.

[20] Hazelbauer, G. L., Falke, J. J., and Parkinson, J. S. 2008. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci*. 33: 9-19.

[21] Kentner, D. and Sourjik, V. 2006. Spatial organization of the bacterial chemotaxis system. *Curr Opin Microbiol*. 9: 619-624.

[22] Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. 1999. Structure of CheA, a signal-transducing histidine kinase. *Cell*. 96: 131-141.

[23] Park, S. Y., Quezada, C. M., Bilwes, A. M., and Crane, B. R. 2004. Subunit exchange by CheA histidine kinases from the mesophile Escherichia coli and the thermophile Thermotoga maritima. *Biochemistry*. 43: 2228-2240.

[24] Quezada, C. M., Gradinaru, C., Simon, M. I., Bilwes, A. M., and Crane, B. R. 2004. Helical shifts generate two distinct conformers in the atomic resolution structure of the CheA phosphotransferase domain from Thermotoga maritima. *J Mol Biol*. 341: 1283-1294.

[25] Mo, G. Y., Zhou, H. J., Kawamura, T., and Dahlquist, F. W. 2012. Solution Structure of a Complex of the Histidine Autokinase CheA with Its Substrate CheY. *Biochemistry*. 51: 3786-3798.

[26] Bhatnagar, J., Borbat, P. P., Pollard, A. M., Bilwes, A. M., Freed, J. H., and Crane, B. R. 2010. Structure of the Ternary Complex Formed by a Chemotaxis Receptor Signaling Domain, the CheA Histidine Kinase, and the Coupling Protein CheW As Determined by Pulsed Dipolar ESR Spectroscopy. *Biochemistry*. 49: 3824-3841.

[27] Park, S. and Crane, B. R. 2011. Crystallization and preliminary X-ray crystallographic analysis of CheW from Thermotoga maritima: a coupling protein of CheA and the chemotaxis receptor. *Acta Crystallogr F*. 67: 504-506.

[28] Park, S. Y., Beel, B. D., Simon, M. I., Bilwes, A. M., and Crane, B. R. 2004. In different organisms, the mode of interaction between two signaling proteins is not necessarily conserved. *P Natl Acad Sci USA*. 101: 11646-11651.

[29] Briegel, A., Li, X., Bilwes, A. M., Hughes, K. T., Jensen, G. J., and Crane, B. R. 2012. Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. *P Natl Acad Sci USA*. 109: 3766-3771.

[30] Greenswag, A.R., Muok, A.M., Li, X. and Crane, B.R. 2015. Conformational Transitions that Enable Histidine Kinase Autophosphorylation and Receptor Array Integration. *J Mol Biol.* 427: 3890-3907.

Chapter 4

TOWARDS ELUCIDATING THE STRUCTURAL AND BIOCHEMICAL PROPERTIES OF THE *DROSOPHILA MELANOGASTER* CIRCADIAN CLOCK PROTEIN PERIOD

4.1 Abstract

The protein Period is an essential component of the *Drosophila melanogaster* circadian clock; along with binding partner Timeless, it forms the main repressor complex in the clock's negative feedback loop. Despite extensive study, there is only structural information available for a relatively small segment of Period that encompasses its ordered PAS domains. However, key functional regions of Period lie outside the PAS modules. In these uncharacterized stretches, point mutations are known to affect the rhythm of the clock, but the underlying mechanism responsible is unclear. Here, we are able to express and purify a construct consisting of the first 700 amino acids (PER 1-700). *In vitro* biochemical assays on PER 1-700 show marked conformational differences between the wild-type protein and mutants thereof which alter circadian rhythms of flies.

4.2 Introduction

4.2.1 The Drosophila Circadian Clock

Many organisms, from cyanobacteria to fruit flies to humans have defined circadian rhythms [1-3]. These circadian rhythms place many biological functions in an approximately 24 hour period, coinciding with the diurnal cycle. While the clock of an organism is not exclusively regulated by the daily cycle of light and darkness, light and other factors such as temperature (known as zeitgebers) do help to entrain the clock and keep its 24 hour rhythm [4, 5].

In general, the eukaryotic circadian clock consists of two main regulatory systems that work in opposition: the positive elements and the negative elements (Figure 4.1). Associated with the positive elements are two types of genes: clock genes and clock-controlled genes. When the positive elements are in abundance, they initiate transcription of the clock genes. These clock genes are then translated into proteins which are the negative elements. The negative elements in turn repress the production of the positive elements, by blocking their transcription. The level of positive elements also modulates the activity of the many clock-controlled genes of the organism. This transcriptional-translational feedback mechanism is the basis for the circadian clock in its simplest form.

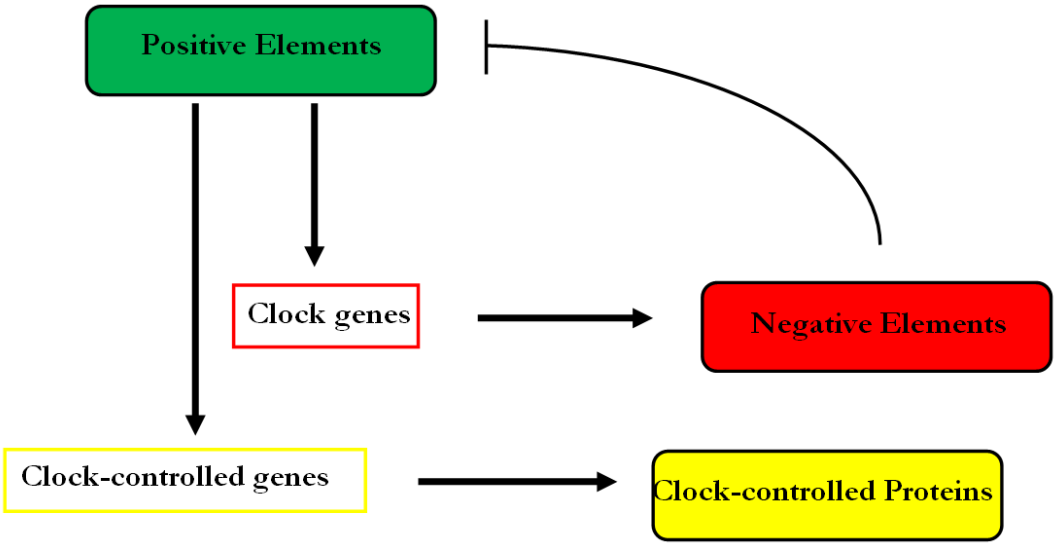

Figure 4.1: The basic feedback loop which governs the circadian clock. Positive elements initiate transcription of clock and clock-controlled genes. The clock genes are then translated into negative elements, which then repress the expression of the positive elements. Figured adapted from [6, 7].

Drosophila melanogaster, the common fruit fly, is (along with the fungus *Neurospora crassa*) the main model organism for studying the circadian clock. Although the mechanism of positive and negative elements outlined above seems fairly simple, there are many components to the positive and negative limbs of the *Drosophila* clock. These components often have complex interactions which are poorly or incompletely understood, especially from a detailed structural standpoint. The positive elements of the *Drosophila* clock are the proteins Clock (CLK) and Cycle (CLC) [8]. Clock and

Cycle then promote the production of the two main negative elements in the clock: Period (PER) and Timeless (TIM) [9]. These two proteins, along with the kinase Doubletime (DBT), form a PER-TIM-DBT complex, which then interacts with a complex of Clock and Cycle [9-12]. Doubletime phosphorylates Clock, which then prevents Clock from binding DNA. This causes the production of PER and TIM to be repressed, which completes the negative feedback loop [9].

Doubletime is not the only active kinase in the *Drosophila* clock. In fact, phosphorylation is one of the main signaling mechanisms of the core clock components. Two other kinases, Shaggy and Casein Kinase 2 (CK2) phosphorylate Timeless and Period, which gates their entry into the nucleus, allowing them to complex with Clock and Cycle [13, 14]. Furthermore, phosphorylated Timeless is ubiquinated by a protein called Jetlag (JET) and this ubiquitination signals for Timeless to be degraded by the proteasome [15-17]. Once Timeless is degraded, the PER-TIM complex no longer forms. Timeless protects Period from phosphorylation, so in its absence, PER is phosphorylated by Doubletime [18, 19]. This signals for PER to be ubiquitinated by the protein Slimb, which in turn causes it to be degraded [20, 21].

Light-dependent regulation of the *Drosophila* clock depends on the protein Cryptochrome (CRY), which senses blue light through its flavin adenine dinucleotide (FAD) cofactor. Light causes the FAD to be reduced, setting in motion a series of

conformational changes in CRY, which ultimately lead to its light-dependent binding of Timeless [22, 23]. When CRY binds TIM, TIM can no longer complex with PER, and as was noted above, PER is then degraded through phosphorylation by Doubletime followed by ubiquitination by Slimb. This allows the clock to be reset by light.

