UNIVERSITY of CALIFORNIA
Santa Barbara

High-field CW EPR with Gd(III) spin labels for structure studies of membrane proteins

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

by

Jessica Ann Clayton

Committee in charge:

Professor Mark S. Sherwin, Chair
Professor Songi Han
Professor Ania Bleszynski Jayich
Professor Philip A. Pincus

September 2017

The dissertation of Jessica Ann Clayton is approved:

Professor Ania Bleszynski Jayich

Professor Philip A. Pincus

Professor Songi Han

Professor Mark S. Sherwin, Chair

September 2017

To my family.

Acknowledgements

It is impossible to identify every single person that has helped me reach the point of writing the acknowledgments section of my Ph.D. dissertation, but many people have undoubtedly played vital roles along the way. First, I would like to thank my advisors, Mark Sherwin and Songi Han, who took me in, taught me to be a real-world scientist, and allowed me a great deal of freedom along the way to explore all the random side projects that crop up during the course of a research project - some fruitful and others not. Blake Wilson and Devin Edwards have been my closest collaborators and teachers, spending countless days in the lab with me running experiments and discussing every measurement, simulation, and interpretation of data presented in this dissertation as well as many other EPR projects not discussed here. Chung-ta Han, Sunyia Hussain, Nikki Schonenbach, Matt Idso, Julia Fisher, and Neil Eschmann from the Han and Chmelka labs taught me a great deal about the world of biochemistry and protein science. I owe thanks to Darren Valovcin and Lucas Brady for their help in using the computing cluster and for teaching me to write better code. While not directly connected to any of the work in this dissertation, I also owe thanks to Hunter Banks and Changyun Yoo for being wonderful lab and office mates. From the Han lab, Ilia Kaminker, Alisa Leavesley, Alicia Lund, Ryan Barnes, Ting Ann Siaw, Tim Keller, and John Franck were very helpful to bounce ideas off of. I've had the opportunity to work with many talented undergraduate students in both the Sherwin and Han labs, including Sirish Narayanan, Evelyn Chang, Naomi Baxter, Sam Aronson, Andrew Pierce, and Mary Lou Bailey. And to everyone else in the Sherwin and Han labs not explicitly named here, thank you - you've been a pleasure to work with.

Adelheid Godt and Mian Qi at Bielefeld University have provided incredibly high quality samples which made much of this work possible. My surprise collaboration with Katharina Keller, Maxim Yulikov, and Gunnar Jeschke at ETH Zurich has been very educational and resulted in what I initially thought of as a small side project turning into a major chapter of this dissertation. Daniella Goldfarb, Akiva Feintuch, Alberto Collauto, Thorsten Bahrenburg, Marie Ramirez-Cohen, Aleksei Litvinov, Arina Dalaloyan, and Yin Yang were extraordinarily welcoming during my stay at the Weizmann Institute in Israel, allowed me to work in their lab for a month, and taught me much about high-field EPR with Gd(III). Steffen Glaser has spent several sabbaticals at UCSB, collaborated with us on projects, and is always a patient and instructive teacher. I'd also like to thank Susumu Takahashi for pioneering the high-field EPR work at UCSB and for useful discussions on various experiments and Stefan Stoll for help with simulating EPR spectra.

The physics department at UCSB has numerous wonderful staff people who in one way or another contribute to helping things run smoothly. In particular, I would like to thank David Enyeart, Rob Marquez, Jennifer Farrar, Mike Deal, Dave Prine, Erin Ferguson, and Guy Patterson.

Many a fun lunch, evening, or weekend off was spent in the company of friends and classmates, whom I would like to thank semi-anonymously: DCV, CBW, HBB, LTB, ZAG, RS, KS, SR, AR, AM, KF, CN, MS, DI, EB, BW, SM, SF, KJS, CT, JP, AF, CN,

SA, GB, SO, BK, CS, AF, SC, NE, KD, TP, EL, SR, GC, AW, KH, CB, AM-A, BL, KS, BF, AL, RB, TK, AL, IK.

Finally, and certainly not last, I would like to thank my wonderful parents and family for always being encouraging, supportive, and putting up with too few visits and phone calls during the past several years.

Without a doubt I have neglected to acknowledge everyone on this page, given the incredibly large number of very talented and genuinely helpful people I've had the privilege of working with during the past several years. If you feel like you should have been acknowledged here, but aren't, please fill in the gaps below as you see fit.

I would like to thank _____ for all of their _____, their _____ _____, and their _____ sense of _____. They have been a _____ _____ in completing this work, and I would have certainly _____ _____ without them. _____ is the _____ _____ I have ever had the _____ of knowning, and I truly appreciate their _____.

Curriculum Vitæ

Jessica Ann Clayton

Education

2017	Ph.D., physics, University of California, Santa Barbara, California
2014	M.A., physics, Univeristy of California, Santa Barbara, California
2011	B.S., physics, University of California, Merced, California
2007	Del Oro High School, Loomis, California

Professional Experience

2012 - 2017	Graduate research assistant, Department of Physics, UCSB
2012 - 2016	Questboards outreach volunteer coordinator, ITST/CNSI, UCSB
2011 - 2012	Teaching assistant, Department of Physics, UCSB
2010 - 2012	Undergraduate research assistant, Department of Physics, UCM

Publications

Clayton, J.A., Qi, M., Godt, A., Goldfarb, D., Han, S., Sherwin, M.S., "Gd^{3+}–Gd^{3+} distances exceeding 3 nm determined by very high frequency continuous wave electron paramagnetic resonance," *Physical Chemistry Chemical Physics* **2017**, 19, 5127-5136.

Ghosh, S., Shcherbatyuk, G., Inman, R., and Clayton, J.A., "Nanostructured Luminescent Solar Concentrators," *SPIE Newsroom: Solar and Alternative Energy (invited article)* **2010**.

Manuscripts in Preparation

Clayton, J.A., Keller, K., Qi, M., Wegner, J., Koch, V., Hintz, H., Godt, A., Han, S., Jeschke, G., Sherwin, M.S., Yulikov, M., "Quantitative analysis of zero-field splittings in Gd(III) complexes."

Wilson, C.B., Aronson, S., Clayton, J.A., Glaser, S., Han, S., Sherwin, M.S., "Phase-cycling and spin relaxation time measurements in a pulsed free-electron-laser-powered electron paramagnetic resonance spectrometer."

Wilson, C.B., Clayton, J.A., Edwards, D.T., Agladze, N., Han, S., Sherwin, M.S., "Demagnetization shifts in very high frequency pulsed electron paramagnetic resonance."

Conference Presentations

July 2017	International Society of Magnetic Resonance Conference/Rocky Mountain Conference on Magnetic Resonance "High-field CW EPR with Gd(III) tags for structure-dynamics studies of proteorhodopsin"
March 2017	American Physical Society March Meeting "High-field CW electron paramagnetic resonance spectroscopy with Gd(III) tags for structure-dynamics studies of proteorhodopsin"
July 2016	Rocky Mountain Conference on Magnetic Resonance "Distance measurements in Gd3+-labeled proteorhodopsin oligomers by 240 GHz CW EPR"
July 2015	Rocky Mountain Conference on Magnetic Resonance "Determination of the zero-field splitting (ZFS) parameters of Gd3+ complexes"
July 2014	Rocky Mountain Conference on Magnetic Resonance "Measurement of Gd^{3+}–Gd^{3+} distances by CW EPR at 240 GHz"
February 2014	UCSB Materials Outreach Symposium "Electron paramagnetic resonance of Gd^{3+} spin labels for studies of structure in biological systems"
July 2013	Rocky Mountain Conference on Magnetic Resonance "Measurement of Gd^{3+}–Gd^{3+} distances by CW EPR at 240 GHz"
October 2010	Society for Advancement of Chicanos and Native Americans in Science Conference "Optical properties of ZnO nanowires"

Recognition

2017	UCSB doctoral student travel grant
2016	Springer best student poster award
2015	Shared EPR/EFEPR summer school travel grant
2011	Grossman scholar fund award
2010 - 2011	UCM renewable energy scholarship
2010 - 2011	UCAGEP scholar
2009 - 2011	UC regents scholarship

Note for the reader

This document is a revised and expanded version of the dissertation which Jessica Ann Clayton submitted for publication under the title "High-field CW EPR with Gd(III) spin labels for structure studies of membrane proteins" to ProQuest and to the Alexandria Digital Research Library at UCSB.

Abstract

High-field CW EPR with Gd(III) spin labels for structure studies of membrane proteins

by

Jessica Ann Clayton

Electron paramagnetic resonance (EPR) in combination with site-directed spin labeling (SDSL) is a powerful tool for elucidating the structure, organization, and dynamics of biomolecules in native-like environments. With EPR and SDSL, we can site-specifically label pairs of sites in a biomolecule and accurately measure the distance, or distribution of distances, between them on length scales ranging from Ångstroms to several nanometers. Of particular interest are membrane proteins and higher-order membrane protein complexes, which have historically resisted traditional biophysical characterization techniques. EPR as a means to measure protein structure becomes even more powerful at high fields and using Gd(III) spin labels, which together provide much improved sensitivity. This work expands on the capabilities of high-field continuous-wave (CW) EPR for distance measurement with spin labels based on Gd(III) complexes. First, we investigate a model system of and show that CW EPR with Gd(III) labels allows for distance measurements in the range of at least 1.2 - 3.4 nm at cryogenic temperatures. We additionally show that distance measurements are possible up to room temperature. Next,

we investigate the zero-field splitting - a property of great importance for determining the EPR lineshape of high-spin systems - for a variety of different Gd(III) complexes. Combining EPR spectra measured at 35 GHz, 95 GHz, and 240 GHz, we compare literature models for the broadly distributed second-order ZFS parameters D and E. We test these results against a superposition model for predicting the magnitude of the ZFS based on knowledge of the structure of a Gd(III) complex, which can potentially be useful for designing new Gd(III) complexes tailored for use as spin labels with high-field EPR. Finally, we apply high-field CW EPR with Gd(III) spin labels to the study of proteorhodopsin (PR), a transmembrane protein that functions as a light-driven proton pump for marine bacteria. Inter-PR CW EPR distance measurements in the range of $\sim$ 1.5 - 3 nm are used to elucidate the functionally relevant oligomeric structure of PR, demonstrating the usefulness of this technique in targeting complex oligomeric systems. Finally, we present the development of methods which will allow CW EPR with Gd(III) to be used as a probe of protein dynamics, by measuring at a distance change induced by motions of the E-F loop region of PR upon light activation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Electron paramagnetic resonance (EPR), also referred to as electron spin resonance (ESR), is a spectroscopic technique for studying systems containing unpaired electrons. These unpaired electrons may be either intrinsic to the system under study, or introduced artificially by the spectroscopist. EPR is often thought of as the less well-known cousin of NMR (nuclear magnetic resonance), sharing many of the same basic concepts. EPR was discovered in the USSR by Yevgeny Zavoisky in 1944, one year before the first observation of NMR was recorded. However, while NMR has been wildly successful since that time - resulting in Nobel prizes in physics, chemistry, and medicine - the development of EPR was long hampered by the lack of necessary microwave components and various experimental challenges related to fast relaxation times and broad lines which make detection difficult. Only in the the 1980s was the first commercial pulse EPR spectrometer available, followed ten years later by the first commercial high-field spectrometer.

In recent years, the availability of equipment and the development of new methods has allowed a wide variety of users access to EPR spectroscopy. This versatile spectroscopic technique has proven particularly useful in the field of structural biology, where the use of EPR methods has expanded as modern EPR spectrometers become more affordable, flexible, and user friendly.

The history of NMR as a spectroscopic tool has shown great benefit from operating at higher magnetic fields or equivalently, since this is a resonance technique, at higher microwave frequencies where spectral resolution and sensitivity are increased. However, while the forward march of NMR is restricted primarily by the lack of commonly available magnets operating above 20 T, EPR is limited by the lack of high-frequency (>100 GHz) high-power microwave sources. This difference is the result of the magnetic moment of electron spins being 660 times larger than that of a proton. Fortunately for the EPR spectroscopist, this also means EPR has much higher spin sensitivity than NMR at a given spin concentration, even allowing for single-spin detection. In the last 5-10 years, high-frequency sources have begun moving from the realm of being very expensive, specialized laboratory equipment, to now having a commercially available EPR spectrometer operating at 263 GHz. Power is still typically limited at these elevated frequencies, with most > 200 GHz sources producing only tens of mW of power. This forces pulsed EPR experiments to use very long excitation pulses, prohibiting the study of fast-relaxing systems [147, 134, 26, 47, 42, 159, 163, 38, 154]. Specialized laboratories operating with a free-electron laser (FEL) [159], extended interaction klystron amplifier (EIKA), traveling

wave tube amplifier (TWT), or other high-power source are exceptions to this, though these can hardly be said to be common. Fortunately, continuous wave (CW) EPR has a much lower power requirement than does pulsed EPR, and much information can be gleaned from CW EPR measurements. CW EPR also has the benefit of being fast, with typical measurement times of order ten minutes, while many pulsed EPR experiments require hours of signal averaging for a single measurement.

This dissertation discusses efforts to expand the capabilities and use of high-field/high-frequency CW EPR. We focus on applications in structural biology, particularly in membrane proteins and their assembly, which often resist other traditional biophysical characterization techniques. Membrane and transmembrane proteins are critical for a variety of cellular functions, including acting as receptors to relay information between the interior and exterior of the cell and as transporters for moving ions and molecules across the cell membrane. Due to these functions, membrane proteins in humans are the primary target for more than half of pharmaceutical drugs currently on the market. Despite their great importance and prevalence in nature, as of the writing of this dissertation only 3.5% of the protein structures stored by the RSCB Protein Data Bank (PDB) are classified as membrane proteins. The vast majority of the protein structures on PDB were acquired using x-ray crystallography, with a smaller number collected using NMR spectroscopy. While these methods are indispensable for mapping complete (or nearly complete) high-resolution protein structures, they provide only snapshot images of the protein. In contrast, EPR provides only sparse structural information, but we are granted

in exchange the flexibility to observe proteins in very native-like environments. Additionally, since EPR measurements can be conducted under ambient conditions, this allows for the possibility of monitoring triggered conformational dynamics, i.e. the ability to watch a protein move in real time.

Since paramagnetic electrons are relatively rare in nature, EPR allows for site-specific investigation of protein structure and dynamics. Some biological systems natively include paramagnetic centers which can be used directly as internal probes. Systems which do not natively contain paramagnetic electrons, or in cases where the native electron is not at the desired location, can be made paramagnetic through the attachment of a spin label. Spin labels vary in the EPR active ion and the means of attachment, but in general can be introduced site-specifically, singly or in pairs, to specific site(s) of interest on a protein. Some of the most commonly used spin labels are those based on the $S = 1/2$ nitroxide radical, which are very effective for structure studies with X-band (9.5 GHz) EPR. When used in pairs, these spin labels have been used to measure distances, though their range is limited and the magnetic properties of nitroxides become unfavorable for distance measurements with high-field EPR. High-field CW EPR, which in principle has higher sensitivity, has been insufficiently utilized as a tool for studying protein structure and dynamics, largely due to the lack of appropriate spin labels for high-field EPR and lack of the experimental and theoretical techniques necessary to effectively use these spin labels. We focus on a new class of spin labels based on the Gd(III) ion, which have shown in recent years to show favorable magnetic properties for use as spin labels for EPR at

high (>100 GHz) frequencies. Additionally, Gd(III) has an effective spin $S = 7/2$, giving rise not only to greater range and sensitivity in EPR distance measurements, but also to interesting and complex spin physics.

We explore a variety of Gd(III) complexes in the course of this dissertation primarily using a 240 GHz CW EPR instrument. We begin with the study of a model system to explore the use of high-field CW EPR with Gd(III) for measuring inter-spin distances - one of the primary means in which CW EPR can provide structural information. We investigate distance limits, temperature dependence, sensitivity, and work towards development of a comprehensive understanding of the line shape of the CW EPR signal. We additionally discuss analysis techniques for teasing out information on the dipolar coupling between two neighboring spins from this signal, necessary for future use of this as a distance measurement tool. We show that high-field CW EPR with Gd(III) labels doubles the distance range accessible by X-band CW EPR with nitroxide spin labels, and furthermore, can be measured from cryogenic temperatures up to room temperature.

Next, we explore a fundamental property of Gd(III) complexes, the zero-field splitting (ZFS), which arises from the high-spin ($S = 7/2$) nature of the Gd(III) ion. The ZFS is an integral component to understanding the measured EPR signal and, to a large degree, determines which Gd(III) complexes are suitable for use with high-field EPR since the ZFS directly impacts the intrinsic linewidth of the Gd(III) complex. Because the development of new Gd(III) complexes for use as spin labels will be an essential component of expanding the use of high-field EPR in the future, we rationalize the

magnitude of the ZFS of a particular Gd(III) complex using a superposition model, before it is synthesized and measured by EPR.

Finally, we take the first steps in applying high-field CW EPR with Gd(III) spin labels to the study of structure and dynamics in a membrane protein. We choose as a model system the membrane protein proteorhodopsin (PR), which has a seven α-helical structure common to many proteins, including those found in humans. PR exhibits functionally-relevant oligomerization and light-induced conformational changes, both of which we investigate by CW EPR distance measurements which are enabled by the long-range and high-temperature capabilities afforded by the use of Gd(III)-based spin labels at high-field.

As a whole, the work in this dissertation elevates high-field CW EPR with Gd(III) spin labels for distance measurement from the first proof-of-principle experiment to a tool which is being actively pursued as an integral part of biophysical studies on the membrane protein PR. However, in terms of application we have only begun to explore a single biological system. As high-frequency EPR spectrometers and Gd(III) spin labels continue to become more advanced and more widely used, the methods developed in this dissertation will provide the background for new applications in the study of membrane protein structure and assembly, in a small part, contribute to the continued advancement of EPR as a whole.

1.1 Fundamentals of Electron Paramagnetic Resonance

Many introductory texts and comprehensive tomes have been written on the theory of magnetic resonance, both EPR and NMR, and therefore we will not delve into great depth here. Rather, this introduction will serve as a very cursory overview of the theoretical ideas and experimental techniques used throughout this dissertation. The more interested reader is referred to the excellent texts by Abragam [2], Abragam and Bleaney [3], Callaghan [36], Levitt [95], Poole [115], Saha and Das [136], Slichter [146], Schweiger and Jeschke [138], and Weil and Bolton [172]. In addition to these standard texts, the thesis of Devin Edwards [51] lays out much of the theoretical and experimental background that the work in this dissertation expands on and *Methods in Enzymology* vols. 563-564 [120, 121] very nicely summarize much of the related recent literature.

The basic description of spin systems and dynamics are the same for EPR and NMR, as well as many of the experimental concepts, with the major differences arising from the difference in magnetic moments between the proton and the electron. As a general statement, experimental schemes are often far simpler for EPR compared to NMR, but the spin dynamics are more complex.

An electron carries an intrinsic angular momentum, called spin. The theoretical description of EPR is typically written out in terms of spin $S = 1/2$ for simplicity, which can be straightforwardly expanded for the high-spin case. The intrinsic angular momen-

tum of an electron, also known as the electron magnetic moment or Bohr magneton, is written as

$$\mu_B = \frac{e}{2m}\hbar S \tag{1.1}$$

where S is the spin quantum number and $\hbar$ is the reduced Planck constant. In an EPR experiment, the electrons are placed in a strong magnetic field B_0, which by convention points along $\hat{z}$. This induces a Zeeman splitting, creating two states of different energy corresponding to the magnetic quantum numbers $m_s = \pm 1/2$ (for $S = 1/2$). The Zeeman splitting, or the energy difference between these spin states, is $\Delta E = E_{|+1/2\rangle} - E_{|-1/2\rangle} = g\mu_B B_0$, where for a free electron $g \approx 2.0023$. The condition for resonance is satisfied when applied microwaves satisfy the condition

$$\Delta E = h\nu = \hbar\omega = g\mu_B B_0, \tag{1.2}$$

that is, when the energy of the applied photons matches the energy separation of the $|+1/2\rangle$ and $|-1/2\rangle$ states, then a photon can be absorbed and transitions between these two states induced. The resonance frequency satisfying Equation 1.2 scales linearly with the strength of the applied magnetic field B_0. Table 1.1 lists some common spectrometer frequencies used for EPR.

Due to the stong magnetic fields used in EPR, the Zeeman interaction is often the dominant energy and other effects can be represented as perturbations on the eigenstates determined by the Zeeman interaction. Other relevant interactions for EPR spectroscopy

Microwave Band	Range (GHz)	Typical (GHz)	Magnetic Field (T)
L	1 - 2	1.5	0.054
S	2 - 4	3	0.11
X	8 - 12	9.5	0.34
K_u	12 - 18	17	0.6
K	18 - 26.5		
K_a	26.5 - 40		
Q	30 - 50	35	1.25
W	75 - 110	95	3.4
D	110 - 170	140	5.0
G	110 - 300	240/263	8.6/9.4

Table 1.1: Common spectrometer frequencies and fields (assuming g=2) used for EPR. These are often referred to by the corresponding microwave band designations - a holdover from WWII RADAR use. Note: In most EPR literature, the letter band designations are not consistently named at frequencies above ∼100 GHz.

include hyperfine interaction (electron spin/nuclear spin), the nuclear Zeeman interaction (nuclear spin/static magnetic field), the nuclear quadrupole interaction (nuclear spin/electric field gradient), the zero-field interaction for $S > 1/2$ (electron spin/electron spin), the exchange interaction (electron spin/electron spin), and dipole-dipole interaction. These interactions range in magnitude from MHz to GHz and, unlike in NMR, it is possible in EPR to break the high-temperature approximation assumed here. For the work presented in this dissertation we will focus primarily on the dipole-dipole interaction and the zero-field splitting interaction, which will be described in more detail in Chapters 2 and 3.

1.1.1 Continuous Wave EPR

The simplest conceivable magnetic resonance experiment is based on sweeping the magnetic field through the resonance condition defined by Equation 1.2 while the applied

microwave frequency is held constant. Equivalently, the magnetic field could be held fixed and the frequency swept, though this is rarely done in EPR due to experimental complications. The applied resonant microwaves are usually equivalently described as a small irradiating magnetic field B_1 which is applied perpendicularly to the large static magnetic field B_0. On resonance, transitions (both absorption and emission) are induced between the $|+1/2\rangle$ and $|-1/2\rangle$ states, with the rates determined by the populations of the states. In principle, for equally populated states no transitions would be induced. However, the application of B_0 and resulting Zeeman interaction gives energetic preference to the $|-1/2\rangle$ state with the result being a net absorption of radiation. Because the spin states obey Boltzmann statistics, we can increase the population of the $-1/2$ state by decreasing the temperature. At high magnetic fields and low temperatures, nearly complete polarization can be achieved, where $> 99\%$ of the spins are in the $|-1/2\rangle$ state. Even for high-spin systems, it is relatively easy to generate $> 90\%$ population of the lowest energy state (Appendix C.2).

A typical CW EPR experiment involves fixing the irradiation frequency and sweeping the magnetic field through resonance with a slight modulation (at 10s - 100s of kHz) applied to the magnetic field . Phase-sensitive lock-in detection of the signal modulation at this frequency yields the derivative of the absorption spectrum. This CW EPR signal encodes information about the various terms present in the spin Hamiltonian. For example, the location of the resonance according to Equation 1.2 gives us the g-value, which often differs from that of a free electron. Because each spin in the system experiences a

slightly different environment, based on the various interactions of each spin with other electrons and nuclei in the sample, the delta-function EPR resonance will become smeared out. The shape and location of the CW EPR signal therefore encodes information about the local environment surrounding paramagnetic centers, on length scales ranging from Ångstrom scale to more than 10 nanometers. For example, the dipole-dipole interaction results in a broadening of the CW EPR lineshape, with the magnitude of the broadening related to the distance separating the two interacting electron spins.

1.1.2 Rapid-passage EPR

A closely related technique to CW EPR is rapid-passage EPR (also sometimes referred to as fast-passage EPR). As it is used in this work, rapid passage EPR is experimentally identical to CW EPR, with the exception that relatively high microwave power and small modulation amplitudes are used. Instead of a derivative lineshape, the absorption lineshape is directly obtained [116, 171, 136, 22]. Note the rapid passage EPR used here is similar to, but distinct from the rapid-scan EPR technique developed by Eaton and Eaton [158]. Rapid passage is an old magnetic resonance technique, dating back to the original writings of Bloch [25], though it has been largely forgotten about in modern EPR literature.

Many measured EPR lineshapes are inhomogeneously broadened and can be thought of as a combined series of individual resonances from spin packets each of which obey the Bloch equations, resulting from a distribution of local magnetic fields around the

spin packets. In such a system, a condition can be reached where the amplitude of the magnetic field modulation is comparable to the homogeneous line width of each spin packet, so that an individual resonance is traversed in a time of the order of or smaller than the mean relaxation time of the spin system. Under rapid passage, the spin system is not allowed to come to equilibrium value at the start of every passage through resonance and so the intensity of the measured line, which is proportional to the transverse magnetization M_x and M_y, will be maximized on resonance [25, 116, 136]. There are a great many different passage regimes that can be observed in an EPR experiment, depending on the various time scales associated with the spin system and the experiment, and are discussed in detail by Weger [171].

1.1.3 Pulsed EPR

In addition to EPR experiments conducted in the frequency/field domain, but EPR experiments can also be carried out in the time domain by using pulses of microwave radiation to coherently manipulate the spins. Pulsed EPR is particularly useful for directly measuring relaxation times and for isolating specific parts of the spin Hamiltonian. Much as pulsed NMR offers many advantages, so to does pulsed EPR, though the technique is comparatively limited due to difficulty in fully exciting the broad EPR spectrum. There are a great variety of pulsed EPR experiments, and the reader is referred to the text by Schweiger and Jeschke [138] for a more complete description.

When an electron spin is placed in a magnetic field, a torque is exerted on the electron,

causing it to precess about the magnetic field. The angular frequency of this precession is called the Larmor frequency and is related to the magnetic field B_0 by $\omega_L = -\gamma B_0$, where γ is the gyromagnetic ratio. For a bulk system of electron spins in a magnetic field B_0, there will be a net magnetization along the z-axis (denoted M_0) with the various transverse components of the magnetization in the $x-y$ plane canceling each other out. The source in the EPR experiment supplies linearly polarized microwaves to produce a field B_1 perpendicular to B_0. These linearly polarized microwaves can be thought of as a magnetic field oscillating at the microwave frequency.

Alternatively, we can consider this situation in a frame rotating at the microwave frequency. In this rotating frame, the magnetization components precessing at the Larmor frequency now appear stationary and the field B_1 is also stationary. In the rotating frame, the magnetization will precess about B_1 at a frequency $\omega_1 = -\gamma B_1$ called the Rabi frequency. For B_1 parallel to the x-axis, this will rotate the magnetization M_0 about the $+x$-axis for as long as microwaves are applied, where the tip angle α depends on the length of the applied pulse t_p according to $\alpha = -\gamma |B_1| t_p$. Pulses are often referred to by their tip angle, e.g. a $\pi/2$ pulse corresponds to a rotation of M_0 by $\pi/2$ or 90°. A $\pi/2$ pulse results in magnetization along the $-y$-axis, while a π pulse results in a rotation of the magnetization M_0 to the $-z$-axis. In this way, combinations of pulses can be applied to manipulate the spin system at the end of which the measured EPR signal comes from any remaining transverse magnetization in the $x-y$-plane. In the time between microwave pulses, the spins are allowed to 'relax'. This relaxation is characterized by

two time scales, T_1 and T_2, where the spin-lattice relaxation time/longitudinal relaxation time T_1 refers to the return of the component of M_0 along B_0 to thermal equilibrium and the spin-spin relaxation time/transverse relaxation time T_2 is the decay of the transverse component of M_0 to equilibrium as a result of dephasing of the electron spins.

One very important pulsed EPR experiment for structure studies is double electron-electron resonance (DEER), also referred to as pulsed electron double resonance (PELDOR). There are many variants of this experiment - here we will focus on 4-pulse DEER since this is the most widely used. This experiment directly measures the dipolar coupling between electron spins separately from other electron spin interactions. Microwave pulses are applied at two nearby frequencies ω_A and ω_B, called the observer frequency and the pump frequency. The observer frequency supplies a Hahn echo pulse sequence with an additional third pulse at a later time to refocus the primary echo ($\pi/2$-τ_1-π-τ_1-(*echo*)-τ_2-π-τ_2-*echo*), where the interpulse delays τ_1 and τ_2 are held fixed. Meanwhile, a π inversion pulse at the pump frequency is applied at various time t with respect to the first echo at the observer frequency. If we consider two electron spins A and B, the pump pulse inverts the spin state of B, thus inverting the local magnetic field felt by spin A. This results in a modulation of the echo as the pump pulse is incremented. The frequency of the modulation is related to the dipolar coupling frequency and thus to the interspin distance, and can be found by Fourier transforming the signal into the frequency domain. In addition to this oscillation, transverse relaxation of spin A and couplings to other electron spins not excited by the observer pulses will lead to an attenuation of the

echo.

1.2 EPR for protein structure studies

EPR in combination with SDSL has become an essential tool for structure studies of a variety of biomolecules, including proteins, synthetic polymers, nucleic acids, and lipids. EPR methods - including pulsed EPR and CW EPR spectroscopy - complement conventional high-resolutions structure measurements, such as X-ray crystallography and NMR spectroscopy. EPR is particularly advantageous in the study of complex protein-lipid assemblies where practical difficulties with low-yield expression and crystallization often preclude other biophysical characterization techniques [72].

1.2.1 Site-directed spin labeling

Site-directed spin labeling (SDSL) allows for the investigation by EPR of local structure and dynamics in proteins by introducing paramagnetic centers into a biomolecule at user-defined sites. Some proteins contain naturally occuring paramagnetic centers, such as metal ions, but most proteins are diamagnetic and are thus EPR silent. This allows for SDSL to provide information on protein structure and dynamics at sites of interest with virtually no background. Spin labels were first developed in 1965 in the lab of McConnell [157] and much pioneering work was done in the lab of Hubbell [8, 10].

Spin labels can be used to report on dynamics, solvent accessibility, polarity of the local environment, and distance distributions between two spin labels in the nanometer

range. This last application will be our focus here. There are a wide variety of spin labels available for EPR and are chosen based on the chemical and spectroscopic properties of the spin label as well as the method of introduction into the protein under study. A spin label typically consists of a paramagnetic center along with a functional group which binds specifically to a given amino acid. Cysteine residues are most commonly used as the amino acid which the spin label binds to, due to their relative scarcity in nature. In this labeling scheme, site-directed mutagenesis is used to first remove any native cysteine residues in the protein and then to reintroduce cysteine residues at the site(s) which are to be spin labeled. The thiol group on the cysteine reacts specifically with the functional groups methanethiosulfonate (MTS) and maleimide, creating a covalent bond which attaches the spin label to the protein.

Labels based on the nitroxide radical, where the paramagnetic center is formed by an electron localized on the N−O bond, have long the most commonly used spin label due to their chemical stability and simple EPR signature [69]. However, nitroxides are prone to reduction and cannot be used e.g. for in-cell applications without extensive modification to supply steric shielding. However, the increase in size of the spin label necessary to achieve this results in an increased potential to perturb the structure of the labeled protein. In recent years, lanthanide spin labels based on complexes of the Gd(III) ion have been shown to be redox-stable, and have been successfully used to demonstrate EPR distance measurement in-cell [118, 103, 162]. Gd(III) spin labels also have drastically different spectroscopic properties than nitroxide labels, making them

much better suited for use with EPR at Q-band frequencies and above [117, 65, 56]. The high-spin nature of Gd(III) gives rise to stronger dipolar interactions than does the $S = 1/2$ nitroxide label, so that longer distances may be measured between pairs of Gd(III) spin labels. Furthermore, the linewidth of the central transition of Gd(III) narrows with increasing magnetic field, giving rise to greater sensitivity to dipolar interactions in the measured EPR signal compared to nitroxide spin labels whose EPR lineshape broadens with increasing magnetic field. Unfortunately, the high-affinity chelating agents necessary to avoid free Gd(III) ions are often bulky because affinity is largely controlled by denticity of the complex, or the number of donor groups in a ligand which bind to the Gd(III) in a coordinating complex. Besides Gd^{3+}, Cu^{2+} [43, 44, 177] and Mn^{2+} [15, 104, 86] have been shown to be interesting paramagnetic centers for structural studies of proteins with SDSL EPR. Spin label pairs can also be of different type - a method referred to as orthogonal spin labeling, where e.g. a nitroxide label is paired with a Mn^{2+} or Gd^{3+} label [60].

Another important property of spin labels is the conformational flexibility of both the label and the linker joining the label to the protein. If the spin label is very flexible, motions of the label will dominate EPR measurements of dynamics, rather than the desired dynamics of the protein under study. Different strategies have been employed to address this, including the use of short, rigid linkers [176] or bi-dentate spin labels [173] which bind to two neighboring amino acids on the protein and thereby reducing motions of the spin label relative to the protein backbone.

Two of the most exciting frontiers of SDSL EPR are *in vivo* measurements and distance measurement at physiological temperatures. In both cases, Gd(III) spin labels have yielded important breakthroughs and have the potential to greatly advance the use of EPR for structural studies in cell biology [135]. In the future, Gd(III) spin labels and high-field EPR together may enable the ability to measure triggered conformational dynamics of membrane proteins.

1.2.2 Distance measurement by EPR

The determination of nanometer-scale distances between spin labeled sites in a biomolecule via EPR spectroscopy can provide information on the structure and organization of biomolecules, and can also be used for the tracking of conformational changes [157, 74, 80, 75, 28]. The basic concept for all EPR distance measurement relies on the simple observation that the strength of dipole-dipole coupling, which can be inferred from various EPR experiements, goes as the inverse of the distance between the spin labels cubed. EPR distance measurements are complementary to Förster resonance energy transfer (FRET), where distances between a donor and an acceptor fluorophore are detected when the labels are positioned near the Förster radius. While this technique is very sensitive, allowing for single-molecule detection, precise distances and distance distributions are often not measured. Additionally, the labels must be carefully chosen based on the expected distance(s). On the other hand, EPR allows for precise, quantitative, determination of distances and distance distributions over a wide range of distances ($\sim$

0.8 - 10 nm). Labeling is also greatly simplified by the ability to use identical spin labels, which are also significantly smaller than typical fluorophores [135].

Nitroxide radicals are very commonly used for distance measurement with EPR at X-band frequencies, with the spin labels placed at user selected sites to generate a system with two proximal electron spins either by doubly spin labeling a biomoledule or by multimer formation of singly spin labeled biomolecules [151, 140]. The distance between two such spin labels may be determined by pulsed EPR techniques such as DEER/PELDOR, which can detect dipolar interactions between nitroxide radicals up to 8.0 nm apart at X-band frequencies, but has limited utility below 2.0 nm [137, 82, 165]. This distance range is limited at the upper end by short phase-memory relaxation times (T_m) and at the lower end by complications of the exchange coupling becoming of comparable or greater magnitude than the dipole-dipole coupling. Distance measurement by pulsed EPR is complemented by lineshape analysis of the CW EPR signal, which, as demonstrated in the seminal work of Rabenstein and Shin [122] and followed by many others [153, 76], allows for the determination of distances in the range of approximately 0.7 - 2.0 nm under favorable conditions for nitroxide radicals attached to a macromolecule. CW EPR finds particular utility in the study of membrane proteins, offering a site-specific probe of structure and dynamics in native or native-mimicking environments under ambient solution conditions.

Despite the great success of nitroxide radicals for structure studies of biomolecules at X-band frequencies, distances in the borderline region of applicability of CW EPR and

pulsed EPR, particularly in the 1.6 - 1.9 nm range, still remain difficult to access [16, 29]. Given the importance of discerning the structure and structural changes in biomolecules within this distance regime, we seek in this work to develop spin labels and EPR techniques which can serve to span this gap, deliver rapid measurement times, and work at ambient temperatures. We choose to focus on CW rather than pulsed EPR methods, due to better sensitivity and easier application to measuring samples at physiologically relevant temperatures. Unlike pulsed EPR, which typically requires cryogenic temperatures and hours to days of signal averaging, CW EPR can be used directly at ambient temperature with often a single, minutes long scan providing sufficient signal for analysis. Emerging trityl spin labels present one possible alternative to traditional nitroxide spin labels, as their very narrow linewidth allows for CW EPR distance measurements in the range of 1.7 - 2.4 nm. These new labels also show promise for single-frequency pulsed EPR experiments, where a long phase memory time and better stability allow for room temperature distance measurements on immobilized biomolecules [133, 91, 141, 142, 6]. In this dissertation we investigate another alternative, namely spin labels based on Gd(III) complexes measured by high-frequency CW EPR.

1.3 Benefits and complications of moving to high frequencies and Gd(III)-based spin labels

Increasing availability of microwave sources and components has allowed routine EPR measurements to move from X-band and Q-band, up to W-band and higher frequencies. Several EPR spectrometers, both home-built and commercial, are now in operation in the range of 100 - 300 GHz [114, 63, 66, 147, 163, 73, 47, 38]. At these high frequencies, sensitivity is increased [138], but the performance of nitroxide spin labels wanes due to strong inhomogeneous line broadening. During the last 10 years, spin labels based on Gd(III) have become an important tool for structure studies in biological systems with EPR at Q-band frequencies and above. High-frequency EPR applications relying on Gd(III) complexes have been demonstrated in peptides [68, 60, 106, 102], nucleic acids [149], proteins [117, 175], and in in-cell environments [118, 103, 162, 105].

Gd(III)-based spin labels possess several spectroscopic advantages over nitroxide-based spin labels for EPR at high fields. These benefits include: (i) a high-spin $S = 7/2$ ground state which confers longer range dipolar interactions than an $S = 1/2$ system and thus increases the accessible distance range in CW EPR measurements, (ii) an isotropic g-value, (iii) a sharp central $|-1/2\rangle \rightarrow |1/2\rangle$ transition that narrows with increasing magnetic field, and (iv) no orientation selection, the latter three of which concentrate the distance information into a single narrow resonance peak, resulting in increased sensitivity.

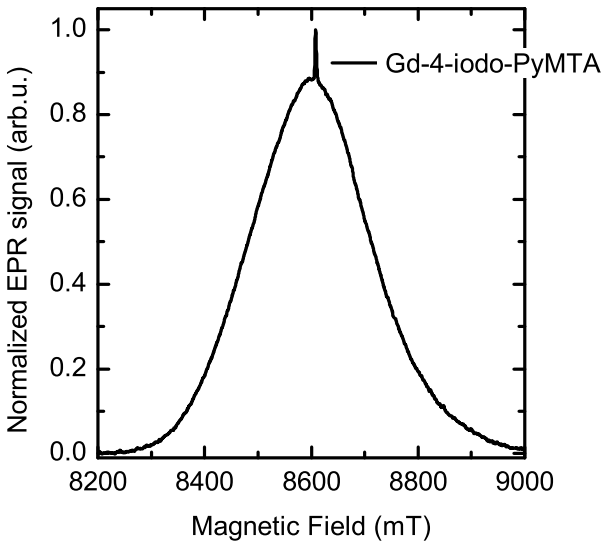

Figure 1.1: Typical Gd(III) EPR spectrum at 240 GHz, in this case for 300 μM Gd-4-iodo-PyMTA (chemical structure in Figure 2.1) in D_2O/glycerol$-d_8$ (60:40, v:v) as measured by rapid passage EPR at 5 K. The sharp central peak primarily results from the $|-1/2> \rightarrow |1/2>$ transition, which to first order in perturbation theory narrows with increasing magnetic field. The broad bell-shaped part of the spectrum is the result of the remaining transitions in the $S = 7/2$ Gd(III) complex, which are broadened in glassy solutions by a broad distribution of ZFS parameters.

For Gd(III) complexes with sufficiently small zero-field splitting (satisfying $D \ll g\mu_B B_0$) and measured at suffiently high frequencies (K_a-band and above), the $|-1/2\rangle \rightarrow |1/2\rangle$ transition appears as an intense central peak in the EPR spectrum on top of a broad background due to all other transitions (Figure 1.1) [125]. This characteristic shape of the EPR spectrum arises from the high-spin ($S = 7/2$) nature of Gd(III) and the broad distributions of zero-field splitting (ZFS) parameters for Gd(III) in frozen glassy solutions. The magnitude and distribution of the ZFS in turn arises primarily from the interactions between the Gd(III) ion and the ligands, from a chelate complex and/or solvent molecules, which bind the Gd(III). The details of the ZFS interaction have been found to influence many of the effects observed when using Gd(III) spin labels for EPR, including distortions of the Gd(III)-Gd(III) distance distributions measured by the DEER experiment at short distance ranges [45, 128, 101], population transfer in Gd(III)-Gd(III) DEER experiments [48], the effect of the reduction of the Gd(III)-nitroxide DEER echo intensity [179, 97], the width and shape of the central Gd(III) transition [41], the absence of orientation selection for Gd(III) in the DEER experiment [127], and the transition-dependent transverse relaxation of the Gd(III) complexes [124]. In order to fully understand these experimental observations, the particulars of the magnitude and distributions of the ZFS parameters in Gd(III) complexes must first be understood. The particulars of the ZFS in high-spin systems also creates a depth of additional interesting spin physics which may be studied and utilized. And furthermore, unraveling how the ligand structure of Gd(III) complexes determines the magnitude of the ZFS could help in the design of new Gd(III)

spin labels that are customized for the best performance with high-field CW or pulsed EPR experiments.

The width of the central transition in CW EPR measurements of Gd(III) scales as $D^2/g\mu_B B_0$. For small D and large B_0, this leads to a very narrow linewidth and therefore greater sensitivity to dipolar broadening of the CW EPR lineshape at high magnetic field as compared to lower magnetic fields where the central transition is prohibitively broad [117, 65]. As a result of this narrow linewidth, it is possible to resolve the contribution of the dipolar interaction with a proximal Gd(III) species to the lineshape of the CW EPR spectrum, manifested primarily as an increase of the linewidth of the $|-1/2\rangle \rightarrow |1/2\rangle$ transition. This was recently shown in a proof-of-principle study by Edwards *et al.* in frozen glassy solutions of $GdCl_3$ in D_2O/glycerol$-d_8$ at 240 GHz and 10 K [53]. In this work, substantial dipolar broadening was seen up to average Gd(III)-Gd(III) distances of 3.8 nm, which is about twice the longest interspin distance resolvable by CW EPR lineshape analysis of nitroxide radicals at X-band. This is an important prospect, given that pulsed EPR methods using Gd(III)-based spin labels are particularly prone to artifacts in the extracted distance distributions below 3 nm [45, 128, 101] and X-band CW EPR with nitroxide is only sensitive to distances of less than 2 nm.

Extending the CW EPR distance measurement sensitivity to above 3 nm is an important milestone because this length scale spans many relevant intra-protein distances, as well as many inter-protomer distances in protein oligomers. For a sense of scale, consider a pair of nitroxide-based spin labels at X-band, which can easily give access to informa-

tion on an interspin distances of 1.5 nm, and ask what fraction of a typical protein can be probed. Most proteins fold into globular domains, consisting of tightly packed atoms with an approximate density of 1.37 g/cm^3. A spherical region of mass M within such a globular protein has a radius R (nm) = 0.6 $\times$ $M^{1/3}$, where the mass is given in Daltons [55]. A sphere of radius R = 1.5 nm encloses a protein mass of 18 kDa, encompassing approximately 160 amino acids that could potentially be spin labeled. Now consider doubling this radius to 3 nm, a distance easily accessible by Gd-based spin labels at 240 GHz. This yields a protein mass of 94 kDa containing 850 amino acids - nearly eight times as many residues in the protein that can potentially be spin labeled for the measurement of intra-protein distances. This increased flexibility in choosing sites for mutagenesis and spin labeling is particularly useful in cases where e.g. little is known about the structure of a protein or protein complex and therefore the distance estimated from a model of the protein may be different from the real distance, or in cases where mutation of certain amino acids is not possible due to the nature of the mutation (e.g. charged amino acids) or is limited by geometric restriction of the site.

Chapter 2

Distance measurement by Gd(III) lineshape analysis

Material in Capter 2 and portions of the introductory material in Chapter 1 have been previously published, and is reproduced from Ref. [41] with permission from the PCCP Owner Societies.

JAC performed the CW EPR experiments and data analysis under the supervision of Mark Sherwin and Songi Han. Mian Qi designed and synthesized the Gd(III) compounds used in this chapter under the supervision of Adelheid Godt.

2.1 Introduction

The first report of using 240 GHz CW EPR lineshape analysis for distance measurement with Gd(III) used as a simple model system random frozen solutions of $GdCl_3$ in

D_2O/glycerol$-d_8$ [53] where the average interspin distance was varied by changing the concentration of $GdCl_3$. However, the distribution of distances in such a system is necessarily very broad, with the result that the measured spectra is a sum of the EPR spectra arising from the many interspin distances present in the sample. The measured CW EPR lineshapes of $GdCl_3$ were found to be nearly Lorentzian, with the primary change with average interspin distance being a broadening of the peak-to-peak linewidth. Additionally, the relatively high concentrations required to achieve short average interspin distances resulted in artifacts in the lineshape arising from refractive broadening due to high spin concentrations, further complicating analysis. In contrast, a single measurement presented in this work of a bis-Gd(III) complex with a flexible bridge (C2-595) and a broad distance distribution centered about 1.6 nm showed a tantalizingly complex lineshape in addition to strong peak-to-peak broadening.

When using site-directed spin labeling for distance measurement in biological systems, the typical use case is a doubly-labeled biomolecule where two spin labels are separated by a relatively fixed distance. Indeed, the motion of the linker connecting the spin label to the biomolecule is often a dominant contributing factor to measured distance distributions by this method. Therefore, in further developing CW EPR with Gd(III)-based spin labels as a distance measurement tool, we relied on a series of model compounds of the type Gd-PyMTA–spacer–Gd-PyMTA, which we call Gd-rulers, to provide known, fixed distances between pairs of spin labels.

In this chapter, we present a fundamental study of the CW EPR lineshape analysis-

based approach to determining Gd(III)-Gd(III) distances at very high frequencies (240 GHz). A series of "Gd-rulers," molecules with well-defined Gd(III)-Gd(III) distances, were used to test the validity of this approach and provide a simple, well-characterized model system with which to further develop data analysis methods. Experiments with these Gd-rulers show that dipolar broadening of the central transition of Gd(III) is detectable in a biradical system at cryogenic temperatures, and that the maximal distance sensitivity is approximately twice what can be measured using X-band CW EPR with nitroxide spin lables, as was suggested in previous measurements of random solutions of $GdCl_3$ [53]. A simple procedure for simulating the dependence of the Gd(III) lineshape on interspin distance was developed, in which the broadening of the central transition is modeled as an $S = 1/2$ spin whose CW EPR lineshape is broadened through electron-electron dipolar interactions with a neighboring $S = 7/2$ spin. The well-known spacer stiffness, and therefore the calculable most probable Gd-Gd distances r of the Gd-rulers, as described below, allows for a careful check on the relationship between the measured EPR line broadening and the distance r, revealing that the peak-to-peak broadening follows the expected $1/r^3$ dependence of the dipolar interaction. Furthermore, we show that this correlation is maintained even at biologically relevant temperatures – from around the protein dynamical transition temperature ($\approx$ 190-220 K) [49], where proteins begin to exhibit biological function, up to room temperature for samples immobilized in a glassy matrix.