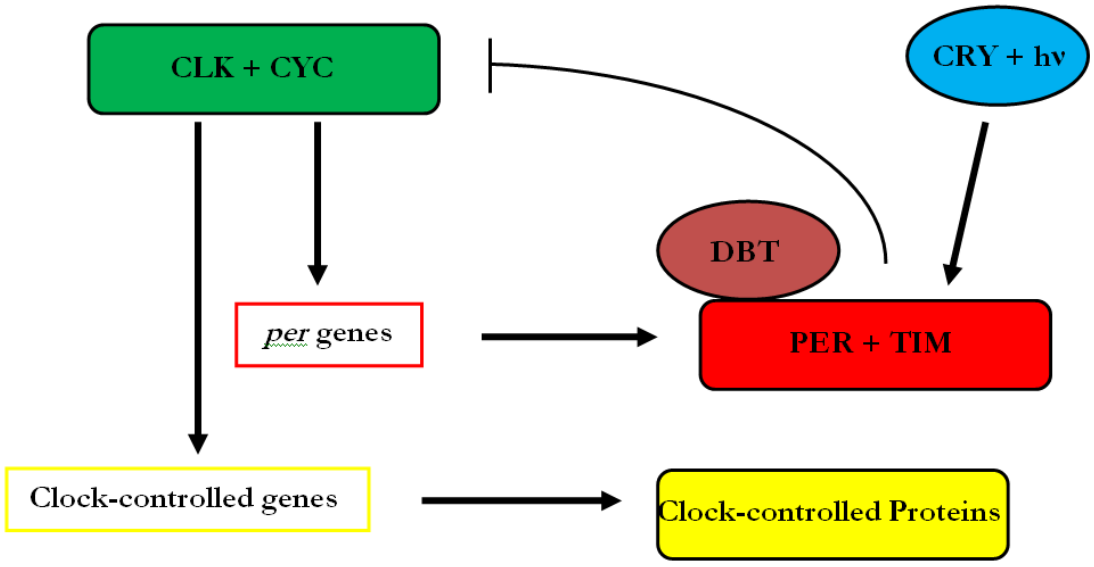

Figure 4.2: The basic feedback loop of the *Drosophila* circadian clock. Clock (CLK) and Cycle (CYC) are the positive elements, while Period (PER) and Timeless (TIM) are the negative elements. The kinase Doubletime plays an important role in circadian regulation and the photo-sensor Cryptochrome (CRY) helps to entrain the clock based on light.

4.2.2 Period

Period is one of the main negative elements (along with Timeless) in the *Drosophila* clock, repressing Clock and Cycle in the nucleus. However, relatively little is

known about the structure of this 1224 residue protein. Much of the current information about PER comes from *in vivo* studies aimed at uncovering the interactions between the various members of the clock. The only region of structural homology in Period is the PAS AB region, residues 240-454. PAS domains are conserved motifs which occur in both prokaryotic and eukaryotic signaling proteins. PAS domains are named for the 3 proteins in which they were first discovered: Period, Aryl Hydrocarbon Receptor Nuclear Translocator, and Single-Minded. Often, they sense light, oxygen, or voltage through a bound flavin cofactor, in which case they are referred to as LOV domains. PAS domains are characterized by 5 antiparallel beta sheets, followed by several alpha helices, which often form N or C terminal caps, and almost always have the ability to dimerize. For some PAS domains, especially LOV domains, the oligomeric state is a signaling mechanism and changes in the presence of stimuli [24, 25].

The PAS AB domain in PER is the only region that has been structurally characterized, with Yildiz et. al. solving a crystal structure of residues 232-599, and then with King et. al. solving a crystal structure of residues 229-575 [26, 27]. The PAS domain is known to be important for dimerization, but its functionality beyond that is unknown. Often, PAS domains are integral parts of sensory proteins by way of a bound cofactor (LOV domains). For PER, this is not the case, as there is no bound cofactor. The PAS domains have been implicated as being important for Period's

interaction with Timeless, but the structural nature of this interaction is also unknown [28].

The PER PAS domains mediate dimerization of the larger PER protein. Residues 229-575 were crystallographically characterized to be a dimer [26], as is often the case with PAS domains. *In vivo* and *in vitro* studies (including in this thesis) both show that PER is a dimer in solution [29, 30]. The interesting aspect about the PAS domain crystal structure is that it shows a head-to-tail dimer, with a helical arm wrapped around the opposite monomer (Figure 4.3) [26]. This head-to-tail directionality may be important for the function of the protein, specifically an interaction between the N and C termini of opposite monomers. (See section 4.4 for further discussion.)

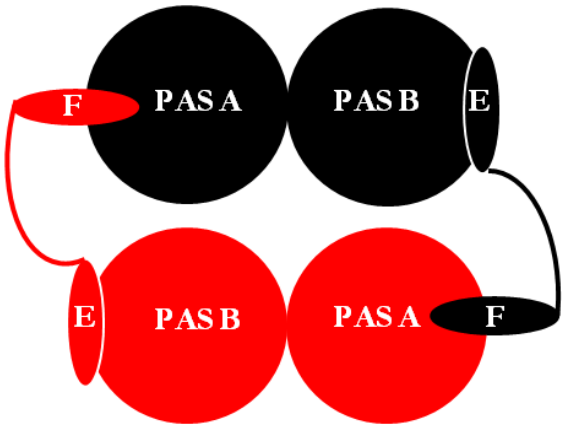

Figure 4.3: Schematic of the head-to-tail dimer of the PER PAS AB region.

Other regions on PER have been shown to be critical for association with its numerous binding partners. Residues 764-1034 interact with the Clock-Cycle complex [31], allowing Clock to be phosphorylated by Doubletime and inhibiting the ability of Clock to bind DNA, including the promoter regions of genes that code for PER and TIM [9]. Residues 765-800 have been implicated in binding Doubletime [21]. Interestingly, these are not the only ranges of residues which are phosphorylated by Doubletime. Many of the first 100 residues in PER, and specifically Ser47, have been shown to be phosphorylation targets of Doubletime [21]. A cluster of residues between 585 and 630 are also phosphorylated and significantly change the rhythmicity of the clock when modified. Mutation of serine 589 to asparagine (S589N) was the first mutation shown to change the length of the clock, shortening its cycle to 19 hours [4]. Mimicking phosphorylation at residue 589 by mutation to an aspartate has the same effect (S589D). This mutation also causes an increased level of repression of the Clock-Cycle complex [32]. *In vivo* studies show that a series of phospho-mimic mutations (generally Ser-to-Asp) in this region cause significantly different behavior in flies depending upon the specific change. For these variants, the level of repressor activity and the length of the circadian cycle are well correlated. For example, mutating serine 629 to aspartate (S629D) causes the opposite behavior as S589D: no repressor activity and a lengthened circadian cycle [32]. Doubletime also phosphorylates a series of residues, S604, S607, T610, and S613, which are usually phosphorylated together as a "cluster" [21]. Here, this cluster is treated as one group,

and along with S589D, continues to have a short circadian cycle, when the Asp substitutions are introduced [32].

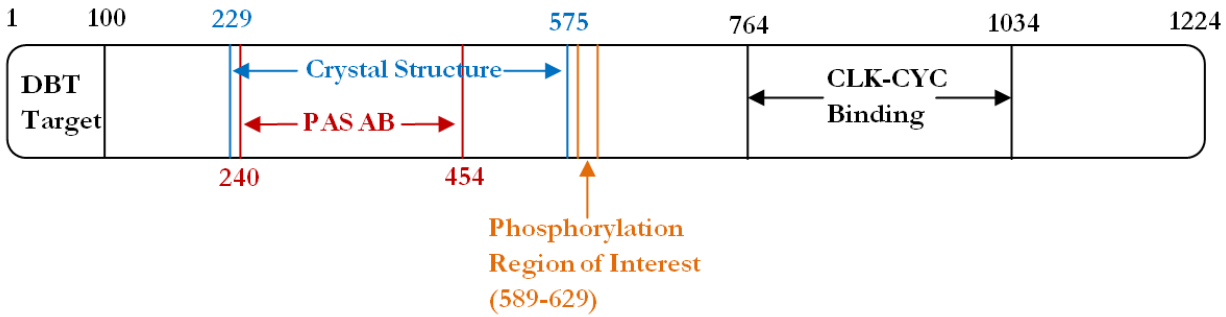

Figure 4.4: PER Features.

Although the phenotypes of the phospho-mutations have been established, there has been little work on their impact on the structural properties of the protein. Here, we have attempted to characterize structural changes caused by these mutations. This work necessitated development of a method to purify a soluble, non-aggregated form of the first 700 amino acids of Period (PER 1-700). PER1-700 and phospho-mimic variants thereof were then conformationally characterized using a number of biochemical and biophysical methods.

4.3 Results

4.3.1 Purification of WT PER 1-700

Previous work in this lab [35] demonstrated the feasibility of expressing and purifying a construct containing the first 700 amino acids of the wild-type Period protein (PER 1-700). Expression at 17˚C using pLysS cells and a GST tag for purification (see section 4.5) yielded a substantial amount of protein. However, size-exclusion chromatography (Figure 4.4) showed that the protein was forming high molecular weight oligomers almost exclusively, most likely as a result of aggregation.

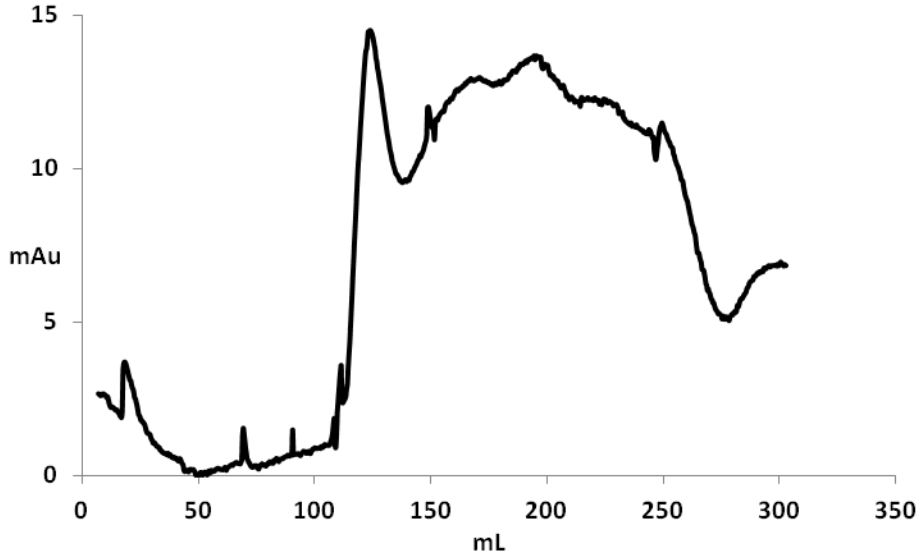

Figure 4.5: SEC trace of aggregated PER 1-700.

It had been shown previously that the N-terminus of PER (residues 1-87) binds RNA, and that this may be a cause of aggregation [30]. A DNA gel confirmed that PER was indeed binding nucleic acids, either RNA or DNA (Figure 4.6). Upon addition of RNase, the nucleic acids were no longer present in the gel, thus corroborating previous results (Figure 4.6).

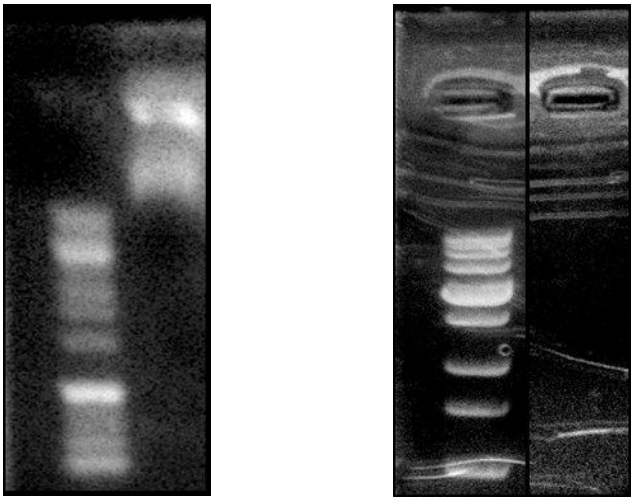

Figure 4.6: DNA gel of PER 1-700 before (left) and after (right) the addition of 328 units of RNase for 3 hours. Bands are visualized using ethidium bromide.

The addition of RNase also allowed for purification of soluble, non-aggregated PER 1-700. Previous protocols (see [33] and section 4.5) were followed up until elution of PER 1-700 from GST beads. Upon elution, 16 uL (328 units) of RNase was added and this solution was rocked at 4˚C for 2-4 hours. The protein was then concentrated to 5 mL and loaded onto a Superdex 200 size-exclusion column. The

resulting trace showed the same aggregation peak that had previously been seen, but also a non-aggregated peak which eluted at the volume corresponding to the molecular weight of a PER 1-700 dimer (Fig. 4.7). This was confirmed using size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS, Fig. 4.8).

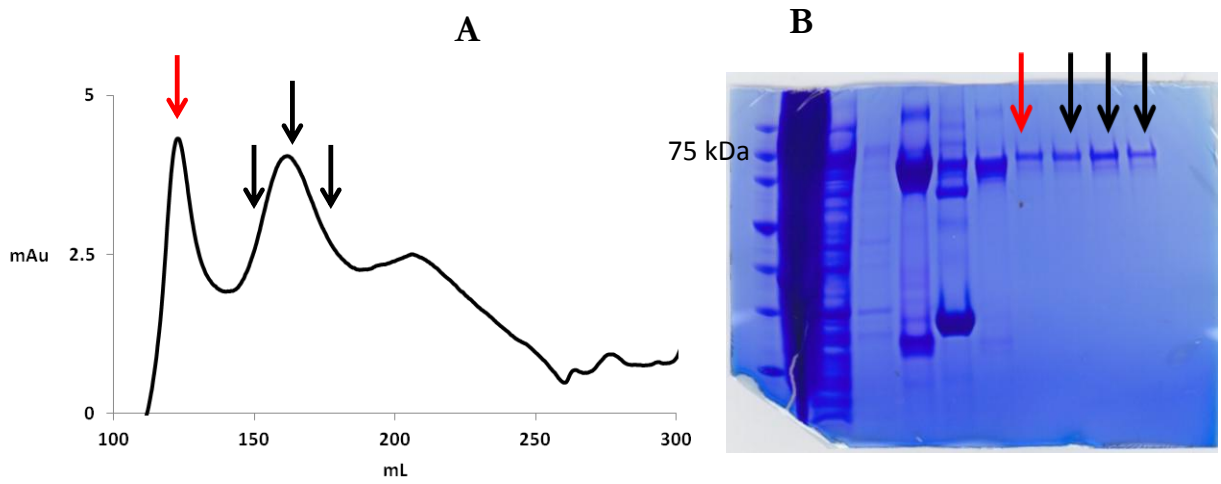

Figure 4.7: (A) Size-exclusion chromatograph of PER 1-700 after treatment with RNase. The first major peak (red arrow) is aggregated protein, while the second peak (black arrows) is soluble protein. (B) A gel of the various purification steps of PER 1-700. Lanes: markers, GST bead flow through, GST bead flow through diluted, GST bead wash, Elution from GST beads, GST beads, after addition of RNase, 4 SEC fractions taken from the corresponding arrows.

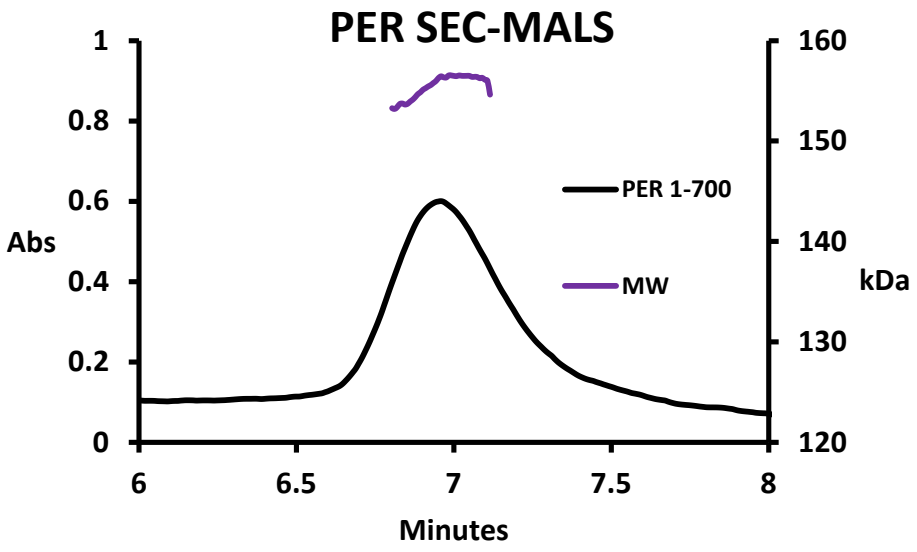

Figure 4.8: SEC-MALS trace of PER 1-700. The molecular weight of ~155 kDa confirms the dimeric nature of PER 1-700.

4.3.2 Characterization of PER 1-700 Mutants which Mimic Phosphorylation

As noted previously, several residues in the first 700 amino acids of PER have been identified as targets for phosphorylation. Serine residues 589, 604, 607, 613, and 629, along with threonine 610, are phosphorylated by the kinase Doubletime, which eventually leads to PER degradation [18, 19]. Mutation of these residues to aspartate in various combinations mimics phosphorylation and causes behavioral changes in fruit flies [32]. The table below lists the various phosphor-mimic mutations and their effect on the behavior of fruit flies. Residues 589 and 629 stand on their own, and 604, 607, 610 and 613 are "clustered" together.

Table 4.1: PER 1-700 variants and their *in vivo* effects

Mutation(s)	Denoted As	Behavior Effect
Wild-type	PER1	None
S589D, S604D, S607D, T610D, S613D, S629D	PER2	Unknown
S589D, S604D, S607D, T610D, S613D	PER5	Stable, super-repressor activity, shortened circadian cycle
S629D	PER6	Stable, no repressor activity, lengthened circadian cycle

Finding the structural and biochemical effects that these mutations impart on PER is key to understanding how they change the *in vivo* behavior of fruit flies.