2.2 The Gd(III) EPR spectrum

The general effective spin Hamiltonian for a pair of two interacting Gd(III) ions A and B is written as [3, 138, 65]

$$\mathcal{H} = \sum_{i=A,B} [g_i \mu_B \overrightarrow{B}_0 \hat{S}_{Zi} + \hbar \cdot \mathbf{A}_i \cdot \hat{I}_i + \hbar \hat{S}_A \cdot \mathbf{D_i} \cdot \hat{S}_B] + \hbar \hat{S}_A \cdot \mathbf{T} \cdot \hat{S}_B \tag{2.1}$$

The first term gives the contribution from the isotropic electron Zeeman interaction, where μ_B is the Bohr magneton, B_0 is the external applied magnetic field, and g ∼1.992. At high magnetic fields this term dominates, resulting in eight Zeeman levels and seven allowed EPR transitions. The second term in Equation 2.1 is the contribution from hyperfine coupling, which arises from the ^{155}Gd and ^{157}Gd isotopes of gadolinium with a 30% combined natural abundance. However, the resulting hyperfine coupling is small - on the order of 16 MHz. For most Gd(III) complexes prepared with the natural abundance of isotopes, including all of those studied in this work, the hyperfine interaction is not resolved in EPR experiments and is typically ignored. However, these unresolved hyperfine interactions may contribute a small amount to the intrinsic CW EPR linewidth of a Gd(III) complex [30].

The third term in Equation 2.1 is the zero-field splitting (ZFS) interaction, which for most Gd(III) complexes used as spin labels is of order 1000 MHz. Typically, the two Gd(III) are the same and therefore their principal D values are the same and the subscript i can be dropped. Theory and measurements of the ZFS interaction related to high-field EPR of Gd(III) complexes will be addressed more fully in Chapter 3. The

magnitude of the ZFS interaction in Gd(III) is small compared to the electron Zeeman interaction and thus can be treated by perturbation theory. To first order in perturbation theory, the EPR transition frequencies are

$$\omega^{(1)}_{M_S \to M_{S+1}} = \frac{1}{\hbar} g\mu_B B_0 + \frac{2M_S + 1}{2} \left[D \left(3cos^2\beta - 1 \right) \right] \tag{2.2}$$

where β is the angle between the static magnetic field $\overrightarrow{B}_0$ and the ZFS tensor. To first order in perturbation theory, the linewidth of the $|-1/2\rangle \to |1/2\rangle$ transition of Gd(III) is independent of the ZFS interaction, while the remaining higher-order transitions scale linearly with the second-order axial ZFS parameter D. Carrying out perturbation theory to second order gives the following contribution to the central transition

$$\omega^{(2)}_{1/2 \to 1/2} = \frac{\hbar D^2}{16 g\mu_B B_0} \left(4S(S+1) - 3 \right) \left[-2 \left(sin^2 2\beta \right) + \left(sin^4 \beta \right) \right] \tag{2.3}$$

To second order, the linewidth of the $|-1/2\rangle \to |1/2\rangle$ transition scales with $D^2/g\mu_B B_0$. Furthermore, if $\frac{4\hbar D^2}{g\mu_B B_0} \ll 1$, the central transition will be narrow. This results in the characteristic high-field Gd(III) EPR spectrum (Figure 1.1) consisting of an intense narrow central peak arising from the $|-1/2\rangle \to |1/2\rangle$ transition and a broad featureless background due to all other transitions which are smeared out by the large distribution of ZFS parameters, which will be discussed fully in Chapter 3. At fields of 8.6 Tesla (240 GHz), this portion of the Gd(III) spectrum can have peak-to-peak linewidths as narrow at ∼0.5 mT while the broad component spans ∼1 T, and can be leveraged to measure dipolar broadening induced by neighboring electron spins at distances of up to several

nanometers.

The final term in Equation 2.1 is the electron spin dipole-dipole coupling, where **T** is a tensor describing the total interaction between two electron spins. The dominant contribution to this term comes from the secular part of the dipolar interaction, which in the point dipole approximation is given by

$$\mathcal{H}_{dd}^{secular} = \omega_{dd}^{0} S_Z^A S_Z^B (3cos^2\theta - 1) \tag{2.4}$$

and where the magnitude of the dipolar coupling is given by

$$\omega_{dd}^{0} = \frac{\mu_0}{4\pi} \frac{g_A g_B \mu_B^2}{\hbar} \frac{1}{\overrightarrow{r}_{AB}^3} \tag{2.5}$$

where θ is the angle between the interspin vector r_{AB} and the static magnetic field B_0, g_A and g_B are the g factors of the two spins, and μ_B is the Bohr magneton. For short interspin distances and small ZFS the pseudo-secular part of the dipolar interaction

$$\mathcal{H}_{dd}^{pseudosecular} = -\frac{\omega_{dd}^{0}}{4} (S_+^A S_-^B + S_-^A S_+^B)(3cos^2\theta - 1) \tag{2.6}$$

will also contribute significantly to the measured central transition.

The secular part of the dipolar interaction is energy conserving, and is therefore the dominant contribution. The pseudo-secular term is energy conserving only if a flip-flop interaction occurs between two spins with identical transition energies. While an energy conserving flip-flop is relatively rare given the inhomogenous broadening present in many systems of interest for CW EPR, this term can become important for $S = 7/2$ spin

systems. The remaining terms in the dipolar interaction are not energy conserving, as the raising and lowering operators in these terms require a spin flip. Therefore, the contributions from these terms will be small and are usually ignored. A more complete description of the terms in the dipole-dipole coupling term, often referred to as the "dipolar alphabet" can be found in the dissertation of Edwards [51].

2.3 Gd-rulers as a model system

The Gd-rulers are based on oligo(*para*-phenyleneethynylene)s (oligoPPEs), which serve as molecular building blocks for tailor-made spacers. Both ends of the rod-like spacer molecule are terminated by the ligand PyMTA (a pyridine-based tetracarboxylate) which binds the Gd(III) ion and is connected to the oligoPPEs via its pyridine ring. The PPE spacers were made highly hydrophilic by the addition of porpargyloxy side chains on the oligoPPEs which attach to water solubilizing poly(ethylene glycol) (PEG) chains through a click chemistry reaction after spacer assembly [119].

A series of six water-soluble Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3** spanning a Gd-Gd distance range of 1.2 nm to 4.3 nm were used in this work (Figure 2.1). Gd-rulers of this type, and rulers formed with the same spacers but different spin labels, have been carefully characterized in the literature [84, 64, 48, 131, 45, 86]. These previous studies confirmed that the Gd-rulers are very rigid, resulting in very narrow distance distributions, and additionally experimentally determined the length and flexibility of the oligoPPE spacers.

The most probable interspin distance for each ruler at a given temperature may be

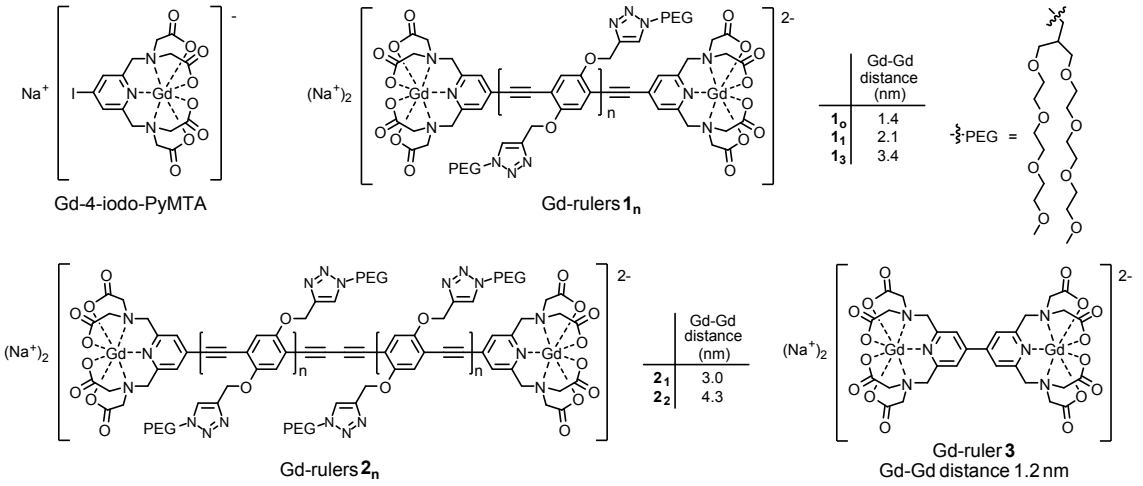

Figure 2.1: Chemical structures of Gd-4-iodo-PyMTA and of Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3**. Gd-4-iodo-PyMTA served as a reference for the intrinsic CW EPR lineshape of the spin label in the absence of dipolar broadening. The listed Gd-Gd distances are the calculated most probable distances at 173 K, i.e. the glass transition of a 60:40 (v:v) mixture of D_2O and glycerol$-d_8$ used as the matrix for the EPR experiments at 30 K.

Gd-ruler / T(K)	**3**	$\mathbf{1_0}$	$\mathbf{1_1}$	$\mathbf{2_1}$	$\mathbf{1_3}$	$\mathbf{2_2}$
173	1.17	1.42	2.09	3.00	3.41	4.30
215	1.16	1.42	2.08	2.98	3.40	4.27
288	1.16	1.41	2.07	2.96	3.36	4.22

Table 2.1: Calculated most probable Gd-Gd distances of the Gd-rulers at 173 K, 215 K, and 288 K using the worm-like chain model. Units are in nm. For the measurements at 30 K, the shape of the Gd-rulers became frozen at the glass transition temperature of the mixture of D_2O and glycerol$-d_8$ (60:40, v:v), which is 173 K. Therefore, the most probable Gd-Gd distances of Gd-rulers at 173 K were calculated for the measurements at 30 K.

calculated by applying a worm-like chain model, as described by Dalaloyan *et al.* [45]. For measurements at 30 K, we assume that the shape of the Gd-rulers became frozen at the glass transition temperature of the solution which was primarily composed of a D_2O/glycerol$-d_8$ (60:40, v:v) mixture. If the sample is cooled slowly, as was done in the data presented in this chapter, the glass transition temperature of this mixture is 173 K as determined by differential scanning calorimetry (DSC). In this slow cooling process, a part of the water mixture crystallizes to ice at 221 K and the residual mixture transforms to glass at 173 K. If instead the samples were shock frozen, the glass transition temperature of the mixture is 157 K. For measurements at temperatures above 173 K, the experimental temperature was used to calculate the most probable distance (Table 2.1). Calculation of the most probable distance is done by considering the Gd-rulers as a poly(*paraphenyleneethynylene*) chain and disregarding deviations in the stiffness of the Gd$-$N bond and the butadiyne moiety in Gd-rulers $\mathbf{2_1}$ and $\mathbf{2_2}$. The persistence length of poly(*para*-phenyleneethynylene) in *ortho*-terphenyl at the glass transition temperature of 246 K is known in the literature to be $l_{ps(246K)}$. Assuming the bending modulus is independent of temperature, the persistence length of poly(*para*-phenyleneethynylene) at the glass transition temperature of 173 K for a D_2O/glycerol$-d_8$ (60:40, v:v) mixture can be calculated by

$$l_{ps(173K)} = l_{ps(246K)} \frac{246K}{173K} \tag{2.7}$$

to be $l_{ps(173K)} = 23.7$ nm. The length of Gd-ruler $\mathbf{1_1}$ is known from crystal structure

Gd-ruler	$\mathbf{1_0}$	$\mathbf{1_1}$	$\mathbf{2_1}$	$\mathbf{1_3}$	$\mathbf{2_2}$
Max Gd-Gd distance	1.43	2.12	3.06	3.49	4.43

Table 2.2: Gd-Gd distances calculated by DFT for the Gd-rulers at maximum extension. Distances are given in units of nm.

to be $L(\mathbf{1_1}) = 2.12$ nm [71]. Taking this as a reference, the contour lengths l_{ct} of the Gd-rulers $\mathbf{1_n}$ and $\mathbf{2_n}$ are calculated by adding up the lengths of the constituent units of the spacers according to

$$l_{ct}(\mathbf{1_n}) = L(\mathbf{1_1}) + (n-1) \cdot L_{PE} \tag{2.8}$$

$$l_{ct}(\mathbf{2_n}) = L(\mathbf{1_1}) + (2n-1) \cdot L_{PE} + L_E \tag{2.9}$$

where n is the number of repeating units as given in Figure 2.1 and the lengths of the constituent units taken from DFT calculations to be $L_{PE} = 0.6864$ nm and $L_E = 0.2579$ nm. The mean Gd-Gd distances are then calculated according to

$$r = \sqrt{2l_{ps(173K)}l_{ct} - 2l_{ps(173K)}^2 \left(1 - exp\left(-\frac{l_{ct}}{l_{ps(173K)}}\right)\right)} \tag{2.10}$$

The Gd-Gd interspin distances of the Gd-rulers at maximum extension was computed by DFT calculations and can be found in Table 2.2. These values, along with the persistence length calculated in Equation 2.7, can be used to calculate the Gd-Gd distance distribution by treating the Gd-rulers as a semiflexible polymer with a large bending rigidity, as described by Wilhelm and Frey [174].

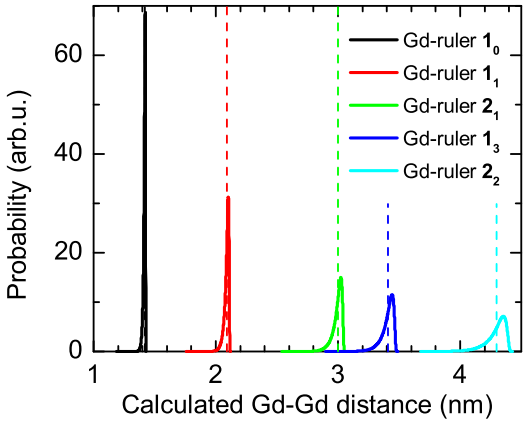

Figure 2.2: Calculated distance distributions (solid lines) and most probable Gd-Gd distance (dashed lines) for the Gd-rulers. These calculations assume the persistence length of the Gd-rulers is that calculated in Equation 2.7 at the glass transition temperature of 173 K for a D_2O and glycerol$-d_8$ (60:40, v:v) mixture.

2.4 Dipolar broadening in 240 GHz CW EPR lineshape of Gd-rulers

To investigate the effects of pairwise dipolar broadeing in Gd(III), a series of six water-soluble Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3** (Figure 2.1) spanning a Gd-Gd distance range of 1.2 nm to 4.3 nm were used. Because the dipolar coupling is determined primarily via a broadening of the CW EPR lineshape, it is necessary to have a measure of the intrinsic lineshape of the spin label in the absence of dipolar broadening [122, 153]. For this purpose, Gd-4-iodo-PyMTA, a Gd(III) complex closely resembling the spin label of the Gd-rulers (Figure 2.1) was included in this study. We assume a negligible effect of the type of substituent at the pyridine ring on the CW EPR lineshape (see also Appendix C.19). These samples were prepared for cryogenic measurements by dilution to 300 μM

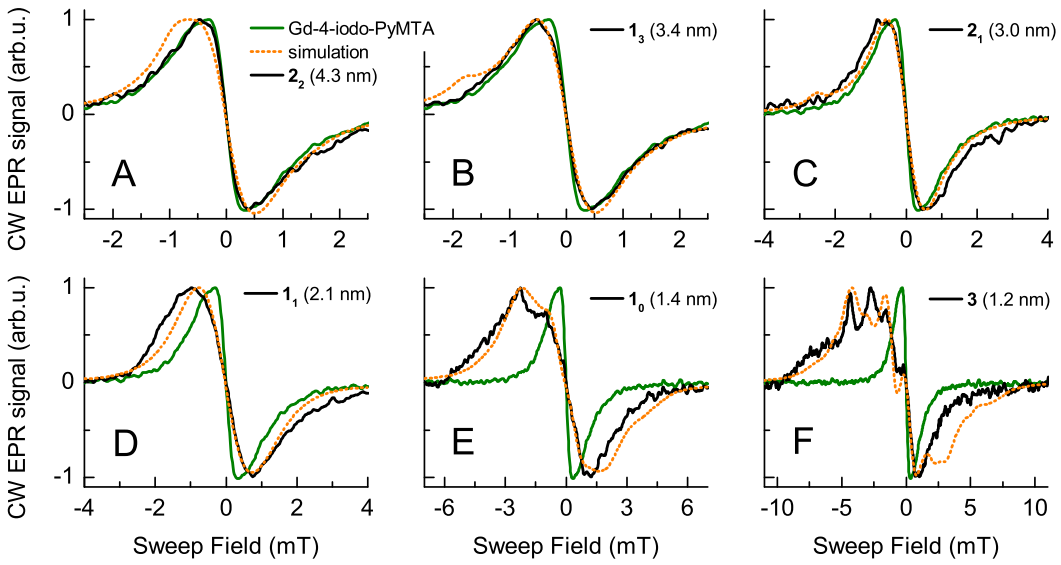

Figure 2.3: Lineshapes of the central transition of the Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3** (A-F, solid black curves) in D_2O/glycerol$-d_8$ measured by CW EPR at 240 GHz and 30 K. Overlaid in (A-F) in green is the measured lineshape of Gd-4-iodo-PyMTA, which serves as a reference of the intrinsic lineshape of the Gd-PyMTA spin label in the absence of dipolar broadening. In dotted orange is plotted the simulated lineshapes for each Gd-ruler resulting from simulations with a simple model consisting of an $S = 1/2$ spin whose CW EPR lineshape is broadened through electron-electron dipolar interaction with a neighboring $S = 7/2$ spin. In these simulations, the magnitude of the dipolar interaction was taken to be that corresponding to the calculated most probable distance for each Gd-ruler at the glass transition temperature of 173 K. (Table 2.1).

with a mixture of D_2O/glycerol$-d_8$ (0.6:0.4, v:v). A high fraction of glycerol was used to ensure good glass formation upon freezing. Deuterated solvents were used to minimize the broadening resulting from hyperfine interactions with neighboring water protons. A final sample concentration of 300 μM was chosen to avoid contributions to the lineshape from refractive broadening [53] and to ensure that the average intermolecular separation was sufficiently large so that only intramolecular dipolar interactions are observed in the lineshape.

For the initial investigation of dipolar broadening of the 240 GHz CW EPR lineshape

of the Gd-rulers with respect to Gd-4-iodo-PyMTA, measurements were conducted at 30 K, where the $|-1/2> \rightarrow |1/2>$ transition is nearly maximally populated (Figure C.2). A sample volume 8 - 10 μL was used, and loaded into a Teflon sample cup backed by a mirror within a modulation coil mounted at the end of the waveguide.

The CW EPR lineshape of the central transition of Gd-4-iodo-PyMTA is a single peak (Figure 2.3 A-F, in green). For the purposes of a simple quantitative measure of the dipolar broadening in the CW EPR lineshape of the Gd-rulers, the peak-to-peak linewidth was defined as the separation of the positive and negative peaks in the CW EPR lineshape (which is typically plotted as the as-recorded signal, which is the first derivative of the absorption lineshape). For Gd-4-iodo-PyMTA measured at 30 K, this peak-to-peak linewidth is ~ 0.77 mT. We take this linewidth to be the intrinsic linewidth of the Gd-PyMTA complex in the limit of no electron-electron dipolar interactions. Of course it is impossible to entirely remove all dipolar interactions in such a sample, however further lowering the concentration of Gd-4-iodo-PyMTA to 100 μM did not decrease the measured linewidth, while concentrations above 1 mM began to show evidence of lineshape broadening resulting from dipolar interactions between neighboring Gd-4-iodo-PyMTA complexes in the solution. The central transition of Gd-4-iodo-PyMTA and the Gd-rulers were also recorded by echo-detected (ED) EPR at W-band (95 GHz) and 10 K, as a corollary measurements to the DEER studies in the work of Dalaloyan *et al.* [45]. While W-band measurements are not the focus of this dissertation, it is worth noting that at this reduced field, the peak-to-peak linewidth of Gd-4-iodo-PyMTA was found to

be ∼1.6 mT. This increase in the intrinsic linewidth of the spin label from 240 GHz (8.6 T) to 95 GHz (3.4 T) is consistent with the expectation that the linewidth of the central transition scales as $D^2/g\mu_B B_0$.

The meausred 0.77 mT peak-to-peak linewidth of Gd-4-iodo-PyMTA is significantly broader than the 0.55 mT intrinsic linewidth of $GdCl_3$ in D_2O/glyercol$-d_8$ at 240 GHz [53]. This is largely the result of contributions of the ZFS term on the intrinsic lineshape of the Gd(III) complex. Very roughly speaking, the higher degree of symmetry in the atoms which coordinate the Gd(III), the smaller the ZFS and hence the narrower the intrinsic linewidth of the complex. $GdCl_3$, which has the nine possible coordination sites of Gd(III) fulfilled by deuterium molecules in solution, presents a narrower lineshape than Gd-4-iodo-PyMTA, in which the Gd(III) makes seven coordinating bonds with the PyMTA ligand with the remaining two possible coordinations being fulfilled by deuterium molecules in solution. This is at first glance concerning, since to extract distances from this measurement one must be able to resolve broadening of the CW EPR lineshape resulting from dipolar interactions on top of the intrinsic linewidth of the Gd(III) complex. If the intrinsic linewidth is to large, as is the case in some commonly used spin labels such as Gd-4MMDPA, the relatively small contribution to the linewidth from dipolar interactions cannot be resolved [53]. We are fortunate in this work that the 0.77 mT intrinsic linewidth of the Gd-4-iodo-PyMTA complex is still sufficiently narrow to resolve dipolar interactions at extended distances up to ∼3.5 nm. In further developing high-field CW EPR with Gd(III)-based spin labels as a distance measurement tool, much can be gained

by choosing existing or designing new spin labels with as small a ZFS as possible given other constraints. Therefore in application of this distance measurement technique in biological systems, as will be discussed in Chapter 4, spin labels based on the Gd-DOTA complex were chosen, a decision based in large part on the narrower intrinsic linewidth of the DOTA family of complexes.

The 240 GHz CW EPR spectra of all of the measured Gd-rulers at 30 K (in black) are overlaid with that of Gd-4-iodo-PyMTA (in green) in Figure 2.3 A-F. It is worth noting that all of the CW EPR lineshapes in Figure 2.3 were acquired with a single scan of less than 10 minute duration. While averaging is in principle possible on the 240 GHz EPR spectrometer, for most measurements it is not necessary to achieve an acceptable SNR. Additionally, this spectrometer does not currently employ a resonant cavity, as is common in most EPR spectrometers, which greatly simplifies sample loading and measurement. The speed and relative ease with which these kind of data are acquired greatly enhances the measurement throughput and broad applicability to biological samples.

For the longest Gd-ruler studied, $\mathbf{2_2}$ (4.3 nm), there is no significant difference between the measured lineshape of the Gd-ruler and of Gd-4-iodo-PyMTA (Figure 2.3 A). This is because broadening resulting from intramolecular dipolar coupling, while certainly present, is too weak at this distance to resolve on top of the intrinsic linewidth of the spin label. For Gd-ruler $\mathbf{1_3}$ (3.4 nm), we begin to resolve the first hints of dipolar broadening, and measure a slight broadening of the peak-to-peak linewidth to 0.97 mT (Figure 2.3 B). This is slightly shorter than the 3.8 nm Gd(III) - Gd(III) distances that Edwards

et al. projected Gd(III)-based spin labels to be sensitive to [53]. However, this is not surprising given that the intrinsic linewidth of Gd-4-iodo-PyMTA is ∼40% larger than that of $GdCl_3$ in D_2O/glyerol$-d_8$ [53], so we expect that our ability to resolve dipolar broadening of the CW EPR lineshape at extended distances using Gd-PyMTA as the spin label will be somewhat reduced compared to solutions of $GdCl_3$. For Gd-ruler $\mathbf{2_1}$ (3.0 nm), the broadening is much more obvious, with a peak-to-peak linewidth of 1.15 mT, an increase of nearly 50% with respect to the intrinsic linewidth given by Gd-4-iodo-PyMTA (Figure 2.3 C). For the Gd-rulers $\mathbf{1_1}$ (2.1 nm), $\mathbf{1_0}$ (1.4 nm), and **3** (1.2 nm), dramatic broadening of the peak-to-peak linewidth with respect to Gd-4-iodo-PyMTA is observed, with linewidths of 1.75 mT (Figure 2.3 D), 3.47 mT (Figure 2.3 E), and 5.22 mT (Figure 2.3 F), respectively. For the Gd-rulers $\mathbf{1_1}$ (2.1 nm), $\mathbf{1_0}$ (1.4 nm), and **3** (1.2 nm), the EPR lineshape becomes distinctly asymmetric with the low-field side becoming wider than the high-field side. For the two shortest Gd-rulers $\mathbf{1_0}$ (1.4 nm) and **3** (1.2 nm), a number of additional peaks are seen to appear in the low-field side of the lineshape, with the high-field side remaining virtually unchanged between these distances. Due to the complex lineshapes of the short Gd-rulers, the peak positions were determined by fitting the region about the peak with a cubic polynomial and determining the location of the maximum (see Section 2.6) [53].

2.5 Dipolar broadening in Gd-rulers at physiologically relevant temperatures

The Gd-rulers were additionally studied at elevated temperatures to investigate the feasibility of applying distance measurement with Gd(III)-based spin labels to structure studies of proteins at physiologically relevant temperatures in the future. The first such measurement was conducted at a temperature of 215 K, which was chosen by virtue of being at or near the protein dynamical transition [49] above which essential protein functions become active. At this temperature, the Gd-ruler samples, which were likewise prepared for 215 K measurements by dilution to 300 μM with a mixture of D_2O/glyercol$-d_8$ (40:60, v:v), should still be sufficiently viscous such that the dipolar coupling is not completely averaged out by molecular tumbling [67].

At 215 K, we measure an increase of intrinsic linewidth from Gd-4-iodo-PyMTA to 1.0 mT. This increase in linewidth with temperature was surprising, given that this was not observed in the work of Edwards et al. [53] nor in our simulations, which will be discussed later in this chapter. The precise cause of this will require further investigation, but may be related to changes in T_2 with temperature. At 215 K we also see a decrease in SNR, which is likely primarily the result of the sample now being well above the glass transition temperature for a mixture of D_2O/glyercol$-d_8$ (40:60, v:v), which will result in significant absorption of microwaves by the water present in the sample. This can be improved upon by the use of thin sample geometries for lossy aqueous samples to

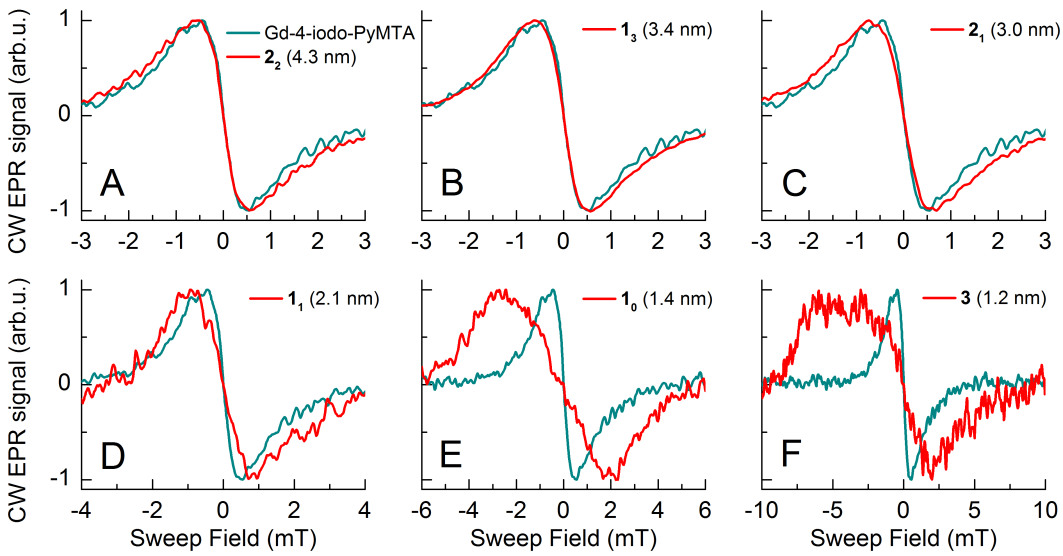

Figure 2.4: Lineshapes of the central transition of the Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3** (A-F, red) in D_2O/glycerol$-d_8$ measured by CW EPR at 240 GHz and 215 K. For reference of the intrinsic lineshape at this temperature, the spectra of Gd-4-iodo-PyMTA is overlaid in each plot (A-F, teal).

maximize B_1 at the sample position and minimize microwave absorption, as described in Appendix A.2.3. Nevertheless, the lineshape of the Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3** still display broadening with respect to Gd-4-iodo-PyMTA. However at this temperature, while it is possible to resolve a change in linewidth of the Gd-ruler $\mathbf{2_1}$ (3.0 nm) with respect to the Gd-4-iodo-PyMTA lineshape, for the Gd-rulers $\mathbf{1_3}$ (3.4 nm) and $\mathbf{2_2}$ (4.3 nm) any dipolar coupling effects fall within the noise of the measurement (Figure 2.4). This reduction in maximum resolvable distance at 215 K to approximately 3.0 nm from 3.4 nm at 30 K is not surprising, given the increase in the intrinsic linewidth and reduction in SNR with elevated temperature, as well as partially decreased dipolar broadening for elevated temperatures as a result of the changing spin populations with temperature.

Finally, the Gd-rulers and Gd-4-iodo-PyMTA were studied at 288 K, which is the

ambient temperature of the cryostat in the 240 GHz EPR spectrometer in the absence of heating or cooling. More importantly, we want to test the viability of lineshape analysis for distance determination near room temperature, which is the ideal environment for investigating protein dynamics and conformational changes. However, trying to extract information about the magnitude of dipolar coupling is difficult at elevated temperatures due to molecular tumbling [11]. Because the dipolar coupling tensor $\mathbf{T}$ (Equation 2.1) is traceless (in the simplest representation of the spin Hamiltonian), the dipolar part of the interaction is averaged to zero for fast and isotropic rotation of the spin label. For this reason, in much of the body of EPR literature looking at dipolar-coupled spins the sample is typically immobilized by freezing, by greatly increasing the viscosity, or by other methods.

In this study, we followed the method of Eaton and coworkers and others [109, 99, 112] where the Gd-rulers were immobilized in dehydrated glassy trehalose. The trehalose forms an amorphous matrix which is solid at room temperature and thus inhibits averaging out of the dipolar interaction by molecular tumbling [109, 112, 145]. The preparation of these samples is described in Appendix A.

At 288 K, the intrinsic linewidth of Gd-4-iodo-PyMTA was again found to be significantly broader (1.39 mT) than at lower temperatures. For the Gd-rulers $\mathbf{1_3}$ (3.4 nm) and $\mathbf{2_2}$ (4.3 nm), it was not possible to record a CW EPR spectrum with sufficient SNR to determine a peak-to-peak linewidth. This was attributed to difficulties in sample preparation, which was not thoroughly investigated and optimized in this work, and

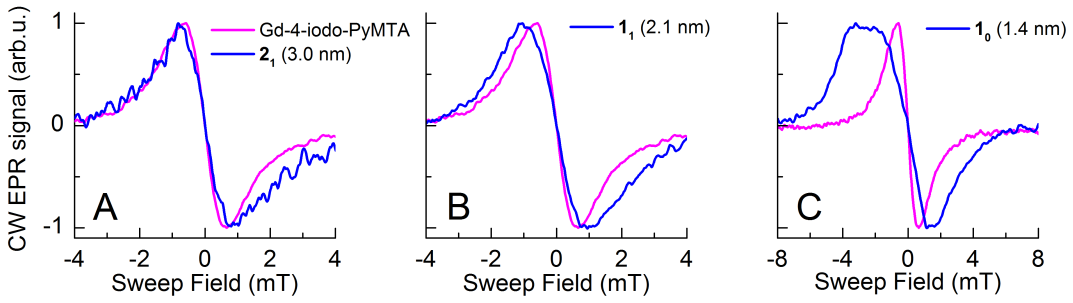

Figure 2.5: Lineshapes of the central transition of Gd-4-iodo-PyMTA (A-C, pink) and the Gd-rulers $\mathbf{1_n}$ and $\mathbf{2_n}$ (A-C, blue) in a dehydrated amorphous trehalose matrix measured by CW EPR at 240 GHz and 288 K. For the Gd-rulers $\mathbf{1_3}$ (3.4 nm) and $\mathbf{2_2}$ (4.3 nm), it was not possible to record a CW EPR spectrum with sufficient SNR for analysis. Gd-ruler **3** (1.2 nm) was not available at the time of these measurements.

not to an intrinsic limitation of the experimental technique. It was observed that the drying process took significantly more time for the longer Gd-rulers, possibly caused by the increased content of the hydrophilic PEG side chains in the longer Gd-rulers 2.1. If there was a partial demixing of the Gd-rulers from the trehalose matrix during the slow drying process, an increase in local concentration of the Gd-rulers may have occurred. The fragile, dehyrated sample was also very prone to static cling, making loading into the Teflon sample bucket difficult. Nevertheless, a broadening of the CW EPR lineshape with respect to Gd-4-iodo-PyMTA was observed for the Gd-rulers $\mathbf{1_0}$ (1.4 nm) and $\mathbf{1_1}$ (2.1 nm). The Gd-ruler **3** (1.2 nm) was not available at the time of these measurements. For Gd-ruler $\mathbf{2_1}$, broadening of the lineshape can no longer be unambiguously determined given the increased intrinsic linewidth and dramatically decreased SNR at this temperature as compared to frozen samples. Additionally, the possible demixing of the Gd-rulers from the trehalose matrix could contribute to the observed increase in linewidth and corresponding decrease in distance sensitivity. SNR of room temperature

EPR measurements can be similarly improved upon by the use of a thin-film aqueous sample as described in Appendix A.2.3 with the addition of a viscogen, such as glycerol or sucrose, to minimze molecular tumbling.

2.6 Quantifying broadening in CW EPR spectra of the Gd-rulers

There are many methods in the literature for extracting distance information from CW EPR data, the majority of which has been developed for $S = 1/2$ spin probes due to the popularity of nitroxide spin labels and relatively simple theory. Perhaps the simplest is a quantification of the peak-to-peak broadening of the CW EPR lineshape in a sample with dipolar coupling relative to the lineshape of the sample in the absence of dipolar broadening. Here, we define the peak-to-peak broadening to be the difference between the peak-to-peak linewidth of a Gd-ruler and the intrinsic linewidth in the absence of dipolar broadening, as given by the lineshape of Gd-4-iodo-PyMTA. While determining the positions of the positive and negative peak in the derivative CW EPR lineshape is straightforward in the case of a single, narrow peak, the more complex lineshape of the Gd-rulers $\mathbf{1_1}$ (2.1 nm), $\mathbf{1_0}$ (1.4 nm), and **3** (1.2 nm) along with the poor SNR in the 215 K and 288 K measurements necessitated slightly more intensive analysis to determine peak-to-peak linewidths as accurately as possible. First, the approximate locations of the positive and negative peaks in the CW EPR spectrum were found by looking for

zero-crossings in the smoothed first derivative of the measured CW EPR spectra that exceeded an input amplitude threshold chosen to exclude any zero-crossings that may be present in the noise in the baseline. The region about these approximate peak locations in the unsmoothed CW EPR spectra were then fit to a third-order polynomial and the extrema of the cubic fits taken as the location of the positive and negative peak positions in the experimental CW EPR spectrum. This procedure was repeated for several measurements of the lineshape of each Gd-ruler and Gd-4-iodo-PyMTA, and a weighted average computed to determine the linewidth of each compound. The uncertainty in measuring the peak-to-peak linewidths, as resulting from both SNR and the somewhat poor approximation of using a cubic fit to determine the positive peak position in e.g. Gd-rulers $\mathbf{1_0}$ (1.4 nm) and **3** (1.2 nm), was also calculated by propagating the errors in the computation of the weighted average [160].

A plot of the peak-to-peak broadening in the CW EPR spectra of the Gd-rulers vs. the calculated most probable Gd(III)-Gd(III) for each measurement temperature is shown in Figure 2.6. For all temperatures (and significantly different sample composition in the case of the 288 K measurements), the decrease of the peak-to-peak broadening with increasing interspin distance r scales as $1/r^3$. Such a correlation is expected from the form of the dipolar coupling term in the spin Hamiltonian (Equation 2.5). We can also estimate from Figure 2.6 that the upper distance limit for measuring significant dipolar broadening with a Gd-PyMTA type spin label is between 3.4 nm and 4.3 nm at 30 K. This is consistent with the $\sim$3.8 nm upper limit estimated by Edwards *et al.* [53] for

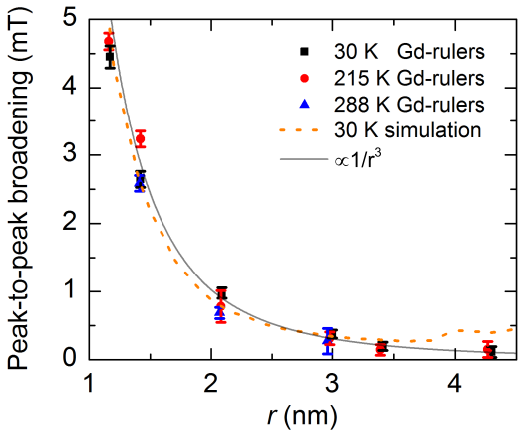

Figure 2.6: Peak-to-peak broadening in the 240 GHz CW EPR spectra of the Gd-rulers measured at 30 K, 215 K, and 288 K plotted as a function of the calculated most probable Gd(III)-Gd(III) distance for each temperature (Table 2.1). For all measured temperatures the peak-to-peak broadening falls off as $1/r^3$, as expected for electron-electron dipolar coupling (Equation 2.5).

resolving dipolar broadening between pairs of Gd(III). This is not expected to represent the absolute upper distance limit for this technique, as the distance sensitivity is highly dependent on the intrinsic linewidth of the Gd(III) complex used. At 30 K, Gd-4-iodo-PyMTA has a linewidth of ∼0.77 mT, which is somewhat broader than other complexes such as the Gd-DOTA family of spin labels whose linewidths are typically in the range of 0.5 - 0.7 mT at 240 GHz. The complex Gd-NO_3Pic has one of the narrowest CW EPR linewidths yet observed for a Gd(III) complex, of ∼0.45 mT at 240 GHz. If we require, as an example, a 25% increase of the intrinsic peak-to-peak linewidth to clearly resolve dipolar broadening (this approximately corresponds to the broadening measured for Gd-ruler $\mathbf{1_3}$ (3.4 nm)), then a distance of ∼4.0 nm should be resolvable for a spin label with a 0.5 mT intrinsic linewidth assuming a $1/r^3$ dependence on distance.

2.7 Modeling the lineshape of the Gd-rulers

In addition to this broadening of the peak-to-peak linewidth, a surprising amount of structure in the CW EPR lineshape begins to emerge for short interspin distances, an effect which was not previously observed in random solutions of $GdCl_3$, but was perhaps hinted at in the lineshape of a bis C2-Gd595 complex [53]. This work quoted an average interspin separation of ∼1.6 nm for the bis complex, and while a rather more flexible compound than the Gd-rulers studied in this dissertation, the complex lineshape of C2-Gd595 qualitatively shares many features in common with the lineshapes of Gd-rulers $\mathbf{1_0}$ (1.4 nm) and **3** (1.2 nm). In the analysis by Edwards *et al.* [53] on the bis C2-Gd595 complex using a Pake convolution method, the dipolar convolved spectrum reproduced well the experimental peak-to-peak linewidth of the experimental data, but failed to capture the finer details of the CW EPR lineshape of C2-Gd595. Here, we take a different approach to analyze the lineshapes of the Gd-rulers.

Exact simulations of the spin Hamiltonian to reproduce the high-field CW EPR lineshape of dipolar coupled Gd(III) spins are not yet available, due to the large computational cost and the requisite *a priori* knowledge of the magnitudes and distributions of ZFS parameters which was not available at the time of this work. However, a simple model was found to qualitatively reproduce many of the features observed in the 240 GHz CW EPR spectra of the Gd-rulers. In this model, we take as a starting point the spin Hamiltonian for a pair of interacting Gd(III) ions A and B as given by Equation 2.1. The observed CW EPR lineshape of the sharp central peak at 240 GHz is dominated by the

$|-1/2\rangle \rightarrow |1/2\rangle$ transition, with all other transitions being sufficiently smeared out by the broad distribution of ZFS parameters in a glassy sample that they are not explicitly resolved in this measurement. Therefore, we can approximate the central peak in the Gd(III) spectra as arising from an $S = 1/2$ spin. In the spectra of the Gd-rulers, we assume that the lineshape of this $S = 1/2$ spin is broadened through dipolar interactions with a neighboring $S = 7/2$ spin. Next, we must consider the impact of the other terms in the Hamiltonian on the lineshape of the central transition, namely the ZFS term and hyperfine interactions. In general, for Gd(III) complexes the magnitude of the ZFS term will be orders of magnitude smaller than the static magnetic field and will only affect the central transition to second order in perturbation theory, contributing primarily to the intrinsic linewidth of the Gd(III) complex. Hence, we introduce a further simplifying assumption that the ZFS interaction can be accounted for by imposing an artificial broadening on the line of the $S = 1/2$ spin.

Simulations of the 240 GHz CW EPR spectra for the Gd-rulers measured at 30 K were carried out based on the above simplifying assumptions using the EasySpin toolbox (version 5.0.16) [155] for MATLAB (Mathworks 2014a). The function `pepper`, which calculates CW EPR spectra for frozen solutions, was used with exact matrix diagonalization. The simulated spin system consisted of an $S = 1/2$ spin and an $S = 7/2$ spin with ZFS was ignored for the $S = 7/2$ spin. Simulations included as an input the experimental temperature of 30 K to account for the Boltzmann distribution of $S = 7/2$ spin populations (Appendix C.2). The lineshape of Gd-4-iodo-PyMTA was first

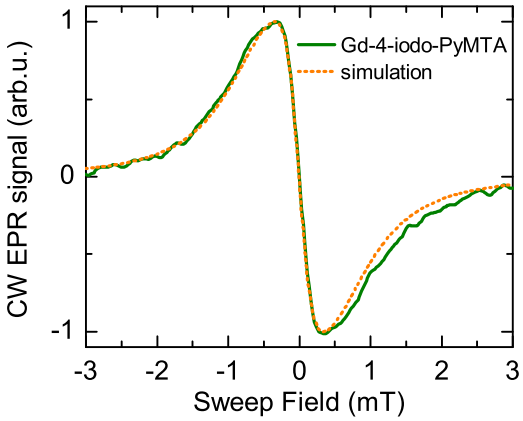

Figure 2.7: Measured 240 GHz CW EPR lineshape of Gd-4-iodo-PyMTA at 30 K with the corresponding simulated lineshape. The simulation is of a $S = 1/2$ spin was matched to the lineshape of Gd-4-iodo-PyMTA by imposing an artificial broadening in the form of a g-strain on the isotropic g-value of the $S = 1/2$ spin.

reproduced so that simulations of the Gd-rulers contained the proper intrinsic linewidth of the Gd-PyMTA complex. This was done by introducing an artificial broadening in the simulations as a strain of the isotropic g-value of the $S = 1/2$ spin. This g-strain was taken to be a Lorentzian distribution of g-values centered at $g = 1.992$ with a FWHM of 0.00028, chosen so that the resulting simulated lineshape of the $S = 1/2$ spin in the absence of dipolar interaction matches as closely as possible the measured CW EPR lineshape and the 0.77 mT linewidth of the central peak of Gd-4-iodo-PyMTA (Figure 2.7).

The isotropic g-value of the $S = 7/2$ spin was also taken to also be g= 1.992. A small Voigtian convolutional line broadening (`lwpp = [0.2 0.5]` in EasySpin) was included so that the simulated derivative CW EPR lineshape produced a smooth line. The g-strain on the $S = 1/2$ spin and the convolutional line broadening were the only free

parameters in these simulations and were determined by the criteria of reproducing the lineshape of Gd-4-iodo-PyMTA. To simulate the broadening in the CW EPR spectra of the Gd-rulers, an electron-electron interaction term was added to the simulation. This function in EasySpin computes the full interaction tensor including contributions from both the secular (Equation 2.4) and pseudo-secular (Equation 2.6) part the dipolar couling term. The through-space dipolar interaction between the two electron spins was input by defining the principal values of the traceless electron-electron interaction matrix in its eigenframe. The magnitude of the dipolar interaction (Equation 2.5) used in these simulations corresponds to that at the calculated most probable Gd(III)-Gd(III) distance for a Gd-ruler at 173 K (Table 2.1). Simulations were also run using the distance distributions for the Gd-rulers shown in Figure 2.2, but the Gd-rulers are sufficiently rigid that the effect of the small distance distributions on the simulated CW EPR lineshapes were negligible and therefore only the most probable distance was used in the final simulations. The results of these simulations are overlaid on the spectra of the Gd-rulers in Figure 2.3. Remarkably, given the many simplifying assumptions these simulations are based on, the measured lineshape of the Gd-rulers are well described by the simulations. Particularly in the case of the short Gd-rulers $\mathbf{1_0}$ (1.4 nm) and **3** (1.2 nm), many of the details of the measured lineshape - including the approximate positions and amplitudes of the splittings resulting from interaction with an $S = 7/2$ spin - are reproduced. Furthermore, the dramatic broadening and complex lineshape observed for Gd-Gd distances of 1.2 nm and 1.4 nm suggest that sub-Ångstrom Gd-Gd distance discrimination within this range

is possible with Gd-based spin labels. While not quantitative at this level, these simulations do validate the assumption that the observed broadening of the measured CW EPR lineshape of the Gd-rulers is arising primarily from the electron-electron dipolar interaction between two Gd(III) complexes. This allows us to confidently determined Gd(III)-Gd(III) distances from the magnitude of the dipolar broadening of the 240 GHz CW EPR lineshape.

Unfortunately, simulations of this type can become rather complicated at temperatures of 215 K or 288 K, i.e. at those temperatures which are physiologically relevant. Simulation of EPR spectra in the rigid-limit, as was used for the above 30 K simulations, is rather well defined in the literature, as is the rapid-tumbling limit. If samples prepared in glassy trehalose could be made such that there is a truly isotropic distribution of spin labels and spin label orientation, then the rigid-limit case may be used for room temperature simulations of CW EPR spectra. However, if we wish to measure some protein or other biological system which has been spin labeled with Gd(III) and is in a liquid solution, we can easily fall into an intermediate regime where the dipolar interaction and the dynamical effects of the protein and the motion of the spin label relative to the protein become important. It has been shown that there are at least some cases, for nitroxide-labeled proteins where the rotational correlation time of the interspin vector is sufficiently slow, where the rigid lattice case may still be used to a good approximation at physiologically relevant temperatures [11]. In this work, the rigid-limit simulation did not satisfactorily reproduce the high-temperature data, however it remains to be tested

whether or not this is the case for a Gd(III)-labeled protein.

2.8 Distance limits

The results of the lineshape analysis of 240 GHz CW EPR spectra of Gd-4-iodo-PyMTA and the Gd-rulers place the upper distance limit for measuring significant dipolar broadening between 3.4 nm and 4.3 nm at 30 K. However, this is not expected to represent an absolute upper limit for this technique, as the distance sensitivity of this measurement is highly dependent on the intrinsic linewidth of the chosen Gd(III) complex, as well as the SNR and general quality of the data. For Gd-4-iodo-PyMTA, the peak-to-peak CW EPR linewidth was measured to be 0.77 mT at 240 GHz and 30 K, while Gd(III) complexes such as Gd-DOTAM and Gd-NO_3Pic, which were previously shown to have very narrow EPR linewidths at X- and W-band frequencies [30], have linewidths of 0.53 mT and 0.45 mT at 240 GHz, respectively. This corresponds to even narrower EPR lines than the 0.55 mT observed for solutions of $GdCl_3$ in D_2O/glycerol$-d_8$ [53], allowing us to project that the distance sensitivity by CW EPR linehsape analysis can be further improved by an appropriate choice of Gd(III) complex for the spin label. If, conservatively, we assume that a 25 % increase of the intrinsic peak-to-peak linewidth can be clearly resolved as line broadening (as for Gd-ruler $\mathbf{1_3}$ (3.4 nm)), then assuming a $1/r^3$ dependence of the broadening on interspin distance, a measured distance of $\sim$ 4.0 nm should be achievable for a Gd(III) spin label with a 0.5 mT intrinsic linewidth. However, the rational design of new Gd(III) complexes is not so straightforward, and it remains to be seen whether

or not the chemical alterations needed to make such complexes suitable for use as spin labels will impact their intrinsic linewidth. This is discussed further in Section 3.9.

Achieving such narrow intrinsic linewidths for Gd(III) complexes in practical application to spin labeling proteins or other biological systems remains a difficult challenge. In order to employ a Gd(III) complex as a spin label it needs to be chemically altered, e.g. by the introduction of an MTS or maleimide linker which binds to a cysteine residue introduced into the protein at the site of interest by site-directed mutagenesis. The addition of such a linker can disrupt the the type, number, and organization of atoms in the ligand which bind the Gd(III) ion, which in turn can alter the intrinsic linewidth of the Gd(III) complex. Furthermore, non-covalent interaction with the compound to which the spin label is bound, such as by proteins or lipid systems, will likely impact the linewidth. It is also desirable to have Gd(III)-based spin labels be as small as possible, so that binding to buried or interfacial sites in a protein is not limited by steric hindrance. This is often difficult to achieve in practice, as relatively small Gd(III) complexes such as Gd-4MMDPA tend to have prohibitively large intrinsic linewidths, while those with very narrow intrinsic linewidths, such as Gd-NO_3Pic or Gd-DOTA, tend to be large in physical size. Furthermore, since small Gd(III) complexes such as Gd-4MMDPA make few coordinating bonds with the Gd(III), additional problems arise from the Gd(III) not being tightly bound to the ligand, resulting in possible separation of the Gd from the ligand or multiple 4MMDPA ligands coordinating a single Gd(III) ion. The intrinsic linewidth of the central transition of is, to a large extent, determined by the second-order

ZFS parameters of the Gd(III) complex. As will be discussed in Chapter 3, it is generally observed that the more symmetric, rigid complexes which satisfy as many as possible of the 9 coordinating bonds of Gd(III) have the smallest second-order ZFS parameter values and thus the smallest intrinsic linewidth. It remains a challenge for this work that the Gd(III) complexes which display the narrowest intrinsic linewidths also tend to be physically large compared to nitroxide spin lables, resulting in difficulties in spin labeling due to steric constraints.

The linewidth of the central transition may be further narrowed by moving to higher measurement frequencies, since the $|-1/2\rangle \rightarrow |1/2\rangle$ transition scales with $D^2/g\mu_B B_0$. This move to higher frequencies would be particularly advantageous for leveraging existing commercially available Gd-based spin labels such as Gd−4 MMDPA, whose intrinsic linewidth of 1.3 mT at 240 GHz is too broad for sensitive CW EPR distance measurements at this frequency [53].

The upper limit for measurable Gd-Gd distances is also expected to improve with improved data analysis methods. In this chapter, distances were extracted using only the peak-to-peak linewidth of the 240 GHz CW EPR spectrum. However, the Gd(III) lineshape can be rather complex, as seen in the case of the short Gd-rulers, and encodes a great deal more information than merely the peak-to-peak separations. Indeed, this simplification of the CW EPR spectra to a peak-to-peak broadening forces rather large error bars on the shortest Gd-rulers due to ambiguity in defining peak positions (Figure 2.6). The complex spectra observed for these short Gd-rulers should, in principle, allow for

more accurate fitting than the rather featureless spectra of the longer Gd-rulers. Other methods for distance determination which rely on such simplistic interpretations of the CW EPR spectra, such as second-moment analysis (Appendix B), present similar limitations. Additionally, the distribution of distances often contains information of equal value to that of the mean interspin distance. For example, distance distributions may inform on the distribution of conformers of a biomolecule or on aggregation resulting in multiple label-to-label distances, making them an important analysis target. CW EPR lineshape analysis should be sensitive to the interspin distance distribution, but the extraction of this information will require a better understanding of the various factors contributing to the CW EPR lineshape. The method of Pake convolution (deconvolution), extended to the $S = 7/2$ case by Edwards *et al.* [53], is one potential analysis method to determine distance distributions or multiple distances. However, while the definition of the Pake pattern may be extended to arbitrarily high-spin, it does not include other factors, such as zero-field splitting, which contribute significantly to high-spin EPR spectra. In order to appropriately account for such additional terms in the spin Hamiltonion, and the complicated interplay of the dipolar interaction with the zero-field splitting terms, one must simulate the full effective spin Hamiltonian. The simulations discussed above provide a start in this direction by capturing the primary contributions to the CW EPR lineshape, but a more detailed analysis would require simulation of the full effective spin Hamiltonian of two interacting $S = 7/2$ spins, including contributions from the zero-field splitting term. Dalaloyan *et al.* have shown that below a distance of 3.4 nm and with

small ZFS - conditions under which the high-field CW EPR technique is most sensitive - inclusion of the effects of the zero-field splitting parameter D on the pseudo-secular part of the dipolar interaction is crucial for extracting accurate distances and distance distributions from DEER of Gd-Gd systems [45, 128]. The interplay of high-spin, zero-field splitting, and dipolar interactions in systems of interacting Gd pairs was recently investigated theoretically by Manukovsky *et al.* [101], confirming the importance of considering the influence of the ZFS term on the resulting EPR spectrum when conducting distance measurements with Gd(III)-based spin labels.

In light of this work, one should consider as a first step to improving the CW EPR simulations presented here the inclusion of contributions from the ZFS term on the secular and pseudosecular parts of the dipolar interaction. In frozen glassy solutions, the second-order ZFS parameters are the most important and may be determined by independent measurements by the method(s) discussed in Chapter 3. However, while a full spin Hamiltonain simulation of the 240 GHz CW EPR spectra would greatly enhance our understanding of the contributions to the lineshape, such simulations are highly computationally intensive and may not be very practical for day-to-day use in distance measurement in e.g. protein systems. In the W-band DEER measurement of dipolar coupling of Gd(III) complexes, this can be mostly avoided by choosing Gd(III) complexes with large ZFS [45, 128, 101]. Unfortunately, this is not possible for the 240 GHz CW EPR measurement, since we rely very heavily on the Gd(III) complex having a narrow intrinsic linewidth which goes hand-in-hand with a small ZFS. Ideally, one would find

some simplifying approximation which would speed computation time while preserving the integrity of the simulation, by e.g. only diagonalizing the portions of the spin Hamiltonian which are most significant [45]. Another approach may be extending the Pake convolution method to include ZFS. While the high-spin extension of the Pake pattern proposed by Edwards *et al.* [53] does not allow for inclusion of the ZFS, other EPR simulation software, such as the EasySpin toolbox, may be able to provide an appropriate broadening function which does include the ZFS interaction.

2.9 Impact and Outlook

Lineshape analysis of 240 GHz CW EPR spectra of rigid Gd-rulers at 30 K demonstrates scaling of dipolar line broadening with $1/r^3$ and distance sensitivity ranging from 1.2 nm up to $\sim$ 3.4 nm when Gd-PyMTA is used as the spin label. The same $1/r^3$ dependence is observed at biologically relevant temperatures of 215 K and 288 K, with the upper distance limit reduced to $\sim$ 3.2 nm and $\sim$ 2.9 nm, respectively. The origin of this reduction in the upper distance limit with increasing temperature is not yet fully understood. Fortunately, we have no indication that these distances represent fundamental limits of this technique. The upper measurable distance by CW EPR could be further increased by using a Gd(III)-based spin label with a narrower intrinsic linewidth and/or moving to higher measurement frequencies. Additionally, further improvements in the analysis of the CW EPR lineshape would serve to greatly enhance the reliability of the measurement, and hopefully in future work be able to provide information on both mean distances and

their distributions. The lineshape simulations presented in this chapter represent a start towards this goal, but need to be further improved in order to understand all the details of the complex lineshape of interacting Gd-Gd pairs, particularly for short interspin distances.

The results in this chapter allow us to project that distance determination by lineshape analysis of CW EPR spectra recorded at very high frequencies using Gd(III) complexes as spin labels will be a highly useful technique for structure studies of complex biological systems where the application of pulsed EPR techniques, such as DEER, is challenging or when measurements above the solvent glass transition temperature are desirable or necessary. Practical applications of this technique will benefit from employing Gd(III) complexes with very narrow central EPR lines, independent of the local environment of the spin label. This calls for Gd(III) complexes with ligands filling all coordination sites of the Gd(III) ion, being resistant to substitution of any coordinating functional group moieties of the biomolecule, and being conformationally fixed as to keep the geometry of the complex independent of the environment. Work is already underway to develop improved Gd(III)-based spin labels for EPR which are more rigidly affixed to the protein, either by using a short rigid linker [1], or by binding the spin label to two nearby residues on the protein [173]. Recent reports indicate these new Gd(III) complexes give very narrow distance distributions when measured by W-band DEER, and would be also be interesting for high-field CW EPR applications.

Chapter 3

Zero-field splitting parameters of Gd(III) complexes

The material in this chapter is being prepared in a different format for publication as: Jessica A. Clayton, Katharina Keller, Mian Qi, Julia Wegner, Vanessa Koch, Henrik Hintz, Adelheid Godt, Songi Han, Gunnar Jeschke, Mark S. Sherwin, and Maxim Yulikov, "Quantitative analysis of zero field splittings in Gd(III) complexes."

MQ, JW, VK, and HH synthesized the Gd(III) complexes under the supervision of AG. KK performed the EPR measurements at Q-/W-band, worked on convergence criteria for the simulations, and performed analysis and calculations related to Model 1 under the supervision of MY and GJ. JAC performed the 240 GHz EPR measurements, worked on the convergence criteria for the simulations, and performed the analysis and calculations related to Models 2 and 3 under the supervision of MSS and SH. GJ performed

the analysis with the superposition model. All authors analyzed the experimental and computational results for all models.

3.1 Introduction

After the study of the CW EPR lineshapes of Gd-rulers presented in Chapter 2, it became evident that precise, detailed knowledge of the zero-field splittings of Gd(III) complexes would be of great importance for further developments in simulating the CW EPR lineshapes of Gd(III) complexes and for the design of new Gd(III) complexes for use as spin labels.

In glassy frozen solutions of Gd(III) complexes, the lineshapes of EPR transitions are dominated by the angle-dependent zero-field splitting (ZFS) term in the spin Hamiltonian. This interaction arises from interaction of the Gd(III) ion with ligands, as well as some relativistic corrections and configuration interaction terms arising from the two electron's spin-orbit coupling operators [111]. Due to the angular dependence of this term, there can arise cases of energy level crossings or resonant conditions when a single microwave frequency corresponds to two different EPR transitions with or without a level in common. Accordingly, several spectroscopic effects observed for Gd(III) complexes are connected to the mean values and distributions of the ZFS parameters. In particular, the following effects are thought to be influenced by the distribution of ZFS parameters: distortions of the Gd(III)-Gd(III) distance distributions measured by W-band DEER for short distances (<3.5 nm) [45, 128, 101], population transfer in Gd(III)-Gd(III) DEER

experiments [48], reduction of the Gd(III)-nitroxide DEER echo intensity [179, 97], the width and shape of the central Gd(III) transition which is relevant for the CW EPR based distance measurements discussed in Chapter 2 [41], the absence of orientation selection for Gd(III) measured by DEER [65], and the transition-dependent transverse relaxation of Gd(III) complexes [124].

A detailed analysis and understanding of these various spectroscopic effects requires knowledge of the distribution of ZFS parameters for the Gd(III) complex under study. Currently, quantum chemistry calculations, including DFT, do not allow for the determination of the ZFS parameters of Gd(III) complexes with the precision necessary for EPR applicaitons [93]. The computation of ZFS parameters is further complicated by the broad distributions of ZFS parameters observed for Gd(III) complexes in frozen solutions due to the dynamical rearrangement of the chelating molecule around the Gd(III) ions in the solution before freezing [179]. As such, the most accurate method of determining ZFS parameters is by the fitting of EPR spectra. Here, the quality of the EPR data, reliability of the model, and the fitting procedures used are of crucial importance. The major developments in this direction to date were in studies focused on the relaxivities of different Gd(III) complexes used for magnetic resonance imaging (MRI)[125, 22], where Gd(III) complexes are often employed as contrast agents. These previous studies used two somewhat different models for the distribution of ZFS parameters, which have since become widely adopted in the literature. However, no study has been undertaken to determine the precision and reliability of the ZFS parameters determined by these methods,

and the two models have never been directly compared.

Due to the importance of the reliable determination of these values, we undertake in this chapter a detailed investigation of the fitting of EPR spectra to extract ZFS parameter values. In this work, we fit a set of multifrequency EPR to the two models provided by literature for the distribution of ZFS parameters and explore important details of fitting procedures, compare the models, and compute typical errors bars for the determined parameters. Finally, we studied correlations between the structures of the Gd(III) complexes and the magnitudes and distributions of the ZFS parameters using a superposition model [111, 125].

3.1.1 Theoretical background

The full effective Gd(III) spin Hamiltonian was introduced in Chapter 2. Here, we focus on the contributions most important to the determination of the ZFS parameters of Gd(III) complexes in frozen glassy solutions.

Two out of six stable isotopes of Gd have magnetically active nuclei with nuclear spin of 3/2: ^{155}Gd (14.8 % natural abundance, $\mu/\mu_N = -0.2582$) and ^{157}Gd (15.65 % natural abundance, $\mu/\mu_N = -0.3385$) [96]. The hyperfine coupling with these nuclei is small, about 16 MHz, and is unresolved in the Gd(III) EPR spectrum [30]. Therefore hyperfine interactions are typically ignored in EPR simulations. The other four stable isotopes of Gd (^{154}Gd, ^{156}Gd, ^{158}Gd and ^{160}Gd) have zero nuclear spins. The main remaining contributions to the spin Hamiltonian of an isolated Gd(III) ion are the electron Zeeman

(EZ) interaction and the zero-field splitting (ZFS) interaction. The general form of this spin Hamiltonian can be written as

$$\hat{H} = \beta \left(\vec{B} \cdot \mathbf{g} \cdot \hat{\vec{S}} \right) + \sum_{k,q} B_k^q \hat{O}_k^q \tag{3.1}$$

where β is the Bohr magneton, $\vec{B}$ is the static magnetic field, $\mathbf{g}$ is the $\mathbf{g}$-tensor, $\hat{\vec{S}}$ is the total spin vector operator, $\hat{O}_k^q$ are spin operator equivalents for corresponding spherical harmonics, and B_k^q are numeric coefficients for each of the spherical harmonics operators using the notation of the extended Stevens operators. In the EPR spectral simulations performed in this work, we assume an isotropic $\mathbf{g}$-tensor that is described by a single $\mathbf{g}$-value of $g = 1.992$ [130, 31]. The sum can be expanded as

$$\sum_{k,q} B_k^q \hat{O}_k^q = B_2^0 O_2^0 + B_2^2 O_2^2 + B_4^0 O_4^0 + B_4^2 O_4^2$$

$$+ B_4^4 O_4^4 + \text{higher-rank terms} \tag{3.2}$$

where the B_2^i $(i = 0, 2)$ are the second-rank ZFS terms ($D \equiv 3B_2^0$ and $E \equiv B_2^2$) and the B_4^k $(k = 0, 2, 4)$ are the fourth-rank ZFS terms for half-integer spins. For $S \geq 3$, there are also sixth-rank ZFS terms (omitted above) [90, 3].

In principle, all of the coefficients B_k^q can be determined from EPR data. Such studies have been performed on Gd(III)-doped single crystals where the angular dependencies of the EPR transitions could be precisely measured and a large number of ZFS coefficients could be reliably determined [3, 34]. However, in all reported EPR studies so far of

Gd(III) complexes in frozen glassy solutions, a broad distribution of the second-order ZFS parameters is always observed [126, 65, 178]. Because of the broad distributions of the second-order ZFS parameters and no access to any angle-resolved information, it is not possible to determine the 4th and 6th order ZFS parameters for Gd(III) complexes in frozen glassy solutions. Furthermore, the higher-order ZFS terms, while present for Gd(III), are substantially smaller than the second-order terms and typically can be safely neglected [58].

We therefore model the EPR spectra of Gd(III) complexes under the simplification that only terms quadratic in the total electron spin operators are left in the spin Hamiltonian. The second-order ZFS term in the spin Hamiltonian is commonly parameterized by two coefficients D and E,

$$\hat{H}_{ZFS} = D \cdot \left(\hat{S}_Z^2 - \frac{1}{3} S(S+1) \right) + E \cdot \left(\hat{S}_X^2 - \hat{S}_Y^2 \right)$$

$$= 2D/3 \cdot \hat{S}_Z^2 + (-D/3 + E) \cdot \hat{S}_X^2 + (-D/3 - E) \cdot \hat{S}_Y^2. \tag{3.3}$$

This approximation has been demonstrated to be physically reasonable on a number of Gd(III) complexes in frozen glassy solutions, where typical fitted ZFS parameter distributions show only a very small fraction of species with nearly axial symmetry ($E \approx 0$) and an even smaller fraction of high symmetry cases with $D \approx 0$ and $E \approx 0$ where higher rank ZFS terms would become important [125].

We can write the eigenvalues of the ZFS operator in its eigenframe as D_X, D_Y and D_Z,

such that the parameters D and E are defined as $D = 3D_Z/2$ and $E = (D_X - D_Y)/2$. Therefore,

$$D_X = -D/3 + E$$

$$D_Y = -D/3 - E \tag{3.4}$$

$$D_Z = 2D/3$$

By convention, the D_Y value should lie between D_X and D_Z, and the relation

$$|D_X| < |D_Y| < |D_Z| \tag{3.5}$$

must hold true. By this definition, $|E| \leq |D/3|$, and D and E must have the same sign. The ZFS tensor is traceless, so we have the additional relation $D_X + D_Y + D_Z = 0$.

In order to determine the second-order ZFS parameters for a particular Gd(III) complex, we need to fit two distributions, $P(D)$ and $P(E)$, to the measured EPR spectra. In many cases, it is convenient to instead of $P(E)$ to fit the distribution $P(E/D)$, since it assumes the same range of values $0 \leq E/D \leq 1/3$ for any D. To further restrict the number of fit parameters, necessary as a result of the rather featureless EPR spectra characteristic of Gd(III) complexes, a simple model for the form of the distributions of $P(D)$ and $P(E/D)$ is typically assumed. This problem was tackled in two different ways in the previous reports of Raitsimring *et al.* [125] and Benmelouka *et al.* [22]. The

models for the ZFS parameter distributions proposed by these authors are summarized next.

3.1.2 Model 1 (*Benmelouka, et al.)*

The simplest model for the distributions of D and E in the ZFS term of the spin Hamiltonian (Eqn. 3.3) was described by Benmelouka *et al.* [22]. The authors assumed that the distributions of D and E in frozen glassy solutions of Gd(III) complexes can be described by two uncorrelated Gaussian distributions (Figure 3.1 (A)), which can be written as

$$P(D) = \frac{1}{\sqrt{2\pi\sigma_D^2}} \cdot \exp\left(-\frac{(D - \langle D\rangle)^2}{2\sigma_D^2}\right);$$

$$P(E) = \frac{1}{\sqrt{2\pi\sigma_E^2}} \cdot \exp\left(-\frac{(E - \langle E\rangle)^2}{2\sigma_E^2}\right). \tag{3.6}$$

The authors reported reasonably good agreement between EPR spectra of Gd(III) complexes and their simulations with this model for spectra measured at 240 GHz, Q-band, and X-band [22, 21]. Since the D and E values are linear combinations of the eigenvalues of the ZFS tensor, this model essentially assumes Gaussian distributions for the D_X, D_Y and D_Z values, with identical widths of distributions for D_X and D_Y. Note that if these distributions are broad, some combinations of D and E values will contradict Equations 3.4 and 3.5. Due to this conflict with the conventional definition of these parameters, the distributions of D and E were redefined by a reordering of the indices to be in line

with Equations 3.4 and 3.5. The properly redefined distributions for D and E appear bimodal, as described in more detail in Section 3.4. This situation is sketched in the Figure 3.1 (B). Furthermore, due to the conventional definition, small variations in the D_X, D_Y and D_Z values can shift the position of a point in the D and/or E distributions from the positive to the negative component of the distribution. The bimodality of the distributions is therefore a consequence of the definitions of these variables, rather than a true physical effect.

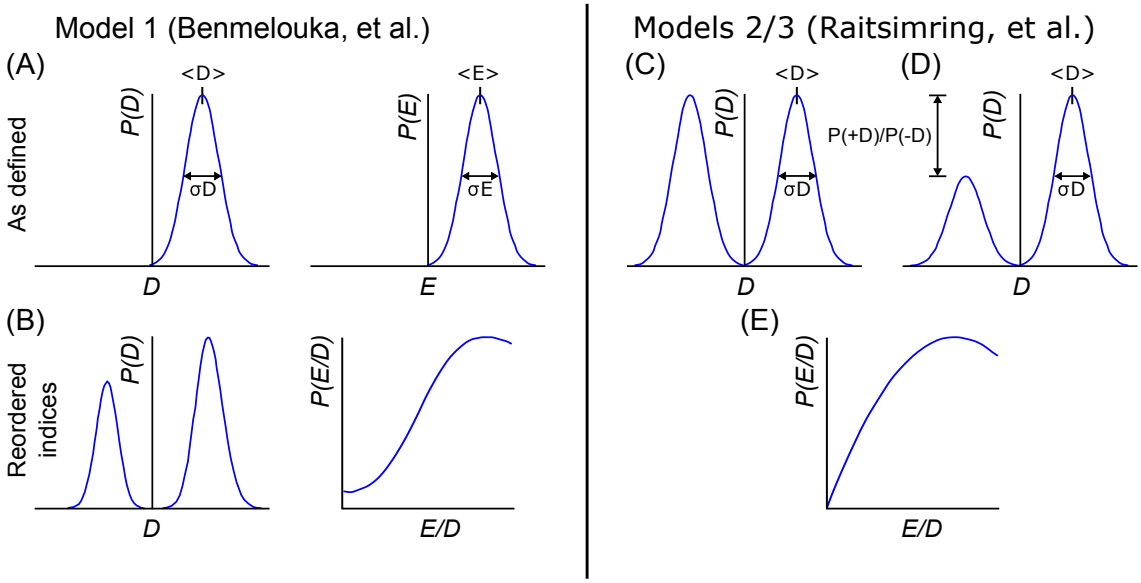

Figure 3.1: Graphical representation of the models for the distributions of the second-order ZFS parameters used in this work. (A) Model 1 assumes $P(D)$ and $P(E)$ are described by two uncorrelated Gaussian distributions. (B) Reshuffling of the indices to correct for inconsistencies in the definition of this model result in an approximately bimodal Gaussian distribution for D and E. (C) Model 2 assumes $P(D)$ is a bimodal Gaussian distribution, where the $+D$ and $-D$ contributions have equal amplitude and width. (D) Model 3 adds an asymmetry parameter $P(+D)/P(-D)$ to Model 2, which allows the relative amplitudes of the $+D$ and $-D$ contributions to vary. (E) For Models 2 and 3, E/D follows a polynomial distribution given by $P(E/D) \propto (E/D) - 2 \cdot (E/D)^2$.

3.1.3 Models 2 and 3 (*Raitsimring, et al.*)

Another approach to model the broad distributions of ZFS parameters D and E was suggested by Raitsimring *et al.* [125, 126]. The distributions of ZFS parameters were built computationally under the approximation that ZFS term can be represented as a linear combination of the ZFS contributions from single coordinating atoms in the ligand, where each coordinating atom is assumed to be identical and contribute an axial ZFS ($E = 0$) of magnitude D directed along the Gd-atom bond. This model was then incorporated into Monte Carlo simulations where the coordinating atoms were assumed to have randomly distributed positions over a metal ion sphere. This Monte Carlo modeling found that the D distributions are bimodal, with the centers of the two approximately Gaussian modes placed nearly symmetrically with respect to $D = 0$. This result was found to be qualitatively similar for many Gd(III) complexes with various numbers (7-9) of ligands binding the metal ion. In application to fitting experimental EPR spectra, this distribution was simplified to a bimodal Gaussian distribution, in which the positive D and negative D peaks in the $P(D)$ distribution were assumed to have equal amplitude and width (Figure 3.1 (C)). The distributions $P(E/D)$ were found to be slightly different for the positive D and negative D modes, but could be approximately described by a polynomial function of the form

$$P(E/D) \propto (E/D) - 2 \cdot (E/D)^2. \tag{3.7}$$

According to Equation 3.7, the maximum of the probability density function $P(E/D)$

corresponds to the value $E/D = 0.25$. At $E/D = 0$ (axially symmetric complexes) the probability density is exactly zero and $P(E/D)$ builds up approximately linearly for small values of E/D (Figure 3.1 (E)).

This model is a phenomenological assumption, as for typical Gd(III) complexes the coordinating atoms about the Gd(III) ion physically cannot assume any arbitrary orientation. Rather, the structure of the ligand dictates the average positions and mobility of the coordinating atoms. For ligands which offer fewer than nine coordinating atoms, the remaining coordination positions can be filled by solvent molecules. Nevertheless, the authors reported that simulations with this model successfully reproduced experimental EPR spectra of a series of Gd(III) complexes and the model has been widely adopted in the EPR literature [125, 126].

In this work, in order to discuss the effect of the bimodality of the distribution of the ZFS parameter D, we shall consider two versions of this model. In Model 2, we assume that the relative weights (amplitudes) of the positive D and negative D components of the $P(D)$ distribution are equal (Figure 3.1 (C)). In Model 3, we allow different relative weights of the positive and negative Gaussian modes in the D distribution, denoted by the ratio of amplitudes $P(+D)/P(-D)$ Figure 3.1 (D). This asymmetry in the bimodal $P(D)$ distribution was observed in the Monte Carlo simulations of Raitsimring *et al.* [125], and was found in this work to be necessary to account for the experimentally observed asymmetry of the Gd(III) EPR spectra at high fields. Both Models 2 and 3 use the definition of the $P(E/D)$ distribution given by Equation 3.7 and sketched in Figure

3.1 (E).

3.2 Gd(III) complexes

A series of six Gd(III) complexes were chosen to be included in this work (Figure 3.2). Gd-DOTA **(2)** was obtained commercially from Macrocyclics, Inc. and was used without further purification. The synthesis details of the remaining Gd(III) complexes, which include Gd-NO_3Pic **(1)**, Gd-maleimide-DOTA **(3)**, R-(Gd-PyMTA) **(4ab)**, Gd-TAHA **(5)**, iodo-(Gd-PCTA-[12]) **(6)**, and Gd-PyDTTA **(7)** will be described in the upcoming publication of this work [40]. The complex R-(Gd-PyMTA) **(4a)** is identical to the Gd-4-iodo-PyMTA complex used in Chapter 2 of this dissertation. Details of the sample preparation, EPR spectrometers used, and other experimental parameters can be found in Appendix A.

3.3 Numerical simulations

The EPR spectra of Gd(III) complexes were simulated in MATLAB with home written scripts based on the EasySpin toolbox [155]. Absorption powder spectra were computed using full matrix diagonalization with the EasySpin function `pepper`. The spin system structure in EasySpin was defined as a single spin $S = 7/2$ with an isotropic g-value of 1.992. The strains for g, D, and E were set to zero. Distributions of the D and E parameters were formed according to each of the three models presented in Sections

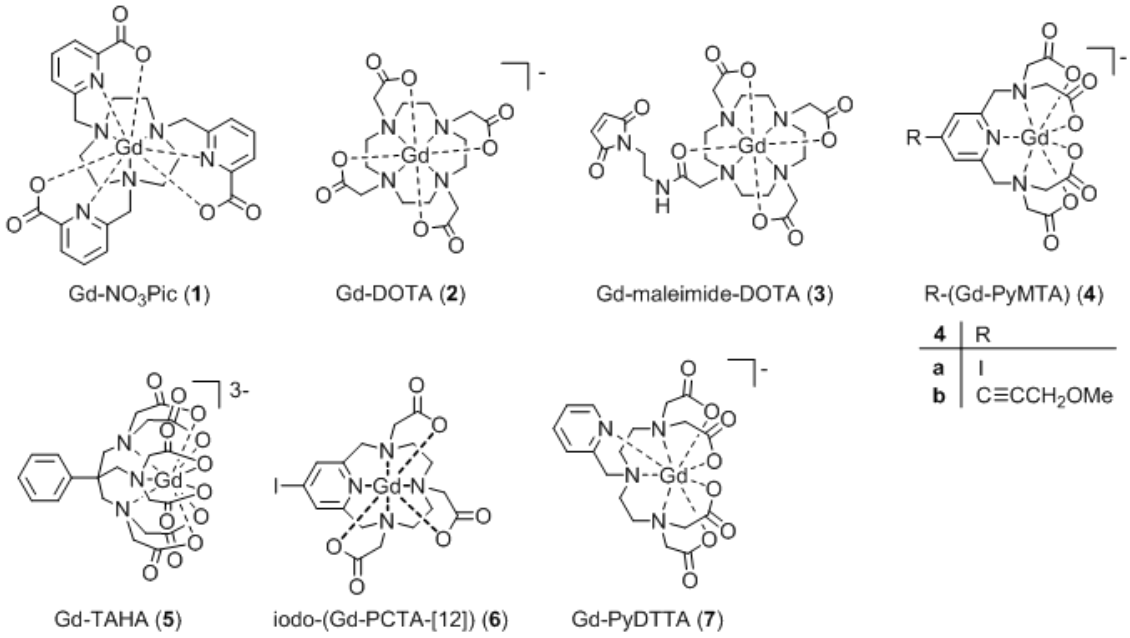

Figure 3.2: Structural formulas and naming of the Gd(III) complexes (**1**) - (**7**) which were studied in this work.

3.4 and 3.1.3 by computing an EPR spectrum for each pair (D, E) with the EasySpin function `pepper` and then summing these spectra with the weights $W(D, E)$ according to the probability products $W(D, E) = P(D) \cdot P(E)$. This was done to allow for additional flexibility in the forms of the D and E distributions beyond that available from the `DStrain` and `EStrain` options provided by EasySpin, as well as to avoid artifacts from these functions which arise in cases where the strain is comparable to the $\langle D \rangle$ or $\langle E \rangle$ values which result from computing the line broadening from the D and E strains in the linear approximation.

Orientational averaging was performed in 3 degree increments and a 10-fold interpolation of the orientation grid. The magnetic field range for simulation was chosen to well cover the experimental field range, as the EasySpin function `pepper` forces the computed

spectra to zero at its boundaries. The number of field points was set to 8000 to reach sufficient convergence. The simulation output was set to separate the subspectra computed for each transition of the $S = 7/2$ Gd(III). For the 240 GHz spectra, whose data was obtained by rapid passage measurements, the contributions of the individual transitions were summed as is to arrive at the final simulated spectra. For the spectra obtained from echo-detected EPR measurements (Q-/W-band data), each allowed transition was computed separately and the final spectra were obtained by summing the contributions of the individual transitions according to their effective flip angles (C.1.5).

Two different approaches to the sampling of the D and E distributions were investigated. First, the distributions of ZFS parameters were sampled by a regular grid of points. Second, a Monte Carlo sampling was used, where a large set of randomly distributed (D, E) pairs is generated and the overall EPR spectrum is computed as a linear combination of the EPR spectra for all (D, E) pairs. It was found in the course of this work that the Monte Carlo sampling of the D and E distributions resulted in the optimal computation cost and avoided unphysical artifacts in the simulated spectra associated with oversampling in the vicinity of the $D = 0$ point of the D value distribution. Both approaches require careful calibration of the number of points sampled in the D and E distributions in order to reach convergence of the simulated EPR spectrum.

Extensive details of the numerical simulations can be found in C.1. For all simulations presented in this Chapter, a Monte Carlo sampling of the D and E distributions was used. Unless otherwise noted, no additional line broadening terms (e.g. `lw` or `lwpp` in

EasySpin, higher-order ZFS terms, unresolved hyperfine coupling, dipolar coupling, etc.) were included.

3.3.1 Spectral features of the simulated EPR spectra

The simulated EPR powder spectra of Gd(III) complexes consists of seven allowed transitions $|m_S\rangle \leftrightarrow |m_S + 1\rangle$, which are broadened by anisotropy of the ZFS. According to Kramers' theorem, for a half integer spin the levels $|\pm m_S\rangle$ are degenerate in zero magnetic field. When the ZFS is weak compared to the electron Zeeman interaction, as is the case for all Gd(III) complexes studied in this work, the subspectrum of the $|-1/2\rangle \leftrightarrow |+1/2\rangle$ transition is much narrower than the other transitions [12]. This central transition is only broadened by ZFS to second-order in perturbation theory, while the other Gd(III) transitions are broadened by ZFS at first-order, as was discussed in Section 2.2. Due to the narrowing of the width of the $|-1/2\rangle \leftrightarrow |+1/2\rangle$ transition with increasing magnetic field, the peak intensity and the relative width of this transition with respect to the full width of the Gd(III) EPR spectrum decreases with increasing magnetic field.

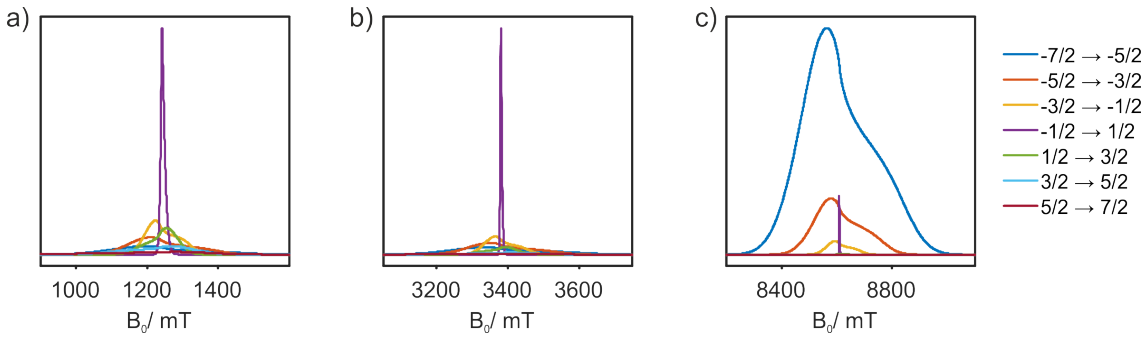

Figure 3.3: Contributions of the seven allowed EPR transitions for Gd(III) to the overall EPR spectrum. (a) Q-band and 10 K, (b) W-band and 10 K, (c) 240 GHz and 5K.

An illustration of these spectral features are shown in Figure (3.3). The $|-1/2\rangle \leftrightarrow |+1/2\rangle$ transition contributes a sharp narrow feature at the center of the EPR spectrum, while the other transitions are broad and contribute primarily to the envelope of the total lineshape. The relative populations of the seven allowed EPR transitions will vary with temperature according to a Boltzmann population distribution. At a temperature of 10 K, the $|-1/2\rangle \leftrightarrow |+1/2\rangle$ transition dominates the overall spectrum at Q-band and W-band. At 240 GHz and 5 K, the central transition is further narrowed and reduced in intensity. Additionally, the low temperature and high magnetic field results in approximately 90% of the population in the lowest energy level (Appendix C.2), such that the broad envelope of the $|-7/2\rangle \leftrightarrow |-5/2\rangle$ transition dominates the overall lineshape. The lineshape of this transition is most asymmetric with respect to the position of the narrow central peak, which remains visible due to its narrow width. This feature allows us to trivially assign the sign of D. In the convention that $0 \leq |E|/|D| \leq 1/3$, if the maxima of the broad component of the spectrum is shifted towards high field D is negative, and if it is shifted towards lower field, D is positive. If both positive and negative modes are present in the D distribution, the anisotropy of the EPR lineshape indicates the difference in the weights of these two modes (as in the $P(+D)/P(-D)$ parameter in Model 3).

3.4 Analysis with Model 1

EPR spectra at Q-band, W-band, and 240 GHz were simulated using Model 1 with the variables D, σ_D, E, σ_E, and a small convolutional line broadening (`Sys.lwpp` in EasySpin)

taken as free parameters. Parameters were varied by visual inspection to obtain an estimate of the relevant parameter space and to roughly gauge the performance of the model in fitting the experimental data. Visually optimized simulations for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) are shown in Figure 3.4. The analogous simulations for all other complexes are found in Appendix C.3.

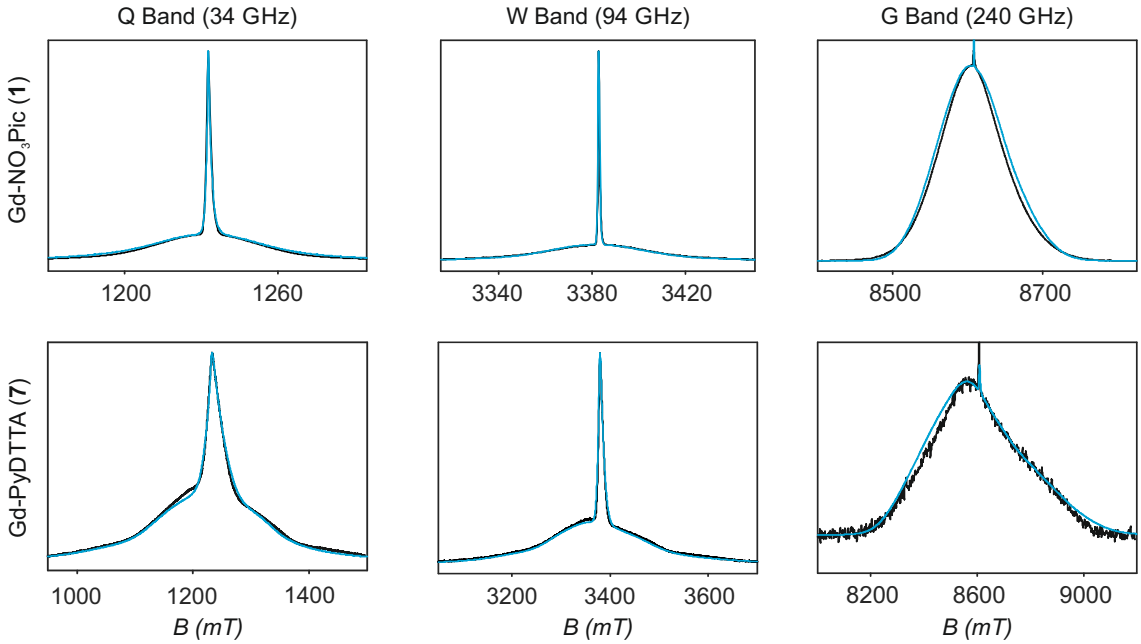

Figure 3.4: Measured EPR spectra (black lines) and corresponding simulations with Model 1 (light blue) for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) at Q-band and 10 K, W-band and 10 K, and 240 GHz and 5 K. The ZFS parameters for these simulations are given in Table 3.1.

In the analysis with Model 1, it was observed that for certain values of the ZFS parameters a conflict could arise in the definitions of the D and E as a pair of uncorrelated Gaussian distributions (Equation 3.6). It has been found, in this work and previously [125, 22, 61], that the widths σ_D and σ_E of the distributions are typically smaller but comparable to the average values $\langle D \rangle$ and $\langle E \rangle$. Therefore, for many Gd(III) complexes

it is possible for a situation to arise where two uncorrelated Gaussian distributions for D and E produce a large fraction of cases where either D and E have different signs, or where the signs of D and E are the same but the relation $|E| < |D/3|$ does not hold (e.g. Figure 3.5 (c, f)). In these cases one can still formally write Equation 3.3 for any pair of values D and E and compute the values D_X, D_Y, and D_Z according to Equations 3.4 and 3.5. However, in order to satisfy the conditions of Equation 3.5 the indexes (X, Y, Z) of the computed D_X, D_Y and D_Z values need to be reshuffled. After rearranging the indices to satisfy Equations 3.4 and 3.5, the D and E values need to be recomputed. The resulting distributions of $P(D)$ and $P(E/D)$ before and after this index rearrangement are sketched in Figure 3.1 (A)-(B).

Complex	D_{init}	σD_{init}	D_{pos}	σD_{pos}	D_{neg}	σD_{neg}	$\frac{P(+D)}{P(-D)}$
Gd-NO_3Pic (**1**)	420	140	472	124	- 418	111	1.4
Gd-PyDTTA (**7**)	1800	514	1845	439	- 1275	271	3.3

Table 3.1: Change in $\langle D \rangle$ and σ_D upon reordering the ZFS parameters in Model 1 according to Equations 3.4 and 3.5 for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**). The input distributions are single-Gaussian (Equation 3.6). Reshuffling of the indices resuls in a bimodal Gaussian distribution of D. Units are given in MHz.

An example of the calculation reordering the $P(D)$ and $P(E)$ distributions for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) is shown in Figure 3.5. The corrected distributions for the D and E parameters are both bimodal with different weights of the positive and negative components. The distribution of newly computed E/D parameter covers the full allowed range from 0 to 1/3. However, for some complexes there remains a significant probability density at $E/D = 0$ (e.g. Figure 3.5 (C)), corresponding to an axial ($E = 0$) ZFS, which is likely unphysical for most Gd(III) complexes. The maximum of

the probability distribution $P(E/D)$ occurs at the value $\langle E \rangle$ / $\langle D \rangle$. Overlaying the newly obtained D distribution by two Gaussians shows that the maxima are slightly asymmetric with respect to zero, shifting towards larger values for the dominant component (Figure 3.5 (a,d)). Additionally, the widths of the two new Gaussian distributions are slightly narrower compared to the width of the input distribution. The input parameters values and the recomputed parameter values after reshuffling of the indices for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) are given in Table 3.1. Table 3.2 summarizes the ZFS parameters for all six Gd(III) complexes determined by visually optimized simulations with Model 1 before reordering of the indices (the values obtained after reordering of the indices are given in Table C.3). Because these reported parameter values are the result of a fit-by-eye, error bars could not be computed.

Comparing the corrected D and E/D distributions for Model 1 to those of Model 3 (sketched in Figure 3.1 (B, D-E)) we note very similar forms of the distributions of ZFS parameters. After learning that these models are so closely identical, with some particular advantages to Model 3 which will be discussed in detail below, we did not pursue any more rigorous fitting with Model 1 and instead turned to make a more detailed analysis with Models 2 and 3.

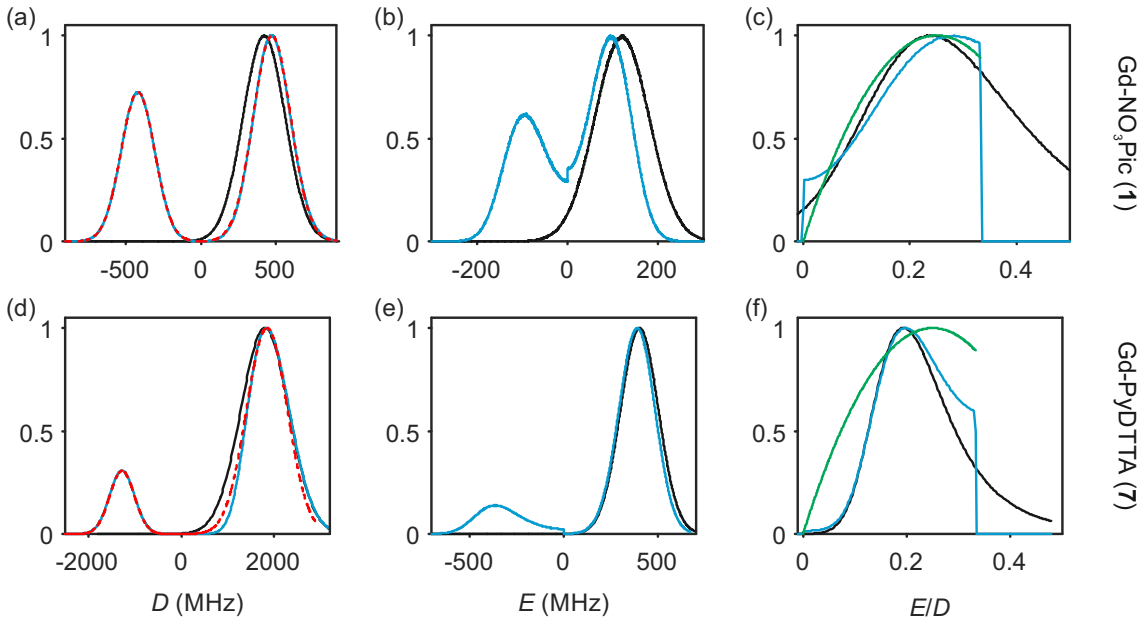

Figure 3.5: Distribution of ZFS parameters for Model 1 as defined in Equation 3.6 (black) and after rearranging of the indexes (X, Y, Z) of the computed D_X, D_Y and D_Z values according to Equations 3.4 and 3.5 (light blue) for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**). Gaussian distributions positioned symmetrically about $D = 0$ are overlaid on the rearranged distribution (dashed red). (a, d) D value distributions, (b, e) E value distributions, and (c, f) E/D distributions. The green line shows $P(E/D)$ distribution from Equation 3.7 which was used in Models 2 and 3.