While a crystal structure would provide important structural information, biochemical assays can, and do, show significant differences between PER mutants (see Appendix 2 for PER crystallization efforts). During purification, the amount of bound RNA was measured for each mutant by DNA gel (Figure 4.9).

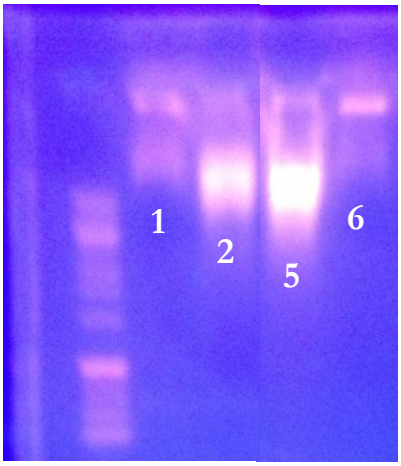

Figure 4.9: (A) The relative amount of RNA bound by each PER mutant during purification was measured by DNA gel.

From Fig. 4.9, the mutant which has super-repressor activity, PER5, binds much more RNA as compared to the WT and PER6. There is a direct correlation between repressor activity and RNA binding: the greater the repressor activity, the more RNA bound by the protein. (Note: PER2's activity is unknown). There is another, more subtle implication from these results, which is that the N and C termini of PER 1-700 are somehow interacting. Previous work has shown that the RNA is bound in the first 87 residues of PER [33]. These results show that mutations near the C terminus of this construct modulate the amount of RNA bound at the N-terminus. This could either be due to long-range effects or a direct interaction between the N and C termini. In the crystal structure of the PAS domains of PER, they are arranged in a head-to-tail alignment, which could explain the interactions of the N and C termini. If the entire 700 amino acid construct is in a head-to-tail dimeric arrangement, the N-terminus of one monomer could be located quite close to the C-

terminal end of the other monomer, allowing for this RNA binding modulation. Further evidence for this model comes from Heather King's thesis [33], which shows that when the N-terminal end is truncated, the C-terminus becomes unstable, supporting the idea of a direct interaction between these segments.

Along with having different RNA binding properties, the various mutants of PER 1-700 also have different aggregation properties. Despite the exact same expression and purification steps, there is a large variance in the amount of aggregation seen on the size-exclusion column at the last step of purification. Like RNA binding, the trend tracks with the behavioral response. As the repressor activity of the variant increases, so does the stability of the construct. PER5 and PER2 have a large soluble, non-aggregated peak and a smaller aggregation peak. The level of the two PER1 peaks is approximately equal, and PER6 is almost entirely aggregated, only yielding a very minor amount of usable protein (Figure 4.10).

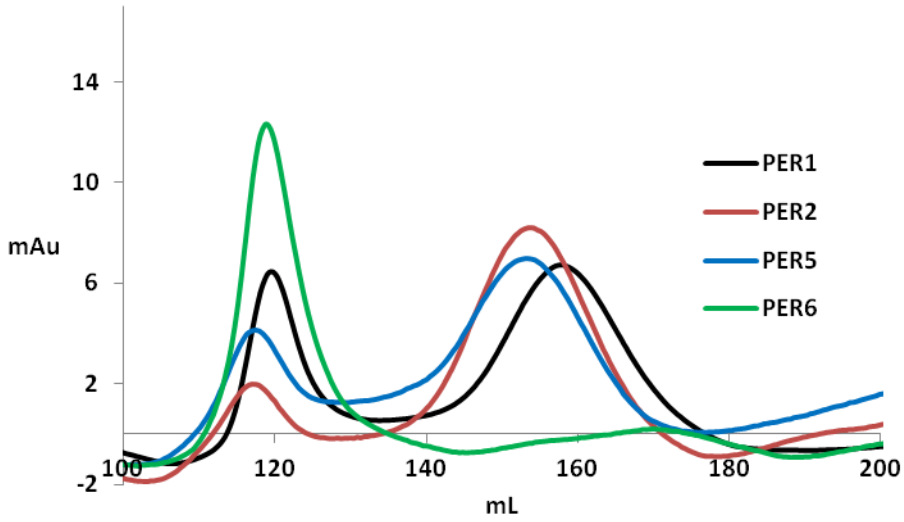

Figure 4.10: SEC traces of PER1, 2, 5, and 6. The shift in the volume of the peaks is most likely due to column repacking, as the samples were run at different times. (See Figure 4.11.)

Despite significant differences during purification, these proteins behave remarkably similarly after purification. SEC-MALS was run on PER2 and PER5 as well as the WT, to investigate if the mutations changed the oligomeric state of PER. The traces in Figure 4.11 show that each variant elutes at the same volume, and the light scattering confirms the dimeric nature of these proteins. SEC-SAXS was also employed, to determine if mimicking phosphorylation causes any large scale conformational changes to PER. From the Kratky Plot below (Figure 4.13), it appears as though the phosphomimics have the same general shape and structure as WT PER. A Kratky Plot, which plots the angle of scattering against the intensity of the

scattering times the square of the angle ($I*q^2$ vs. q), is used to measure the amount of folding vs. random coil structure in a sample.

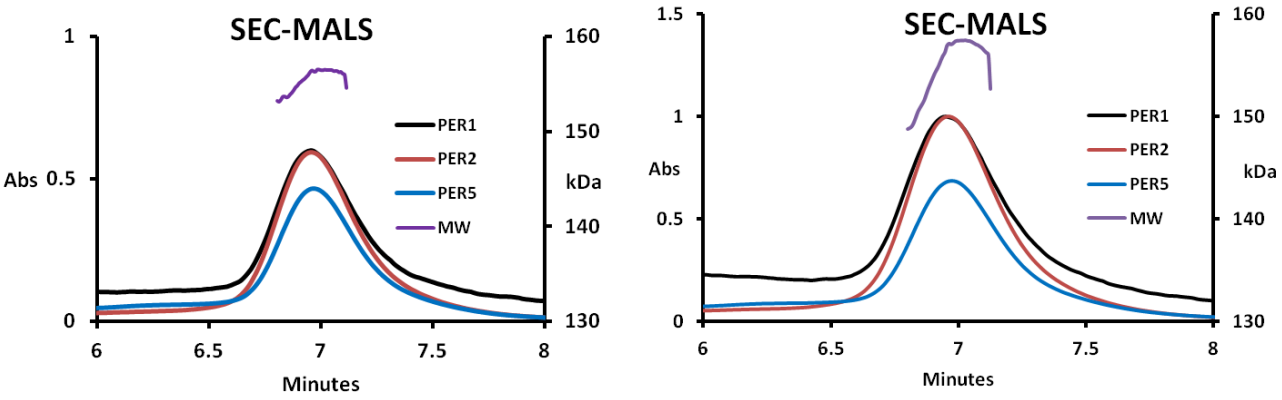

Fig. 4.11: SEC-MALS traces for PER1, PER2 and PER5 at 4 mg/mL (left) and 6 mg/mL (right). Each variant is a dimer in solution.

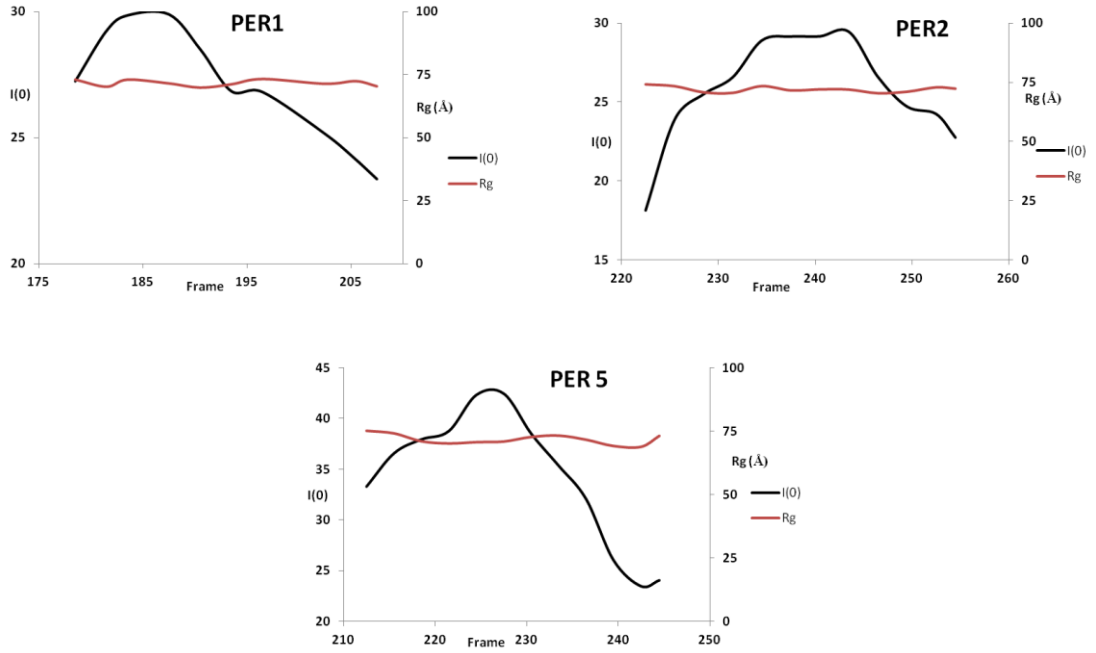

Fig 4.12: I(0) and Rg by frame of SAXS data collected.