3.5 Analysis with Models 2 and 3

Models 2 and 3 were also initially fit by visual inspection, and it was observed that for rather broad ranges of the ZFS distribution parameters the correspondence between experimental and simulated data was quite good. Additionally, the criteria for convergence of the simulations (see Section 3.3 and Appendix C.1) required run-times of more than 24 hours on a standard desktop computer to simulate a single EPR spectrum, so that standard least-squares minimization routines (e.g. Nelder-Mead/downhill simplex, Levenberg-Marquardt, etc.) for fitting were not practical. Therefore, in order to formal-

ize the finding of the best-fit solution and the determination of error bars on the resulting ZFS parameters for Models 2 and 3, we generated a large library of simulated spectra for each measurement frequency and temperature that map out a region of the parameter space of D and σ_D containing the values of these parameters that we expect for the Gd(III) complexes studied here. This library of simulations samples a parameter space spanning $D = 300 - 1950$ MHz and $\sigma_D = 50 - 600$ MHz in steps of 50 MHz. In order to have a common library to query all Gd(III) complexes studied in this work, typical values for the measurement frequency (at Q-/W-band) and temperature (at 240 GHz) were used in place of the exact experimental values for each Gd(III) complex, as detailed in Table C.2. The small measurement-to-measurement deviations from these typical values were found to not significantly impact the line shape of the simulated EPR spectra, and hence are not expected to significantly alter the final determined ZFS parameter values. For this library of simulations, the contributions to the line shape from each transition and from the $+D$ and $-D$ modes of the bimodal D distribution for Models 2 and 3 were saved separately. This allows for the same library of simulations to be used for both Models 2 and 3, by either summing the contributions as is, or by adding a weighting term denoted $P(+D)/P(-D)$ which introduces the asymmetry in the D distribution for Model 3. Further details of the inputs used to generate the library of simulated spectra can be found in Appendix C.1.6.

Each simulated EPR spectrum in the library described above was compared to the data at the corresponding measurement frequency by fixing the baseline, with a baseline

constant offset removed if necessary, and scaling the amplitude of the simulation to best fit the data in a least-squares sense. The RMSD between each simulation and the data was then computed according to

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(sim(i) - dat(i))^2} \tag{3.8}$$

where N is the number of points of the measured EPR lineshape.

For all three models, the fraction of the $P(D)$ distribution near $D = 0$ was significant enough to produce a sharp feature in the lineshape in the vicinity of the Gd(III) **g**-value position. This results in the models predicting a sharper feature in the region of the central peak than is observed experimentally. The effect of this overly sharp peak in the simulated lineshape is sometimes in the literature reduced by the addition of an intrinsic linewidth, as was done here in the analysis with Model 1. However, the addition of an intrinsic linewidth of this form, while indeed reducing the overall RMSD of the fit, does not have any physical interpretation and is used only as a beautifying parameter. Because this linewidth parameter increases the parameter space but without helping in determining the ZFS parameter values to any greater precision, we chose to leave out any intrinsic linewidth terms in the analysis with Models 2 and 3. However, due to the ambiguity in the models in correctly reproducing the region of the central peak, as well as this region being poorly sampled in the measured 240 GHz EPR spectra, we chose to initially exclude this region from the analysis. This was done by visually identifying the 'kinks' in the EPR spectrum where the sharp central peak meets the broad

envelope of the spectrum, and excluding points interior to this from the calculation of the RMSD. This should result in approximately excluding from the fit contributions of the $|-1/2\rangle \leftrightarrow |+1/2\rangle$ transition, leaving only the remaining field ranges in the left and right wings of the spectrum for computation of RMSD errors of the fit. The dependence of the RMSD on the ZFS parameters was visualized as RMSD error maps. For Model 2, contour plots of the RMSD error as a function of the ZFS parameters D and σ_D are shown for Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) in Figure 3.6 (full results can be found in Appendix C.4).

The minimum RMSD values in these contour plots, as well as for all simulations with best-fit parameter values computed in this Chapter, always exceed the noise level of the experimental data. Additionally, the position of the minimum RMSD value in the contour plots is not always the same for the three tested microwave bands. This is a result of attempting to describe the ZFS interaction in an ensemble of Gd(III) complexes using a simplified model for the distributions of ZFS parameters. While the various models available in literature seem to be relatively accurate, given the quality of the fits to experimental data in our work and in the work of others, this does not necessarily mean that the model is accurately describing the physical system. Therefore, caution must be taken in attributing the best-fit ZFS parameter values for a particular model to the true physical values.

In order to obtain an objective estimate of the precision of the determined ZFS parameter values, we take as a range of acceptable fit values as those which fall within the

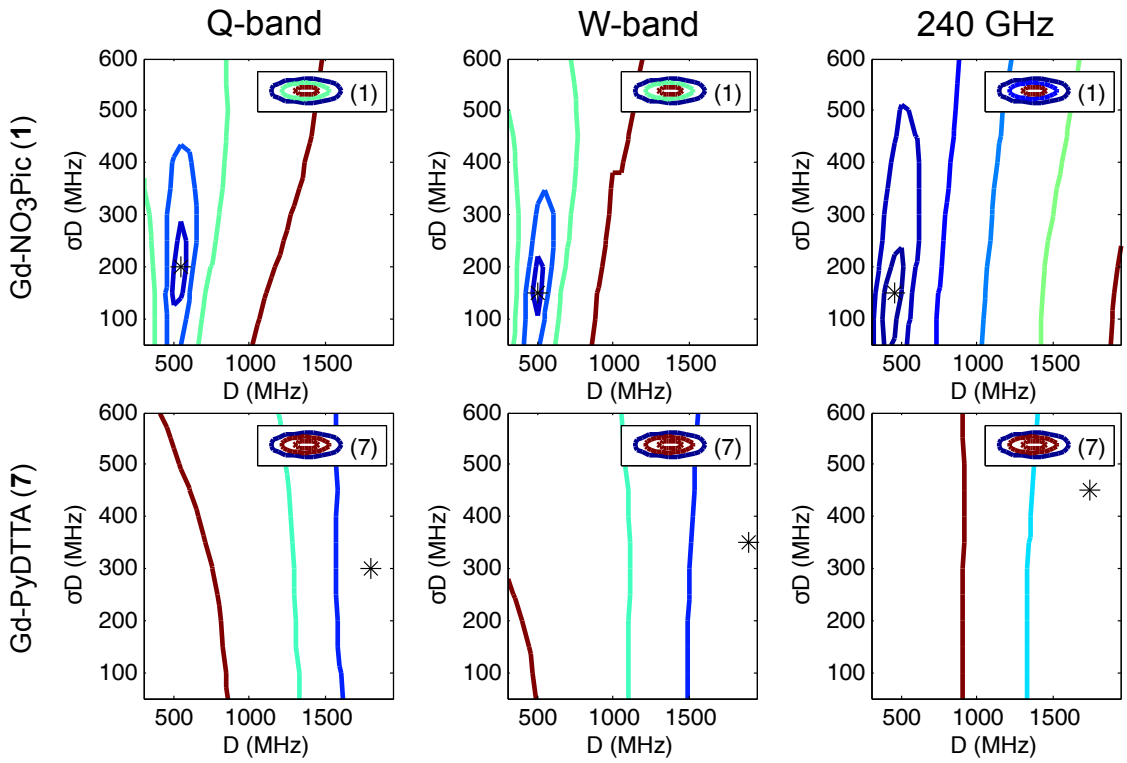

Figure 3.6: Contours of constant RMSD as a function of (D, σ_D) parameter values using Model 2 and for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) at Q-band and 10 K, W-band and 10 K, and 240 GHz and 5 K. Simulated EPR spectra for each (D, σ_D) pair were normalized to the experimental data using only the outer shoulders of the EPR spectra and the region of the sharp central peak was excluded from calculation of RMSD values. The asterisk denotes the set of parameter values available in the library of simulated spectra which has the minimum RMSD value for each measurement frequency. Each contour line represents a doubling of this minimum RMSD value.

contour of twice the minimum RMSD value. This region is bounded by the first contour line in Figure 3.6 and elsewhere. This is, if anything, expected to be a conservative estimate of the errors of the ZFS parameter values reported here. It can be further observed from Figure 3.6 that the 50 MHz grid sampling of the (D, σ_D) parameter space in our library of simulations is perhaps too coarse of a sampling for complexes with very small ZFS, as is the case for Gd-NO_3Pic (**1**). Therefore, we interpolate the ZFS parameter

values on this grid by making the assumption that the first contour line bounding the region of twice the minimum RMSD should be smooth if we had arbitrarily fine sampling of the (D, σ_D) parameters, and that the minimum RMSD should lie at the center of this contour. To find this position, the first contour is fit by an ellipse, from which the best-fit (D, σ_D) values are taken to be at the center of this ellipse, with the errors on these parameters given by the lengths of the semiminor and semimajor axes. Taking a weighted average of the determined values for D and σ_D and their associated errors at each frequency gives the final best-fit results for Model 2, which are summarized in Table 3.2.

Two examples of the EPR spectra simulated at the three microwave bands using the best-fit ZFS parameters (D, σ_D) for Model 2 (Table 3.2) and the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) are shown in Figure 3.7. Full results can be found in Appendix C.4. For the EPR spectra measured at Q-band and W-band, Model 2 gives very reasonable fits to the experimental data, despite the fixed equal ratio between the positive and negative components of the bimodal D distribution (Figure 3.1 (C)). The simulation also reproduces the position and width of the central peak, despite this region being excluded from the fit. However, the simulations at 240 GHz show significant deviations from the simulations with Model 2.

Model 3 is identical to Model 2, but with an additional allowance for the optimization of the relative contributions from the $-D$ and $+D$ components of the bimodal D distribution. The asymmetry of the D distribution is defined as the ratio between the amplitudes

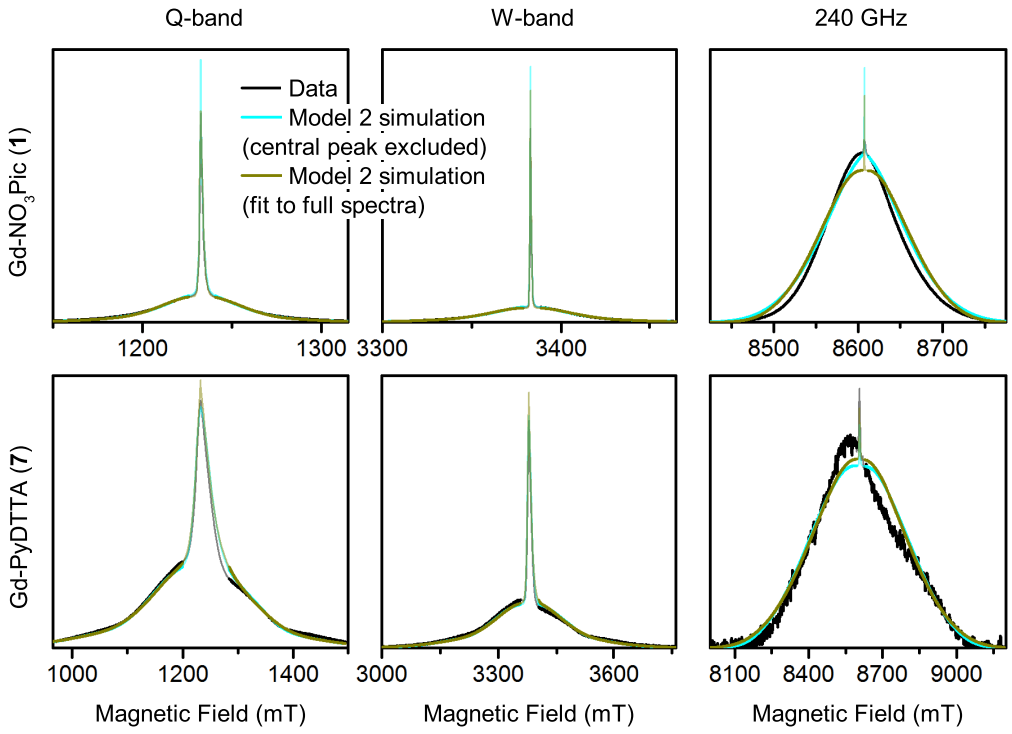

Figure 3.7: Simulations with Model 2 using the best-fit D and σ_D parameter values, determined with and without the region of the sharp central peak included in the RMSD error map analysis, for the complexes Gd-NO3Pic (**1**) and Gd-PyDTTA (**7**). The faded region indicates the portion of the spectra which were excluded from analysis in the prior case.

Model	Complex	D (MHz)	σ_D (MHz)	$\frac{\sigma_D}{D}$	$\frac{P(+D)}{P(-D)}$	E (MHz)	σ_E (MHz)	$\frac{\sigma_E}{E}$	$\frac{E}{D}$	lwpp (mT)
1	Gd-NO3Pic	420	140	0.33	—	120	60	0.50	0.29	[0.5 0]
	Gd-DOTA/ Gd-maleimide-DOTA	600	240	0.40	—	150	75	0.50	0.25	[1.0 0.1]
	R-(Gd-PyMTA)	1070	357	0.33	—	306	153	0.50	0.29	[0.8 0]
	Gd-TAHA	1250	417	0.33	—	357	119	0.33	0.29	[0.6 0]
	iodo-(Gd-PCTA-[12])	1780	508	0.29	—	396	132	0.33	0.22	[1.0 0.1]
	Gd-PyDTTA	1800	514	0.29	—	400	133	0.33	0.22	[2.0 0.1]
2	Gd-NO3Pic	500 ± 19	169 ± 52	0.34	—	—	—	—	—	[0 0]
	Gd-DOTA/ Gd-maleimide-DOTA	652 ± 44	409 ± 200	0.63	—	—	—	—	—	[0 0]
	R-(Gd-PyMTA)	1263 ± 112	350 ± 109	0.28	—	—	—	—	—	[0 0]
	Gd-TAHA	1335 ±75	408 ± 178	0.31	—	—	—	—	—	[0 0]
	iodo-(Gd-PCTA-[12])	1845 ± 194	400 ± 289	0.22	—	—	—	—	—	[0 0]
	Gd-PyDTTA	1810 ± 173	367 ± 289	0.20	—	—	—	—	—	[0 0]
3	Gd-NO3Pic	**485** ± 20	**155** ± 37	0.32	**1.8**	—	—	—	—	[0 0]
	Gd-DOTA/ Gd-maleimide-DOTA	**714** ± 43	**328** ± 99	0.46	**0.3**	—	—	—	—	[0 0]
	R-(Gd-PyMTA)	**1213** ± 60	**418** ± 141	0.34	**1.6**	—	—	—	—	[0 0]
	Gd-TAHA	**1361** ± 69	**457** ± 178	0.34	**1.9**	—	—	—	—	[0 0]
	iodo-(Gd-PCTA-[12])	**1861** ± 135	**467** ± 292	0.25	**3.6**	—	—	—	—	[0 0]
	Gd-PyDTTA	**1830** ± 105	**390** ± 242	0.21	**3.6**	—	—	—	—	[0 0]

Table 3.2: Extracted magnitudes of D, σ_D, E, and σ_E values using Model 1 including an additional Voigtian convolutional line broadening (`lwpp = [Gaussian Lorentzian]`), extracted magnitudes of D and σ_D values using Model 2, and extracted magnitudes of D, σ_D, and $P(+D)/P(-D)$ values using Model 3. All ZFS parameter values in this table were determined with the region about the central peak excluded from analysis. The values reported for Model 1 are the ZFS parameter values before reordering of the indices. The bold font indicates the overall best-fit parameters from the three models which were used for calculations with the superposition model.

of the positive and negative component, $P(+D)/P(-D)$ (Figure 3.1 (D)). The asymmetry parameter $P(+D)/P(-D)$ was determined by fixing $\langle D \rangle$ to the closest available value in the library of simulations to that determined using Model 2 (Table 3.2) and then varying $P(+D)/P(-D)$. Only the spectra at 240 GHz were used for the determination of $P(+D)/P(-D)$, since these were the available measurement conditions under which the asymmetry in the experimental spectra were most prominent (Figure 3.3). We additionally attempted to determine $P(+D)/P(-D)$ using the Q-/W-band data, but these spectra were not sufficiently sensitive to variations in this parameter to assign a best-fit value. The effect of this parameter was visualized with contour plots of RMSD errors as a function of $P(+D)/P(-D)$ and σ_D, and is shown for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) in Figure 3.8. Once the asymmetry parameter $P(+D)/P(-D)$ was determined via the minimum RMSD value in this error map, that value was fixed and the (D, σ_D) RMSD error maps were recomputed for the three microwave bands to find the best-fit values of these parameters. The resulting ZFS parameter values for Model 3 are given in Table 3.2 with the corresponding simulations plotted with the full dataset in Figures 3.9 and 3.10. RMSD contour plots for all six Gd(III) complexes analyzed with Model 3 are plotted in Appendix C.5.

In nearly all cases, with the exception of Gd-maleimide-DOTA (**3**), it appeared that the $P(+D)/P(-D)$ value is rather poorly constrained when the criterion of twice the minimal RMSD is used as the measure of goodness-of-fit. This criterion, while used here for consistency, is evidently not a reasonable error estimate for the $P(+D)/P(-D)$

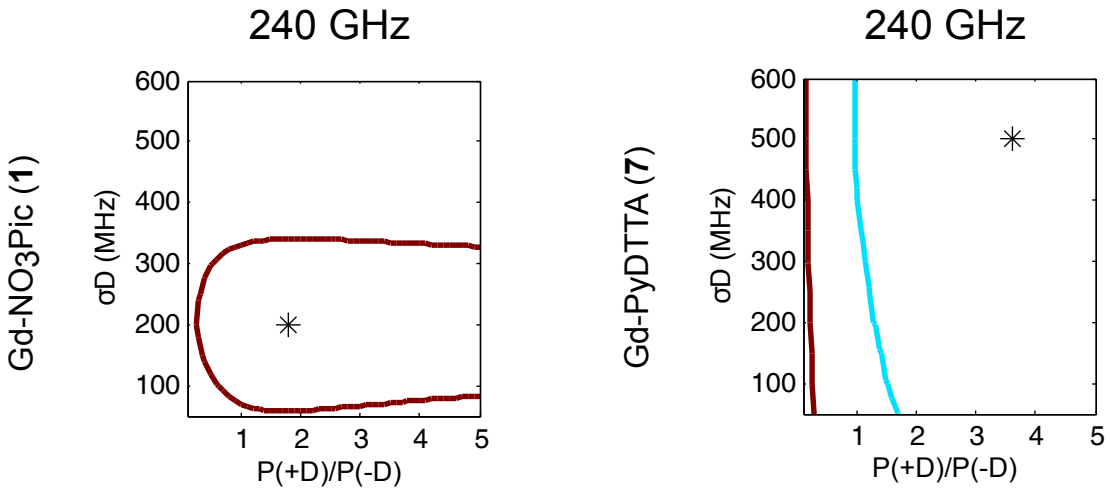

Figure 3.8: Contours of constant RMSD as a function of $P(+D)/P(-D)$ and σ_D parameter values using Model 3 and the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) at 240 GHz and 5 K. The mean of the ZFS parameter D was set to D = 500 MHz and D = 1800 MHz, respectively, corresponding to the closet D value available in the library of simulations to the D value as determined by Model 2 for this complex (Table 3.2). The asterisk denotes the position of minimum RMSD.

parameter in Model 3. The most obvious effect of this parameter on the simulated high-field EPR spectra is to set the relative positions of the broad component of the spectrum with respect to the sharp central peak corresponding to the $|-1/2\rangle \rightarrow |1/2\rangle$ transition. However, because the width of this central peak is so narrow compared to the broad component of the spectrum at 240 GHz, it has a relatively small impact on the overall RMSD of the fit. It was found that the separation between the sharp central transition and the peak of the broad component of the 240 GHz EPR spectra varies linearly with the $P(+D)/P(-D)$ values determined from the RMSD contour plots, and was used to estimate a typical deviation of 0.34 for the value of the $P(+D)/P(-D)$ parameter (Appendix C.8), though this varied between the different Gd(III) complexes. Practically, due to the various approximations inherent to the model, it is difficult by

either method to assign an accurate value for $P(+D)/P(-D)$. However, as was discussed in Section 3.3.1, it is trivial to assign the overall sign of D based on the positioning of the broad spectral component with respect to the sharp peak of the $|-1/2\rangle \rightarrow |1/2\rangle$ transition. If D is positive ($P(+D)/P(-D) > 1$), then the broad feature will be shifted towards lower fields. If D is negative ($P(+D)/P(-D) < 1$), then the broad feature will be shifted towards higher fields. This condition places a further constraint on the value of $P(+D)/P(-D)$ which is not reflected in the RMSD contour plots.

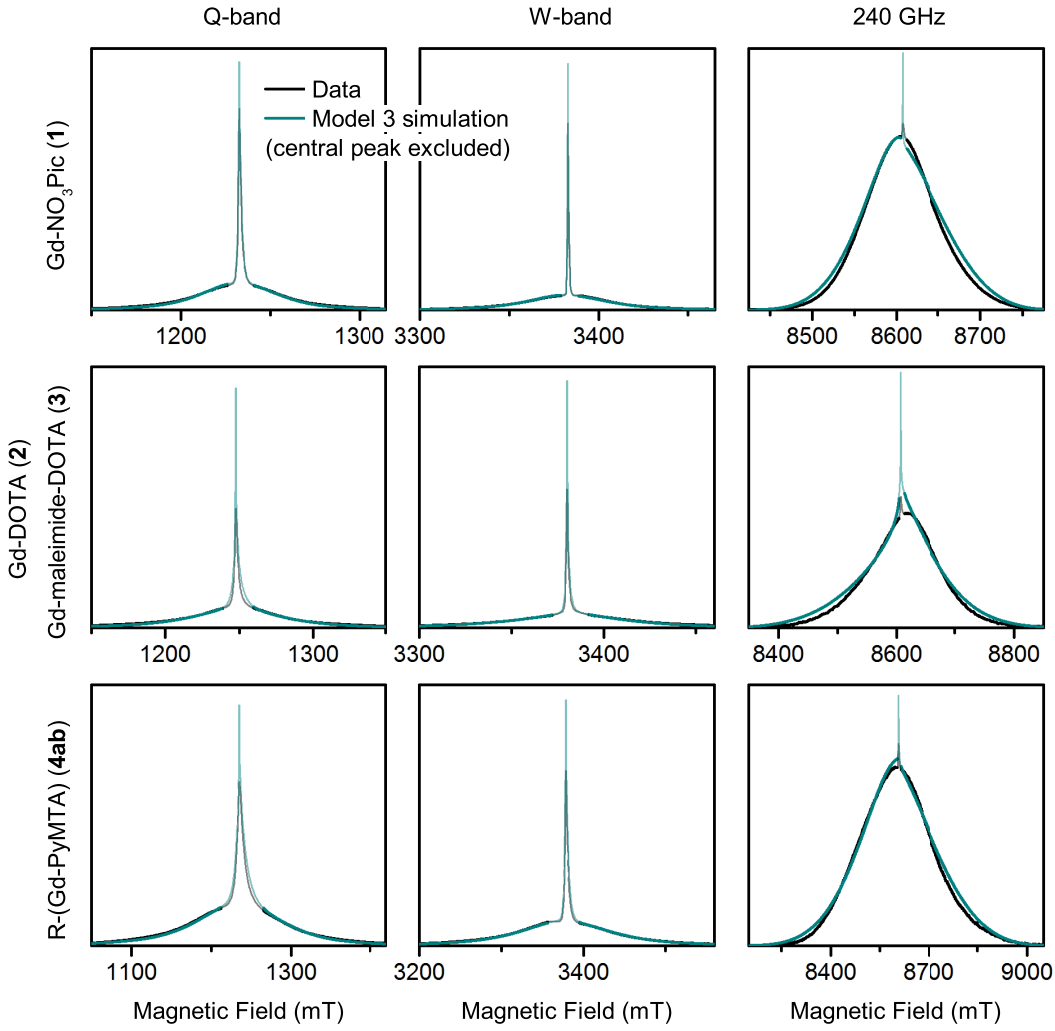

Figure 3.9: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-NO_3Pic (**1**), Gd-DOTA (**2**) (240 GHz sepctra)/Gd-maleimide-DOTA (**3**) (Q-/W-band spectra), and R-(Gd-PyMTA) (**4a**) (240 GHz spectra)/R-(Gd-PyMTA) (**4b**) (Q-/W-band spectra). Overlaid are simulations with Model 3 using the best-fit ZFS parameters presented in Table 3.2. The faded regions indicate the portion of the spectra about the central transition which was excluded from the RMSD error map calculations.

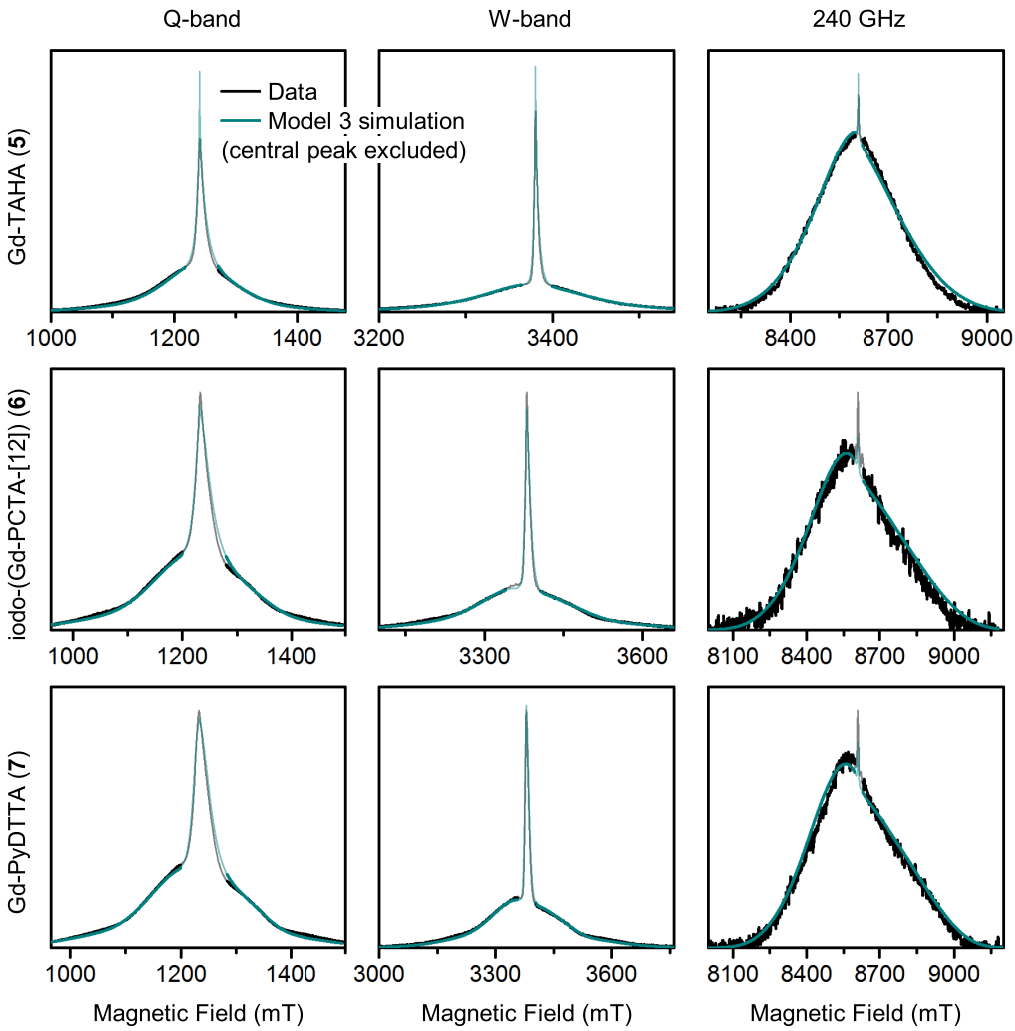

Figure 3.10: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-TAHA (**5**), iodo-(Gd-PCTA-[12]) (**6**), and Gd-PyDTTA (**7**). Overlaid are simulations with Model 3 using the best-fit ZFS parameters presented in Table 3.2. The faded regions indicate the portion of the spectra about the central transition which was excluded from the RMSD error map calculations.

3.6 Comparison of the models for ZFS parameter distributions

For all six Gd(III) complexes studied, there is very good agreement in the mean value of D between the three models. The determined D value from Model 1 initially appeared to have the greatest deviation from Models 2 and 3, in some cases lying outside of the range of uncertainty. However, after correction of the index ordering in Model 1, the newly computed D value, taken as the mean of the dominant component in the bimodal distribution, is in very good agreement with Models 2 and 3. The best-fit ZFS parameters D and σ_D for all three models are compared in Figure 3.11.

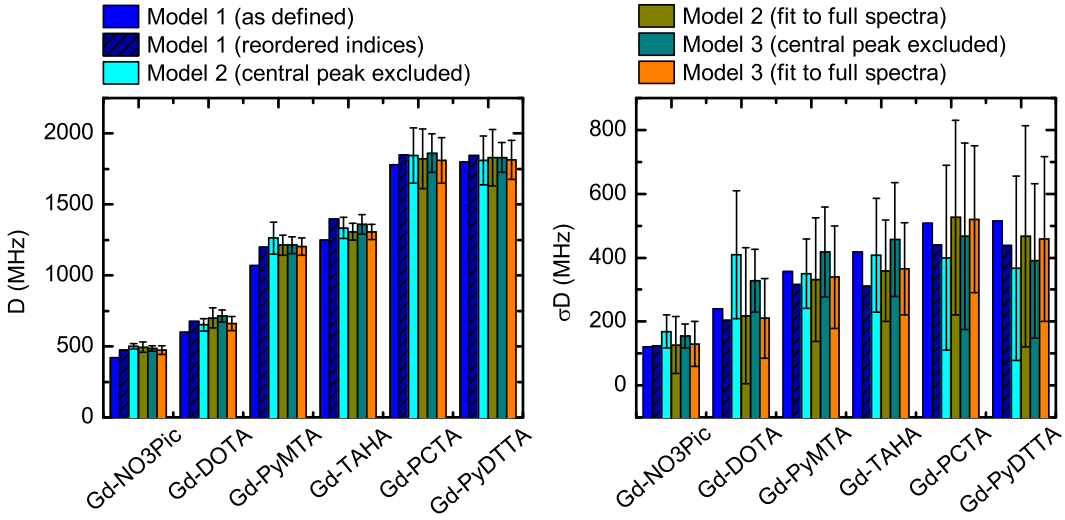

Figure 3.11: Comparison of the extracted values for the mean and width of the ZFS parameter D for the three models. Model 1 was fit by visual inspection, and therefore error bars on the ZFS parameters D and σ_D were not computed. For Models 2 and 3, mean values and error bars for D and σ_D were computed by combining results from RMSD error maps which compare a library of simulated spectra to the data at the three measurement frequencies. Models 2 and 3 were fit with the region about the central transition excluded from analysis, and with the full EPR spectra included in the analysis.

In principle, both Models 1 and 2/3 appeal to physical intuition. Model 1 follows from the central limit theorem, however this would require the presence of a virtually unlimited number of different random ligand contributions to the D and E distributions in order to be strictly valid. Models 2/3 appeal to the near equality of the coordinating atoms in the ligand and to the non-directional character of the bonds in the Gd(III) complex. Model 2 is fundamentally limited by the restriction that the D distribution must be fully symmetric. This appears to be an adequate description of the EPR spectra acquired at Q-/W-band, but fails to capture the asymmetries present in the 240 GHz data. Model 3 has two primary advantages over Model 1: (i) better cancellation in the regions $D = 0$ and $E = 0$ which represent fully symmetric (i.e. unphysical) Gd(III) complexes, and (ii) flexibility in varying the relative weights of the $+D$ and $-D$ modes in the D distribution, which are fixed in Model 1. Additionally, Model 3 only has three free parameters $(D, \sigma_D, P(+D)/P(-D))$, while Model 1 has four free parameters (or six if `lwpp` is included). Given the relatively featureless EPR spectra characteristic of Gd(III) complexes in frozen glassy solutions, having few free parameters is highly desirable in order to have confidence in the fit.

A plot of the ZFS parameter distributions determined by Models 1 and 3 for all six Gd(III) complexes is given in Appendix C.6. The ZFS parameter values determined with Model 1 (Table 3.2) indicate that the width σ_D of the distribution is within 29-40 % of $\langle D \rangle$ and σ_E is within 33-50 % of $\langle E \rangle$. After reshuffling of the indices of Model 1 according to the conventional parameter definitions, we noted that the corrected D distribution

for Model 1 becomes bimodal with the two approximately Gaussian modes placed nearly symmetrically about zero. Further, these two modes have approximately the same width, as in Models 2/3. The widths of the two modes in the corrected D distribution for Model 1 are slightly narrower than the width of the non-corrected single Gaussian distribution. If the value of $\langle D \rangle$ is taken from the dominant mode, this value is approximately a 12 % increase from the initial single-Gaussian D value for complexes (**1**-**5**). For complexes (**6**-**7**) with the largest ZFS, there is less than a 5 % increase in the corrected D value. However, in these two cases the negative mode is shifted much closer to $D = 0$ than for the other Gd(III) complexes. The corrected $P(E)$ distribution for Model 1 similarly becomes bimodal. Calculating the $P(E/D)$ distribution from the corrected values for Model 1 gives a result that is qualitatively similar to the $P(E/D)$ distribution in Models 2/3, with a minimum probability density at $E/D = 0$ and a maxima around $E/D = 0.25$. However, the overall similarity of the two E/D distributions is not as good as for the D distributions, with the maximum of the $P(E/D)$ distribution of Model 1 varying by about 15 % with respect to the maxima of the $P(E/D)$ distribution of Models 2/3. Additionally, the value $E/D = 0$ is exactly excluded in Model 3, but retains a significant probability in Model 1.

For Models 2 and 3, the (D, σ_D) contour plots show that the value of D is rather well constrained with error bars in the range of ± 5-10 %. By comparison, the value of σ_D is much less well constrained by Models 2/3. In particular, for iodo-(Gd-PCTA-[12]) (**6**) and Gd-PyDTTA (**7**), the contour plots suggest that the σ_D value can assume

essentially any allowed value. For the Gd(III) complexes with weaker ZFS - Gd-NO3Pic (**1**), Gd-DOTA (**2**)/Gd-maleimide-DOTA (**3**), and R-(Gd-PyMTA) (**4ab**) - the σ_D value is somewhat better constrained by the fit. But even in the best case of Gd-NO3Pic (**1**), the σ_D value still varies by approximately $\pm 30\%$ within the area encompassed by the doubled minimum RMSD contour curve (Figure C.4). Including an asymmetry of the D value distribution helped to slightly better constrain the range for the σ_D values, but did not significantly alter the best-fit values for D and σ_D. For the Q-band and W-band spectra, the minimum RMSD value of the (D, σ_D) contour plots was not largely altered by the addition of the asymmetry parameter which has minimal effects on the simulated EPR lineshape at these frequencies. For the 240 GHz spectra, which displays the greatest degree of asymmetry in the measured EPR spectra, the minimal RMSD value decreased by more than a factor of two in some cases with the addition of the $P(+D)/P(-D)$ parameter in Model 3 compared to the fits using Model 2 (Appendix C.9).

3.7 Effects of including the central peak in the analysis with Models 2/3

For completeness, the analysis with Models 2 and 3 was repeated using the full range of the EPR spectra, including the region of the sharp central transition. In this analysis, the simulated spectra were scaled by a least-squares method to the experimental traces using the full range of the experimental EPR spectra, and the resulting RMSD values

were computed. The RMSD error contour plots, best-fit ZFS parameter values, and corresponding simulated EPR spectra are given for both models in Appendix C.7. A subset of these results, for the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**) analyzed with Model 2, are included in Figure 3.7 for comparison. The inclusion of the central peak leads to RMSD values that are up to an order of magnitude larger for the Q-/W-band spectra, where this region represents a significant fraction of the overall EPR spectrum (Appendx C.9). For complexes with small ZFS parameter values measured at Q-/W-band, this region can dominate the RMSD value of the fit. At 240 GHz, there is only about a factor of two change in the minimum RMSD value of the (D, σ_D) contour plots when the sharp peak is included due to the relatively small contribution of the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition to the overall EPR spectrum at 240 GHz and 5 K. However, despite the overall increased RMSD values, we noted enhanced stability of the fit when the region of the central peak was included. This is particularly evident in the case of Gd-DOTA (**2**)/Gd-maleimide-DOTA (**3**), whose fit at 240 GHz is significantly improved when the full EPR spectrum is considered in the analysis. Additionally, since the position of the broad background with respect to the sharp central transition in the 240 GHz spectra is one of the primary hallmarks of the $P(+D)/P(-D)$ parameter, this value is more readily determined if the full 240 GHz EPR spectrum, including the central peak, is used.

Importantly, despite uncertainties in the best treatment of the region of the central transition, the predictions for the best-fit D and σ_D values did not change significantly

upon inclusion of the region of the central transition (Figure 3.11). However, the range of acceptable D values within the doubled minimum RMSD contour was found to increase upon the inclusion of the central peak due to the increased minimum RMSD value. This effect is much more notable for Gd(III) complexes with small ZFS than those with larger ZFS parameter values.

While the final results for the value of the ZFS parameter D do not significantly change based on the treatment of the region of the central transition, the interpretation of the RMSD contour plots becomes difficult when it is included. For the broad component of the EPR spectrum, the spectral width increases linearly with D, as long as the high-field approximation holds. However, the central $|-1/2\rangle \leftrightarrow |1/2\rangle$ subspectrum is broadened only to second order in perturbation theory by the ZFS term, so the spectral width increases quadratically with D. This difference in scaling of spectral components with respect to the magnitude of the ZFS splitting makes it difficult to simultaneously fit both regions well by the methods described here. Additionally, the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition, due to its narrow linewidth, is very sensitive to other potential sources of line broadening such as higher-order ZFS terms, unresolved hyperfine couplings, dipolar interactions, etc. These considerations in the treatment of the central transition make little difference in the analysis of the 240 GHz data, where the width of the sharp central peak represents merely a fraction of a percent of the total linewidth of the full EPR spectrum and thus contributes only a very small fraction to the overall RMSD value of the fit. At lower microwave frequencies such as Q-band, these considerations become

much more important. However, given the quality of agreement in the fits regardless of the treatment of the central transition, it appears that the results are not largely skewed by the inclusion of the central peak and therefore it is not necessary to include an additional convolutional line broadening term in the fit. Rather, uncertainties in the simple models in describing the true ZFS parameter distributions seem to be the dominant source of error.

3.8 Superposition model to predict ZFS of Gd(III) complexes

The measurement of second-order ZFS parameters for frozen glassy solutions of Gd(III) complexes by analysis of EPR spectra is currently one of the most accurate means of accessing this information. However, measurements can only be carried out on existent Gd(III) complexes and requires a high-frequency EPR spectrometer (or, preferably, spectrometers at multiple frequencies). While such spectrometers are becoming increasingly available, they are by no means commonly laboratory equipment, especially at frequencies above W-band. Due to the demand for new Gd(III) complexes with specific ZFS characteristics that are suitable for bioconjugation, there is a need to be able to predict the ZFS characteristics of a proposed ligand structure before beginning chemical synthesis of a new Gd(III) complex. There has been effort in the realms of quantum chemistry calculations and DFT in predicting ZFS parameters of Gd(III) complexes, however these

predictions are not yet accurate enough for EPR application. Here, we make use of a superposition model [111, 125] to simulate ZFS parameter distributions for Gd(III) complexes.

In the superposition model, the ZFS tensor is expressed as a sum of ligand-field contributions from individual nuclei in the coordination spheres of an s state ion [111]. Here we use the simplification for Gd(III) complexes in frozen glassy solutions that was previously introduced by Raitsimring *et al.* [125], where only directly coordinating nuclei are considered and only the first order contribution to the ZFS Hamiltonian is computed. The contribution of each of the coordinating nuclei can be parameterized by the ZFS parameters D and E/D. We follow Raitsimring *et al.* in first building a ZFS tensor,

$$\mathbf{D} = \sum_k \left(\frac{r_{0,k}}{r_k}\right)^{\tau} \mathbf{R}(0,\theta_k,\phi_k) \begin{pmatrix} d_k & 0 & 0 \\ 0 & d_k & 0 \\ 0 & 0 & -2d_k \end{pmatrix} \mathbf{R}^{\mathrm{T}}(0,\theta_k,\phi_k) \ , \tag{3.9}$$

where $r_{0,k}$ is a reference ligand atom to Gd(III) distance, r_k is the actual ligand atom to Gd(III) distance, τ is a scaling exponent, $\mathbf{R}(0,\theta_k,\phi_k)$ is an Euler rotation matrix in $zy'z''$ notation and $\mathbf{R}^{\mathrm{T}}(0,\theta_k,\phi_k)$ its transpose, and the d_k are single-ligand ZFS contributions that are assumed to have axial symmetry with the unique axis being along the vector from the coordinating atom to the Gd(III). The parameters θ_k, ϕ_k, and r_k are spherical coordinates of the ligated atom in the frame where the ZFS tensor $\mathbf{D}$ is expressed. The ZFS parameters D and E are obtained by diagonalization of the traceless symmetric tensor $\mathbf{D}$ and ordering of the principal values as described in Section 3.1.1.

In contrast to Raitsimring *et al.*, we rely on coordination geometries from known crystal structures of lanthanide complexes. We also allow for a distance dependence of

the single-atom contributions and for atom-type dependent ZFS magnitudes d_k at atom-type dependent reference distances $r_{0,k}$. In particular, we distinguish between coordinated oxygen atoms with $r_O = 2.42$ Å and nitrogen atoms with $r_N = 2.65$ Å. Our model thus has three fit parameters: the scaling exponent τ and the reference ZFS magnitudes d_O and d_N. Note that the choice of r_O and r_N, which were taken as typical ligand atom to Gd(III) distances for these complexes, is not critical since for a given τ changes in these reference distances merely result in a well-defined change in d_O and d_N. We have also tried to fit a model which does not distinguish between coordinating oxygen and nitrogen atoms, but the fits were significantly worse and gave an unphysical negative scaling exponent τ (data not shown).

3.8.1 Geometry of the Gd(III) complexes

The required ligation polyhedra were taken from crystal structures obtained from the Cambridge Crystallographic Data Centre and converted to .xyz files using the Mercury software. Homewritten MATLAB scripts were used for further processing. Oxygen and nitrogen atoms closer than 3 Å to the lanthanide ion were considered as belonging to the first coordination shell. For the structures of Gd-NO_3Pic [62], of Gd(III) ligated to a modified DOTA ligand that roughly resembles Gd-maleimide-DOTA [5], and of a Gd-PyMTA-linker-Gd-PyMTA construct (the same as the 2.1 nm Gd-ruler in Chapter 2) [71], a full set of nine coordinating bonds were detected. For the latter two complexes, one of the ligands was water.

No structure was found for a lanthanide ion coordinated by the PCTA-[12] ligand. Instead, we used the structure of Ho(III) coordinated by a similar ligand with three phosphonate groups substituted for the carboxylate groups [94], which we presume has a very similar coordination polyhedron. Although this complex was synthesized in water and the crystals contain nine water molecules per two Ho(III) complexes, none of the water molecules are coordinated to the Ho(III) ion and the coordination number is only eight. The same coordination type is observed for Lu(III). We tried to place an additional water ligand at a typical lanthanide-oxygen distance for such ligation (2.43 Å), but this led to a situation where the oxygen atom came at least as close as 2.13 Å to another coordinating atom. Since no distance between two directly coordinating atoms shorter than 2.62 Å was detected in any other complex, we assume that the lanthanide complexes of PCTA-[12] have low affinity for water as a ninth ligand.

No structures were found for a lanthanide complex with TAHA or PyDTTA as the ligand. Hence, our model with three free parameters can be fit to only four of the six Gd(III) complexes discussed in this chapter. As a fit criterion, we used the mean square relative deviation $\sum_i (1 - |D_{\mathrm{model},i}| / \bar{D}_{\mathrm{exp},i})^2$ of the modeled ZFS magnitude from the mean experimental ZFS magnitude D determined by the fit with Model 3 as given in bold in Table 3.2.

3.8.2 Superposition model fit results

The best fit was obtained for $\tau = 1.102$, $d_{\mathrm{N}} = 991.3$ MHz, and $d_{\mathrm{O}} = 915.9$ MHz and reproduces very closely the experimental values for the four complexes with known geometries (Table 3.3). The fit value for PCTA-[12] was the only complex of the four which showed deviation from the experimental value, however the exact coordination geometry for iodo-(Gd-PCTA-[12]) is not known. The positive scaling coefficient τ and similar reference values for the ZFS contributions by N and O ligands are physically plausible and confirm that the ZFS is dominated by the symmetry of the first ligand shell.

Ligand	D_{exp} (MHz)	D_{model} (MHz)
NO_3Pic	485	485
maleimide-DOTA	714	714
PyMTA	1213	1213
PCTA-[12]	1861	1684

Table 3.3: Experimentally determined magnitudes of the ZFS parameter D by Model 3 and magnitudes determined by a fit with a superposition model.

The model was further tested with related structures. For Gd-DOTA without a linker in the ligand [37], we find a similar D_{model} value of 666 MHz and similar values for the coordination geometry of DOTA complexes of other lanthanide ions [20], assuming that Gd(III) takes the position of the other lanthanide ion. For the geometry of the Pr(III) of DOTA, we find $D = 689$ MHz, for Nd(III) 688 MHz, for Dy(III) 679 MHz, but for the coordination geometry of Ce(III) as strongly different ZFS with $D = 301$ MHz is found.

3.8.3 Superposition model for the distributions of ZFS parameters

Next we consider the form of the distributions of ZFS parameters. In the superposition model, a distribution of the ZFS is caused by a spatial distribution of the coordinated atoms. Raitsimring *et al.* [125] allowed for a random distribution of ligands on a sphere with the Gd ion in the center. This broad distribution of ligands is unrealistic given steric constraints of the ligands. Here we assume that the ligand atom positions are distributed about the mean positions found in the crystal structures. In the simplest approximation, distributions of the individual atoms are independent and correspond to a Boltzmann equilibrium distribution in an isotropic three-dimensional harmonic potential. This approximation leads to an isotropic three-dimensional Gaussian distribution of the atom positions that can be characterized by the standard deviation σ_{xyz} of the atom positions along the x, y, and z coordinates. A distribution of atom position of this form corresponds to the Debye-Waller factor (B factor) in crystal structure determination.

First, σ_{xyz} was varied for the maleimide-DOTA model. The experimentally observed relative standard deviation σ_D/D of $\approx 1/3$ was matched for $\sigma_{xyz} \approx 0.1$ Å. For some of the crystal structures, σ_{xyz} can be estimated from Debye-Waller factors to be in the range of 0.15 - 0.25 Å at ambient temperature [20, 27]. It is not surprising that similar values are found in frozen glassy solutions, where they likely correspond to the thermal distribution of ligand atom positions at the glass transition temperature, but may also be additionally influenced by strain in the glass.

However, this model led to a larger $\langle D \rangle$ than obtained with the same model parameters for $\sigma_{xyz} = 0$. This is expected, since the spatial distribution of the atom position on average causes more asymmetry of the ligand field. We corrected for this effect by reducing d_{N} and d_{O} each by a factor of 0.845. This model successfully reproduced D and σ_D/D for maleimide-DOTA and provided a mean value of 0.195 for E/D, which is in reasonable agreement with the experimental value of 0.25 obtained using Model 1 (Table 3.2). This model also reproduced the trend in D among the four tested ligands for which there were both experimentally determined ZFS parameter values and crystal structures available (Table 3.3). However, the variation between the ligands was much weaker than observed experimentally and relative distribution width σ_D/D decreased more strongly with increasing D than was experimentally observed.