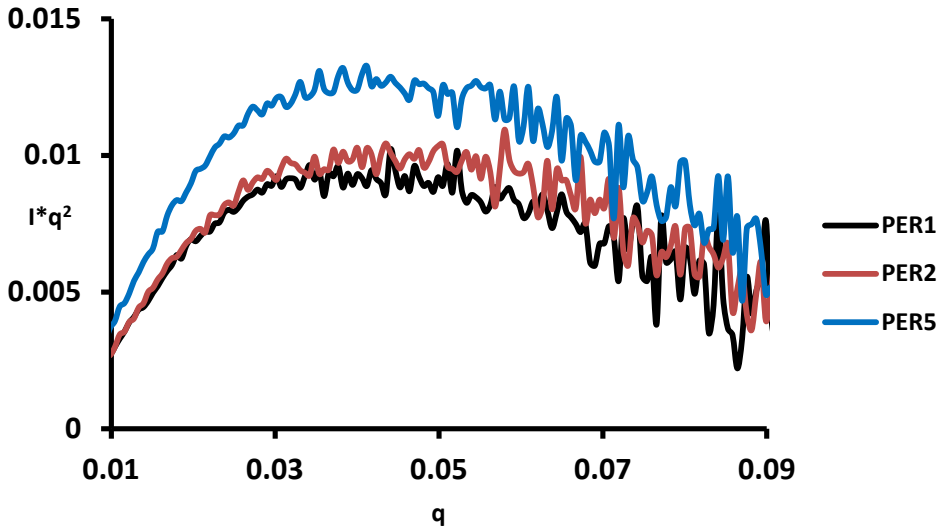

Fig 4.13: Kratky Plot of PER1, PER2 and PER5 generated from SAXS data. A Kratky Plot assesses the "foldedness" of a protein. In this case, there is no observable difference between the 3 variants of PER.

One assay that has proven useful in the past at detecting modest conformational changes in large proteins has been proteolytic cleavage with the enzyme trypsin [23]. Trypsin is a serine protease which cleaves peptide bonds on the C-terminal side of positively charged residues lysine or arginine. In the first 700 amino acids of PER, there are 40 lysines and 36 arginine residues, affording many of sites for trypsin to cleave the protein. A gel is run on the digested protein, and bands can be analyzed by mass spectrometry. In general, trypsin will cleave sites that are solvent

exposed before ones that are buried on the interior of the protein. By comparing the gel bands between PER variants, it is possible to determine differences in regions of the protein that are accessible. Furthermore, greater digestion by trypsin usually indicates more flexibility, as flexibility allows more time-averaged exposure of lysine and arginine residues.

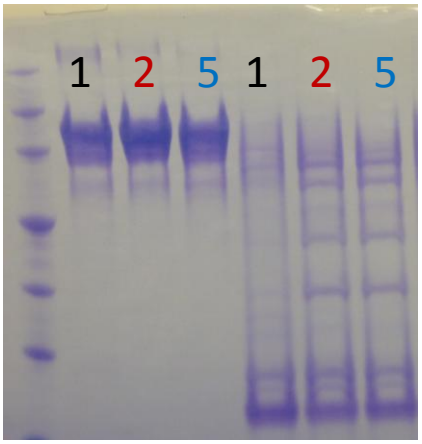

Fig. 4.14: Trypsin digest on PER1, PER2 and PER5. Left 3 lanes are undigested protein (control) and the right 3 lanes are the digested protein. Each sample contained 2 uL of 80 uM protein, digested for 75 seconds with 2 uL of 80 uM trypsin and quenched with 2 uL of 160 uM trypsin inhibitor and 10 uL of SDS loading dye.

It is clear that there are several intermediate bands in PER2 and PER5 which are either not present or very faint in PER1. The top bands also appear to be weaker in PER1, and the bottom band appears stronger. Relevant band intensities were quantified with ImageJ to confirm these observations (Figure 4.15).

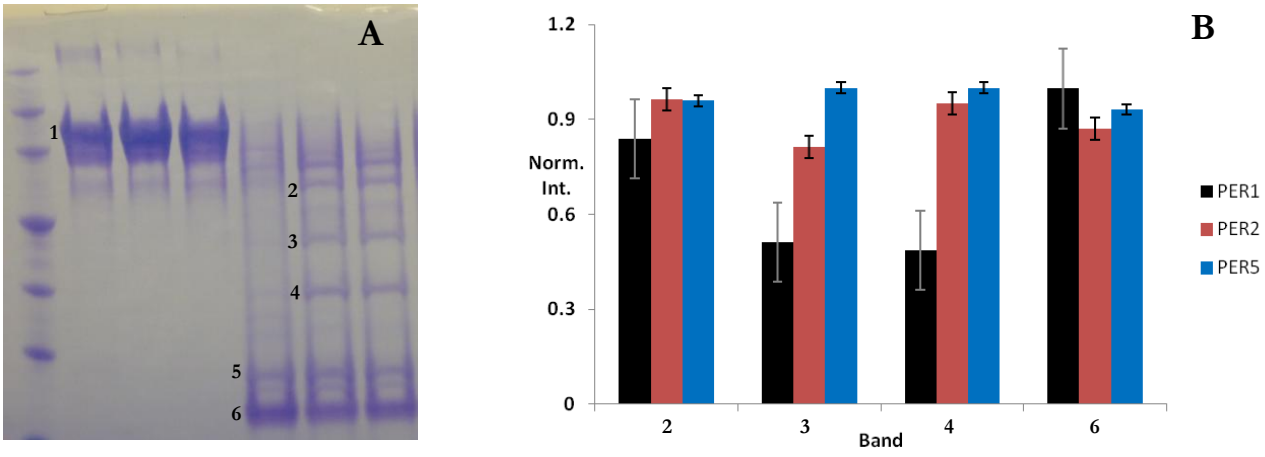

Figure 4.15: (A) Gel from Fig. 4.14, labeled with bands cut out and sent for sequence analysis by mass-spectrometry. (B) Relative intensities of several band comparing PER1, PER2, and PER5. PER2 and PER5 have much stronger intermediate bands, while PER1 has a stronger bottom band, suggesting greater conformational flexibility in PER1.

Major bands were analyzed by mass-spectrometry to determine the sequence of each fragment. The six bands labeled in Fig. 4.15 (A) were sent in, including the undigested protein (1) as a control. Table 4.2 lists the fragments from each band.

Table 4.2: PER Fragment Analysis by Mass Spectrometry

Band	Residues
1	1-700
2	L39-R649
3	G79-R587
4	Mixture of Fragments spanning G79-R587
5	Mixture of Fragments spanning G79-R587
6	Mixture of Fragments spanning G79-K605