Ligand	D_{exp}	D_{sim}	$\left(\frac{\sigma_D}{D}\right)^{\mathrm{exp}}$	$\left(\frac{\sigma_D}{D}\right)^{\mathrm{sim}}$	$\left(\frac{P(+D)}{P(-D)}\right)^{\mathrm{exp}}$	$\left(\frac{P(+D)}{P(-D)}\right)^{\mathrm{sim}}$
NO_3Pic	485	514	0.32	0.30	1.8	1.2
DOTA	714	698	0.46	0.24	0.3	1.5
maleimide-DOTA	714	736	0.46	0.23	0.3	4.0
PyMTA	1213	1261	0.34	0.16	1.6	85
PCTA-[12]	1830	1654	0.25	0.13	3.6	5.2

Table 3.4: Fit of the distribution of the ZFS parameter D by a superposition model. D is given in units of MHz. Experimental data analyzed with Model 3 (Table 3.2) are used for comparison. Note that in the experimental results, DOTA and maleimide-DOTA were analyzed as the same compound, while in the superposition model analysis the ZFS parameter values were computed using the known structures for each complex.

Closer inspection of the structures with Debye-Waller factor information [20, 27] shows that the thermal ellipsoids of directly coordinating atoms usually have a smaller extension along the lanthanide ion to ligand atom bond than perpendicular to it. An attempt to fit models with different Gaussian distributions σ_r and $\sigma_{\theta,\phi}$ for spherical coordinates r on

the one hand and θ and ϕ on the other hand did not significantly improve the situation. For the final distribution model, we returned to the σ_{xyz} parameterization, but reduced σ_{xyz} to 0.05 Å in order to obtain a compromise between reproducing the mean values and distribution widths of D for the four Gd(III) complexes. We also tested $\sigma_{xyz} = 0.03$ Å and $\sigma_{xyz} = 0.07$ Å, but these choices provided worse agreement with experimental data when considering both D and σ_D/D. The results for this model with $\sigma_{xyz} = 0.05$ Å are given in Table 3.4. The model parameters used for final calculations were $d_{\mathrm{N}} = 989$ MHz, $d_{\mathrm{O}} = 943.5$ MHz, and $\tau = 0.100$.

The probability density distributions for D and E/D for Gd-DOTA computed with this model are shown in Figure 3.12. The distribution of D is bimodal, arising from the definition of D via the principal value with the largest magnitude, irrespective of sign. The asymmetry of $P(E/D)$ arises from the sorting of the principal values that ensures $|E/D| \leq 1/3$.

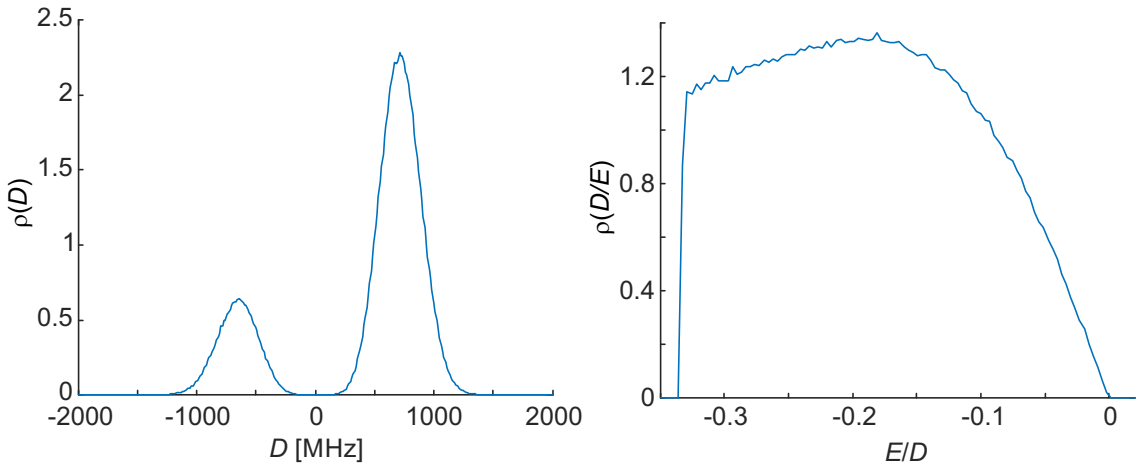

Figure 3.12: Distributions of ZFS parameters for Gd-DOTA predicted by the superposition model. (a) Distribution of D and (b) distribution of E/D.

3.8.4 Further predictions

Together with existing crystal structures of Gd(III) complexes, the superposition model described above can now be used for prediction of ZFS values for additional Gd(III) complexes (Table 3.5). The superposition model was used to predict ZFS parameter values for an additional seven Gd(III) complexes (Figure 3.13) for which there were structures available. The predicted values, given in Table 3.5 are mostly within the range of the values measured in this work with the exception of the HAM2 complex, for which a larger ZFS is predicted. Note that the uncertainties of the predictions for EDTA and HAM2 may be particularly large, because water coordination geometry is likely to differ between the crystal and aqueous solution if three free coordination sites are available.

Ligand	EDTA	DOTAM	DTMA	DO4Py	DO3A	DTPA	HAM2
D_{sim} (MHz)	520	751	781	968	1120	1213	2163
$(\sigma_D/D)_{\mathrm{sim}}$	0.30	0.23	0.22	0.19	0.16	0.16	0.07
$(E/D)_{\mathrm{sim}}$	0.20	0.18	0.21	0.20	0.21	0.24	0.27
Reference	[161]	[27]	[23]	[110]	[37]	[89]	[24]

Table 3.5: ZFS parameters predicted by the superposition model for Gd(III) with additional ligands (Figure 3.13) and references for the crystal structures used.

3.9 Discussion and remaining questions

In this chapter, we have endeavored to determine the ZFS parameters for six Gd(III) complexes of interest for use as spin labels in future EPR studies. We explored the use of simple models which describe the distributions of the second-order ZFS parameters

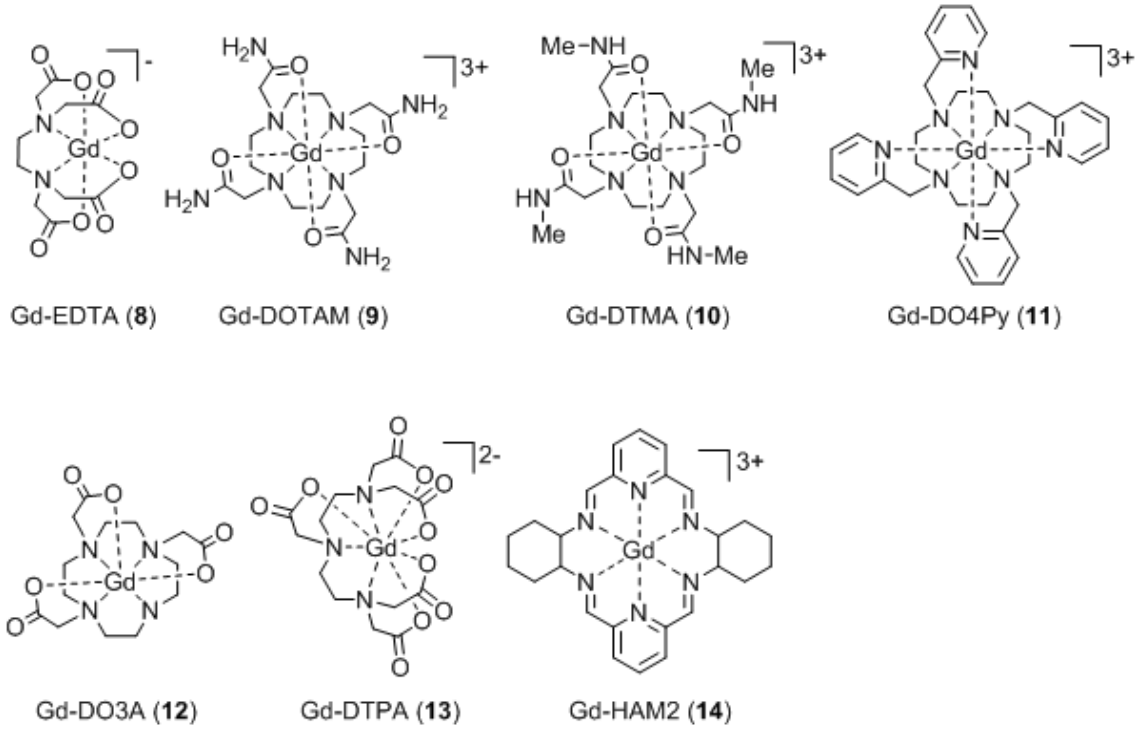

Figure 3.13: Structural formulae and naming of the Gd(III) complexes **8** - **14** considered in further predictions of ZFS with the superposition model.

D and E (or E/D) for Gd(III) complexes in frozen glassy solutions. These models were evaluated by comparing simulated EPR lineshapes with experimental spectra recorded at Q-band ($\sim$ 35 GHz), W-band ($\sim$ 95 GHz), and 240 GHz. Finally, we developed a superposition model, calibrated with our determined ZFS parameter values for four Gd(III) complexes, to provide predictions for additional Gd(III) complexes which can be experimentally verified at a later date.

The three models investigated were found to give equivalent results for the magnitude of the ZFS parameters D and σ_D to within our determined uncertainty on these parameter values. For all three models, the distribution of D was found to be a bimodal Gaussian distribution, with the $+D$ and $-D$ contributions placed approximately

symmetrically about $D = 0$. For spectra recorded at Q-band, the EPR lineshape can be reasonably well reproduced under the assumption that the bimodal D distribution is symmetric about $D = 0$ (Model 2). However, the spectra recorded at 240 GHz displayed a significant asymmetry which in order to reproduce in the simulations required the distribution of D to be asymmetric. In Model 3, this was introduced as an additional free parameter $P(+D)/P(-D)$ to set the relative amplitudes of the $+D$ and $-D$ modes in the D distribution. Model 1 was originally defined as two uncorrelated single-Gaussian distributions for D and E. However, after correcting these distributions according to the standard definitions of the D and E parameters this model also produced a bimodal distribution very similar to that of Model 3, though in Model 1 the relative amplitudes of the $+D$ and $-D$ modes in the D distribution are fixed. Additionally the recomputed bimodal D distributions from Model 1 have, for some Gd(III) complexes, a significant probability at $D = 0$. This is unphysical and would lead in the simulated lineshape to an artificially sharp and narrow central peak at the g-value position.

Models 2 and 3 employed an E/D distribution which was broadly distributed in the range $0 \leq |E/D| \leq 1/3$ and described by the polynomial $P(E/D) \propto (E/D) - 2 \cdot (E/D)^2$. After the correction to Model 1, a similar distribution of E/D was recovered, though the agreement was not as good as for the distributions of D. In particular, the distribution of E/D for Model 1 has a significant probability at $E = 0$, corresponding to full axial symmetry, which is likely not physical for Gd(III) complexes in frozen glassy solutions. The maximum of $P(E/D)$ in Model 1 also deviates from that in Models 2/3, particularly for

the complexes iodo-(Gd-PCTA-[12]) (**6**) and Gd-PyDTTA (**7**). For these two complexes, the maxima is shifted towards smaller E/D compared to the $P(E/D)$ distribution of Models 2/3. Interestingly, this result goes with physical intuition. Depending on the overall sign of D, either the low-field or high-field side of the 240 GHz EPR spectrum can be prominent. For non-axial ZFS ($E \neq 0$), the EPR spectrum becomes more symmetric as E/D approaches 1/3. At the other extreme, for an axial ZFS ($E = 0$) the EPR lineshape is extremely asymmetric. Therefore, it makes sense that for iodo-(Gd-PCTA-[12]) (**6**) and Gd-PyDTTA (**7**), which have the most dramatically asymmetric spectra of the six complexes measured in this study, that the peak of the E/D distribution would be shifted towards zero. Similarly, the complexes Gd-NO_3Pic (**1**), R-(Gd-PyMTA) (**4ab**), and Gd-TAHA (**5**), which have the most symmetric 240 GHz EPR spectra of the complexes measured, have the peak of the E/D distribution shifted towards $E/D = 1/3$ with respect to the $P(E/D)$ distribution used for Models 2/3 (Appendix C.6).

For the six Gd(III) complexes considered in this chapter, the magnitude of the ZFS was found to be in the range of approximately D = 450 - 2000 MHz. Of these, only Gd-DOTA (**2**)/Gd-maleimide-DOTA (**3**) was found to have a negative D value. For all three models, we find that $\sigma_D/D \approx 1/3$, which is comparable to that which has been assumed in the literature. Gd-TAHA (**5**) displayed the most symmetric EPR spectrum at 240 GHz, which resulted in a fairly satisfactory fit with Model 2. In contrast, the other Gd(III) complexes displayed a greater degree of asymmetry in the 240 GHz EPR spectrum and the quality of the fit was found to improve greatly with the addition of

the asymmetry parameter $P(+D)/P(-D)$ in Model 3. The effect of this asymmetry parameter was less evident at the lower microwave bands, but still slightly improved the quality of the fit at Q-/W-band.

Of the three models tested in this work, the distribution of ZFS parameters described by Raitsimring *et al.* [125], with the addition for the allowance of an asymmetry in the bimodal D distribution and no additional line broadening terms, provides the best compromise between a small number of fit parameters and good match between the simulated and experimental EPR spectra. All three models investigated in this chapter seem to be realistic approximations of the true ZFS parameter distributions, but do not perfectly reproduce the experimental EPR spectra of the Gd(III) complexes. This remaining discrepancy between the simulated and experimental lineshapes reflect the limitations of such simplistic models to a complex physical system, and forces us to set relatively large error bars for D and particularly for σ_D. Because the model-related error is so large, we do not expect a significant contribution to the RMSD of the fit from noise present in the EPR spectra, with perhaps the exception of the 240 GHz spectra of iodo-(Gd-PCTA-[12]) (**6**) and Gd-PyDTTA (**7**). The addition of a small convolutional line broadening helps to reduce the RMSD of the fit, indicating that the model fails to capture all of the contributions to the lineshape in the region of the central peak. This region of the EPR spectrum is dominated by the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition, which is only broadened to second-order by the ZFS. This results in this transition remaining narrow at high fields and thus very sensitive to other terms in the effective spin Hamiltonian

which may be present, such as higher-order ZFS terms, hyperfine interactions, dipolar interactions, etc. However, the inclusion of a phenomenological line broadening term makes the parameter space too large to confidently assign parameter values given the relatively featureless EPR spectra typical of Gd(III) complexes in frozen glassy solutions. More careful assignment of other contributions to the width of the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition is beyond the scope of this work.

The combining of fit results from EPR data acquired at Q-band ($\sim$ 35 GHz), W-band ($\sim$ 95 GHz), and 240 GHz served to greatly enhance the confidence of the determined ZFS parameter values, as has been noted by previous studies of a similar nature [39, 21]. Beyond simply providing more data with which to compare fit parameters, the different microwave bands give slightly different information on the ZFS due to the varying contributions of the EPR transitions at different measurement frequencies and temperatures. However, the RMSD error contour plots for Models 2 and 3 imply that a rough estimation of the ZFS parameters D and σ_D would be possible with data from only a single measurement frequency at Q-band or above. The exception to this is the determination of the $P(+D)/P(-D)$ parameter, which required the use of 240 GHz data recorded at low temperature. At frequencies below Q-band, the determination of ZFS parameters becomes difficult due to the magnitude of the ZFS becoming comparable to the measurement frequency. The report of Benmelouka *et al.* [21] indicated that the ZFS parameters determined using EPR measurements at 240 GHz could successfully simulate measurements carried out at X-band ($\sim$ 9.5 GHz). However, when we attempt to

simulate X-band data recorded for a subset of the six Gd(III) complexes studied in this work, we fail to reproduce the low-frequency lineshape using the determined best-fit ZFS parameters for Model 3 (Appendix C.10). We attribute this to the dominant contribution of the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition in the EPR spectra recorded at X-band and 30 K, which are simulations fail to accurately reproduce even for the high-field data. However, this result is still surprising given the success of Benmelouka *et al.* in simulating low-frequency EPR data [21].

Several of the Gd(III) complexes studied in this Chapter were relatively recently synthesized and so do not have ZFS parameter values available in the literature for comparison. Fortunately, there is a large body of work regarding the ZFS of the DOTA family of complexes, due to their more common use as MRI contrast agents. There are also a couple of reported values for the PyMTA ligand due to the recent use of this complex as a spin label for distance measurement by high-field CW EPR (see Chapter 2) and pulsed EPR techniques such as DEER and RIDME at high-field. Raitsimring *et al.* [125] reported values of $D = 669$ MHz and $\sigma_D = 334$ MHz for Gd-DOTA. This work also quotes a value of $D = 641$ MHz for Gd-DOTA from the work of Clarkson *et al.* [39]. Benmelouka *et al.* give ZFS parameter values for Gd-DOTA of D = -570 MHz, σ_D = 570 MHz, $E = 0$ MHz, $\sigma_E = 390$ MHz, with an estimated error of 60 MHz on each of these values [22]. Note that the values reported by Benmelouka *et al.* were determined using what we refer to as Model 1. If these values were corrected as discussed above, the D value would be in much better agreement with that determined by Raitsimring, *et al.*.

More recently, ZFS parameter values were reported for the complexes Gd-DOTA and Gd-PyMTA based on EDEPR measurements at W-band. When the W-band data was analyzed using what we refer to as Model 2, this gave ZFS values for Gd-DOTA of D = 500 and σ_D = 190 MHz [101]. A previous report [128] analyzed the W-band spectrum of Gd-DOTA using a model similar to our Model 1 and gave values of D = 500 MHz, σ_D = 450 MHz, E = 167 MHz, σ_E = 100 MHz. However, this analysis differs from our analysis with Model 1 in that the distribuitons of D and E were computed using the EasySpin [155] functions `DStrain` and `EStrain`, which are input as the FWHM of a Gaussian distribution of the parameters D and E. These functions compute the strains in an approximate way, corresponding to the first term in a Taylor expansion or to first-order perturbation theory, and are valid only as long as the width of the distribution is much smaller than the parameter itself. Additionally, the analysis in this report assumed that the D and E are correlated. Due to these differences, it is unclear how to interpret the reported values for σ_D and σ_E, though the value for D is similar to that reported by Manukovsky *et al.* [101]. Reported ZFS parameter values for W-band EDEPR spectra of Gd-PyMTA analyzed with Model 2 gave D = 1150 MHz, σ_D = 575 MHz [45] and D = 1150 MHz, σ_D = 300 MHz [101].

The available literature values for Gd-PyMTA are slightly lower than the values for D that we find in this work, though still within our determined error bars for Models 2/3. The literature values for Gd-DOTA are also in reasonable agreement with our findings, though again generally somewhat smaller values for D are found in the literature. This is

likely in part due to our use of two distinct complexes, Gd-DOTA (**2**) (at 240 GHz) and maleimide-Gd-DOTA (**3**) (at Q-/W-band), which were not differentiated in our analysis. However, inspection of the contour plots of RMSD error seem to indicate a smaller D value for Gd-DOTA (**2**) than for Gd-maleimide-DOTA (**3**). This trend is also supported by the analysis with the superposition model, which indicates a value of $D = 698$ MHz for Gd-DOTA and $D = 736$ MHz for Gd-maleimide-DOTA (**3**). Not differentiating these two complexes in the analysis perhaps also explains the larger degree of variation in the best-fit ZFS parameters for this complex compared to the other Gd(III) complexes. Of the ZFS parameter values we calculate for Gd-DOTA (**2**)/Gd-maleimide-DOTA (**3**), the closest to literature value is the $D = 660$ MHz determined by Model 3 with the full EPR spectra included in the analysis. If we instead separate the results for the two complexes, we arrive at values of $D = 600$ MHz $\pm$ 198 MHz for Gd-DOTA (**2**) and $D = 750$ MHz $\pm$ 200 MHz for Gd-maleimide-DOTA (**3**), which is in much better agreement with literature values and also to the predictions from our superposition model. This increase in the magnitude of the ZFS for Gd-maleimide-DOTA (**3**) compared to Gd-DOTA (**2**) is most likely due to the addition of the maleimido linker, which slightly lowers the symmetry of the complex. The two measured R-(Gd-PyMTA) complexes did not show a similar difference in the magnitude of the ZFS parameter D (see also Appendix C.11), likely because the atoms which directly coordinate the Gd(III) are identical for both complexes, with the change in functional group being well separated from the first coordination shell.

The change to the ZFS upon the addition of a maleimido, MTS, or other linker will

be an important consideration when designing new Gd(III) complexes for use as spin labels. It is, at present, difficult to predict what effect the addition of a linker will have on the EPR spectrum of a Gd(III) complex. The superposition model discussed in this chapter appears to predict reasonably well the change in ZFS between Gd-DOTA and Gd-maleimide-DOTA, and so may be a useful tool in predicting the ZFS for new Gd(III) complexes that are functionalized for use as spin labels for EPR.

Chapter 4

Structure studies of proteorhodopsin

The PR samples studied in this chapter were prepared and spin labeled by several graduate and undergraduate students working in the Han lab at UCSB. Evelyn Chang and Naomi Baxter prepared the PR samples for the W-band DEER measurements, under the supervision of Sunyia Hussain and Nikki Schonenbach. The PR samples for CW EPR measurements were prepared by Chung-ta Han and Sirish Narayan. EPR measurements and analysis in this chapter were performed by JAC. The capillary sample holders for room temperature measurements were developed by C. Blake Wilson.

4.1 Introduction

As was shown in Chapter 2, 240 GHz CW EPR with Gd(III)-based spin labels has extended the upper measurable distance limit of the CW EPR technique to more than 3.5 nm, depending on the instrinsic linewidth of the spin label. CW EPR is readily

applicable to proteo-liposomes and above the protein dynamical transition temperature, and thus has become an important tool for providing targeted structural information on the nanometer scale on biological systems in native-like environments. In this chapter, we present an application of Gd(III)-based spin labels with high-field EPR to the study of the structure, dynamics, and oligomerization of the membrane protein proteorhodopsin (PR), which functions as a light-driven proton pump and is found in many marine bacteria [19, 18]. PR presents a rather ideal biological system for exploring the benefits and limitations of high-field CW EPR distance measurement with Gd(III) spin labels in a biological system. PR natively spans the lipid membrane of the cell, forms functionally-relevant oligomers, is light-activated, and is structurally homologous to other 7-alpha-helical transmembrane proteins such as the more well-known bacteriorhodopsin and human proteins such as G-protein coupled receptors (GPCRs).

PR was first isolated from a marine bacterioplankton and is similar in form and function to archaeal rhodopsins [19]. Many variants of PR have been found in nature which are spectrally tuned to their environment, and are broadly classified into green- or blue-absorbing PR (G-PR or B-PR) [169, 100]. The work presented in this chapter will focus on the G-PR variant (BAC31A8), which is the most commonly studied proteorhodopsin [18]. G-PR absorbs light at between 520 - 540 nm depending on the protonation state of the primary proton acceptor residue (D97), which can be tuned experimentally by controlling the pH [59].

Oligomers have been shown to be functionally relevant in a variety of proteins, yet the

precise function of oligomerization remains in debate. Oligomeric proteins are generally difficult to study due to their size, varying rigidity, and complex interactions both between proteins which form the oligomer and interactions between the oligomer and lipid membrane (or membrane-mimicking environments such as detergent micelles or nanodiscs). Further, because their organization and function may be tuned by the local environment, being able to study oligomeric organization in the native lipid membrane environment is an important goal. PR has been shown to oligomerize into different states depending on a variety of factors, predominantly organizing into hexameric oligomers with a smaller fraction of pentameric oligomers as identified by AFM [88], which has been shown to have functional implications [79]. The orientation of the individual PR within the oligomer has been previously determined by short- and long-range distance contraints provided by CW EPR and DEER experiments using traditional nitroxide spin labels and Gd(III)-based spin label Gd-4MMDPA [156, 52].

The first goal of the work presented in this chapter is to extend this work with additional distance constraints determined by W-band DEER and 240 GHz CW EPR using Gd(III) spin labels from the DOTA family of ligands. In this work, we studied oligomeric organization, or association of multiple PR molecules, in PR reconstituted into a membrane-mimetic detergent micelle. In the future, the CW EPR technique may allow for the study of the oligomerization of PR in different detergent environments or even within the lipid membrane. Recent work has demonstrated that the function of PR can be tuned by oligomerization, and that the distribution of monomeric and oligomeric

forms of PR may be altered by controlling detergent composition [98, 81]. Further, certain mutations, such as E50Q, have also been shown to disrupt oligomerization [98, 81]. Given the complexity of this system, and for protein oligomers in general, a handful of carefully chosen distance constraints determined by CW EPR could provide key clues to unraveling the structural arrangement of the PR oligomer in response to various protein-protein and protein-environment interactions.

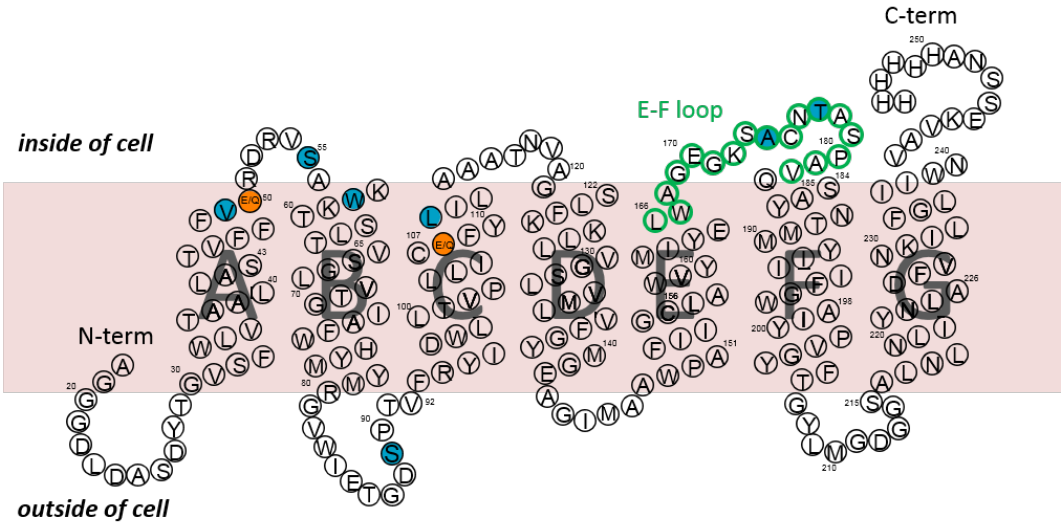

Figure 4.1: Secondary structure of PR as determined by solid-state NMR studies [143, 144, 170], with regions of the protein addressed in this chapter and important functional residues highlighted. The proton donor E108 when mutated to Q results in "slow-photocycle" PR. Residue E50 when mutated to Q enhances monomeric population [98, 81]. Residues 55, 174, and 177 served as probes for studying conformational change of the EF-loop during the photocycle of PR [78]. Residues 55, 58, 89, and 177 served as probes for studies of PR oligomerization [52]. The signal peptide (residues 1-17) is cleaved during expression of PR and is therefore not included in this figure.

The second aim of this work is to develop 240 GHz CW EPR as a tool to provide information on the structural changes which occur in PR during its photocycle. PR can be activated with light to trigger a synchronized cascade of conformational changes

corresponding to different stages in the photocycle, which take place on time scales ranging from <1 μs to more than 1 s. For the purpose of testing CW EPR distance measurements to resolve conformational changes in a protein, we focus on the function-relevant E-F loop segment of PR. This region, along with other residues discussed in this chapter, is highlighted in the secondary structure model of PR shown in Figure 4.1. Upon light activation, an opening of the PR towards the cytoplasmic surface by the outward movement of the E-F loop and/or the F helix has been proposed to be a critical step in facilitating proton pumping in PR, and is thought to be a conserved motif in the activation of other 7TMs. This region is so significant for the function of PR that a single mutation (A178R), which may change the plasticity of the E-F loop has been shown to result in significant changes in the absorption maxima and in the timescale of the photocycle of PR, despite the residue being well-separated from the chromophore [107]. Previous work has shown that the E-F loop is a partially α-helical structure, tucked against the cytoplasmic surface [78, 107]. The E-F loop, upon light activation, twists and lifts outwards as a rigid body concurrent with the decay of the M-state of PR [78]. However, the extent and timing of this motion, and whether or not it is coupled to upward or outward motions of the E- or F-helices, is unknown. These questions could potentially be answered by a carefully selected handful of distance measurements, in a similar fashion to the study carried out by Altenbach *et al.* on rhodopsin [9].

4.2 Pulsed EPR measurements of PR oligomers

In addition to developing Gd(III)-spin probes for distance measurements with 240 GHz CW EPR, extensive work has been undertaken by others to develop Gd(III) distance measurement techniques using high-field pulsed EPR. To date, W-band experiments with the DEER pulse sequence has been the most widely used, and more recently experiments based on the RIDME pulse sequence have been explored on model systems [131]. The advantages of using Gd(III) spin labels with high-field pulsed EPR are similar to the advantages described in the previous chapters for 240 GHz CW EPR - primarily increased sensitivity [65].

In particular, W-band (95 GHz) DEER for distance measurement in biological systems with Gd(III) spin labels has been extensively investigated, and has become the 'gold standard' in the field. Therefore, before embarking on 240 GHz CW EPR measurements on PR, we first conducted a series of W-band DEER experiments on PR oligomers, reproducing the work of Edwards *et al.* [52] but with maleimide-Gd-DOTA as the spin label instead of Gd-4MMDPA. The Gd-DOTA spin label is expected to have several advantages over the previously used Gd-4MMDPA. The DOTA family of spin labels has some of the narrowest EPR lineshapes observed for Gd(III) complexes. Because this complex was originally developed for use as an MRI contrast agent, there are a wide variety of DOTA-type complexes available, all extensively characterized, displaying high binding affinities, and are commercially available. Several Gd-DOTA complexes have been functionalized for use as spin-labels, most commonly by the addition of a maleimide

or MTS linker which binds to cysteine residues on the protein.

The work in this section had the following primary goals: (i) verify successful spin-labeling of PR with maleimide-Gd-DOTA, (ii) demonstrate distance measurement in PR oligomers with maleimide-Gd-DOTA as the spin label reproduces the distance constraints on PR oligomer organization previously determined by Stone *et al.* [156] and Edwards, *et al.* [52], (iii) explore if the narrow linewidth and longer relaxation time T_m of maleimide-Gd-DOTA confers any greater sensitivity to the W-band DEER distance measurement as compared to the results with Gd-4MMDPA, and (iv) explore the viability of conducting these measurements in liposome environments. For this purpose, we generated three mutants of PR singly-labeled with maleimide-Gd-DOTA at residues 55, 58, and 177. Spin dilutions of 1:1 and 1:3 (spin labeled PR:unlabeled WT-PR) were prepared for each sample to control for potential multi-spin effects, which occur when there are significant dipolar interactions between more than two spins. However, based on previous literature, multi-spin effects are expected to be small for Gd(III)-based spin labels compared to what has been observed for nitroxide spin labels due to the small modulation depth when using Gd(III) labels [83, 85, 52, 65]. W-band DEER measurements were carried out at 10 K and distance distributions were extracted from the DEER trace using the MATLAB based software *DD* [150, 32], and are presented in Figure 4.2. In all cases, a two-component Gaussian distance distribution was found to give the best fit. The background decay was assumed to be exponential and was included as a fit parameter. Background-subtracted DEER traces for PR labeled with maleimide-Gd-DOTA at sites 55, 58, and 177 are

presented in Figure 4.2A, and the corresponding best-fit distance distributions in Figure 4.2B. The spin dilutions of 1:1 and 1:3 gave comparable results for sites 58 and 177, suggesting that multi-spin effects were negligible with the maleimide-Gd-DOTA spin label. The 1:1 spin dilution for site 55 was not measured due to time constraints.

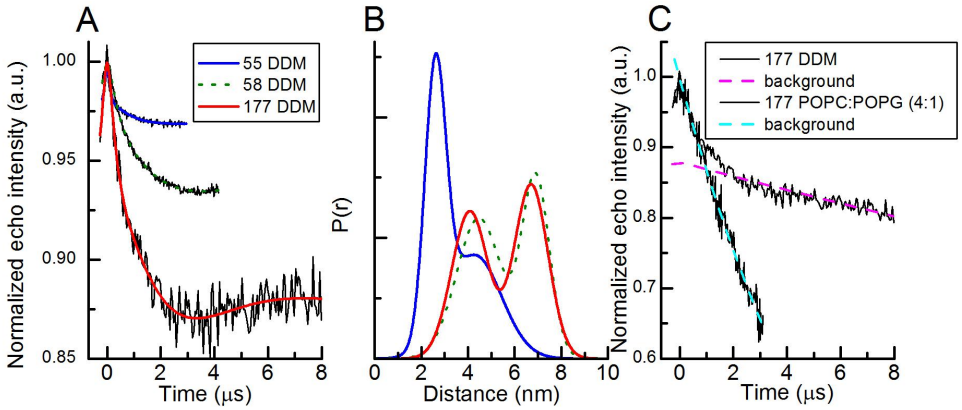

Figure 4.2: Results of W-band DEER measurements of PR oligomers in DDM detergent micelles labeled 1:1 with maleimide-Gd-DOTA at sites 55, 58, and 177. (A) Background-corrected time domain DEER traces and calculated best-fits, (B) distance distributions corresponding to the best-fit calculated DEER traces in (A) using a two-component Gaussian distribution, and (C) raw DEER traces for PR oligomers labeled at site 177 in DDM detergent micelles and in POPC:POPG (4:1) lipid vesicles.

For a hexameric structure, we would expect two nearest-neighbors (r_1), two next-nearest neighbors (r_2), and the diametrically opposed protein (r_3). For hexagonal symmetry, these should occur with probability ratios of 2:2:1 and inter-protein distances should have ratios of $r_1 : r_2 : r_3 = 1 : 1.73 : 2$. For site 55, the nearest-neighbor distance was not resolved, as the expected 1.6 nm inter-protein distance for this residue is below the lower-limit resolvable by W-band DEER with Gd(III) spin labels. The next-nearest neighbor distance for site 55 was found to be $r_2 = 2.8$ nm with 55% population, which is in agreement with the findings of Stone *et al.* [156] that $r_1 = 1.6$ nm and assuming

a hexameric organization of the oligomer. The second peak in the distance distribution was very broad and ill-resolved, with a mean distance of 4.2 nm - well beyond the 3.2 nm distance predicted by a hexameric structure. This discrepancy was somewhat surprising, as the 3 μs dipolar evolution time of this measurement should allow for accurate distance determination up to $\sim$ 4.6 nm. This assumes the usual convention (used in e.g. the DEER Analysis software) that for a 2 μs evolution time, the assignment of mean distance and distribution width is reliable out to 4 nm, the assignment of mean distance along is reliable out to 5 nm, and that these scale with $(T_{evo}/2\mu s)^{1/3}$ [52]. For PR labeled at site 177, the two-Gaussian fit for the distance distribution gave mean distances of r_1 = 4.1 nm and r_2 = 6.7 nm, with 48% and 52% populations, respectively. This is in good agreement with a hexameric assembly and with the findings of Edwards *et al.* [52], who reported a nearest-neighbor distance of r_2 = 3.9 nm. The furthest distance r_3, expected to be at $\sim$8 nm, could not be resolved in the 8 μs dipolar evolution window of this experiment. W-band DEER measurements of PR labeled at site 58 gave two distances centered at 4.4 nm and 6.8 nm, which is not consistent with the expected values of r_1 = 2.3 nm, r_2 = 3.9 nm, and r_3 = 4.6 nm. Rather, this distribution is nearly identical to that expected for PR labeled at site 177. This discrepancy was attributed to a mislabeling of the "PR-58" sample tube, which we believe based on the results of the DEER measurement actually contained and sample of PR-177.

W-band DEER measurements were additionally attempted on PR oligomers labeled at site 177 with maleimide-Gd-DOTA and reconstituted into POPC:POPG (4:1) lipo-

somes, but no oscillations were observed in the DEER trace. It was observed that the lipid sample had a much shorter phase memory time (T_m) compared to the detergent sample and that the background decay of the DEER signal was much steeper, suggesting a high local spin concentration possibly due to a high local concentration of the spin label within the liposome.

These measurements successfully verified that the maleimide-Gd-DOTA spin label reproduced the distances determined previously [52, 156] in the PR hexamer labeled at residues 55 and 177. However, in light of recent publications regarding the analysis of W-band DEER spectra with Gd(III) spin labels with small ZFS and separated by short distances [45, 128, 101], these extracted distance distributions are likely not be reliable with the current analysis, despite the agreement with literature values. The measurement in liposome was unsuccessful in this attempt, and will need future work to improve sample preparation protocols to reduce the local concentration of spin label in the lipid membrane. This could perhaps be accomplished by using a very low protein to lipid ratio, or by preparing PR in nanodiscs. These PR samples in DDM were additionally measured by 240 GHz CW EPR at 30 K, but no peak-to-peak broadening of the lineshape was observed, even for the PR-55 sample which the expected 1.6 nm inter-PR distance and is expected to have a very large broadening of the CW EPR lineshape. This was attributed to poor spin labeling efficiency, as will be discussed next.

4.3 Quantitative estimation of spin labeling efficiency

Poor spin labeling efficiency and possible non-selective binding with the maleimide-Gd-DOTA spin label were identified as a key roadblock in studies of distance measurements in PR with 240 GHz CW EPR lineshape analysis. Unlike the DEER technique, whose signal comes only from dipolar coupled spins, the CW EPR lineshape is a combination of contributions from all Gd(III) species present in the sample, including pairs of dipolar-coupled spin labels, single non-dipolar-coupled spin labels, and Gd(III) ions or Gd(III) spin label free in solution (i.e. not bound to the protein). Of these, the Gd(III) complexes which do not have a close neighboring Gd(III) spin will have the narrowest CW EPR lineshape and therefore dominate the overall measured CW EPR spectrum. Because low-temperature 240 GHz CW EPR distance measurements are relatively time-consuming and costly due to cooling time and liquid helium consumption, we sought a fast, simple method of estimating labeling efficiency to quickly ascertain (i) which samples would be most suitable for CW EPR distance measurement, (ii) identify new sites in PR which display high labeling efficiencies, and (iii) optimize spin labeling protocols and test new spin labeling chemistry.

The concentration of PR in a sample can be easily measured by a UV-VIS absorption measurement and the application of Beer's law. This measurement informs what the Gd(III) concentration of the sample should be in the case of 100% spin labeling efficiency.

The measurement of Gd(III) concentration was done by a method similar in concept to a spin-counting experiment [50], though significantly less refined. In this measurement, a calibration series of samples of known volume and concentration consisting of the Gd(III) spin label in the PR buffer solution are measured by room temperature 240 GHz CW EPR. As one would expect, the double integral of the CW EPR signal (i.e. the area of the absorption lineshape) increases linearly with Gd(III) concentration (Figure 4.3). The same measurement is then performed on an equal volume of spin-labeled PR sample of unknown labeling efficiency, compared to a line fit to the calibration curve, and an estimate of Gd(III) concentration is extracted and used to compute labeling efficiency. The results of these spin labeling efficiency measurements for PR spin labeled at various sites with maleimide-Gd-DOTA are summarized in Table 4.1.

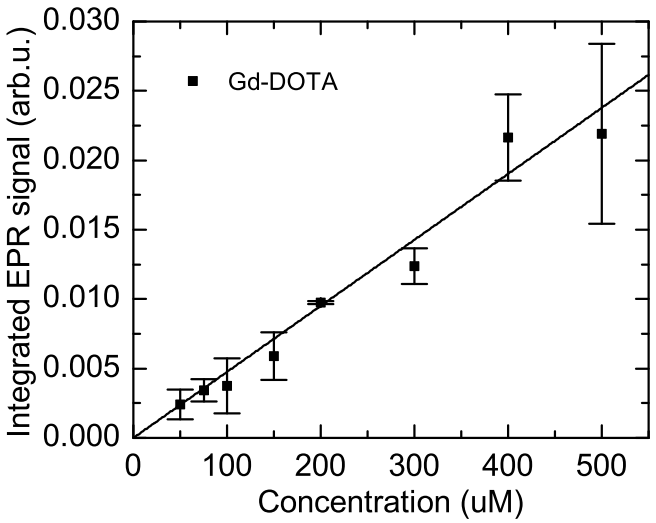

Figure 4.3: Concentration calibration series of Gd-DOTA in PR buffer prepared at various known concentrations. Plotted is the mean of the double integral of the CW EPR signal taken from several measurements of the same sample. Error bars are two standard deviations of the mean of those measurements. Measurements were carried out at $\sim$ 290 K.

Unlike a true spin-counting experiment, the Gd(III) concentrations determined by this method can only be considered as a rough estimate, accurate to within maybe a factor of two. There are many sources which contribute to this inaccuracy, the primary of which is repeatability of sample loading and stability of the spectrometer bridge alignment over time. Additionally, the calibration measurement (e.g. in Figure 4.3) used Gd-DOTA and not the actual spin label, since the spin labels maleimide-Gd-DOTA and Gd-MTS-ADO3A were both found to partially dimerize or aggregate in solution. The CW EPR lineshapes of these three Gd(III) complexes are similar enough to be approximately accurate in a spin counting measurement, but obviously this poses another source of error in the measurement. In order to conduct these measurements at room temperature, a new sample holder geometry was developed to allow for the acquisition of high-SNR room temperature EPR spectra (see Appendix A.2.3), which was not possible using the bucket-style sample holder used for measurements at cryogenic temperatures. The amplitude of the EPR signal resulting from measurements with this sample holder geometry is highly sensitive to not only the sample volume and concentration, but also the positioning of the sample within the spectrometer and the material backing the sample. Nevertheless, the ability to quickly estimate the efficiency of labeling PR with Gd(III)-based spin labels has been invaluable in troubleshooting sample preparation, despite the rather large error bars on the measurement.

4.4 Improving Gd(III) spin labeling of PR

Several residues were identified in PR as potential candidates for Gd(III) spin labeling and study with 240 GHz CW EPR. Candidate residues were chosen to have inter- or intra-PR distances in the range of 2.0 - 3.8 nm, as estimated from the PDB structure (2L6X) of G-PR [132]. These residues were chosen with the primary goal of studying movement of the E- and F-helices and conformational changes of the E-F loop that occur during the photocycle of PR. The secondary goal of these candidate residues for spin labeling was to study the oligomeric organization of PR by conducting inter-PR distance measurements. Further, an attempt was made to identify residues which would be most ameanable to spin labeling with Gd(III)-based spin probes, based on the known protein structure, solvent accessibility of the residue, and prior spin labeling experience. Finally, the choice of residues for spin labeling is further limited by the restriction that the mutation should not otherwise affect the structure or function of the protein.

In particular, residue A174 at the solvent-protein interface on the E-F loop helix has been previously shown to display a large conformational change upon light activation, with the spin labeled residue becoming much more immobile in the light state compared to the dark state [77]. Residue T177 on the E-F loop is also known to show changes with light activation [78, 77]. The A- and C-helices have been hypothesized to not move significantly during the photocycle, and so residues S55 (A-B loop on oligomer interface), V49 (A-helix), and L111 (C-helix) were chosen as partner sites at appropriate distances for studying motion at residues A174 and T177. Residues S55, W58, and S89 were

identified as potential candidates for studying the oligomeric organization of PR. The locations of these various residues are shown in Figures 4.1 and 4.4.

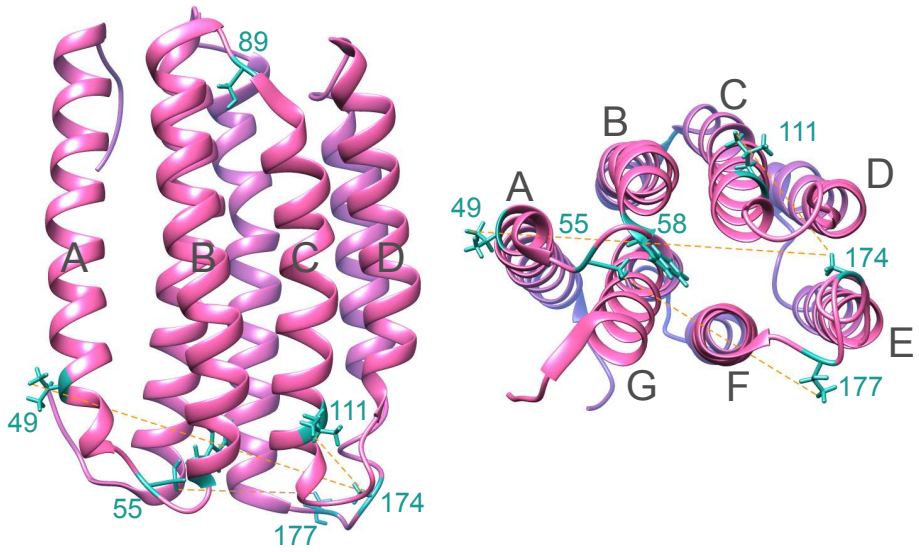

Figure 4.4: Structural model of PR based on PDB 2L6X [132]. Candidate sites for spin labeling with Gd(III)-based spin probes are highlighted. The dotted lines represent intra-protein distances of interest to resolve the motion of the E-F loop. These pairs have estimated intra-PR distances of 2.0 nm (55-177), 1.5 nm (111-174), and 3.5 nm (49-174). The addition of a spin label extends these distances by 0.2 - 0.3 nm. Figure was created using the UCSF Chimera package [113].

For the samples listed in Table 4.1, various of these identified residues in PR were tested for spin labeling efficiency with maleimide-Gd-DOTA and where the sample preparation and spin labeling protocol was not changed significantly between these samples. It was observed, in the small number of cases tested, that monomeric PR had the highest labeling efficiencies, hexameric PR somewhat less, and high-order oligomers or aggregate had very poor spin labeling efficiencies. This has been hypothesized to be due to better accessibility of the residues for monomeric PR compared to the hexamer and higher-order oligomers/aggregate. In particular, residue S55 may be difficult to spin label in

the hexameric state due to its position at the oligomeric interface. The population of monomeric PR in a sample can be enhanced either by size-exclusion chromatography to remove oligomers and aggregates from the sample, or by the introduction of the E50Q mutation. The E50Q mutation results in a greatly enhanced monomeric population compared to wild type seemingly without any effects on the optical properties of G-PR, which has been proposed to be the result of disruption of the H-bonding network that stabilizes the hexamer [98, 81]. Recent work has also shown that different surfactant conditions can be used to stabilize different equilibrium populations of monomeric and hexameric PR [81].

Sample	Labeling efficiency (%)
PR-S89C slow hex 1:1	144
PR-S89C slow hex 1:3	69
PR-S89C slow hex 1:6	90
PR-A174C-L111C hex 1:6	19
PR-A174C-V49C hex 1:6	41
PR-A174C hex 1:1	24
PR-V49C hex 1:1	46
PR-T177C slow hex 1:1	36
PR-S55C slow hex 1:1	92
PR-S55C-T177C slow mono 1:1	26
PR-S55C-T177C slow hex 1:1	26
PR-S55C-T177C slow agg 1:1	$\sim$5

Table 4.1: Estimated spin labeling efficiencies for various PR mutants spin labeled with maleimide-Gd-DOTA. The concentration of PR was determined by UV-VIS absorption at 520 nm. The concentration of maleimide-Gd-DOTA was estimated by a spin-counting type experiment using room temperature 240 GHz CW EPR and compared to a Gd-DOTA calibration. Estimated labeling efficiencies were calculated from these two numbers, taking into account the dilution with WT-PR for the 1:3 and 1:6 samples. Note: due to various uncertainties in the measurement, these values have rather large error bars - perhaps as large as a factor of two.