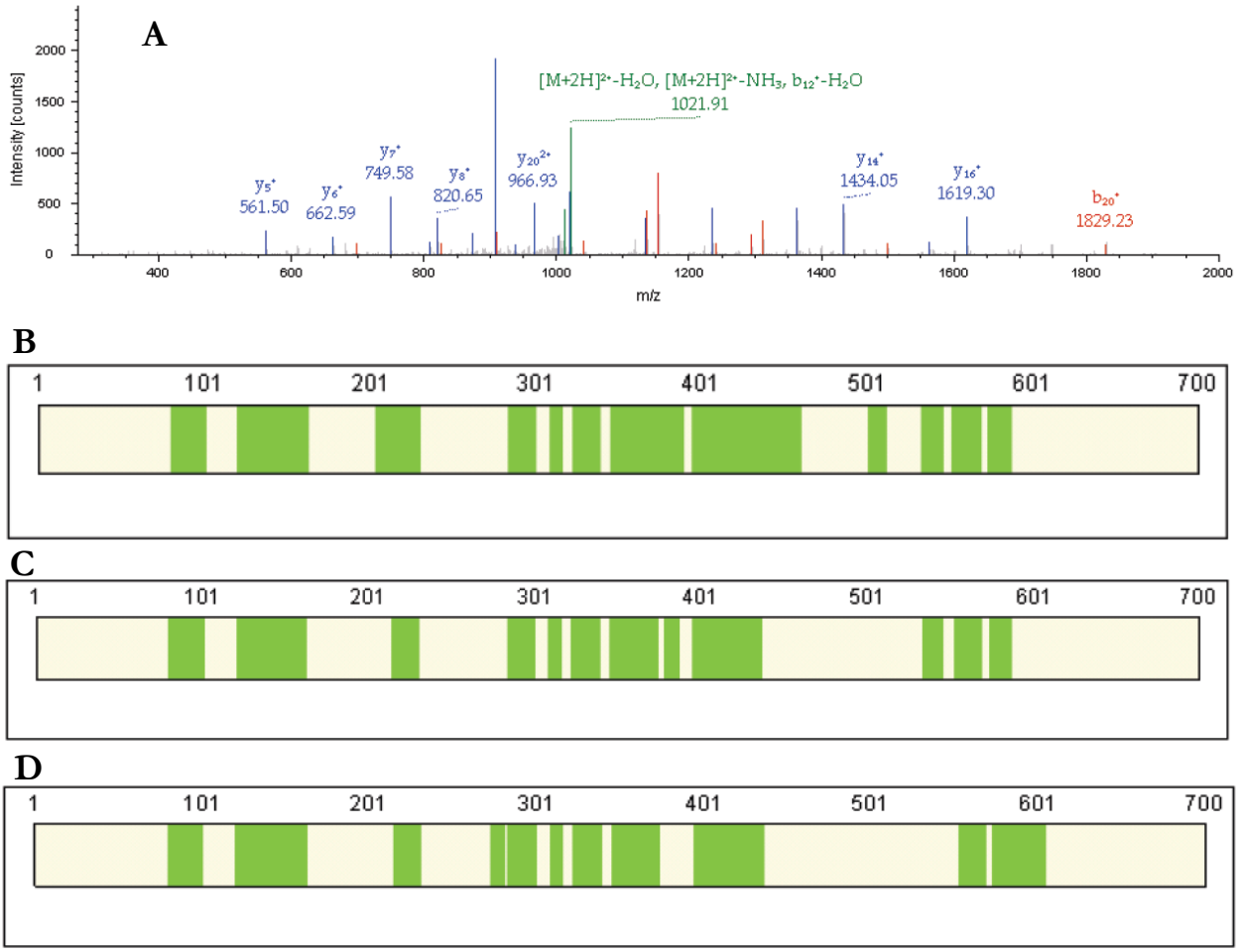

Figure 4.15: (A) Example of a mass-spectrum of 1 fragment of band #4. (B-D) Fragments identified from bands 4 (B), 5 (C), and 6 (D) are shown in green.

4.4 Discussion

The ability to purify WT PER 1-700 and several mutants which mimic *in vivo* phosphorylation represents a significant step forward in the ability to study Period structure, as well as the structural changes caused by phosphorylation. When analyzing all of the different experiments as a whole, several observations become apparent. First, no major structural rearrangements occur as a result of mimicking phosphorylation, even for the most extreme case, PER2, where 6 negative charges are

added to the protein matrix, and the calculated isoelectric point decreases by almost half of a pH unit (6.54 to 6.07) The SEC-MALS data shows that all variants are dimeric in nature, and SAXS confirms that no large conformational changes are occurring.

Secondly, there are subtle structural changes occurring and at least some of these changes are at the N-terminus, far away from the mutation sites, as evidenced by differential RNA binding in the first 87 residues during expression and purification. This may be happening in one of two ways: either through a series of small, long-range conformational changes propagating through the protein, or through a direct interaction between the N and C termini. While there is no direct evidence to support one case over the other, a direct interaction is feasible and the more straightforward explanation. The crystal structure of the PER PAS domains shows them forming in an antiparallel dimeric arrangement. If the entire 700 residue construct is an antiparallel dimer, this could position the N and C termini quite close together. The increased negative charge could then affect conformation at the N-terminus directly.

Figure 4.16: Possible arrangement of an antiparallel PER 1-700 dimer. The phosphor-mimic sites near the C-terminus (yellow P's) lead to conformational changes near the N-terminus (yellow arrows).

There also seems to be a correlation between the ability of the PER variant to repress the Clock-Cycle complex, the length of the circadian rhythm of the variant, and the stability of the protein *in vitro*. PER5, which shows the greatest repressor activity *in vivo* and lengthens the clock, is much less prone to aggregation during purification than either the WT or PER6, the mutant with no repressor activity and which shortens the clock. Furthermore, limited proteolysis by trypsin reveals that PER6 is more easily degraded than WT, and that the WT is more easily degraded than PER5, suggesting that there is increased flexibility and instability as repressor activity decreases. Exactly how the mutations affect the repressor activity is unknown as of yet, and is currently the subject of further investigation.

Table 4.3: Studied mutations of PER 1-700 and their characteristics

Mutation(s)	Phenotype	Oligomeric State	RNA Binding	Aggregation Propensity	Digestion Profile
WT	Normal	Dimer	Low	Medium	Normal
S589D, S604D, S607D, 610D, S613D, S629D	Super-repressor, short rhythm	Dimer	High	Low	Less
S589D, S604D, S607D, 610D, S613D	Super-repressor, short rhythm	Dimer	High	Low	Less
S629D	No repressor, long rhythm	Unknown	Low	High	More

4.5 Methods

4.5.1 Expression and Purification of PER 1-700

Drosophila Melanogaster PER 1-700 (and variants) DNA in the pGEX expression vector was provided by Michael Young at Rockefeller University. Proteins were expressed in *E. coli* BL21(DE3)pLysS cells. Cells were grown at 37˚C until they reached an OD_{600} of .5, at which time they were cooled to 17˚C and then induced with .4 mM IPTG. Additional growth for 18 hours at 17 ˚C was followed by harvesting by centrifugation. Cells were lysed chemically, using 60 mL (per 8L of pellet) of the following solution: 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 3 mM DTT, 4 uL RNase (82 Units, Ambion), 2 uL benzonase nuclease (500 units, Santa Cruz Biotechnology), 1 tablet Serine-Threonine Protease inhibitor (Thermo Scientific), and

5% v/v 10x BugBuster Solution (EMD Millipore). The solution was then passed through a syringe to shear DNA. It was then incubated at room temperature for 90 minutes. This was followed by centrifugation at 21,000 x g for 45 minutes at 4˚C. The supernatant was collected, and 2 uL RNase and 2 uL benzonase nuclease was added, as well as 2 uL of 25% poly-ethylene imine per 1 mL of supernatant. This solution was incubated at 4˚C for 20 minutes, during which time a white precipitate formed. The precipitate was pelleted by centrifugation at 12,000 x g for 20 minutes. The resulting supernatant was filtered through a .42 micron syringe filter, and incubated with glutathione beads overnight at 4˚C

The glutathione beads were then centrifuged at 2,100 rpm for 10 minutes. The supernatant was decanted, and the beads were washed twice with a buffer of 50 mM Tris-HCl pH 7.5, 500 mM NaCl, and 3 mM DTT. The beads were then incubated overnight at 4˚C with 15 mL of 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM DTT, 10% glycerol and 48 units of HRV-3C protease (Takara) to remove the GST tag. The following day, the cut protein was eluted from the beads using 30 mL of 50 mM Tris-HCl pH 7.5, 500 mM NaCl, and 3 mM DTT. RNase (15 uL) was added to the protein for 3 hours at 4˚C, after which the protein was concentrated to 5 mL and further purified on a Superdex 200 (Pharmacia) size-exclusion column.

4.5.2 SEC-MALS

Size-exclusion chromatography coupled with multi-angle light scattering (MALS) was used to study the molar mass of WT PER 1-700 and various mutants. Proteins (2-6 mg/mL) were run at room temperature on an SEC column (WTC050N5 - Wyatt) pre-equilibrated with 50mM Tris-HCl pH 7.5, 500mM NaCl and 3 mM DTT. Analysis and molecular weight determination was carried out with Wyatt technologies ASTRA. Bovine serum albumin (Sigma) was used as a control for data quality.

4.5.3 SEC-SAXS

Size-exclusion chromatography coupled with small-angle X-ray scattering (SAXS) was used to study the conformation of WT PER 1-700 and various mutants. Data was collected at the G1 line of CHESS, using a Pilatus 100K detector [34]. PER1, PER2 and PER5 samples were concentrated to 8-10 mg/mL in 50 mM Tris pH 7.5, 500 mM NaCl and 3 mM DTT and 100uL of sample was loaded on to an analytical size-exclusion column (15/150 Superdex 200). Continuous elution from the column was exposed to x-rays. Buffer blanks were collected for background subtraction. RAW and PRIMUS [35, 36] were used for data analysis and generation of Kratky plots.

4.5.4 Trypsin Digestion Assay and Mass-Spectral Analysis

For the trypsin digestion assay, the following method was used: 2 uL of 80 uM trypsin was added to 2 uL of 80 uM PER and allowed to react for 75 seconds. The reaction was then quenched with 2 uL of 160 uM trypsin inhibitor and 10 uL SDS gel loading buffer. The samples were then heated to 90˚C and run on a gel.

Mass spectral analysis was done as previously described [23]. Briefly, gel bands were cut out and then dissolved in a denaturing solution of 6 M guanidine-HCl and 50 mM Tris-HCl pH 8.0 at 56˚C for 45 minutes. The samples were then reduced with 10 mM TCEP and alkylated with 10 mM iodoacetimide for 1 hour at room temperature. They were then further digested with trypsin (1:25 wt:wt) for 16 hours at 37˚C, at which time then were quenched with .5% (vol/vol) TFA, dried by evaporation, and redissolved in a solution of .1% formic acid and 2% acetonitrile (both vol/vol). Electrospray ionization-MS/MS analysis was done using an UltiMate3000 MDLC system (Dionex) coupled with an LTQ-Orbitrap Velos (Thermo-Fisher Scientific) mass spectrometer equipped with "plug and play" nano-ion source device (CorSolutions LLC). An aliquot of digested peptide (1.0 μL) was injected onto a PepMap C18 trap column (5 μm, 300 μm × 5 mm; Dionex) at a flow rate of 20 μL/min for online desalting and then separated on a PepMap C-18 RP nano column (3 μm, 75 μm × 15 cm; Dionex) and eluted in a 90-min gradient of 5–40% acetonitrile in 0.1% formic acid at 300 nL/min, followed by a 3-min ramping to 95% acetonitrile/0.1% formic acid and a 5-min holding at 95% acetonitrile/0.1% formic

acid. The eluted peptides are detected by the Orbitrap Velos through the nano-ion source containing a 10 μm analyte emitter (NewObjective). Raw spectra were processed using Proteome Discoverer 1.1 (Thermo Scientific) utilizing the *Drosophila melanogaster* Period sequence provided by the author.