The measurements in Table 4.1 seem to indicate that the PR-S55C hex sample had

very good labeling efficiency, however 240 GHz CW EPR measurements at 30 K did not show the very dramatic lineshape broadening expected for the nearest-neighbor distance of 1.6 nm in the hexamer. This effect was observed on multiple preparations of this sample and the origin was never fully explained. However, there were hints from the 240 GHz CW EPR data of PR labeled with maleimide-Gd-DOTA as well as unrelated experiments on X-band DEER of Tau-187 G272C/S285C labeled with maleimide-TEMPO (a standard nitroxide spin label) that the maleimide-Gd-DOTA may be non-selectively binding (i.e. attaching to residues other than mutated cysteine at the residue of interest) to the PR, possibly reacting with the primary amine on the protein [33, 4]. Oxidation of the cysteine residues which the maleimide-Gd-DOTA spin label binds to could also be a potential cause of poor spin labeling efficiency.

To test this hypothesis, two samples of PR-T177C 1:1 slow with maleimide-Gd-DOTA spin label were prepared, with and without TCEP added during the spin labeling step. TCEP (tris[2-carboxyethyl]phosphine) is a reducing agent which can prevent the formation of disulfide bridges and consequential inactivation of the cysteines [166, 87]. However, estimation of spin labeling efficiencies of these two sample preparations showed that TCEP somewhat reduced the labeling efficiency from $\sim$ 36 % in the original sample preparation to $\sim$ 21 % with the addition of TCEP during the spin labeling reaction. This reduction in labeling efficiency could be the result of the reducing agent interfering with the reaction between the cysteine and the spin label. This might be improved by ensuring complete removal of the reducing agent before conjugation with the spin label

so that the thiol group of TCEP does not compete with the target thiols in the protein [87].

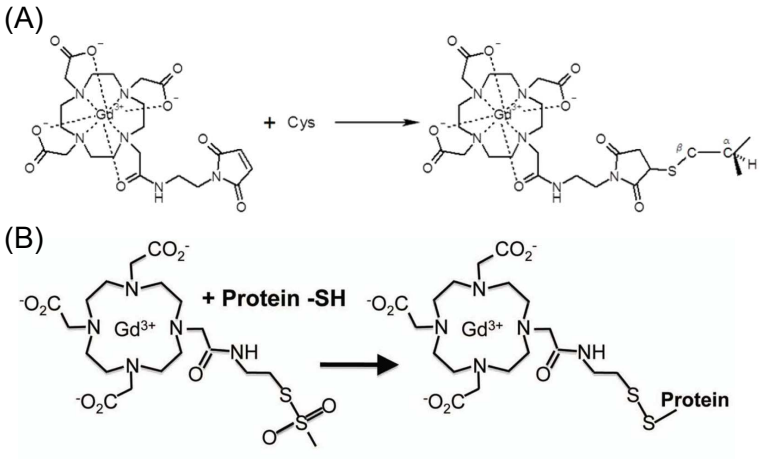

Figure 4.5: Labeling reaction of (A) maleimide-Gd-DOTA [103] and (B) Gd-MTS-ADO3A [92] with cysteine residues.

Due to concerns of non-selective binding of the maleimide-Gd-DOTA spin label to cysteine residues in PR, a situation which may be aggrevated at high pH [57], we did not attempt any further optimization of the spin labeling protocol with maleimide-Gd-DOTA and instead moved on to test a different labeling chemistry. We chose to use the spin label Gd-MTS-ADO3A (Figure 4.5), which has a ligand structure similar to the DOTA family of complexes that results in a narrow EPR lineshape and has been previously used for W-band DEER distance measurements [106]. This spin label reacts with cysteine residues in the protein via the methanthiosulfonate (MTS) group, which provides specific conjugation to the targeted cysteine residue via a disulphide bridge [92].

Estimated labeling efficiencies with the Gd-MTS-ADO3A spin label for various PR samples are listed in Table 4.2. The change in spin label from maleimide-Gd-DOTA

to Gd-MTS-ADO3A increased the labeling efficiency of PR-T177C from 36% to 76%. Again, we observe that monomeric samples have higher spin labeling efficiencies than oligomeric samples. We made several attempts to increase the labeling efficiency further. Increasing the pH from 5.8 to 7.5 during the reaction step of binding the spin label to PR did not affect the labeling efficiency, nor did increasing the ratio of spin label to PR during this reaction step. However, letting the spin labeling reaction go overnight at room temperature did increase labeling efficiency compared to letting the reaction run for a couple hours at 4°C. Similarly to the studies with the maleimide-Gd-DOTA spin label, the overall labeling efficiency can be increased by removing aggregate PR from the sample, either by size-exclusion chromatography or by introduction of the E50Q mutation.

Sample	Effic. (%)	Notes
PR-S55C-T177C slow FPLC mono 1:1	47	
PR-S55C-T177C slow FPLC hex 1:1	32	
PR-S55C-T177C slow E50Q 1:1	46	pH 7.5
PR-T177C slow mix oligomer states 1:1	76	∼21°C labeling
PR-S55C hex 1:1	35	∼21°C labeling

Table 4.2: Estimated spin labeling efficiencies for various PR mutants spin labeled with MTS-Gd-ADO3A. The concentration of PR was determined by UV-VIS absorption at 520 nm. The concentration of MTS-Gd-ADO3A was estimated by a spin-counting type experiment using room temperature 240 GHz CW EPR and compared to a Gd-DOTA calibration. Note: due to various uncertainties in the measurement, these values have rather large error bars.

For the PR-S55C-T177C sample, which will be interesting for future studies on the motion of the EF-loop region, we have increased labeling efficiency from ∼25% to ∼50%. Correspondingly, the CW EPR spectra for 50% labeling efficiency shows significantly

more dipolar broadening of the lineshape in the doubly labeled sample than did the sample with 25% labeling efficiency. This level of spin labeling efficiency may be sufficient to begin estimating distances from dipolar broadening of the 240 GHz CW EPR lineshape, as will be shown in Section 4.5, but is still far below our target of >80% labeling efficiency. Future approaches to optimize spin labeling efficiency with Gd-MTS-ADO3A may include (i) the addition of a reducing agent (DTT) during the PR purification step and careful removal of the reducing agent during the spin labeling step so that cysteine oxidation is prevented but the reducing agent does not interfere with the labeling reaction and/or (ii) increasing the temperature (possibly to ∼30°C) during the spin labeling reaction.

4.5 CW EPR measurements of PR oligomers

After confirming that we can reproduce the previously determined PR oligomer organization [156, 52] with W-band DEER using the new maleimide-Gd-DOTA spin labeling scheme (Figure 4.5A), we moved on to replicating and expanding on these results with the 240 GHz CW EPR distance measurement technique developed in Chapter 2. In the DEER studies, the PR-55 sample provided a connection between the CW EPR studies of Stone *et al.* at X-band with a nitroxide spin label [156] and the W-band DEER studies with Gd(III)-based spin labels [52]. PR-58, while providing an ideal inter-PR distance of ∼ 2.3 nm, where the 240 GHz CW EPR measurement is expected to be very sensitive, was found to be too difficult to spin label likely due to geometric restriction of the Gd-DOTA spin label and general inaccessibility of the residue. PR-177, while of great

interest for study of the motion of the EF-loop region of PR as will be discussed later, presents too long of an inter-PR distance in the hexamer ($\sim$ 4.0 nm) to be measured by 240 GHz CW EPR. Therefore, several new mutations were generated for the purpose of measuring additional inter-PR distances to refine the PR oligomer structure with high-field CW EPR, and for potential future use in studying motions of the EF-loop region (Section 4.4).

Of the various residues in PR where spin-labeling was attempted with maleimide-Gd-DOTA, the PR-S89C samples displayed the highest binding efficiencies (Table 4.1). For PR in a hexameric oligomer, the inter-PR distance at site 89 should be within the range of distances measurable by 240 GHz CW EPR with Gd(III)-based spin labels. Three samples of hexameric PR-89 were prepared with spin dilutions of 1:1, 1:3, and 1:6 with WT-PR ("wild-type" PR, which has all native cysteine residues removed so that the spin label does not bind to the PR) for measurement by 240 GHz CW EPR. The 1:6 spin dilution has, on average, only one spin labeled PR per hexamer and should therefore display approximately no dipolar broadening of the CW EPR lineshape. This sample therefore serves as a reference of the unbroadened, or intrinsic, linewidth of the maleimide-Gd-DOTA spin label bound to residue 89 on PR. At the other limit, the 1:1 sample in principle has every PR in the hexamer spin labeled, though in practice this may be somewhat less depending on the labeling efficiency. Similarly, the 1:3 spin dilution has on average two of the six PR in a hexamer spin labeled.

CW EPR measurements at 240 GHz and 30 K of PR-89 hexamer showed significant

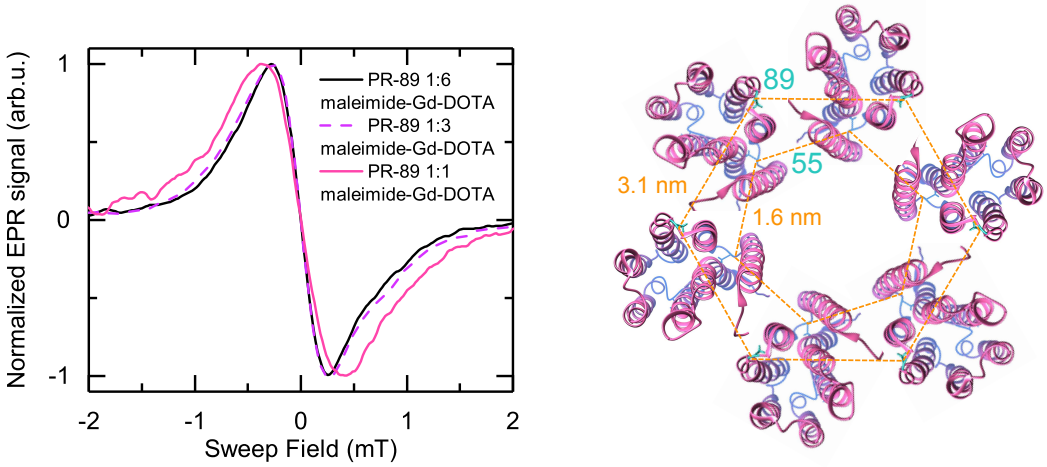

Figure 4.6: CW EPR at 240 GHz and 30 K of PR hexamer labeled at residue 89 with maleimide-Gd-DOTA at spin dilutions of 1:1, 1:3, and 1:6 with unlabeled WT-PR (left). Peak-to-peak broadening of the CW EPR lineshape from the 1:6 to 1:1 spin dilutions indicate an inter-PR distance at residue 89 of ∼3.1 nm in the hexamer. Model of the structure of the G-PR hexamer (right) based on the B-PR crystal structure [129], G-PR monomer structure [132], X-band CW EPR [156], W-band DEER [52] (also from this work), and 240 GHz CW EPR (from this work), drawn using the UCSF Chimera package [113]..

lineshape broadening between the 1:6 and 1:1 labeled samples, with the 1:3 spin dilution displaying a small amount of intermediate broadening in the wings of the spectrum (Figure 4.6). Analysis of the peak-to-peak broadening of the Gd(III) lineshape of the 1:1 sample compared to the 1:6 sample according to the method described in Section 2.6 gives in inter-PR distance of 3.1 nm, which is in good agreement with the 2.8 nm distance estimated from the crystal structure of homologous B-PR hexamer [129]. These measurements represent the first successful use of Gd(III)-based spin labels with 240 GHz CW EPR in measuring a distance in a biological system well beyond that which is accessible by traditional X-band CW EPR with nitroxide spin labels.

Measurements of the oligomeric structure were attempted with 240 GHz CW EPR

at 30 K on PR-55 hexamer labeled with maleimide-Gd-DOTA, including on the sample successfully measured by W-band DEER (Figure 4.2 A-B), however despite nominally good labeling efficiency at this site (Table 4.1), peak-to-peak broadening of the CW EPR lineshape was not observed. Some small amount of broadening was observed in the wings of these spectra, suggestive of poor spin labeling. In contrast, measurements of PR-55 hexamer spin labeled with Gd-MTS-ADO3A, which had approximately double the spin labeling efficiency as for maleimide-Gd-DOTA, showed significant dipolar broadening of the lineshape (Figure 4.7 A). Still, even with improved spin labeling efficiency, the peak positions did not change significantly. This is further indication that while the spin labeling efficiency was significantly improved by the use of the Gd-MTS-ADO3A spin label, it was still not good enough to apply the peak-to-peak broadening analysis developed in Chapter 2 to extract inter-PR distances.

In general, perfect spin labeling is not possible and, therefore, more sophisticated analysis methods are required which can fit more than a single parameter of the lineshape (such as the peak-to-peak linewidth as in Chapter 2, or the second-moment as in Appendix B), account for contributions of contaminant with singly labeled sample resulting from imperfect spin labeling, and ideally be able to fit multiple distances and distributions. Some progress towards this goal was made in simulating the lineshapes of the Gd-rulers in Section 2.7 using a simplified Hamiltonian. However, these simulations are at present very time consuming and the resulting simulated lineshapes are imperfect as a result the various approximations taken.

Another approach to simulating dipolar broadened CW EPR lineshapes was developed in the work of Edwards *et al.* [53] in which the method of Pake convolution/deconvolution, commonly used for the analysis of interacting $S = 1/2$ spin labels at X-band [122, 153], was extended to the $S = 7/2$ case for use with Gd(III)-based spin labels measured at high field. In the work of Edwards *et al.* [53], a single-Gaussian distance distribution was sufficient to describe the broadening of the CW EPR lineshape of random solutions of $GdCl_3$ in D_2O/glycerol$-d_8$. Further, this model successfully reproduced the expected average interspin distance for various concentrations representing different average interspin distances. For the case of spin labeled proteins, a slightly more sophisticated version of this analysis is required to account for the contribution of singly labeled contaminant in the sample as a result of imperfect spin labeling and for the possible multiple distances which may be observed, e.g. in an oligomer. A variety of software tools for conducting Pake convolution/deconvolution analysis to extract distance distributions have been used in the literature including ShortDistances [7], CWdipFit [139], DipCon [16], and DipBF [16], among others. In this work, we employed the Pake convolution method using MATLAB scripts based on the CWdipFit software [139]. This analysis software assumes a sum of Gaussian shaped distance distributions between dipolar coupled spins, and utilizes Monte Carlo/SIMPLEX curve fitting to the data. CWdipFit is provided as source code, and could thus be altered to use a $S = 7/2$ Pake pattern instead of the default $S = 1/2$ Pake pattern. The only other change made to the original software was to change the default range of distances which are included in the fit. The $S = 7/2$

Pake patterns were generated using the EasySpin toolbox [155], though high-spin Pake patterns could alternatively be generated using the method described by Edwards *et al.* [53].

The Pake convolution method was used to analyze 240 GHz CW EPR spectra of PR hexamer labeled 1:1 at residue 55 with Gd-MTS-ADO3A (Figure 4.7). Spin labeling efficiency has not been measured for this sample, though it is expected to be better than the previously determined ∼35% labeling efficiency for a different preparation of the same sample that did not have the hexamer isolated by size-exclusion chromatography. However, the labeling efficiency was not sufficiently high to allow for the extraction of interspin distances via peak-to-peak broadening of the CW EPR lineshape. Additionally, the nearest-neighbor distance at site 55 in the PR hexamer is only 1.6 nm [156, 52], meaning that 240 GHz CW EPR lineshape should also have contributions resulting from the next-nearest neighbor, which has an inter-PR distance of 2.8 nm assuming a perfect hexamer. Therefore, an analysis scheme was required that could account for multiple distance distributions. Spin dilutions of PR-S5C were not available at the time of these measurements, and so a sample of PRT177C labeled with Gd-MTS-ADO3A was used as an approximate reference of the unbroadened lineshape. A fit to these spectra using CWdipFit with a $S = 7/2$ Pake pattern gave as the best-fit a two-Gaussian distance distribution with mean distances of 1.6 nm and 2.9 nm. The prior is in good agreement with the previously determined r_1 distance (Section 4.2) [156, 52] for PR-55 hexamer and the latter is a very reasonable r_2 distance assuming a hexameric organization of the

oligomer. Further, the reconstructed broadened lineshape, simulated by convolving the "monomer" spectrum with a broadening function generated by the determined distance distribution and corresponding $S = 7/2$ Pake patterns, captures both the peak-to-peak linewidth of the measured PR-S55C hexamer spectrum as well as the additional hump on the low-field side of the spectrum (Figure 4.7).

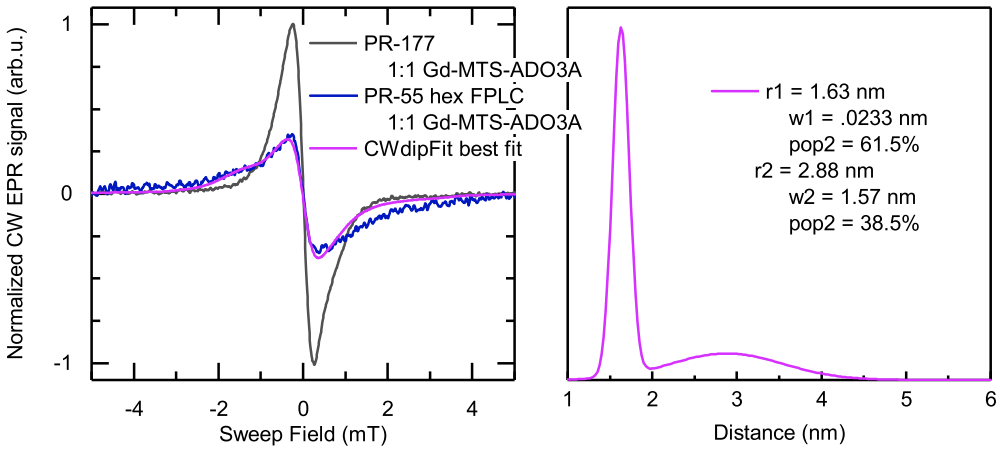

Figure 4.7: CW EPR at 240 GHz and 30 K of PR hexamer labeled 1:1 at site 55 with Gd-MTS-ADO3A (blue) and PR labeled at site 177 (mix of oligomer states) with Gd-MTS-ADO3A (grey), normalized to the double integral. Overlaid is the best-fit dipolar broadened lineshape resulting from Pake convolution analysis (pink) with the corresponding distance distribution shown on the right. The best fit was a two-Gaussian distance distribution having mean distances of 1.6 nm and 2.9 nm, with a ∼20% monomer contribution.

While this analysis is still preliminary, it is rather encouraging that not only do we recover the correct nearest-neighbor distance for the PR-55 hexamer, we also measure the next-nearest neighbor distance. Further encouraging is that this was in a sample with imperfect spin labeling. The CWdipFit analysis indicates that this sample has a ∼20% contribution from spins which are isolated enough to not display dipolar broadening.

However, it is yet to be seen how accurate this estimation of the fraction of non-dipolar-coupled spins from the data analysis is since in this case a proper reference sample was not available, though a face-value interpretation would seem to indicate >50% labeling efficiency for the PR-S55C hexamer labeled with Gd-MTS-ADO3A. The effectiveness of the analysis software in determining monomer fraction in the broadened spectra was tested on a sample of Gd-ruler $\mathbf{1_1}$ (2.1 nm) mixed with Gd-4-iodo-PyMTA (2:1 by volume), where a sample of pure Gd-4-iodo-PyMTA was used as the reference. In this case, analysis with CWdipFit recovered the Gd-ruler distance of 2.1 nm and gave a monomer fraction of 26%, which is reasonably close to the expected 33%.

From a practical use perspective, the Pake convolution analysis has great advantages over both the analysis via peak-to-peak broadening and analysis by first-principles simulation of the effective spin Hamiltonian. The peak-to-peak broadening analysis, while effective for the model system of the Gd-rulers and for other sufficiently clean systems (such as the PR-S89C hexamer), is much too simplistic for generalized analysis of 240 GHz CW EPR spectra of a Gd(III) labeled protein. In these complex biological systems, there are the additional complications to data analysis resulting from imperfect spin labeling and possible multiple distances and/or broad distance distributions, none of which can be properly addressed by looking only at the peak-to-peak linewidth of the CW EPR spectrum. On the other hand, a first principles simulation of the CW EPR lineshape is highly desirable for understanding the complexities of the spin physics involved, but is much too complex and time consuming to be practical for most data analysis. The

Pake convolution method provides a nice intermediate between these two extremes of analysis schemes, being grounded in spin physics and yet sufficiently simple to allow for fast, flexible data analysis. In the two cases tested so far (PR-S55C hexamer and mixed Gd-ruler $\mathbf{1_1}$ (2.1 nm)/Gd-4-iodo-PyMTA), the Pake convolution analysis seems to pick out the expected mean distances, though the distance distributions were broader than expected, particularly for the Gd-ruler. It is also evident from the imperfect reproduction of the broadened lineshape in Figure 4.7, particularly on the wings of the high-field side, that this analysis does not properly capture all the details of the measured lineshape. One main source of error in the Pake convolution analysis is the noise in the measured CW EPR spectra and the quality of the baseline. The latter could be particularly challenging for analyzing Gd(III) spectra since the baseline is not truly zero, but rather the broad component of the full EPR spectrum resulting from transitions other than the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition. Due to the broadness of this component of the EPR spectrum the CW EPR signal is approximately zero (the CW EPR measurement can be thought of as the derivative of the EPR absorption lineshape), yet this could be contributing uncertainty in the analysis. Furthermore, while the Pake pattern implicitly includes contributions from the secular and pseudo-secular terms of the dipolar interaction, it does not take into account contributions from the ZFS. Because 240 GHz CW EPR is sensitive to distances below 4 nm, we know from the literature that contributions from ZFS play in important role in the EPR spectrum [45, 128, 101].

4.6 Towards monitoring conformational change of the E-F loop with CW EPR

One of the primary research thrusts motivating the work of this dissertation is to study protein structure and dynamics in as close to a native environment as possible. So far, we have seen preliminary evidence that 240 GHz CW EPR with Gd(III)-based spin labels can give structural information on the organization of PR oligomers in detergent micelles, based on measurements at 30 K that show dipolar broadening of the CW EPR lineshape. Next, we present the first preliminary evidence of a measurement of conformational change in PR using 240 GHz CW EPR with Gd(III)-based spin labels.

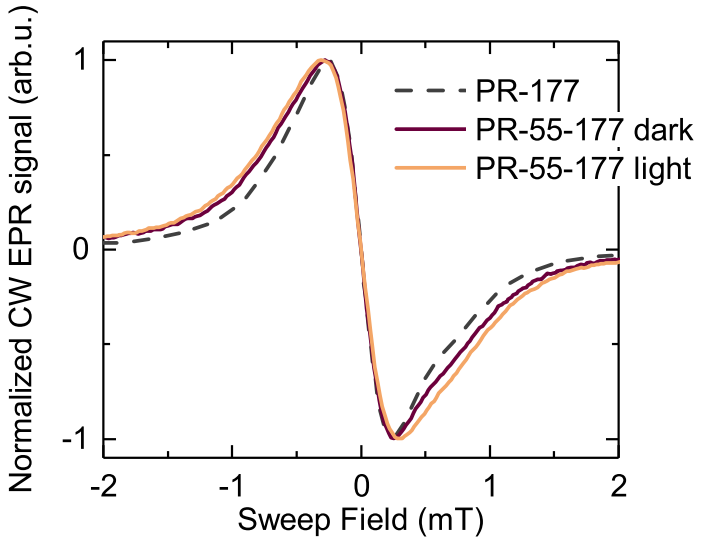

Figure 4.8: CW EPR at 240 GHz and 30 K of PR-S55C-T177C E50Q slow labeled with Gd-MTS-ADO3A measured in the dark and under illumination with a 535 nm LED light source. The lineshape of PR-T177C slow (mix oligomer states) labeled 1:1 with Gd-MTS-ADO3A is overlaid for comparison.

Based on the measurements of spin labeling efficiency discussed in Sections 4.3 - 4.4,

we chose PR-S55C-T177C as a target sample for observing motion of the E-F loop upon light activation. These two residues showed comparatively high spin labeling efficiencies compared to other residues that we have so far attempted to spin label with Gd(III). Further, residue 177 on the E-F loop has been previously shown to exhibit conformational changes upon light activation [78] and is separated by ∼2 nm from residue 55 on the A-B loop (Figure 4.4), which is within the distance range where 240 GHz CW EPR is most sensitive to distance changes. The sample was additionally prepared with the E50Q to enhance monomeric population and was spin labeled with Gd-MTS-ADO3A, resulting in approximately 50% labeling efficiency. CW EPR measurements at 240 GHz were carried out at a temperature of 30 K in the usual manner, giving the lineshape for PR-55-177 with the protein in the dark state (Figure 4.8). When compared to the lineshape of PR-T177C slow (mix oligomer states) labeled 1:1 with Gd-MTS-ADO3A, which approximates the intrinsic linewidth of PR labeled at residue 177 with Gd-MTS-ADO3A, the doubly labeled PR-S55C-T177C sample shows clear dipolar broadening. Again, most of the broadening is seen in the wings of the CW EPR spectrum while the peak-to-peak linewidth shows little change as a result of poor spin labeling efficiency. The sample was then warmed up and allowed to equilibrate at room temperature under illumination of a ∼535 nm LED (Appendix A.2.3) before being cooled back down to 30 K. Upon freezing, there should be a distribution of PR conformations present in the sample corresponding to the various conformations explored by PR during the photocycle. A 240 GHz CW EPR measurement at 30 K of this "light state" of PR is shown in Figure 4.8. Compared to the

dark measurement, this preliminary light measurement shows a slight broadening of the CW EPR lineshape, which suggests a decreased distance between residues 55 and 177 when PR is activated by illumination with green light. While very preliminary at this stage, this measurement suggests we have, for the first time, observed a conformational change in a membrane protein using 240 GHz CW EPR with a Gd(III) spin label.

4.7 Impact and outlook

In this chapter, we explored the application of distance measurement using 240 GHz CW EPR and Gd(III)-based spin labels in studying the structure and conformational changes in the membrane protein proteorhodopsin (PR). These were some of the first experiments of this nature, and so focused primarily on first reproducing known results using the new experimental techniques developed in this dissertation. First, we explored the use of a different Gd(III) spin label, maleimide-Gd-DOTA, which according to the work presented in Chapter 3 has a small ZFS and therefore narrow EPR linewidth, making it a much better candidate for distance measurement by 240 GHz CW EPR than the Gd-4MMDPA spin label that was used in a previous PR study. We began with pulsed EPR measurements at W-band on PR oligomers, and showed that we can indeed replicate previous results of the hexameric organization of PR in detergent micelles [156, 52] using a different Gd(III) spin label. The same hexameric organization of PR oligomers was then reproduced by measurements of inter-PR distances using 240 GHz CW EPR and maleimide-Gd-DOTA or Gd-MTS-ADO3A as the spin label. Additionally, we showed

preliminary evidence that the CW EPR measurement can be used to measure multiple distances in the PR oligomer simultaneously. Finally, we presented first progress towards measuring protein dynamics using high-field CW EPR, by demonstrating a change in distance between a residue on the E-F loop and the A-B loop of PR upon light activation.

High-field CW EPR with Gd(III) spin labels is a very young experimental technique which is only now beginning to show its usefulness and limitations in the study of protein structure and dynamics. For PR, there are many excited directions which could be explored in the future. In this chapter and in previous studies [156, 52], the oligomeric organization of PR was only studied in DDM detergent micelles. However, we know that there are complex protein-protein and protein-lipid interactions which occur in the oligomer and may alter or tune the function of PR. Therefore, it would be interesting to see how the oligomeric organization changes in different detergent and lipid membrane environments. For studies of conformational changes in PR, the ultimate goal of this project is to film the protein in action. Measurement of intra-PR distance changes in the active vs. inactive state of the protein, such as the preliminary data presented here, provide merely a snapshot of the distribution of various PR conformations by virtue of the protein being frozen during the measurement. There is the possibility of doing this measurement in a step-wise time-resolved fashion, where the protein is flash frozen at various increments of time after light activation and the distance changes associated with conformational changes are measured. A perhaps more attractive option would be to conduct similar measurements above the protein dynamical transition, or even at

room temperature, such that the motions of the protein during the photocycle could be observed in real time.

These are ambitious goals, and many experimental hurdles will need to be overcome to accomplish them. In particular, the primary current bottleneck in further expanding the use of 240 GHz CW EPR to study PR and other proteins is difficulties in achieving reliable, high-efficiency spin labeling of the protein with Gd(III)-based spin labels. The development of a method to estimate spin labeling efficiency has helped immensely in troubleshooting problems in spin labeling and identifying samples in which spin labeling was most successful. However, if Gd(III) spin labeling cannot be sufficiently improved, this could severely limit the application of this technique. Fortunately, it seems that for a minimum of $\sim$50% labeling efficiency, further advances in data analysis may be able to partially compensate. Even when spin labeling efficiency is rather poor, we often still observe a broadening of the CW EPR spectrum in a doubly labeled sample, though the broadening is restricted to the outer wings of the spectrum which are less affected by monomeric contributions to the signal. This means that for sufficiently short interspin distances, we may still be able to resolve distance changes, even if the mean distance and distance distribution cannot be quantitatively measured.

Chapter 5

Conclusion

The work presented in this dissertation endeavors to expand the capabilities of CW EPR for structural studies of biological systems, particularly for membrane proteins which often resist other traditional biophysical characterization techniques. Our approach was to use high-frequency (240 GHz) CW EPR combined with spin labels based on Gd(III) complexes to increase the sensitivity, and hence the range, of distance measurements. Studies of a model system, composed of two Gd(III) complexes connected by a rigid molecular linker, showed that we can measure interspin distances ranging from at least 1.2 - 3.4 nm at a measurement temperature of 30 K when Gd-PyMTA is used as the spin label. It was found that the peak-to-peak broadening of the CW EPR signal, relative to the intrinsic linewidth of the Gd(III) complex, follows a $1/r^3$ dependence. Furthermore, this $1/r^3$ dependence persisted at temperatures of 215 K and 288 K, showing that this technique is applicable at physiologically relevant temperatures, and may even be used

at room temperature given that the sample can be somehow immobilized. We developed a very simple model to simulate the lineshape of this model system, where the dipolar broadening of the central transition of Gd(III) is modeled as an $S = 1/2$ spin whose CW EPR lineshape is broadened through electron-electron dipolar interactions with a neighboring $S = 7/2$ spin. This very simple description, which ignores ZFS and that both Gd(III) spins are $S = 7/2$, nevertheless captures many details seen in the experimental CW EPR spectra, including the overall width and the smaller peaks and asymmetries observed at short distances. Given that this model has no free parameters, apart from matching the intrinsic linewidth of the monomeric Gd(III) complex, this suggests that the majority of features observed in the broadened CW EPR lineshape can be directly explained by dipolar interactions.

As a next step, we sought to both further improve on these simple simulations and to also to potentially identify other Gd(III) complexes which are better suited for high-field CW EPR distance measurement by virtue of having a narrow intrinsic linewidth. Both of these goals rely on an accurate knowledge of the zero-field splitting (ZFS) parameters of the Gd(III) complex. We conducted an extensive study of six different Gd(III) complexes of interest for use in EPR studies. EPR spectra of these complexes were recorded at frequencies of Q-band, W-band, and 240 GHz in order to provide the best-possible determination of the ZFS parameters. These multi-frequency spectra were fit using the most common models found in literature for the distribution of the second-order ZFS parameters D and E. We found that, despite their apparently different descriptions,

that the models converged to the same finding. Namely, that the distribution of D is a bimodal Gaussian distribution, where the components centered at $-D$ and $+D$ may occur with different probabilities. The distribution of E/D was found to be broadly distributed between 0 and 1/3, with a minima at $E/D = 0$ and a maxima at approximately $E/D = 0.25$. Thus satisfied that we understand the relative performance of the various models presented in the literature, and with multi-frequency EPR spectra at hand, we were able to with reasonable confidence assign values of the ZFS parameters to each of the six studied Gd(III) complexes. Of these, we found that Gd-PyMTA had a mid-range ZFS magnitude while Gd-DOTA had a smaller ZFS magnitude, suggesting that the latter would lend better sensitivity to CW EPR distance measurements using this spin label. Finally, we attempted to relate the determined ZFS parameter values to the structure of the Gd(III) complexes. We found that a superposition model can, at the very least, provide an estimate of the magnitude of D for a given complex and hopefully in the future will help inform the design of new, tailored Gd(III) spin labels for both CW and pulsed EPR.

Finally, as a first real-world application of 240 GHz CW EPR distance measurement with Gd(III) spin labels, we began structural studies of the membrane protein proteorhodopsin (PR). In light of our ZFS study of Gd(III) complexes, we transitioned to using spin labels based on the Gd-DOTA complex. With this new spin label we were able to replicate previous findings on the hexameric organization of PR oligomers using pulsed EPR at W-band, and then also with 240 GHz CW EPR. The first CW EPR mea-

surement of PR oligomers determined found an inter-PR distance of $\sim$ 3 nm at residue 89, which is well beyond the distance measurable by traditional CW EPR methods. The next CW EPR study on PR oligomers probed a much shorter inter-PR distance of $\sim$ 1.6 nm, such that our CW EPR could successfully measure both the nearest-neighbor and next-nearest-neighbor distance in the hexamer simultaneously. And last, we conducted preliminary measurements of a conformational change in PR, by seeing a distance change between a residue on the mobile E-F loop and a residue on the stationary A-B loop when the protein is in the inactive state vs. a light-induced active state. We saw tantalizing evidence that there is indeed a distance change and hence motion between these two regions of the protein when activated by light. These measurements only begin to scratch the surface of what this technique can contribute to the understanding of PR, as well as to other systems. There are still many details to work out in improving data analysis and spin labeling efficiency, but these look to be surmountable hurdles.

Given these very promising glimpses, the hope is that one day high-field CW EPR with Gd(III) spin labels will be able to give both static structural information and real-time dynamic information on conformational changes in a protein that allow it to do its job. Additionally, it is the hope that this technique will also move into the realm of structural studies of proteins in more native-like systems - including ambient temperatures and in lipid membrane or in-cell environments. Both CW EPR and Gd(III) spin labels have independently proven successful in these various realms - all that remains is to bring them together. While high-field EPR has long been a somewhat esoteric sub-field of

magnetic resonance, the march towards higher fields has done nothing but accelerate in recent years as the technique matures, costs decline, and broader audiences become convinced of this as a useful tool to solve a wide variety of problems. The understanding of biological processes is a particularly interesting application of high-field EPR and, given the never-ending variety that nature provides, will no doubt be puzzling us for generations to come.

Appendix A

Experiments and Methods

A.1 EPR spectrometers

The majority of experiments in this dissertation were performed on a home-built 240 GHz EPR spectrometer in the Sherwin group at UCSB. The Q-/W-band measurements presented in Chapter 3 were performed on spectrometers in the Jeschke group at ETH Zurich, and the X-band measurements in Appendix C were performed on an X-band spectrometer in the Han lab at UCSB. The DEER measurements carried out in Chapter 4 utilized a home-built W-band spectrometer in the Goldfarb group at the Weizmann Institute of Science. These instruments are briefly outlined below.

A.1.1 240 GHz spectrometer - Sherwin group, UCSB

The 240 GHz EPR spectrometer operates in both CW and pulsed modes using a low power solid state source. The spectrometer can alternately operate as a high-power

pulsed Free Electron Laser-powered EPR spectrometer, which is used for experiments beyond the scope of this dissertation. This spectrometer has been described in detail elsewhere [159, 53, 54, 41], and is similar in operation to other spectrometers [167]. In low-power operation, the spectrometer uses a solid state frequency-multiplied source (Virginia Diodes, Inc.), which multiplies a 15 GHz source 16x to achieve an output frequency of 240 GHz with CW power of order 50 mW.

The spectrometer operates in induction mode detection with optical transport of 240 GHz radiation achieved via a quasi-optical bridge. Superheterodyne detection is used, with a Schottky subharmonic mixer (Virginia Diodes, Inc.) and a home-built intermediate-frequency (IF) stage operating at 10 GHz, which is then mixed down to baseband for detection. For CW and rapid passage measurements, field modulation at 20 kHz was used and the 0° and 90° signal components from the IF stage were passed through a pair of pre-amplifiers followed by a pair of lock-in amplifiers.

The radiation from the 240 GHz source may be attenuated via a user-controlled voltage attenuation of the source, or by a pair of crossed wire-grid polarizers. This radiation is then passed to the sample which sits at the end of a 1.25 m long overmoded waveguide (Thomas Keating, Inc.). The probe design allows for a wide variety of sample geometries (Figure A.1). Typically for measurements of frozen glassy samples, or for powder samples, a Teflon bucket-style sample holder of approximately 10 μL capacity is used. For room temperature measurements of liquid samples, a rectangular glass or quartz capillary containing approximately 0.5 μL of sample is used. For the CW and rapid

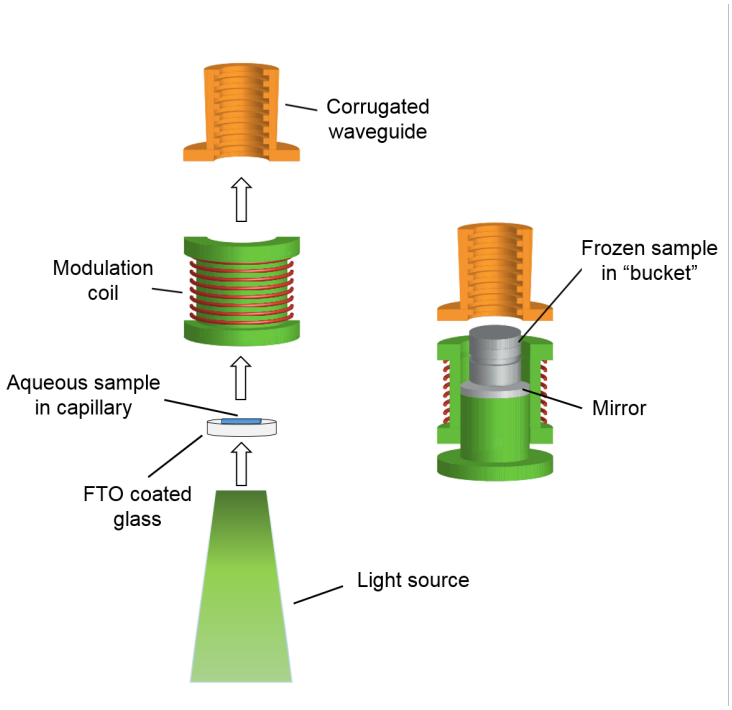

Figure A.1: Schematic overview of various sample holder configurations, adapted from reference [38]. Liquid samples which will be frozen for measurement are loaded into a cylindrical Teflon "bucket" and placed on a mirror inside the modulation coil at the end of the waveguide. Aqueous samples are measured in a thin, rectangular glass capillary. For light-activated measurements, the sample can be backed by glass coated with a transparent conductor (e.g. FTO or ITO) and illuminated from below through a window at the bottom of the cryostat.

passage measurements presented in this dissertation, the sample was placed inside one of several modulation coils (depending on the required sample geometry and modulation amplitude) mounted at the end of the waveguide. The calibration of these modulation coils is discussed below. The waveguide assembly is mounted in a continuous-flow cryostat (Janis Research Company, LLC) which sits in the room-temperature bore of a sweepable 12.5 T superconducting magnet (Oxford Instruments plc). The main coil of the magnet can be swept through a range of field positions from 0 - 12.5 T. Alternatively, this main coil may be clamped at a fixed field position and a smaller sweep coil used, which covers

a range of ± 0.06 T about the main field position. Both the main coil and the sweep coil of the magnet display hysteresis effects. If the true field position needs to be known, e.g. for measurements of the g-value of a sample, a field-standard based on the hyperfine splitting of Mn^{2+} has been used at room temperature [14]. Measurements of a Mn^{2+} field standard showed that the as-recorded sweep coil position requires an additional linear correction, as described below.

The spectrometer can be operated at sample temperatures ranging from 1.5 - 300 K. The sample temperature is recorded with a Cernox temperature sensor (Lakeshore Cryogenics, Inc.), mounted on the outside of the waveguide just above the sample position. A second temperature sensor, attached near the heater and helium intake at the top of the cryostat, was used to control the sample temperature with a PID feedback optimized temperature controller (Model 335 Lakeshore Cryogenics, Inc.). For operating temperatures of 4.2 - 300 K, the cryostat can be run under atmospheric pressure or under vacuum. For temperatures of ∼2.3 - 4.2 K, the pressure of the sample space is reduced using a rotary vain pump. For operating temperatures below ∼2.3 K the cyrostat is operated in batch mode (∼1 hour at 1.5 K), where liquid helium is collected in the cryostat under low pressure.

Further details of the operation and maintenance of the spectrometer can be found in the spectrometer user's manual available from the Sherwin group at UCSB.

Sweep coil correction

The recorded position of the sweep coil of the 12.5 T magnet was found to require a linear correction in data processing, as determined by measurements with a Mn^{2+} field standard. The appropriate correction factor is the recorded sweep coil field vector multiplied by 0.953654765362. The main coil of the magnet does not require this correction, however the main coil does display hysteretic effects. If an accurate field position is required, a field standard such as Mn^{2+} must be measured along with the sample of interest [14].

Digital attenuator correction

After detection of the EPR signal by the Schottkey subharmonic mixer, the 10 GHz signal is passed through an amplifier and a digitally controlled attenuator before being fed through the IF stage. If accurate EPR signal amplitudes are required (e.g. for the spin label concentration estimation discussed in Chapter 4), the effect of attenuation of the measured signal amplitude by the digital attenuator must be corrected for in the data processing. The response of the EPR signal amplitude as a function of the digital attenuator setting is shown in Figure A.2.

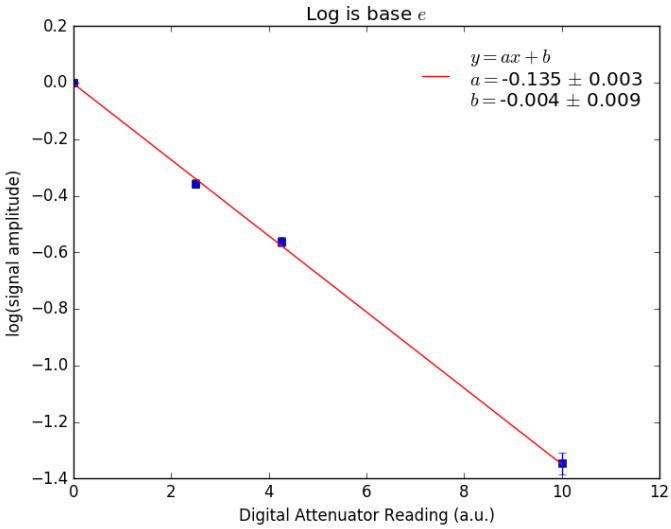

Figure A.2: Response of measured EPR signal amplitude on the setting of the digital attenuator.

Modulation coil calibration

CW EPR spectra are acquired by applying a small sinusoidal modulation to the B_1 field. In general, a higher modulation amplitude will increase SNR. However, too large a modulation amplitude will result in broadening of the peak-to-peak linewidth, and will also affect the peak height. This is potentially very problematic in the CW EPR distance measurements, which rely heavily on the accurate determination of peak-to-peak linewidths. For the accurate lineshapes, the modulation amplitude should be kept to $< 1/10$ of the intrinsic linewidth. If the area under the peak is the information required, as is the case in a concentration measurement, then overmodulation is acceptable as the area is linearly proportional to the modulation amplitude [50].

The most accurate way to determine the modulation amplitude at the sample position is by a CW EPR experiment on a sample with a sufficiently narrow linewidth to show

effects of overmodulation. An example of such a measurement on a Gd(III) sample is shown in Figure A.3, where the current driving the modulation coil was successively increased.

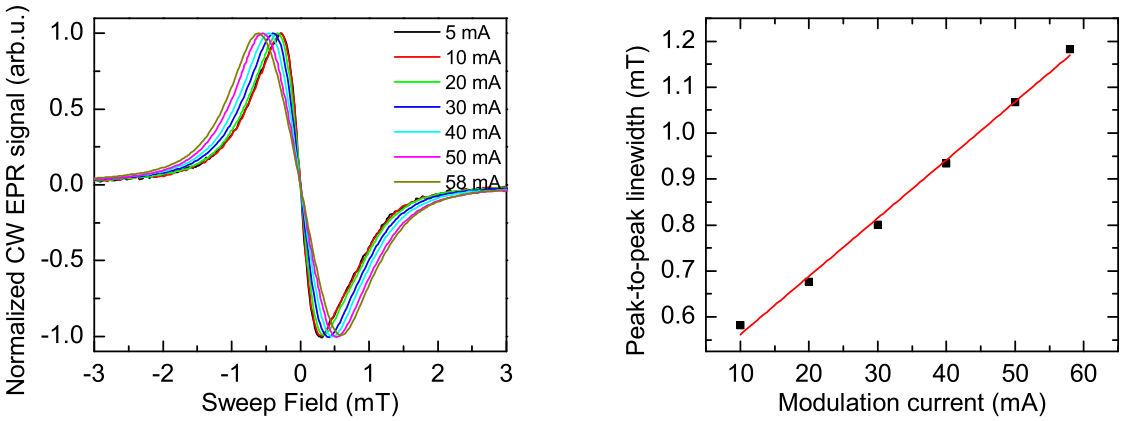

Figure A.3: Overmodulation of the central transition of Gd(III) using the light accessible modulation coil (left). The peak to peak linewidth is linearly proportional to the current driving the modulation coil (right). This data was acquired at 30 K on $\sim 8\mu$L of slow PR E50Q doubly labeled with Gd-MTS-ADO3A at sites 177 and 55. The sample was loaded into a Teflon sample cup which was placed on a sapphire substrate inside the modulation coil.

The modulation amplitude can be calibrated by this overmodulation broadening with the following expression [115, 46]

$$\Delta H_{pp(obs)} = \Delta H_{pp} \left\{ \left(\frac{H_m}{\Delta H_{pp}} \right)^2 + 5 - 2 \left[4 + \left(\frac{H_m}{\Delta H_{pp}} \right)^2 \right]^{1/2} \right\}^{1/2} \tag{A.1}$$

where ΔH_{pp} is the intrinsic peak-to-peak linewidth without modulation broadening, $\Delta H_{pp(obs)}$ is the measured peak-to-peak linewidth of the modulation broadened spectrum, and H_m is the modulation amplitude in Gauss.

Inserting the observed peak-to-peak linewidths in Figure A.3 into Equation A.1 gives a

value of 0.235 Gauss/mA for the wide-bore light accessible modulation coil. The narrow-bore modulation coil used for experiments not requiring light activation was calibrated in a similar fashion using a sample of Mn:MgO, and gave a value of 0.92 Gauss/mA.