REFERENCES

1. Vaidya, A.T. 2012 . *Characterization of Light-Signaling in the Circadian Photoreceptor Proteins Vivid and Cryptochrome*, in *Chemistry and Chemical Biology*. Cornell University: Ithaca, NY. 111.

2. Dunlap, J.C. 1999. Molecular bases for circadian clocks. *Cell*. 96: 271-290.

3. Mitsui, A., Cao, S., Takahashi, A., and Arai, T. 1987. Growth Synchrony and Cellular-Parameters of the Unicellular Nitrogen-Fixing Marine Cyanobacterium, Synechococcus Sp Strain Miami Bg 043511 under Continuous Illumination. *Physiologia Plantarum*. 69: 1-8.

4. Konopka, R.J. and Benzer, S. 1971. Clock Mutants of Drosophila-Melanogaster. *P Natl Acad Sci USA*. 68: 2112-2116.

5. Jurgen Aschoff, R.W. 1979. The Circadian System of Man, in *Biological Systems*. Springer: USA. 311-331.

6. Crosthwaite, S.K., Loros, J.J., and Dunlap, J.C. 1995. Light-Induced Resetting of a Circadian Clock Is Mediated by a Rapid Increase in Frequency Transcript. *Cell*. 81: 1003-1012.

7. Albrecht, U., Sun, Z.S., Eichele, G., and Lee, C.C. 1997. A differential response of two putative mammalian circadian regulators, mper1 and mper2, to light. *Cell*. 91: 1055-1064.

8. Peschel, N. and Helfrich-Forster, C. 2011. Setting the clock - by nature: Circadian rhythm in the fruitfly Drosophila melanogaster. *FEBS Lett*. 585: 1435-1442.

9. Lee, C., Bae, K., and Edery, I. 1999. PER and TIM inhibit the DNA binding activity of a Drosophila CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: a basis for circadian transcription. *Mol Cell Biol*. 19: 5316-5325.

10. Saez, L., Meyer, P., and Young, M.W. 2007. A PER/TIM/DBT interval timer for Drosophila's circadian clock. *Cold Spring Harbor Symposia on Quantitative Biology*. 72: 69-74.

11. Dunlap, J.C. 2006. Running a clock requires quality time together. *Science*. 311: 184-186.

12. Meyer, P., Saez, L., and Young, M.W. 2006. PER-TIM interactions in living Drosophila cells: An interval timer for the circadian clock. *Science*. 311: 226-229.

13. Ko, H.W., Kim, E.Y., Chiu, J., Vanselow, J.T., Kramer, A., and Edery, I. 2010. A Hierarchical Phosphorylation Cascade That Regulates the Timing of PERIOD Nuclear Entry Reveals Novel Roles for Proline-Directed Kinases and GSK-3 beta/SGG in Circadian Clocks. *J Neurol.* 30: 12664-12675.

14. Lin, J.M., Kilman, V.L., Keegan, K., Paddock, B., Emery-Le, M., Rosbash, M., and Allada, R. 2002. A role for casein kinase 2 alpha in the Drosophila circadian clock. *Nature*. 420: 816-820.

15. Sathyanarayanan, S., Zheng, X., Kumar, S., Chen, C.H., Chen, D., Hay, B., and Sehgal, A. 2008. Identification of novel genes involved in light-dependent CRY degradation through a genome-wide RNAi screen. *Gen Dev*. 22: 1522-1533.

16. Koh, K., Zheng, X.Z, and Sehgal, A. 2006. JETLAG resets the Drosophila circadian clock by promoting light-induced degradation of TIMELESS. *Science*. 312: 1809-1812.

17. Peschel, N., Chen, K.F., Szabo, G., and Stanewsky, R. 2009. Light-Dependent Interactions between the Drosophila Circadian Clock Factors Cryptochrome, Jetlag, and Timeless. *Curr Biol.* 19: 241-247.

18. Cyran, S.A., Yiannoulos, G., Buchsbaum, A.M., Saez, L., Young, M.W., and Blau, J. 2005. The double-time protein kinase regulates the subcellular localization of the Drosophila clock protein period. *J Neuro*. 25: 5430-5437.

19. Kloss, B., Rothenfluh, A., Young, M.W., and Saez, L. 2001. Phosphorylation of PERIOD is influenced by cycling physical associations of DOUBLE-TIME, PERIOD, and TIMELESS in the Drosophila clock. *Neuron*. 30: 699-706.

20. Ko, H.W., Jiang, J., and Edery, I. 2002. Role for Slimb in the degradation of Drosophila Period protein phosphorylated by Doubletime. *Nature*. 420: 673-678.

21. Chiu, J.C., Vanselow, J.T., Kramer, A., and Edery, I. 2008. The phospho-occupancy of an atypical SLIMB-binding site on PERIOD that is

phosphorylated by DOUBLETIME controls the pace of the clock. *Gen Dev*. 22: 1758-1772.

22. Emery, P., Stanewsky, R., Hall, J.C., and Rosbash, M. 2000. Drosophila cryptochromes - A unique circadian-rhythm photoreceptor. *Nature*. 404: 456-457.

23. Vaidya, A.T., Top, D., Manahan, C.C., Tokuda, J.M., Zhang, S., Pollack, L., Young, M.W. and Crane, B.R. 2013. Flavin reduction activates Drosophila cryptochrome. *P Natl Acad Sci USA*. 110: 20455-20460.

24. Zoltowski, B.D. and Crane, B.R. 2008. Light activation of the LOV protein Vivid generates a rapidly exchanging dimer. *Biochemistry*. 47: 7012-7019.

25. Conrad, K.S., Bilwes, A.M., and Crane, B.R. 2013. Light-Induced Subunit Dissociation by a Light-Oxygen-Voltage Domain Photoreceptor from Rhodobacter sphaeroides. *Biochemistry*. 52: 378-391.

26. King, H.A., Hoelz, A., Crane, B.R., and Young, M.W. 2011. Structure of an Enclosed Dimer Formed by the Drosophila Period Protein. *J Mol Biol*. 413: 561-572.

27. Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S, Urbanke, C., Sassone-Corsi, P., and Wolf, E. 2005. Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. *Mol Cell*. 17: 69-82.

28. Saez, L. andYoung, M.W. 1996. Regulation of nuclear entry of the Drosophila clock proteins period and timeless. *Neuron*. 17: 911-920.

29. Hennig, S., Strauss, H.M., Vanselow, K., Yildiz, O., Schultze, S., Arens, J., Kramer, A., and Wolf, E. 2009. Structural and Functional Analyses of PAS Domain Interactions of the Clock Proteins Drosophila PERIOD and Mouse PERIOD2. *Plos Biol*. 7: 836-853.

30. Landskron, J., Chen, K.F., Wolf, E., and Stanewsky, R. 2009. A Role for the PERIOD:PERIOD Homodimer in the Drosophila Circadian Clock. *Plos Biol*. 7: 820-835.

31. Chang, D.C. and Reppert, S.M. 2003. A novel C-terminal domain of Drosophila PERIOD inhibits dCLOCK : CYCLE-mediated transcription. *Curr Biol*. 13: 758-762.

32. Top, D. 2014. Unpublished Correspondance.

33. King, H. 2008. A Structural and Biochemical Analysis of the Drosophila Protein Period, a Conserved and Essential Component of the Circadian Clock, in *Chemistry*. 2008, Rockefeller University: New York, NY. 114.

34. Skou, S., Gillilan, R.E., and Ando, N. 2014. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nature Protocols*. 9: 1727-1739.

35. Nielsen, S.S., Toft, K.N., Snakenborg, D., Jeppensen, M.G., Jacobsen, J.K., Vestergaard, B., Kutter, J.P., and Arleth, L. 2009. BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *J App Crystallogr*. 42: 959-964.

36. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J App Crystallogr*. 36: 1277-1282.

Appendix 1

TOWARDS *IN VIVO* PULSED DIPOLAR ESR SPECTROSCOPIC MEASUREMENTS

A1.1 Introduction

Pulsed dipolar ESR spectroscopy (PDS) measures the dipolar coupling between 2 unpaired spins, and from the strength of the dipolar coupling, a distance distribution is extracted. Generally, these unpaired spins are nitroxyl radicals which are covalently attached to cysteines at desired sites in the protein of interest. These radicals have been shown to be unstable in the reducing cellular environment [1, 2] and thus PDS measurements have been limited to *in vitro* samples. Recently, it was shown that nitroxides are stable in *E. coli* long enough for a labeled protein to be extracted from the cell [3]. The advent of metal-based PDS measurements also shows promise in working toward use of this technique on intact whole cells.

A1.2 Results

Attempts were made to make *in vivo* PDS measurements using the protein Superoxide Dismutase 1 (SOD1), a homodimer with 2 Cu^{2+} ions approximately 32 Å apart (see Chapter 2). SOD1 was expressed in *E. coli* BL21-DE3 cells. Cells were grown at 37 °C until they reached an OD_{600} of 1.0, at which point they were induced with 0.3 mM IPTG and 0.25 mM $CuSO_4$. After 18 hours of growth at room

temperature, an aliquot of the cell culture was pipetted into an ESR tube. The tube was spun down at 3,000 rpm until a pellet formed at the bottom and the supernatant was clear. The supernatant was then decanted, and this process was repeated until the cell pellet had a volume of approximately 30 uL in the bottom of the ESR tube. PDS measurements were then attempted, but were unsuccessful. Despite overexpression, there was not a high enough concentration to overcome background from numerous other copper-containing proteins in *E. coli* cells. Furthermore, the abundance of metal ions greatly increases relaxation times, further complicating PDS measurements.

A1.3 Discussion

For *in vivo* PDS measurements, using common biological metals such as copper, iron and manganese as paramagnets may not be the ideal choice. However, gadolinium is a lanthanide metal which has recently been shown to be a suitable spin probe for PDS measurements, and which may afford some advantages relative to other metals [4-8]. First, gadolinium is not native to *E. coli* cells, so there should be no background interference from exogenous gadolinium. Furthermore, it has been shown that Gd^{3+}-Gd^{3+} distances can be measured at high-field, which increases the sensitivity of the experiment, leading to better signal to noise ratios. A PDS measurement has been in cells using Gd^{3+} as a spin probe, but the experiment was not done strictly *in vivo*, as a label was covalently attached outside of the cells [9].

One approach to *in vivo* Gd^{3+} PDS measurements might be to design a Gd^{3+} binding site in a protein. Gd^{3+} has been used as an MRI contrast agent, and a Gd^{3+} binding peptide has been developed for this purpose [10]. By attaching this peptide to a protein which expresses very well, such as GFP, it may be possible to reach a high enough concentration of Gd^{3+} ions for PDS measurements *in vivo*.

REFERENCES

1. Krstic, I., Hansel, R., Romainczyk, O., Engels, J.W., Dotsch, V., and Prisner, T.F. 2011. Long-Range Distance Measurements on Nucleic Acids in Cells by Pulsed EPR Spectroscopy. *Angew Chem Int Edit.* 50: 5070-5074.

2. Belkin, S., Mehlhorn, R.J., Hideg, K., Hankovsky, O., and Packer, L. 1987. Reduction and Destruction Rates of Nitroxide Spin Probes. *Arch Biochem Biophys.* 256: 232-243.

3. Schmidt, M.J., Borbas, J., Drescher, M., and Summerer, D. 2014. A Genetically Encoded Spin Label for Electron Paramagnetic Resonance Distance Measurements. *J Am Chem Soc.* 136: 1238-1241.

4. Raitsimring, A.M., Atashkin, A.V., Poluektov, O.G., and Caravan, P. 2005. High-field pulsed EPR and ENDOR of Gd3+ complexes in glassy solutions. *App Mag Res.* 28: 281-295.

5. Gordon-Grossman, M., Kaminker, I., Gofman, Y., Shai, Y., and Goldfarb, D. 2011. W-Band pulse EPR distance measurements in peptides using Gd3+-dipicolinic acid derivatives as spin labels. *Phys Chem Chem Phys.* 13: 10771-10780.

6. Yagi, H., Banerjee, D., Graham, B., Hubert, 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.

7. Song, Y., Meade, T.J., Astashkin, A.V., Klein, E.L., Enemark, J.H., and Raitsimring, A., 2011. Pulsed dipolar spectroscopy distance measurements in biomacromolecules labeled with Gd(III) markers. *J Mag Res.* 210: 59-68.

8. Garbuio, L., Bordignon, E., Brooks, E.K., Hubbell, W.L., Jeschke, G., and Yulikov, M. 2013. Orthogonal Spin Labeling and Gd(III)-Nitroxide Distance Measurements on Bacteriophage T4-Lysozyme. *J Phys Chem B.* 117: 3145-3153.

9. 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.

10. Daughtry, K.D., Martin, L.J., Sarraju, A., Imperiali, B., and Allen, K.N. 2012. Tailoring Encodable Lanthanide-Binding Tags as MRI Contrast Agents. *Chembiochem*. 13: 2567-2574.

Appendix 2

TOWARDS CRYSTALLIZATION OF THE FIRST 700 RESIDUES OF THE *DROSOPHILA MELANOGASTER* CIRCADIAN CLOCK PROTEIN PERIOD

A2.1 Introduction

The protein Period is an essential component of the *Drosophila melanogaster* circadian clock; along with binding partner Timeless, it forms the main repressor complex in the clock's negative feedback loop. Despite extensive study, there is structural information available for only a relatively small segment of Period that encompasses its ordered PAS domains [1, 2]. Here we have attempted to crystallize a construct consisting of amino acids 1-700 of Period.

A2.2 Results

Broad robotic screening was done on WT PER 1-700, as well as phosphomimics PER2 and PER5 (see Table 4.1) using JCSG I-IV (Qiagen). Previous successful crystallization of the PAS domains of PER was done in a protein solution consisting of 20 mM HEPES pH 7.7, 200 mM NH_4HCO_3 and 5 mM DTT, and so PER variants were buffer exchanged into this buffer before trays were set up [2]. Out of the 384 possible conditions for each variant, there was 1 hit: a crystal formed in a drop containing PER1 and 200 mM $NH_4H_2PO_4$, 100mM Tris-HCl pH 8.5, and 50%

(v/v) 2-methyl-2,4-pentane-diol (MPD) at room temperature. Finer screens were then set up to optimize crystal growth. Crystals grew in 1-3 weeks with the pH of 100 mM Tris-HCl varying from 7.75 to 9.0 and the $NH_4H_2PO_4$ concentration varying from 150 mM to 220 mM. Outside of this range, no crystals grew. Reduction in MPD concentration to 40% (v/v) yielded no crystals. The concentration of PER1 in the same buffer as above was 8-10 mg/mL, and the crystals formed in 3 uL drops containing 2 uL of protein solution and 1 uL of well solution. Varying this ratio to 6:1, 3:1, 1:1 and 1:2 yielded no crystals. Crystals appeared small and needle-like, and were very difficult to pick up and work with (Fig. A2.1). Attempts to mount them on a loop and shoot them have been unsuccessful.

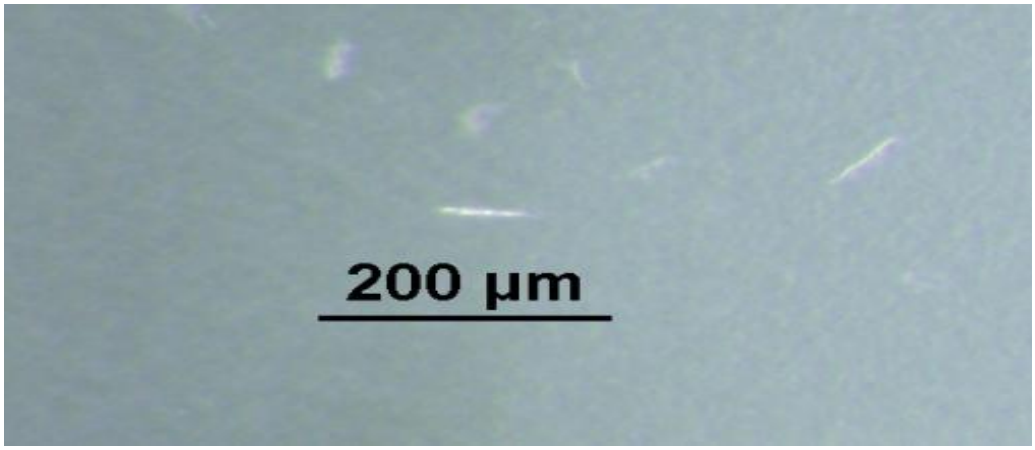

Figure A2.1: Image of PER 1-700 crystals.

A2.3 Discussion

Structural characterization of a large portion of Period would represent a major step forward in the study of the *Drosophila* circadian clock, and thus all possibilities for

obtaining a structure should be exhausted. In obtaining crystals, there are many variables which have not been tested for optimization. The temperature could be adjusted to 17˚C or 4˚C. Seeding was attempted once, and yielded no crystals, but this could be re-explored. Other PER variants should be tested under these conditions to see if they also yield crystals.

REFERENCES

1. Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S, Urbanke, C., Sassone-Corsi, P., and Wolf, E. 2005. Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. *Mol Cell*. 17: 69-82.

2. King, H.A., Hoelz, A., Crane, B.R., and Young, M.W. 2011. Structure of an Enclosed Dimer Formed by the Drosophila Period Protein. *J Mol Biol*. 413: 561-572.