A.1.2 W-band spectrometer - Jeschke group, ETH Zurich

The W-Band spectra presented in Chapter 3 were measured on a Bruker Elexsys E680 X-/W-band spectrometer using a EN 680-1021H resonator. The measurement temperature was stabilized by a He-flow cryostat (ER 4118 CF, Oxford Instruments).

A.1.3 Q-band spectrometer - Jeschke group, ETH Zurich

The Q-band measurements presented in Chapter 3 were performed on a home-built high-power Q-band pulse EPR spectrometer [70] equipped with a rectangular cavity accommodating for oversized 3 mm outer diameter cylindrical samples [164, 114].

A.1.4 W-band spectrometer - Goldfarb group, WIS

The W-band DEER measurements presented in Chapter 4 were performed on a home-built W-band spectrometer described previously [108, 66].

A.1.5 X-band spectrometer - Han group, UCSB

The X-band measurements presented in Appendix C were recorded on an Elexsys E580 (Brucker Biospin) pulsed EPR spectrometer.

A.2 Sample preparation and experimental details

In the following, further details of the samples used in this dissertation are described, along with further details of experimental parameters. This information is organized by chapter.

A.2.1 Chapter 2

The synthesis of the Gd-rulers $\mathbf{1_1}$ and $\mathbf{1_3}$ were previously published [119]. The synthesis of Gd-rulers $\mathbf{2_1}$ and $\mathbf{2_2}$ were reported in reference [41]. The synthesis of Gd-rulers $\mathbf{1_0}$ and **3** will be reported elsewhere.

For sample preparation, stock solutions of Gd-rulers and Gd-4-iodo-PyMTA in D_2O were used. These solutions contained additional compounds remaining as a result of the synthesis, as detailed in Table A.1. For measurements at cryogenic temperatures, stock solutions of the Gd-rulers and Gd-4-iodo-PyMTA were diluted with a 60:40 (v:v) mixture of D_2O and glycerol$-d_8$ (Cambridge Isotopes Laboratories, Inc.) to a concentration of 300 μM.

For measurements near room temperature, the Gd-rulers were immobilized in dehydrated amorphous trehalose [109]. For this purpose, stock solutions of the Gd-rulers and Gd-4-iodo-PyMTA were diluted with a 0.2 M solution of trehalose dihydrate (Sigma Aldrich) in D_2O (Cambridge Isotopes Laboratories, Inc.). The resulting samples had a mole ratio of 40:1 trehalose:Gd-ruler and 40:1 trehalose:Gd-4-iodo-PyMTA. These mixtures were then deposited onto a glass slide, allowed to dry for several days under a flow

Compound	**Content of stock solution in D_2O**	**pH**
Gd-4-iodo-PyMTA	3 mM Gd-4-iodo-PyMTA, 5 mM F_3CCO_2H/F_3CCO_2Na, 15 mM NaCl, 120 mM H_2O	7
$\mathbf{1_0}$	4 mM $\mathbf{1_0}$, 48 mM NaCl, 148 mM H_2O	5
$\mathbf{1_1}$	5 mM $\mathbf{1_1}$, 37 mM NaCl, 100 mM H_2O	8
$\mathbf{2_1}$	5 mM $\mathbf{2_1}$, 0.5 mM F_3CCO_2H/F_3CCO_2Na, 30 mM NaCl, 180 mM H_2O	7
$\mathbf{1_3}$	5 mM $\mathbf{1_3}$, 28 mM NaCl, 100 mM H_2O	8
$\mathbf{2_2}$	5 mM $\mathbf{2_2}$, 30 mM NaCl, 100 mM H_2O	8
3	5 mM **3**, 32 mM NaCl, 122 mM H_2O	

Table A.1: Details of the content of stock solutions of Gd-4-iodo-PyMTA and Gd-rulers $\mathbf{1_n}$, $\mathbf{2_n}$, and **3** used to prepare samples for CW-EPR experiments, including pH and the concentrations of additional compounds left over as a result of the synthesis.

of dry nitrogen at room temperature, and finally placed under vacuum for at least 24 hours before measurement. The fragile solid was removed from the glass slide, crumbled into a powder, and transferred to a Teflon sample cup for measurements.

The phase of the CW EPR spectra was set in post-processing. Each lock-in amplifier produces two components which are at angles ϕ and $\phi + 90°$ with respect to a reference signal taken from the field modulation frequency. The angle ϕ for each lock-in amplifier was chosen such that the quadrature signal was minimized using a least-squares criterion, giving maximum SNR in the in-phase signal [13]. The resulting in-phase signals after this phase nulling procedure were the real and imaginary components of the CW EPR spectrum. These real and imaginary components were then rephased to determine the derivative CW EPR lineshape either by equalizing the positive and negative peaks of the derivative lineshape, with a linear background correction applied [53].

A.2.2 Chapter 3

A series of six Gd(III) complexes were chosen to be included in the work presented in Chapter 3. Gd-DOTA **(2)** was obtained commercially from Macrocyclics and used without further purification. The synthesis details of the remaining Gd(III) complexes, which include Gd-NO_3Pic **(1)**, Gd-maleimide-DOTA **(3)**, R-(Gd-PyMTA) **(4ab)**, Gd-TAHA **(5)**, iodo-(Gd-PCTA-[12]) **(6)**, and Gd-PyDTTA **(7)**, will be reported in the forthcoming publication of this work [40]. The synthesis of Gd-NO_3Pic has been previously published [62].

For 240 GHz measurements, stock solutions of the Gd(III) complexes were diluted to a final concentration of 300 μM in 0.4:0.6 (v:v) D_2O/glycerol$-d_8$. Sample solutions of 10 μL volume were loaded into a Teflon sample cup of $\sim$ 3.5 mm i.d. and $\sim$ 5 mm height and subsequently flash frozen in liquid nitrogen under ambient conditions. 240 GHz measurements were performed at a temperature of approximately 5 K. EPR spectra were acquired using a rapid passage technique, which is similar in practice to CW EPR but records an absorption-like lineshape rather than a derivative lineshape [22, 171, 116]. Rapid passage EPR measurements were carried out with field modulation at 20 kHz with $\sim$ 0.3 mT modulation amplitude. The main coil of the superconducting magnet was used to carry out measurements at a sweep rate of 0.1 T/min. The rapid passage regime was achieved by simply increasing the power applied to the sample by the 240 GHz microwave source.

The phase of the 240 GHz EPR spectra was set in post-processing. Each lock-in

amplifier produces two components which are at angles ϕ and $\phi + 90^{\circ}$ with respect to a reference signal taken from the field modulation frequency. The angle ϕ for each lock-in amplifier was chosen such that the quadrature signal was minimized using a least-squares criterion, giving maximum SNR in the in-phase signal [13]. The resulting in-phase signals after this phase nulling procedure were the real and imaginary components of the EPR spectrum. These real and imaginary components were then rephased to determine the absorption rapid passage EPR lineshape by maximizing the integral of the real component and minimizing the imaginary component, which in the fast passage regime should be approximately zero. 240 GHz spectra were normalized to the envelope resulting from the outer EPR transitions, as the relatively high powers and fast sweep rate necessary to collect data in the rapid passage regime artificially broaden the very narrow central transition of Gd(III).

For Q-/W-band measurements, stock solutions of the Gd(III) complexes were diluted to a final concentration of 25 μM in 1:1 (v:v) D_2O/glycerol$-d_8$. Sample solutions were filled into 3 mm o.d. quartz capillaries for Q-band or 0.5 mm i.d./0.9 mm o.d. quartz capillaries for W-band and subsequently flash frozen in liquid nitrogen under ambient conditions. Q-/W-band measurements were performed at a temperature of 10 K. Field-swept EDEPR spectra were acquired using the Hahn-echo pulse sequence $t_p - \tau - 2t_p - \tau$ and pulse length t_p of 12 ns. The inter-pulse delay τ was set to 400 ns. The power to obtain 12 ns - 24 ns, $\pi/2 - \pi$ pulses was set at the central transition of the Gd(III) spectrum by nutation experiments. The Q-/W-band spectra had a constant field offset

removed and were normalized to a maximum value of 1.0.

For X-band measurements, stock solutions of the Gd(III) complexes were diluted to a final concentration of 300 μM in 0.4:0.6 (v:v) D_2O/glycerol$-d_8$. Sample solutions of 4 μL volume were filled into quartz capillaries and subsequently flash frozen in liquid nitrogen under ambient conditions. X-band measurements were performed at a sample temperature of 30 K. Field-swept EDEPR spectra were acquired using the standard Hahn-echo pulse sequence $t_p - \tau - 2t_p - \tau$ with a pulse length of t_p of 16 ns. The inter-pulse delay τ was set to 200 ns. A 2-step phase cycle was used.

A.2.3 Chapter 4

Production and spin labeling of PR samples

Site-directed mutagenesis and spin labeling (SDSL) methods were used to introduce cysteine mutations at select sites in PR which form covalent bonds with applied spin labels. The template green absorbing PR gene used for protein expression (referred to as "wild-type" in the main text) is a triple mutant version with substitution of the three naturally occurring cysteine residues (107, 156, 175) with serines to prevent non-specific binding of spin labels. Further mutagenesis was conducted on select samples to create the mutants (i) E108Q which extends the M-intermediate state of the photocycle by eliminating a proton acceptor site (referred to as "slow" PR), and/or (ii) E50Q which breaks down interactions the stabilize oligomer formation to result in an increased population of monomeric PR.

Desired mutations of specific residues of PR were generated using standard site-directed mutagenesis methods. The PR gene template was modified by applying primers with mutations codons and polymerase chain reactions (PCR). A two-stage PCR technique was used to prevent the primers from forming self-dimers instead of binding to the PR template during PCR. The PR gene template with the desired mutations incorporated was then recloned into a pET26b (+) vector (Novagen) for expression of the protein. Protein expression was carried out by transforming the pET26b (+) vector with the PR gene into *E. Coli* strain BL21(DE3) competent cells (Agilent), growing bacteria in liter-scale flasks, and inducing overexpression by adding isopropyl-β-D-thiogalactopyranoside (IPTG) al all-trans-retinal (Vitamin A aldehyde). The expressed PR in *E. coli* cells, which were pink-colored in appearance, were then collected by centrifugation. This was followed by purification of the PR, which involves cell lysis, freeze-fracture, and sonication to break the *E. coli* cell membrane such that the PR in cell lysate can later be extracted and solubilized by β-D-dodecyl-maltoside (DDM) surfactants. The surfactant-solubilized PR was then further purified by a metal affinity resin which selectively binds to PR with a hexahistidine tag at its C-terminus. The resin is washed several times with a buffer of a mild concentration of imidazole to remove other impurities. The surfactant-solubilized PR bound to resin is then eluted by a buffer with a high imidazole concentration. A Sephadex desalting column was then used to remove excess salts and exchange PR into a desired buffer (50 mM MES, 150 mM sodium chloride, 0.05 % DDM, pH 5.8) for storage or measurement.

For some samples, hexameric species of solubilized PR were separated by injecting samples onto a HiLoad 16/6000 SuperdexTM 200pg column (GE Healthcare) connected to a BioRad Duoflow FPLC, and 0.5 mL fractions were collected with a BioFrac fraction collector (BioRad; Hercules, CA, USA). Fractions were analyzed for oligomeric state by BN-PAGE.

For the CW EPR and DEER measurements presented in this dissertation, maleimide-Gd-DOTA spin labels were attached to selected cysteine mutation sites on PR. Gadolinium chloride ($GdCl_3$) and maleimido-monoamide DOTA spin tags were first dissolved in MilliO water with a 1:1 molar ratio. To drive the chelating reaction between Gd^{3+} ions and the DOTA spin tags, the solution was continuously adjusted to a pH of 5.8 by titrations of basic buffer at room temperature for $\sim$ 2.5 hours. The pH of the solution becomes stable after the formation of the Gd-DOTA complex reaches equilibrium. A 10x molar mass excess of the maleimide-Gd-DOTA complex solution was then added to PR, and the mixture shaken overnight to allow for the conjugation reaction between cysteine residues on PR and the maleimide-Gd-DOTA complex. The spin labeled PR was later exchanged into the same MES buffer, but made with deuterium oxide (D_2O), using the Sephadex desalting column. The excess unreacted maleimide-Gd-DOTA was also be removed during this step. For cyrogenic EPR measurements, 30% glycerol$-d_8$ by volume was added to promote glass formation.

W-band DEER experiments and samples

For W-band DEER measurements, six PR samples in DDM were prepared with 30% (by volume) glycerol$-d_8$ as a glassing agent. These samples are detailed in Table A.2. The amount of free Gd(III) ion in solution during the labeling step was monitored by the xylenol orange test [17] to ensure no free Gd(III) ions were present in the final samples.

Sample	spin dilution (labeled:WT-PR)	PR conc (μM)
PR-55 hex	1:1	90
PR-55 hex	1:3	365
PR-58 hex	1:1	52.5
PR-58 hex	1:3	171
PR-177 hex	1:1	124.6
PR-177 hex	1:3	348

Table A.2: PR samples prepared in DDM detergent micelles for W-band DEER measurements. All samples contained 30% glycerol$-d_8$ (by volume) as a glassing agent. Note: the DEER data suggests that the PR-58 samples were mislabeled.

Additionally, three PR samples were prepared in a similar fashion and reconstituted into POPC:POPG (4:1) liposomes. These are detailed in Table A.3.

Sample	spin dilution (labeled:WT-PR)	PR conc (μM)
PR-58 hex	1:3	324
PR-177 hex	1:3	248
PR-177 hex	1:4	123

Table A.3: PR samples reconstituted into POPC:POPG (4:1) liposomes for W-band DEER measurements. All samples contained 30% glycerol$-d_8$ (by volume) as a glassing agent. Listed are PR concentrations measured before spin dilution and before addition of glycerol$-d_8$. No DEER oscillations were observed for these samples.

Samples were loaded into a quartz capillary of 0.6 mm I.D./0.84 mm O.D. and flash frozen in liquid nitrogen under ambient conditions. EDEPR experiments were carried out using a standard Hahn-echo pulse sequence $\pi/2 - \tau - \pi - \tau$ with $\pi = 15$ ns and τ

= 550 ns. For DEER experiments, the observation frequency was set to the maximum of the Gd(III) EPR lineshape at 94.85 GHz and the observation lineshape was set to the low-field side of the central peak at 94.95 GHz. The pump pulse was of 15 ns duration and the probe pulse used a Hahn-echo pulse sequence with $\pi = 15$ ns and $\tau = 550$ ns.

Improved sample geometry for high SNR room temperature EPR

Conducting measurements of protein structure and dynamics under ambient conditions will require high SNR room temperature 240 GHz EPR measurements on liquid samples of small volume and low spin label concentration. While PR is surprisingly amenable to being concentrated to several 100s of μM, many membrane proteins of interest, cannot be concentrated to more than $\sim$100 μM. Coupled with imperfect spin labeling and the need to spin-dilute for certain experiments, this means working with Gd(III) spin label concentrations of maybe 10s of μM. Additionally, the expression and purification of protein samples often has low-yield, so we wish to use as little sample volume as possible (at least $<$ 10 μL, preferably less). The sample geometry described below has enabled the measurement of 240 GHz CW EPR spectra of 50 μM Gd-DOTA samples using 0.5 μL sample volume at 290 K with reasonable SNR in a single measurement. The low-concentration and low-volume sensitivity limits of this sample geometry for use in the 240 GHz EPR spectrometer has not yet been fully explored.

High SNR measurements of liquid samples is difficult, due to significant dielectric losses of water at millimeter-wave frequencies. Typical samples used for cryogenic EPR measurements, consisting of $\sim$10 μL of 100μM - 1 mM sample in a Teflon bucket-style

sample cup backed by a mirror, give very little if any measurable signal at room temperature where the sample is now a lossy liquid instead of ice. For measurements on liquid samples at room temperature, thin samples have been shown to give the best signal [148]. A thin, flat sample is easily achieved in practice by loading the aqueous sample into a rectangular glass or quartz capillary. To achieve the greatest signal possible from the small volume of sample available in such a geometry, the sample should be placed as close as possible to the beam waist. Additionally, the B_1 field should be maximized at the sample position and the electric field minimized. This may be achieved by mounting the rectangular capillary filled with aqueous sample on top of an appropriate dielectric material and backed by a mirror. A variety of dielectric materials and geometries are appropriate for such a sample loading scheme. Performance of a particular mirror-dielectric-sample stack may be predicted by transfer matrix calculations, as outlined by Budil and Earle [35]. For the sample geometries used in this work, the sample was assumed to be water, since accurate values of ϵ at 240 GHz needed for the transfer matrix calculations are difficult to find for many samples and materials. However, for more accurate calculations it may be necessary to consider the solute of the sample, since it has been suggested that solutes such as sucrose and glycerol, which are commonly added to biological samples to modulate viscosity or promote glass formation, significantly alter ϵ at mm-wave frequencies. Macromolecules in a sample may also modify the effective ϵ [35].

Sample handling for labeling efficiency measurements

For the measurements of Gd(III) concentration which were used to calculate spin labeling efficiency of PR samples, 0.5 μL of sample was pipetted into a rectangular (0.10 x 2.00 mm I.D.) glass capillary (Vitrocom 5012-300). For reproducible CW EPR signal amplitudes, it was found to be important that the sample very nearly fill the capillary, with a minimum of air space available for later wax sealing. The rectangular capillary may be easily cut to length by scoring a line with a diamond scribe and carefully breaking the capillary on the scored line. Glass capillaries were found to be the most readily available commercially in the required dimensions. However, they were found to contain an EPR active contaminant near $g = 2$ which overlaps with the position of the nitroxide EPR signal. This contaminant is sufficiently separated from the Gd(III) EPR signal in room temperature CW EPR measurements to not overlap the desired signal. Subsequently, quartz capillaries were purchased (Vitrocom 50105-100) which did not have a contaminant at the $g = 2$ position and would thus be better suited for measurements with nitroxide spin labels, though the different size of the capillary (0.10 x 1.00 mm I.D.) necessitates a slightly different dielectric backing material and/or geometry.

The rectangular capillaries after being filled with sample were sealed with wax to prevent rapid sample evaporation according to the following procedure:

1. Light a paraffin wax candle and let burn to melt a small amount of wax. The candle should be pure paraffin without any added color or scent to prevent unwanted contaminant EPR signals.

2. Gently heat the end of a glass Pasteur pipette over the candle flame and dip into the melted wax. A small amount of wax should be sucked up into the pipette by capillary action.

3. Gently heat the wax-loaded pipette over the handle flame, holding the pipette horizontally, until a faint wisp of clear (not black) smoke is seen. The wax must be sufficiently melted to flow easily, but not so hot so as to overheat or boil off the very small amount of sample in the capillary.

4. Quickly and gently touch the end of the wax-loaded pipette to the end of the filled capillary to seal with wax. When doing this, the pipette should be held nearly horizontally (rather than vertically oriented) to prevent the wax from flowing out too fast.

5. Repeat the above procedure to seal the other side of the capillary.

6. Visually inspect the capillary to ensure a good seal has formed, and repeat the above if necessary. Be careful to avoid a large amount of wax building up on the outside of the capillary, as this can cause the sample to not lay flat.

The Gd-DOTA concentration calibration measurement and measurements of labeling efficiency of Gd(III) labeled PR samples presented in Chapter 4 were carried out at ∼290 K, the ambient temperature of the cryostat in the absence of heating or cooling. The sweep coil was used for CW EPR measurements with the correction described in Section A.1.1 applied in post-processing of the data. CW EPR spectra were recorded at a sweep

rate of 0.2 mT/sec and using ∼30 Gauss of field modulation. These measurement parameters were optimized for speed and signal amplitude, and not for accurate measurement of the CW EPR lineshape. The experimental parameters used for these measurements likely distort (broaden) the measured lineshape, but so long as all measurements in the calibration series and the samples compared to the calibration are recorded under the same experimental conditions, the distortion of the EPR lineshape does not change the linear relationship of the double integral of the CW EPR signal with Gd(III) concentration in the sample. The filled and sealed glass capillary was mounted on top of a dielectric stack backed by a mirror, in which the index of refraction and thickness of the dielectric material was chosen to result in an enhanced B_1 field at the sample position.

Experimental setup for light activated measurements

For measurements of PR under light activation, a LED (Thorlabs M530L3) was used as the light source, followed by a set of collimating lenses. The output wavelength of the LED is fairly broadband and centered at ∼535 nm (Figure A.4), with a minimum output power of 350 mW. No heat controller was used, so the peak wavelength may shift ± 0.5 nm about this measured value due to heating.

A mirror was used to direct the green light from the LED up through the window at the bottom of the cryostat. The sample was backed by glass coated with FTO, a transparent conductor. In this manner, green light may pass from below through the FTO coated glass to the sample, while the 240 GHz radiation from above is reflected by the FTO which acts approximately as a mirror at this frequency (Figure A.1).

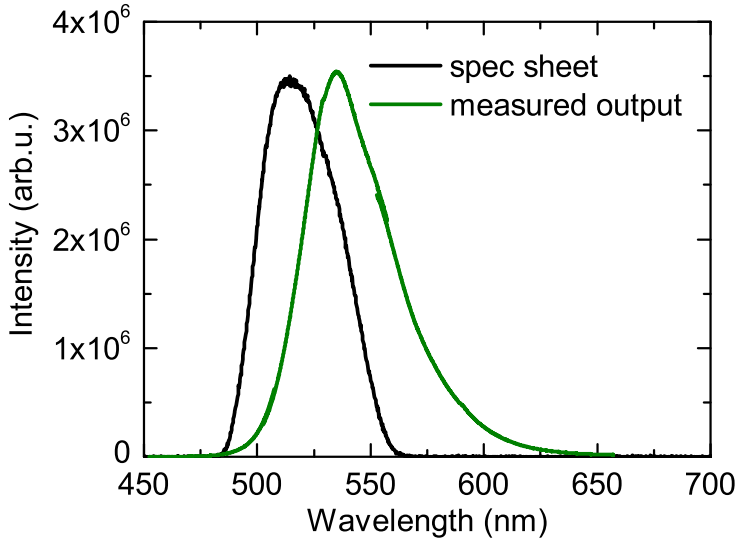

Figure A.4: Output of the LED (Thorlabs M530L3) used for PR light activation experiments.

Appendix B

Second Moment Analysis of Gd-rulers

One of the simplest ways to quantify lineshape broadening of CW EPR data resulting magnetic dipole-dipole interactions of nearby spins is through second moment analysis [168, 2]. The second moment, μ_2, can be used as an empirical parameter to characterize the extent of broadening resulting from spin-spin interaction and related to a pairwise distance between spin labels. In practice, this simplified inter-spin distance determination is typically done by calculating the difference of spectral second moments $\langle \Delta B_D^2 \rangle$ between the absorption spectrum of a doubly spin labeled sample $S_D(B)$ and the corresponding singly labeled reference spectrum $S_S(B)$ which does not have spin-spin interactions,

$$\langle \Delta B_D^2 \rangle = \frac{\int (B - \bar{B}_D)^2 S_D(B) dB}{\int S_D(B) dB} - \frac{\int (B - \bar{B}_S)^2 S_S(B) dB}{S_S(B) dB} \tag{B.1}$$

where B is the magnetic field and $\bar{B}_i$ is the magnetic field corresponding to the mean of each spectra (i.e., the spectral first moment). For the case of a powder sample of like spins, the difference of spectral second moments can be related to the distance between the two spins according to

$$\langle \Delta B_D^2 \rangle = \frac{3}{5} g^2 \beta^2 S(S+1) \sum_j r_{ij}^{-6} \tag{B.2}$$

where g is the gyromagnetic ratio and β is the Bohr magneton [123]. This value is typically quoted for $S = 1/2$ nitroxide spin labels as $r = 2.32 \times 10^8 / \langle \Delta B_D^2 \rangle \; nm$ if $\langle \Delta B_D^2 \rangle$ is given in T^2. For $S = 7/2$ Gd(III) spin labels, assuming $g = 1.992$, this relation becomes $r = 3.85 \times 10^8 / \langle \Delta B_D^2 \rangle \; nm$ if $\langle \Delta B_D^2 \rangle$ is given in T^2.

The results of applying second moment analysis to the series of Gd-rulers is summarized in Figure B.1, which shows the calculated most probable distance for each Gd-ruler at 30 K plotted against the interspin distance determined by second moment analysis. For Gd-rulers **3** (1.2 nm), $\mathbf{1_0}$ (1.4 nm), and $\mathbf{1_1}$ (2.1 nm) there is relatively good agreement with the expected interspin distance. However, for Gd-ruler $\mathbf{2_1}$ (3.0 nm) there is a pronounced disagreement with the expected interspin distance and for the longer Gd-rulers the difference in spectral second moments $\langle \Delta B_D^2 \rangle$ could not be reliably determined. This failure of the second moment analysis to capture longer interspin distances is unfortunate, but not surprising. Steinhoff quotes the upper limit using the method of second moments to be between 1.5 and 1.7 nm for $S = 1/2$ nitroxide spin labels [152]. If we apply the scaling argument of Edwards *et al.* [53], we can estimate the upper limit for $S = 7/2$

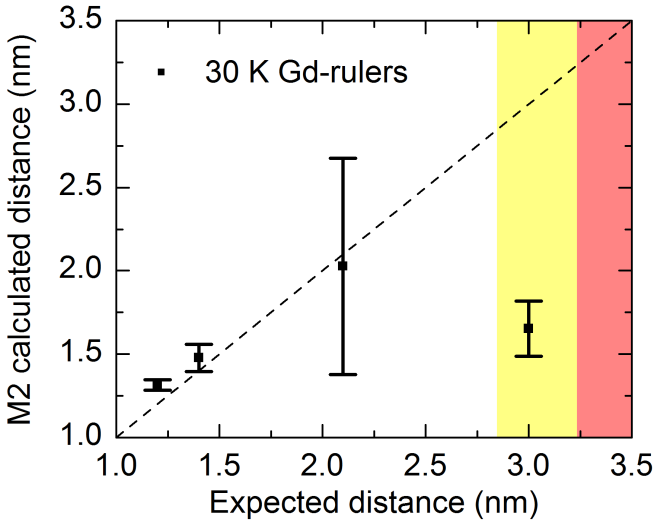

Figure B.1: Second moment analysis of the Gd-rulers. The calculated most probable distance for each Gd-ruler at 30 K plotted against the interspin distance determined by second moment analysis. The second moment determined interspin distance for each Gd-ruler shown here was taken as the average over tens of scans for each point, with error bars taken to be one standard deviation from the mean. The dotted line is a guide to the eye of perfect agreement. The yellow and red bands show approximate confidence regions extrapolated from [152] via a scaling argument.

Gd(III) spin labels to be between 2.8 and 3.2 nm, as shown by the yellow band in Figure B.1. This upper limit, and in general the reliability of this method, is strongly dependent on the quality of the baseline in the experimental data. This is particularly difficult for Gd(III), as the baseline of CW EPR spectra is not truly zero due to contributions from transitions other than the $|-1/2\rangle \rightarrow |1/2\rangle$ transition.

Appendix C

Additional results and details for Chapter 3

This appendix contains additional results and details related to Chapter 3: Determining ZFS parameters of Gd(III) complexes.

C.1 Further details of numerical simulations

Initial fit by eye simulations with Models 1 and 3 were run on a PC. Simulations for Models 2 and 3, including the full library of simulations used to generate RMSD error maps and all final simulations with the determined best-fit ZFS parameters, were run on the Knot cluster at the Center for Scientific Computing at CNSI/MRL, UCSB. The use of the Knot cluster for running simulations scripts written in MATLAB is outlined in Appendix D.

C.1.1 Range and number of magnetic field points

The number of points in the magnetic field range was set to 8000 to reach sufficient convergence, and is especially important to avoid deviations for the 240 GHz simulations. The EasySpin function `pepper` forces the computed spectra to zero at its boundaries, therefore the simulated magnetic field range was chosen to be sufficiently broader than the experimental spectra so as to not induce numerical artifacts. Both the magnetic field range and number of points within these ranges had minimal impact on computation time.

C.1.2 Orientation averaging

Orientation averaging was performed in 3° increments and a 10-fold interpolation of the orientation grid (`Opt.nKnots = [31 10]` in EasySpin). Reducing the orientation resolution to 5° increments had only a very small influence on the simulated EPR spectra given a sufficient number of points in the D and E distributions. Computation time is reduced by reducing the resolution of the orientation averaging.

C.1.3 Regular grid vs. Monte Carlo sampling of D and E (or E/D) distributions

Two different approaches to the sampling of the D and E distributions were investigated. First, the distributions of ZFS parameters were sampled by a regular grid of points. The advantage of this approach is that it allows for uniform sampling of all relevant D and E

values. However, significant computational cost is spent to compute spectra with D and E parameters far away from the maxima of the probability distributions. Additionally, when sampling the vicinity of $D = 0$ point a grid sampling generates a sharp peak corresponding to the smallest sampled value of D and is particularly problematic if the point $D = 0$ is included in the sampling grid. Such a sharp line in the center of the Gd(III) EPR spectrum is unphysical, since for nearly symmetric cases ($E \approx 0$) the width of this central peak would be dominated by inhomogeneous broadening mechanisms (e.g. higher-order ZFS terms, small g anisotropy, unresolved hyperfine couplings, etc.) rather than by the intrinsic homogeneous line width. Because the deviation in the simulated spectra with grid sampling is only found at the peak of the $|-1/2\rangle \leftrightarrow |1/2\rangle$ transition, this could be compensated for by the addition of an intrinsic linewidth.

An alternative to a regular grid sampling is a Monte Carlo sampling scheme, in which a large set of randomly distributed (D, E) pairs is generated and the overall EPR spectrum is computed as a linear combination of the EPR spectra for all (D, E) pairs. For the D distribution, which is Gaussian in all three models, this was generated directly by the MATLAB function `randn()`. The generation of the polynomial distribution for the ratio E/D in Models 2 and 3 was computed by numeric calculation of the corresponding $E/D = A$ values from the uniformly distributed cumulative probability $X = P(E/D \leq A)$. In such a sampling scheme the number of points in a particular range of D and E parameters is in line with the relative weight of this range in the overall EPR spectrum and therefore corresponds optimal computational cost.

For both sampling methods it is necessary to calibrate the convergence of the simulated EPR spectrum with respect to the number of points in the grid or the number N of random steps. The number of points required for convergence will change with the strength of the ZFS because the overall width of the spectrum scales approximately linearly with the strength of the ZFS. It was found in this work that a set of $N = 40000$ random points was sufficient to reach convergence for any of the ZFS values within the range typically observed for Gd(III) complexes.

C.1.4 Isotropic convolutional line broadening

The effect of an isotropic convolutional line broadening parameter (`Opt.lwpp` in EasySpin) was investigated in the simulations with Model 1. This spectral broadening is computed by convoluting the simulated stick spectrum with a Gaussian or Lorentzian lineshape of a given width (the use of both results in a Voigtian lineshape). This broadening does not assume any physical origin, and is therefore considered phenomenological. Additionally, it is recommended by the authors of the EasySpin toolbox to only use `Sys.lwpp` for $S = 1/2$ systems. However, in some cases it can be used to visually adjust the broadening of a simulation to better match experimental data, though no physical conclusions should be drawn from the value of parameter.

It was observed that the addition of the broadening parameter has the greatest influence on the sharp feature at the central peak of the EPR spectrum, while the shape of the outer envelope remains largely unaffected. Increasing the Gaussian and/or Lorentzian

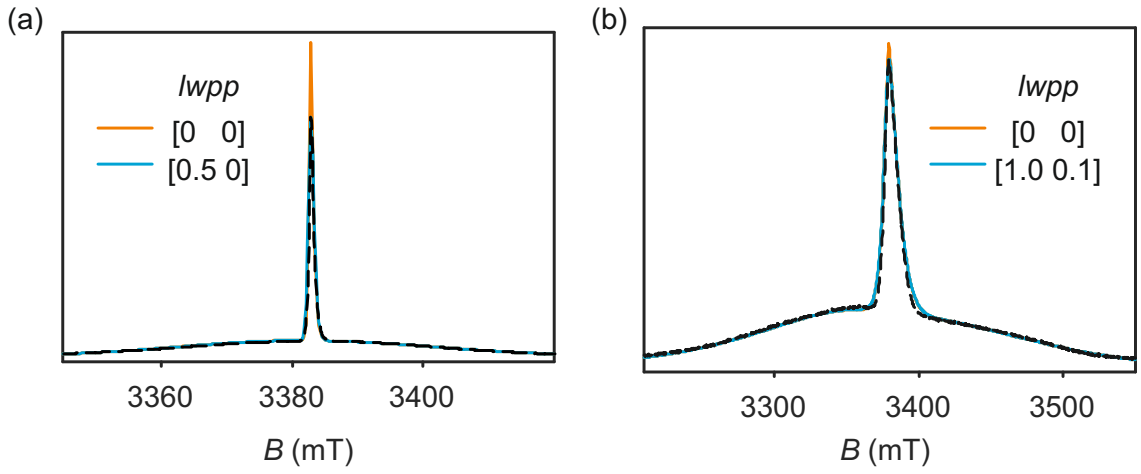

Figure C.1: Influence of the addition of an isotropic convolutional line broadening (in EasySpin, `Sys.lwpp = [Gaussian Lorentzian]`) on the line shape of simulated data for the complexes (a) Gd-NO_3Pic (**1**) and (b) Gd-PyDTTA (**7**) at W-band using the parameters for Model 1 given in Table 3.2. Simulated spectra are scaled to the outer envelope of the experimental data (black dashed lines). `Sys.lwpp = [Gaussian Lorentzian]` is input as the peak-to-peak width of the first derivative lineshapes, in units of mT.

convolutional broadening results in broadening of the sharp central peak. The effect of an isotropic convolutional line broadening parameter in the form of `Sys.lwpp = [Gaussian Lorentzian]` on the simulated spectra is demonstrated in Figure C.1 for W-band spectra of the complexes Gd-NO_3Pic (**1**) and Gd-PyDTTA (**7**). Adding a small broadening to the simulation helps to improve the match between the experimental and simulated Gd(III) EPR spectra.

C.1.5 Flip angle correction for Q-/W-band simulations

For the simulated Q-/W-band EPR spectra, each allowed transition was computed separately and the final simulated spectra were obtained by summing the contributions of the individual transitions according to their effective flip angles. The relative contribution of

each transition P_{m} to the normalized echo intensity with a $\pi/2$ - τ - π - τ pulse scheme is given by [138]

$$P_m = \sin(\alpha\pi/2)\cos(\alpha\pi),$$

$$\alpha = \langle m_S|\hat{S}^+|m_S+1\rangle / \langle -1/2|\hat{S}^+|+1/2\rangle$$

$$= \frac{\sqrt{S(S+1)-m_S(m_S+1)}}{\sqrt{S(S+1)+1/4}} = \sqrt{\frac{63-4\cdot m_S(m_S+1)}{64}}. \tag{C.1}$$

which results in the normalized transition probabilities given in Table C.1.

Transition(s)	P_m
$\|-7/2\rangle \leftrightarrow \|-5/2\rangle$, $\|5/2\rangle \leftrightarrow \|7/2\rangle$	0.419
$\|-5/2\rangle \leftrightarrow \|-3/2\rangle$, $\|3/2\rangle \leftrightarrow \|5/2\rangle$	0.893
$\|-3/2\rangle \leftrightarrow \|-1/2\rangle$, $\|1/2\rangle \leftrightarrow \|3/2\rangle$	0.994
$\|-1/2\rangle \leftrightarrow \|1/2\rangle$	1

Table C.1: Flip angle correction applied to the different allowed transitions of Gd(III) according to their normalized transition probabilities P_m.

Note that the weight coefficients listed in the Table C.1 are only valid with good precision when the electron spin echo is tuned at the maximum of the Gd(III) EPR spectrum. If the spin echo is set up at the high-field or low-field end of the spectrum the microwave pulse power will be adjusted for optimal probability of the $|-7/2\rangle \leftrightarrow |-5/2\rangle$ and $|5/2\rangle \leftrightarrow |7/2\rangle$ transitions, resulting in over flipping of the other Gd(III) transitions.

C.1.6 Parameters used to generate simulation library for Models 2/3

The library of simulations generated for Models 2/3 used typical values for measurement frequency (Q-/W-band) and measurement temperature (240 GHz), as detailed in Table C.2. These values differ slightly from the exact experimental parameters due to tuning of the resonant cavity for the Q-/W-band measurements and difficulty in stabilizing the sample temperature at precisely 5 K for the 240 GHz measurements. The magnetic field ranges of the simulations were set to well cover all of the experimental traces in order to avoid numerical artifacts, and are also listed in Table C.2.

	`Exp.mwFreq`	`Exp.Range`	`Exp.Temperature`
Q-band	34.50	[219 2219]	10
W-band	94.25	[2380 4380]	10
240 GHz	240	[7608 9608]	5

Table C.2: Experimental parameters used as EasySpin inputs to generate a library of simulations for determination of ZFS parameters with Models 2/3. Frequencies are given in units of GHz, magnetic field ranges in mT, and temperatures in K.

The simulation library for Models 2/3 employed Monte-Carlo sampling of the D and E/D distributions, with the number of random steps set to `N = 20000`. No additional line broadening terms were included. The number of points on the magnetic field axis was set to `Exp.nPoints = 8000`, and orientation averaging was set to `Opt.nKnots = [31 10]`.

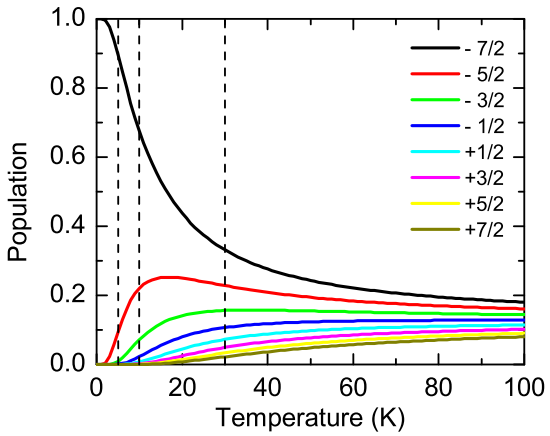

Figure C.2: Boltzmann population of states as a function of temperature at 240 GHz. The dashed lines indicate temperatures of 5 K, 10 K, and 30 K.

C.2 Boltzmann population of states at 240 GHz

Spin populations of states follow a Boltzmann distribution. At high-fields and low temperatures, this can result in $> 90\%$ population of the ground state. Figure C.2 shows the population of each of eight states for a $S = 7/2$ system as a function of temperature.

C.3 Additional results from Model 1

The Gd(III) spectra from the three measurement frequencies were fit by visual inspection using Model 1. Random sampling of the D and E value distributions was used with a convolutional line broadening parameter (`lwpp` in EasySpin) included. Selected fit values, as determined by eye and before reordering of the indices, are presented in Table 3.2 of the main text and Table C.3 below, with the corresponding simulated spectra plotted in Figure C.3. Table C.3 additionally gives the ZFS parameter values D and σ_D for Model

1 after reordering of the indices according to Equations 3.4 and 3.5.

The reording of the indices results in a bimodal distribution for D and so the asymmetry $P(+D)/P(-D)$ can now be computed for Model 1. These values, listed in Table C.3, are quite comparable to the results from Model 3. This is especially comforting since in Model 1 the asymmetry $P(+D)/P(-D)$ is fixed by the definition of the model, while for Model 3 it is taken as a free parameter in the fit.

Compound	D_{init}	D_{pos}	D_{neg}	σD_{init}	σD_{pos}	σD_{neg}	$\frac{P(+D)}{P(-D)}$
Gd-NO_{3Pic} (**1**)	420	472	- 418	140	124	111	1.4
Gd-DOTA (**2**)/ maleimide-DOTA (**3**)	-600	515	- 675	240	160	204	0.6
Gd-PyMTA (**4ab**)	1070	1200	- 1065	357	316	288	1.4
Gd-TAHA (**5**)	1250	1400	- 1117	417	311	272	1.4
Gd-PCTA (**6**)	1780	1850	- 1370	508	440	318	3.3
Gd-PyDTTA (**7**)	1800	1845	- 1275	514	439	271	3.3

Table C.3: Change in $\langle D \rangle$ and σ_D upon reordering the ZFS parameters in Model 1 according to their definitions convention (Equations 3.4 and 3.5) for all studied Gd(III) complexes. Units are given in MHz.

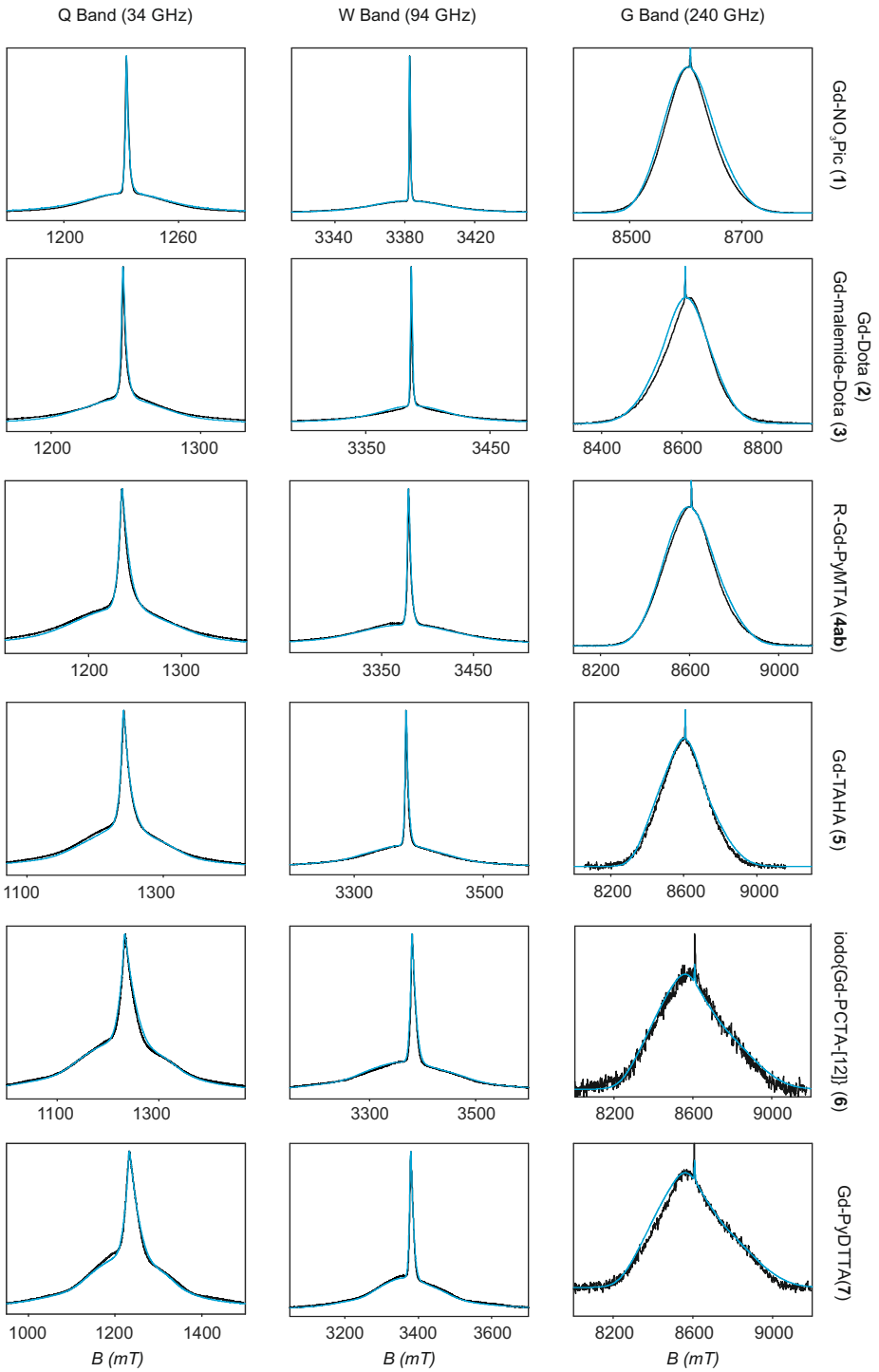

Figure C.3: Selected fits of Model 1 (light blue lines) as determined by eye compared to experimental spectra (black lines) for all measured frequency bands and Gd(III) compounds.

C.4 Additional results from Model 2

Figure C.4 displays the RMSD error as a function of (D, σ_D) parameter values computed using Model 2. In these calculations, the region about the central transition was excluded from the scaling of the simulation to the data and from the RMSD error calculations. Figures C.5 and C.6 show simulations using the best-fit ZFS parameters D and σ_D for Model 2, as given in Table 3.2 in the main text.

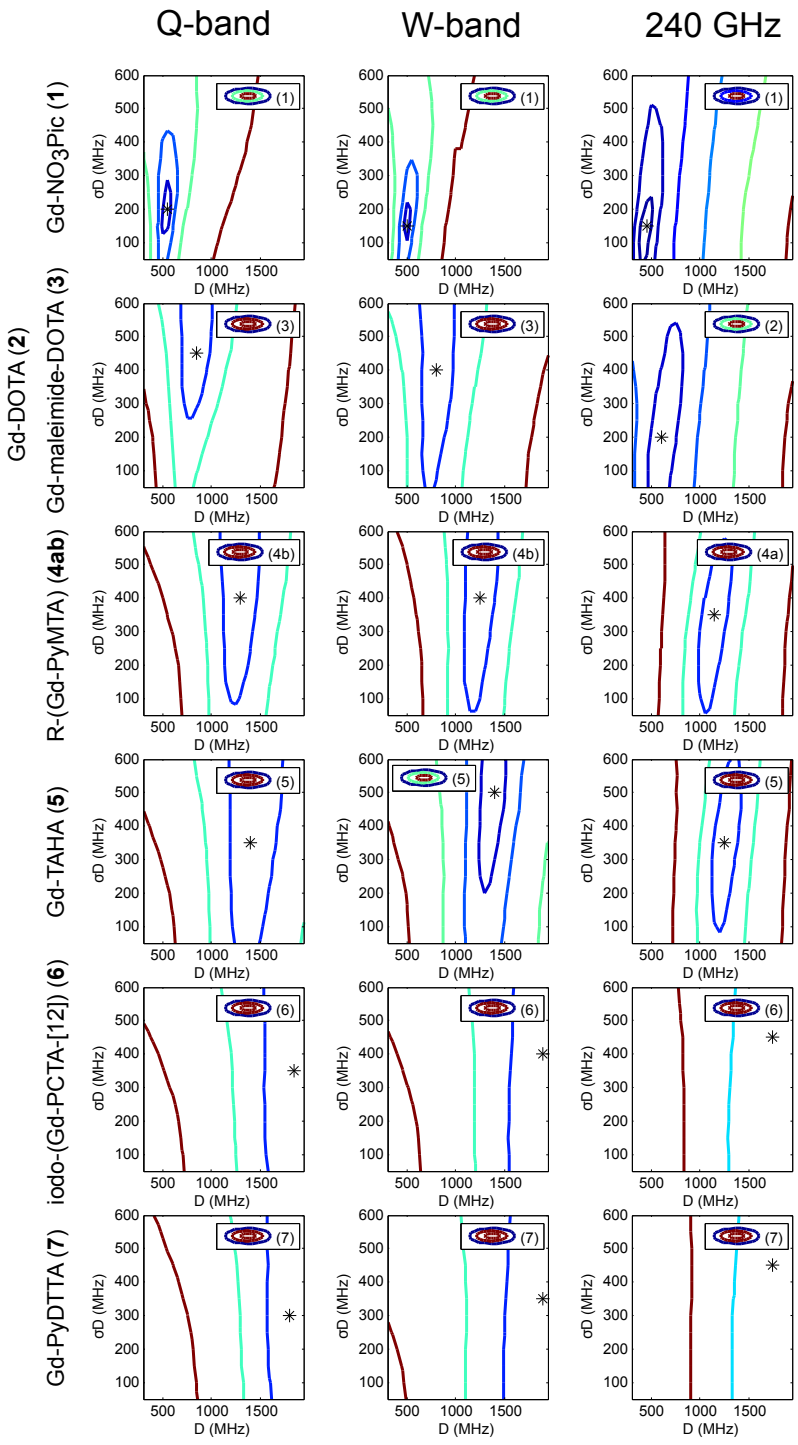

Figure C.4: Contours of constant RMSD as a function of D and σ_D parameter values using Model 2 with the region about the central peak excluded from the calculation of RMSD errors. Simulated spectra were normalized to the experimental data using only the shoulders of the spectra. The asterisk denotes the position of minimum RMSD on the 50 MHz grid of parameter values available in the simulation library.

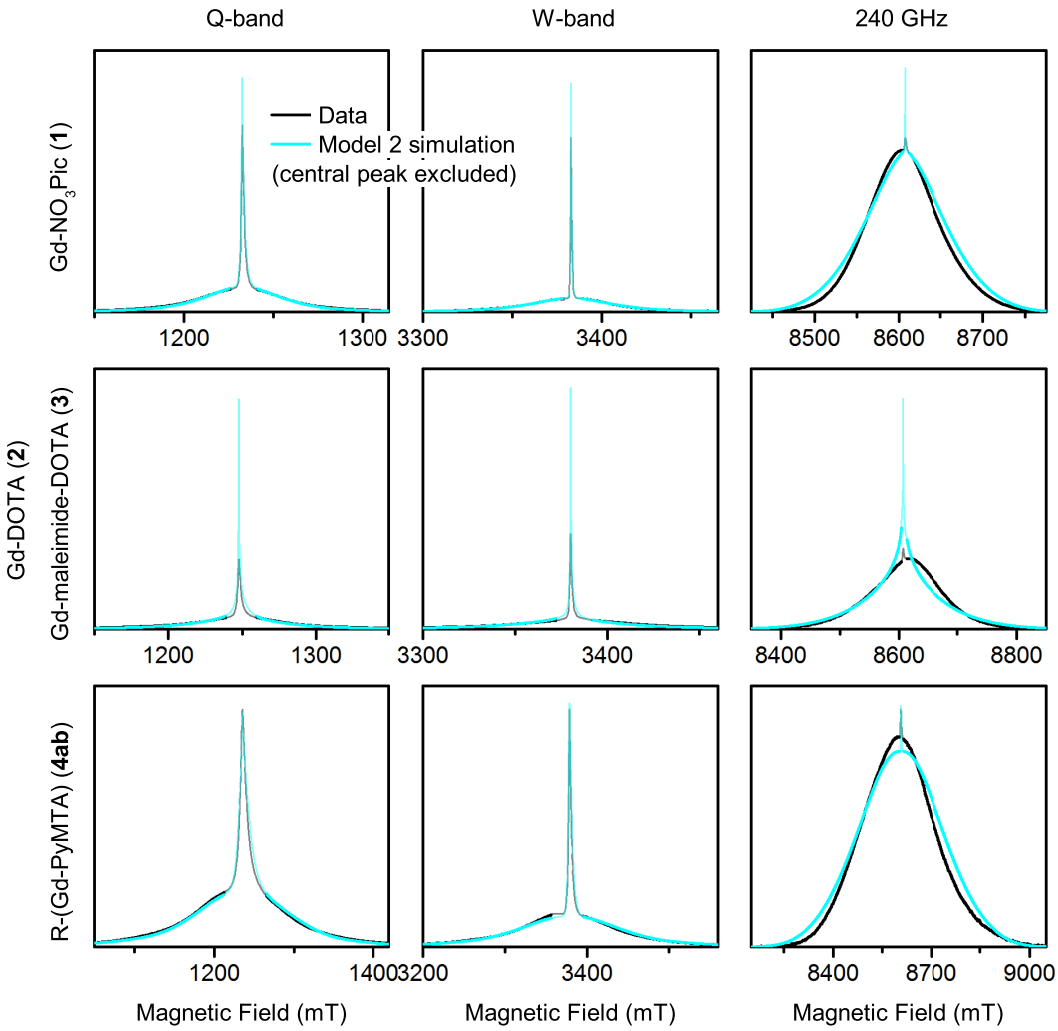

Figure C.5: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-NO_3Pic (**1**), Gd-DOTA (**2**) (240 GHz sepctra)/Gd-maleimide-DOTA (**3**) (Q-/W-band spectra), and R-(Gd-PyMTA) (**4a**) (240 GHz spectra)/R-(Gd-PyMTA) (**4b**) (Q-/W-band spectra). Overlaid are simulations with Model 2 using the best-fit ZFS parameters determined with the region about the central peak excluded from the analysis (Table 3.2). The faded region indicates the portion of the spectra which were excluded from analysis.

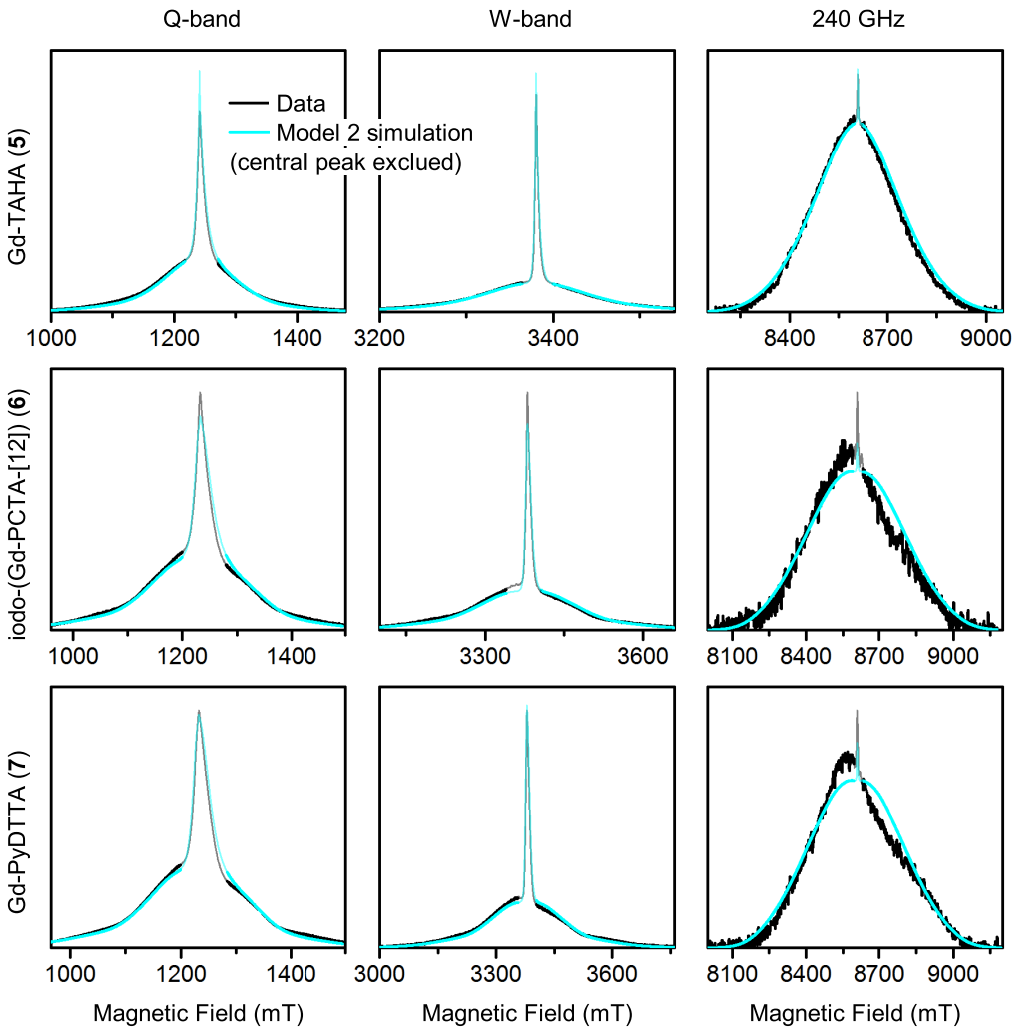

Figure C.6: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-TAHA (**5**), iodo-(Gd-PCTA-[12]) (**6**), and Gd-PyDTTA (**7**). Overlaid are simulations with Model 2 using the best-fit ZFS parameters determined with the region about the central peak excluded from the analysis (Table 3.2). The faded region indicates the portion of the spectra which were excluded from analysis.

C.5 Additional results from Model 3

Figure C.7 displays the RMSD error as a function of (D, σ_D) and $(P(+D)/P(-D), \sigma_D)$ parameter values computed using Model 3. In this calculation, the region about the central transition was excluded from the scaling of the simulation to the data and from the RMSD error calculations.

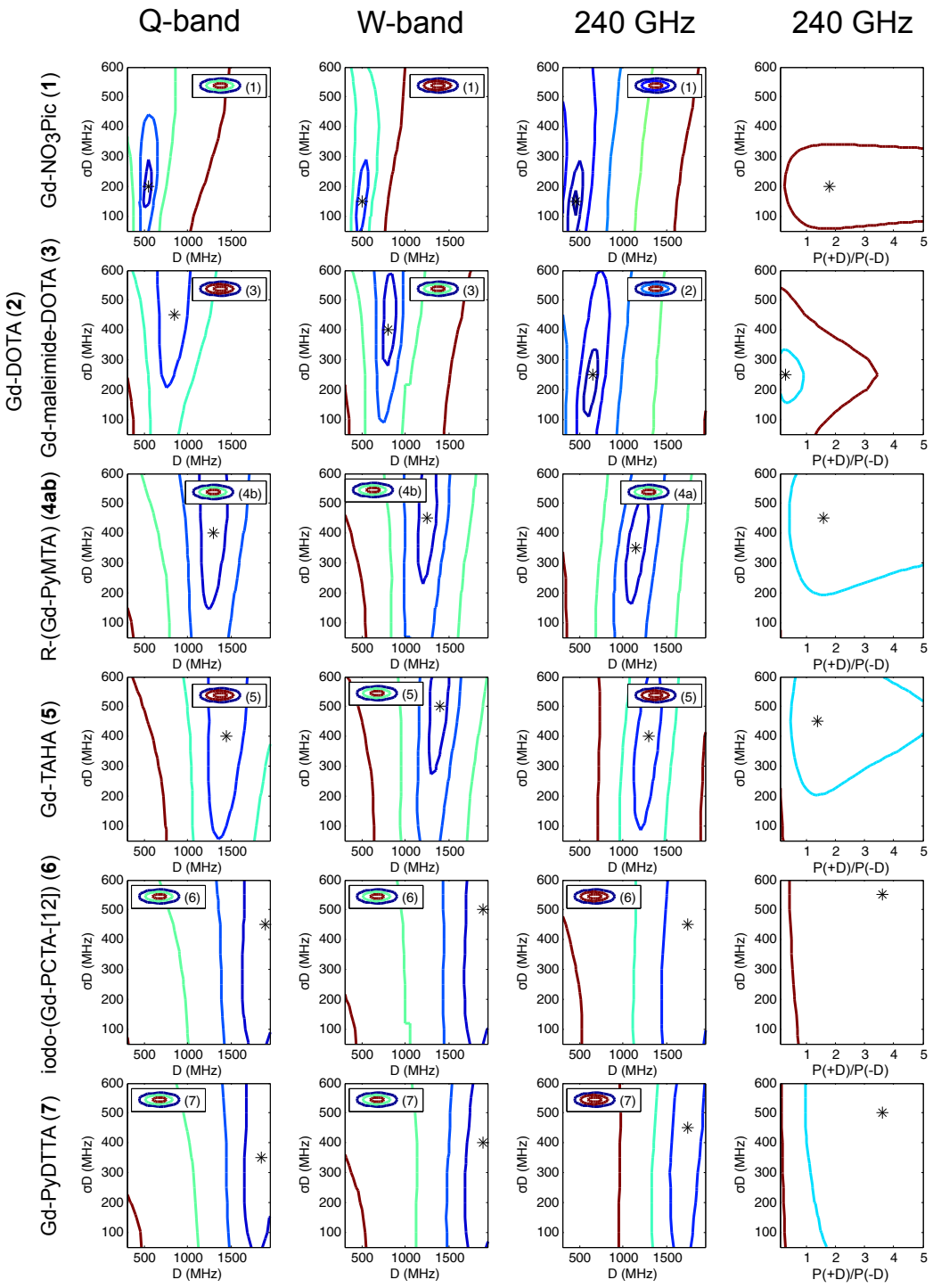

Figure C.7: Contours of constant RMSD as a function of D and σ_D parameter values, and as a function of σ_D and $P(+D)/P(-D)$ parameter values, using Model 3 with the region about the central peak excluded from the calculation of RMSD errors. Simulated spectra were normalized to the experimental data using only the shoulders of the spectra. The asterisk denotes the position of minimum RMSD on the 50 MHz grid of parameter values available in the simulation library.

C.6 Comparison of the D and E/D distributions from Models 1 and 3

Models 1 and 3, after correction to ordering of the indices in Model 1, produced very similar forms of the distributions for the ZFS parameters D and E/D. Figure C.8 gives a comparison for the ZFS parameter distributions determined by Models 1 and 3. The corresponding ZFS parameter values used to generate these plots are given in Tables 3.2 and C.4.

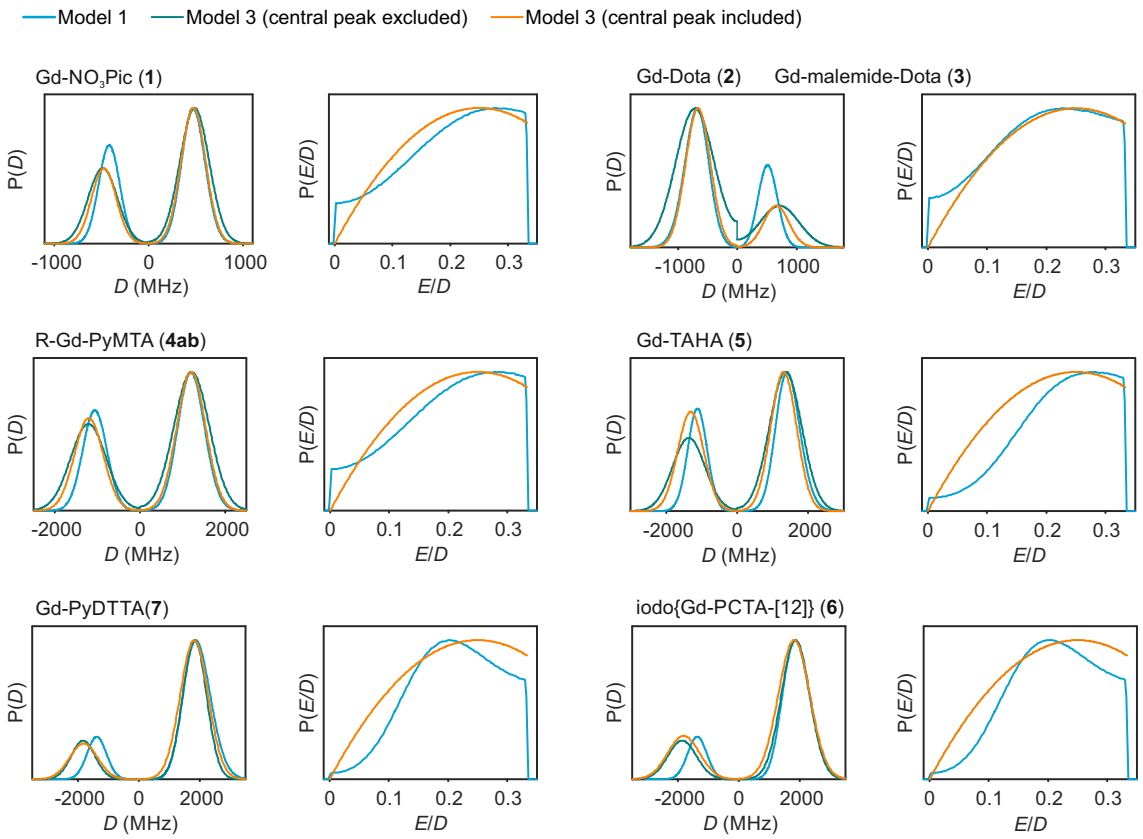

Figure C.8: Comparison of D and E/D distributions for Model 1 (light blue lines), Model 3 excluding the region of the central peak (dark cyan lines), and Model 3 including including the region of the central peak (orange lines) for the different Gd(III) complexes.

C.7 Additional results from Models 2/3 with the central peak included

In the main text of Chapter 3, results presented for Models 2 and 3 were calculated with the region about the central transition excluded from the analysis. Here we present contour plots of the RMSD errors computed using the full EPR spectra including the region of the central peak, the corresponding calculated best-fit ZFS parameters, and the resulting simulated spectra compared to the experimental dataset for both Models 2 and 3.

Model	Complex	D (MHz)	σ_D (MHz)	$\frac{\sigma_D}{D}$	$\frac{P(+D)}{P(-D)}$
2	Gd-NO3Pic	494 ± 36	126 ± 89	0.26	—
	Gd-DOTA/ Gd-maleimide-DOTA	700 ± 71	218 ± 213	0.31	—
	R-(Gd-PyMTA)	1214 ± 71	331 ± 194	0.27	—
	Gd-TAHA	1307 ± 59	359 ± 159	0.28	—
	iodo-(Gd-PCTA-[12])	1821 ± 210	526 ± 305	0.29	—
	Gd-PyDTTA	1829 ± 198	467 ± 346	0.26	—
3	Gd-NO3Pic	473 ± 30	130 ± 71	0.28	1.8
	Gd-DOTA/ Gd-maleimide-DOTA	660 ± 50	210 ± 125	0.32	0.3
	R-(Gd-PyMTA)	1203 ± 60	339 ± 160	0.28	1.5
	Gd-TAHA	1307 ± 54	365 ± 144	0.28	1.4
	iodo-(Gd-PCTA-[12])	1812 ± 160	520 ± 230	0.29	3.2
	Gd-PyDTTA	1814 ± 138	459 ± 258	0.25	3.9

Table C.4: Extracted D and σ_D values using Model 2 using the full EPR spectra for analysis, and extracted D, σ_D, and $P(+D)/P(-D)$ values using Model 3 using the full EPR spectra for analysis.

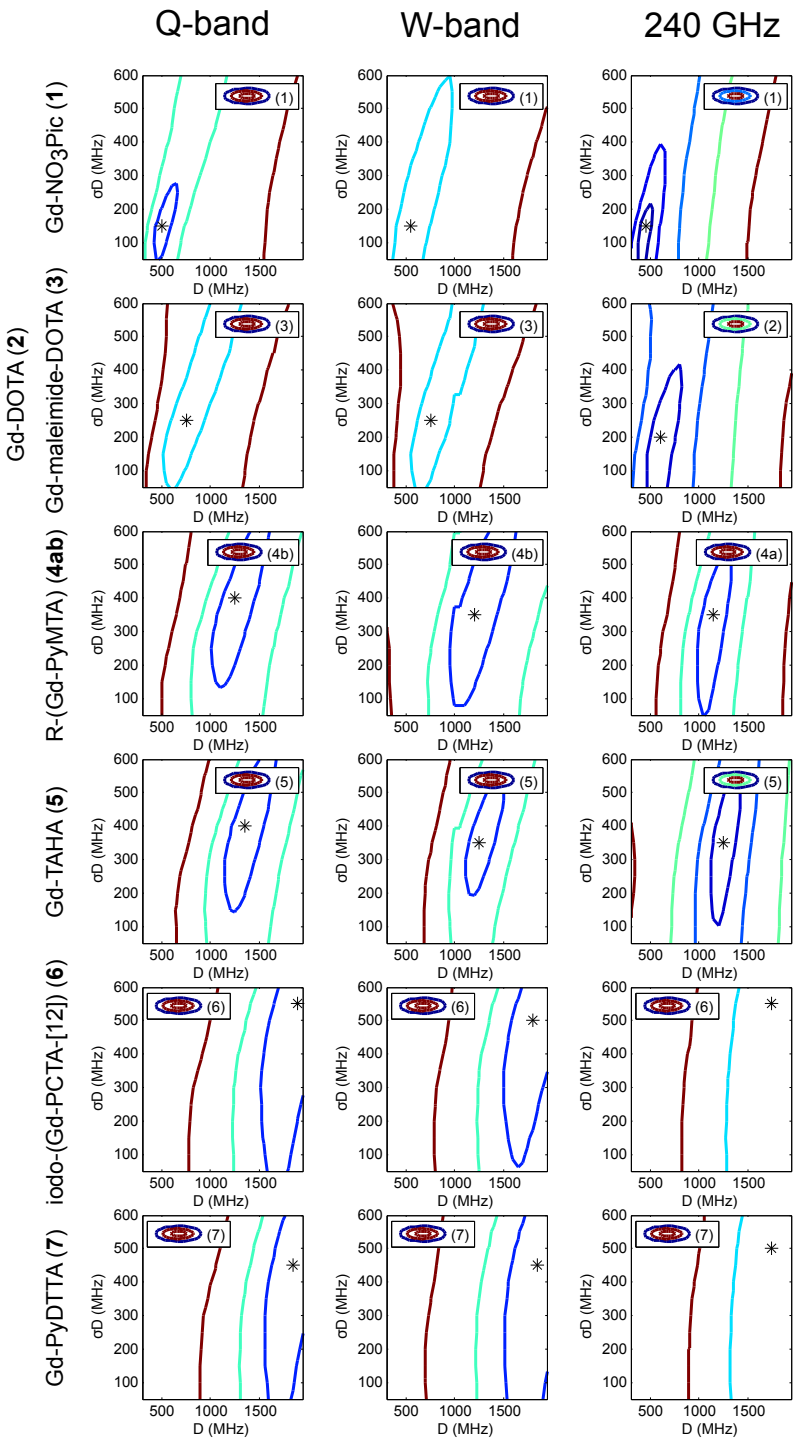

Figure C.9: Contours of constant RMSD as a function of D and σ_D parameter values using Model 2 with the full EPR spectra used for calculation of the RMSD errors. Simulated spectra were normalized to the experimental data using the full spectrum, including the region of the central peak. The asterisk denotes the position of minimum RMSD on the 50 MHz grid of parameter values available in the library of simulations.

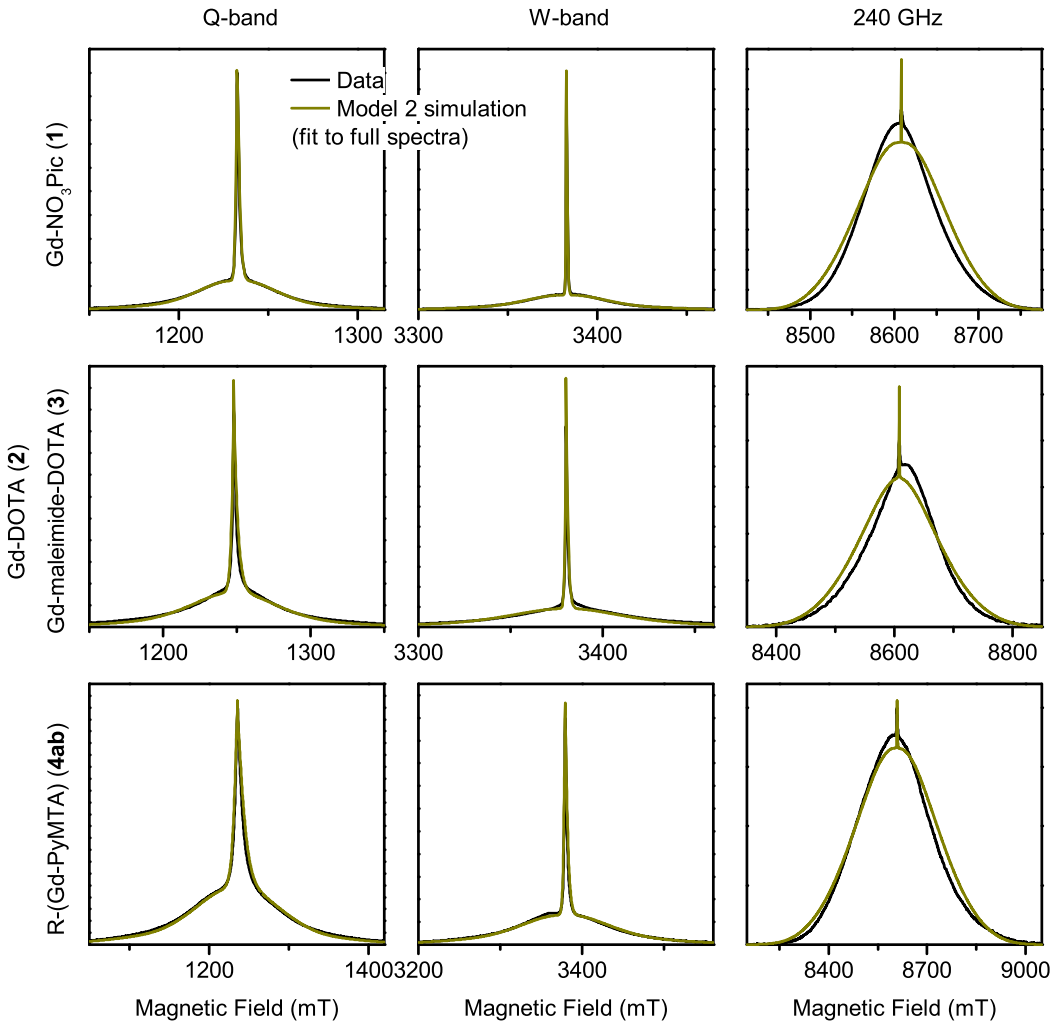

Figure C.10: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-NO_3Pic (**1**), Gd-DOTA (**2**) (240 GHz sepctra)/Gd-maleimide-DOTA (**3**) (Q-/W-band spectra), and R-(Gd-PyMTA) (**4a**) (240 GHz spectra)/R-(Gd-PyMTA) (**4b**) (Q-/W-band spectra). Overlaid are simulations with Model 2 using the best-fit ZFS parameters determined using the full EPR spectra in the calculation of RMSD error maps, given in Table C.4.

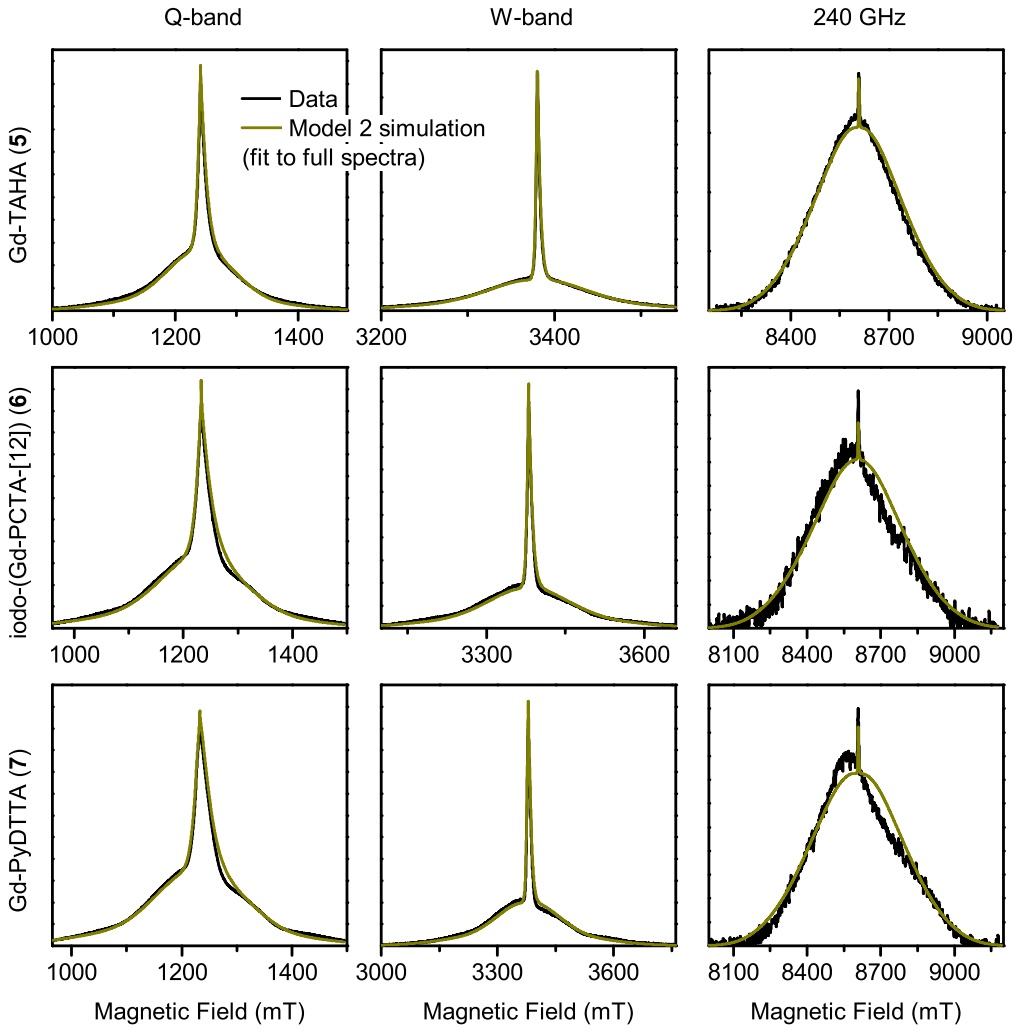

Figure C.11: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-TAHA (**5**), iodo-(Gd-PCTA-[12]) (**6**), and Gd-PyDTTA (**7**). Overlaid are simulations with Model 2 using the best-fit ZFS parameters determined using the full EPR spectra in the calculation of RMSD error maps, given in Table C.4.

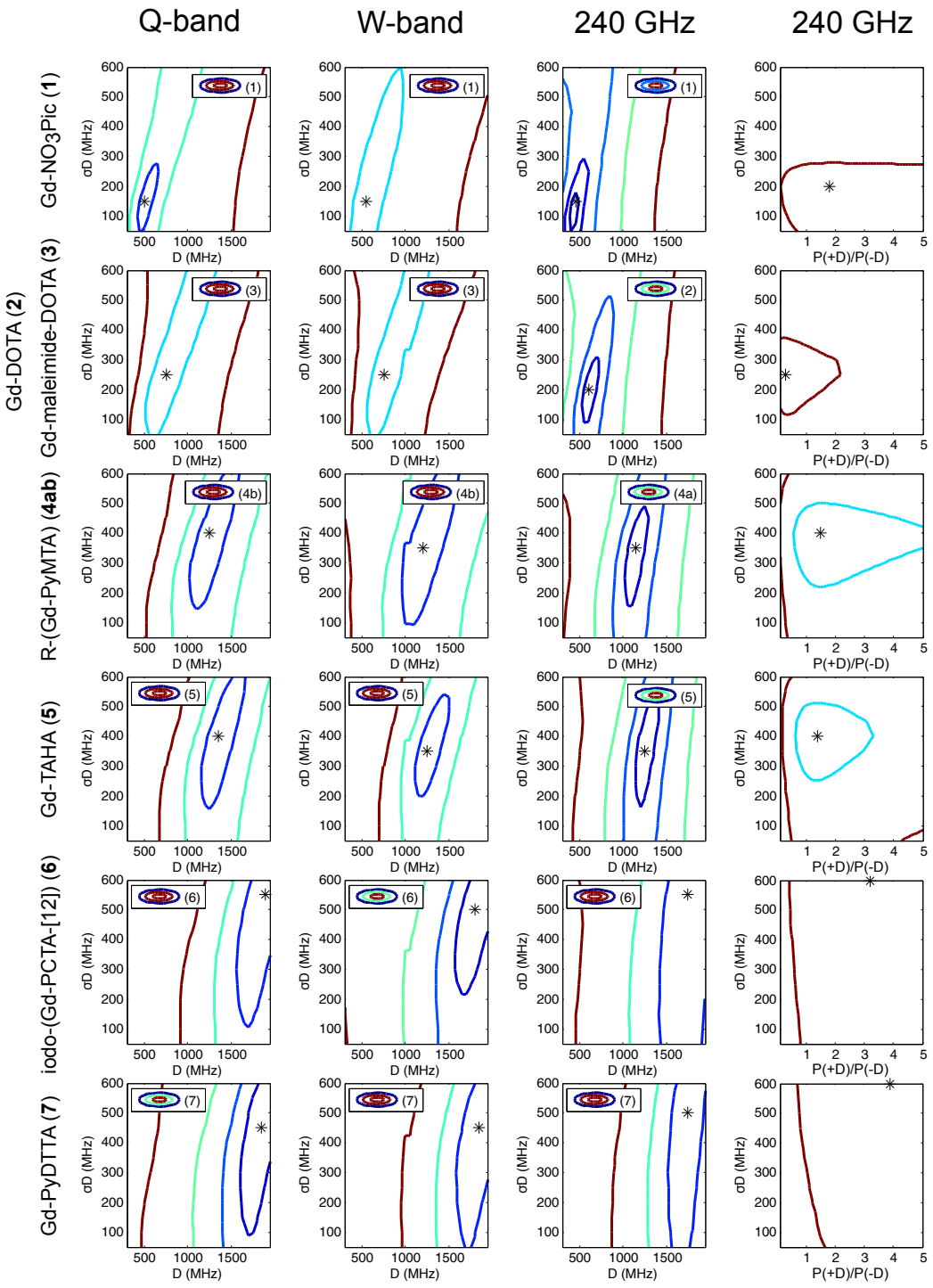

Figure C.12: Contours of constant RMSD as a function of D and σ_D parameter values, and $P(+D)/P(-D)$ and σ_D parameter values, using Model 3 with the full EPR spectra used for calculation of the RMSD errors. Simulated spectra were normalized to the experimental data using the full spectrum, including the region of the central peak. The asterisk denotes the position of minimum RMSD on the 50 MHz grid of parameter values available in the library of simulations.

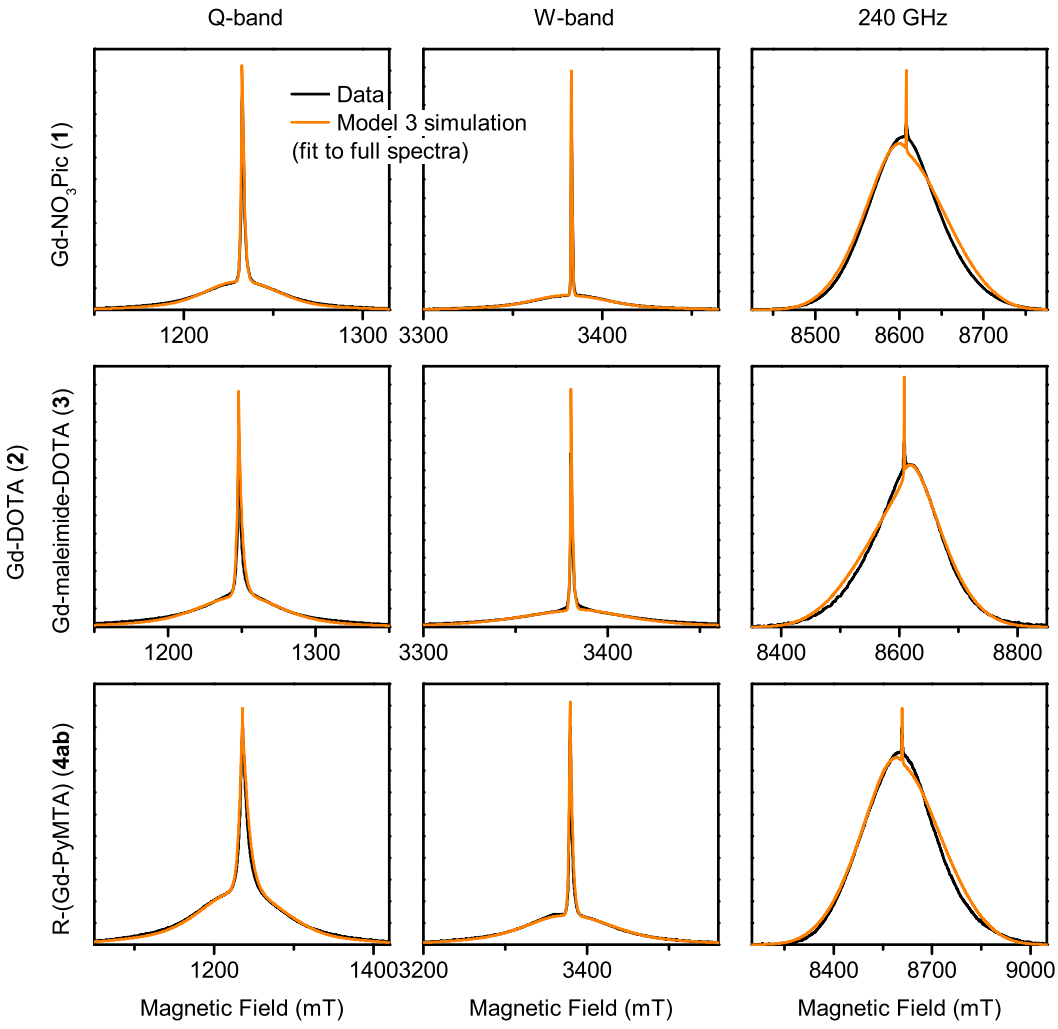

Figure C.13: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-NO_3Pic (**1**), Gd-DOTA (**2**) (240 GHz sepctra)/Gd-maleimide-DOTA (**3**) (Q-/W-band spectra), and R-(Gd-PyMTA) (**4a**) (240 GHz spectra)/R-(Gd-PyMTA) (**4b**) (Q-/W-band spectra). Overlaid are simulations with Model 3 using the best-fit ZFS parameters determined using the full EPR spectra in the calculation of RMSD error maps, given in Table C.4.

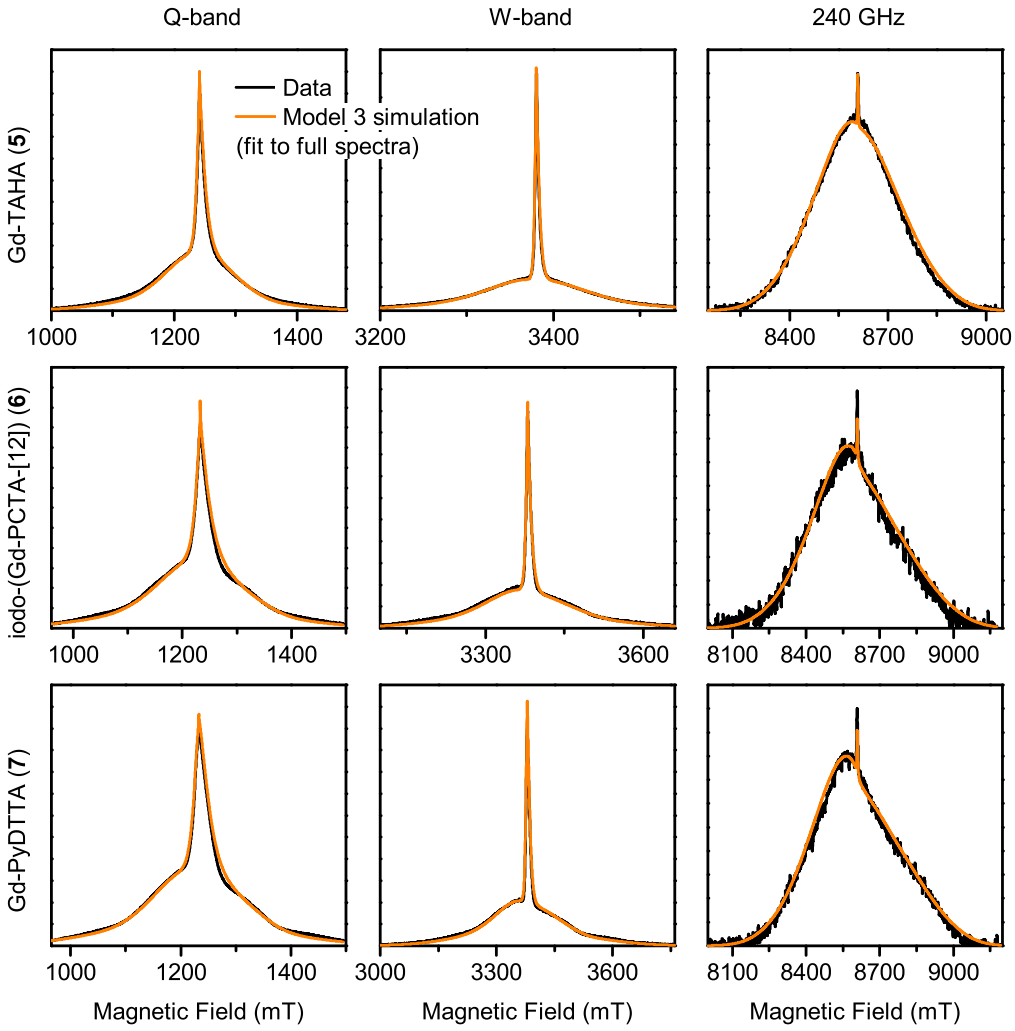

Figure C.14: Measured EPR spectra at Q-/W-band and 240 GHz for the Gd(III) complexes Gd-TAHA (**5**), iodo-(Gd-PCTA-[12]) (**6**), and Gd-PyDTTA (**7**). Overlaid are simulations with Model 3 using the best-fit ZFS parameters determined using the full EPR spectra in the calculation of RMSD error maps, given in Table C.4.

C.8 P(+D)/P(-D) parameter error estimation for Model 3

The criterion of an acceptable fit as being those values which are within the region bounded by a contour of twice the minimum RMSD is not a reasonable estimate of the error on the asymmetry parameter $P(+D)/P(-D)$ in Model 3. The most obvious effect of this parameter on the simulated spectra is to set the relative position of the broad component of the EPR spectrum with respect to the sharp central peak corresponding to the $|-1/2\rangle \leftrightarrow |1/2\rangle$ subspectrum, particularly in the high field data. The width of this central peak is so narrow compared to the broad component of the 240 GHz EPR spectrum that it has a relatively small impact on the overall RMSD of the fit, though there is enough of an effect on the RMSD to assign a position of minimum RMSD in a contour plot of $P(+D)/P(-D)$ and σ_D, as was done to determine the other parameter values for Models 2 and 3.

In order to estimate an error on this parameter value, we plot the separation of the position of the $|-1/2\rangle \rightarrow |1/2\rangle$ transition and the peak of the broad component of the 240 GHz EPR spectra with the value determined for $P(+D)/P(-D)$ via contour plots of the RMSD error as a function of $P(+D)/P(-D)$ and σ_D. This is shown in Figure C.15 for analysis conducted with and without the region about the central transition included in the calculation of RMSD error maps. The plot of the determined $P(+D)/P(-D)$ parameter value vs. the peak separation for the various Gd(III) complexes falls on a line.

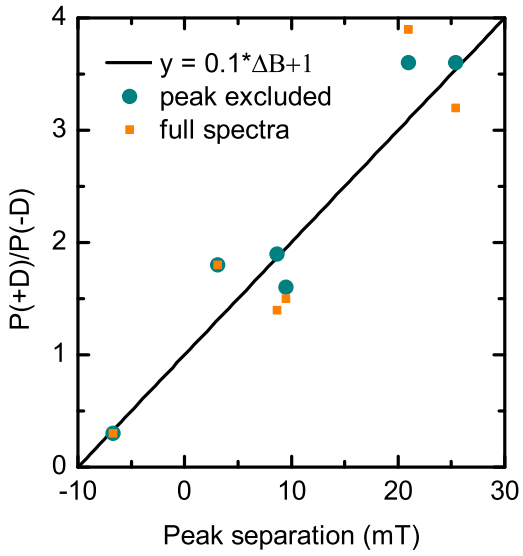

Figure C.15: Plot of determined $P(+D)/P(-D)$ parameter values vs. the separation in field between the broad component and the sharp central peak of the 240 GHz EPR spectra, for analysis in which the central transition was excluded from the calculation of the RMSD error maps and analysis using the full 240 GHz EPR spectra.

The average deviation in the determined $P(+D)/P(-D)$ values from this line is 0.24 for the analysis excluding the central transition and 0.43 for analysis using the full 240 GHz EPR spectra.

C.9 Minimum RMSD errors for Models 2 and 3

Figure C.16 shows the minimum RMSD errors calculated for each of the contour plots for Models 2 and 3 (denoted by an asterisk in the RMSD contour plots). The RMSD errors for the 240 GHz spectra are about an order of magnitude larger than for the Q-/W-band spectra. For the Q-/W-band spectra, the minimum RMSD error does not change significantly whether Model 2 or Model 3 is used, but for the 240 GHz spectra the

addition of the asymmetry parameter P(+D)/P(-D) significantly reduces the minimum RMSD error. The asymmetry parameter has the largest effect on the minimum RMSD value at 240 GHz for Gd-DOTA (**2**), iodo-(Gd-PCTA-[12]) (**6**), and Gd-PyDTTA (**7**). In general, the minimum RMSD error is smaller if the region about the sharp central peak is excluded, particularly for the Q-/W-band spectra where the $|-1/2\rangle \rightarrow |1/2\rangle$ transition makes up a significant portion of the EPR spectra and is particularly sensitive to any additional broadening terms (e.g. higher-order ZFS or hyperfine interactions) which may be present but are not accounted for in the simulations with Models 2 and 3.

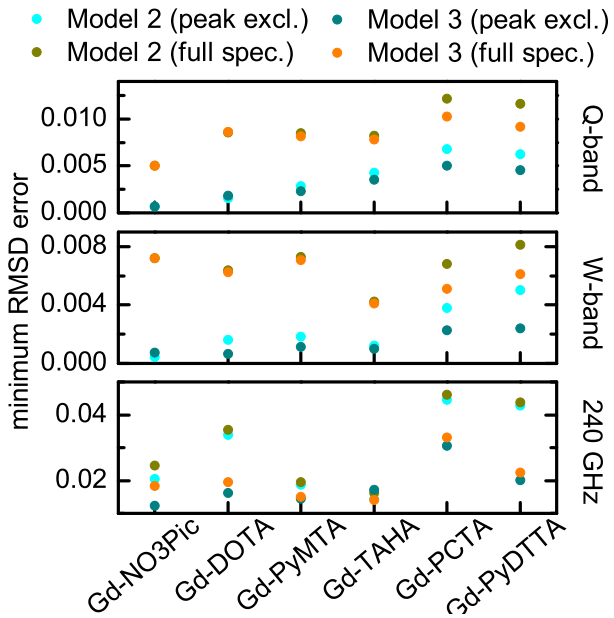

Figure C.16: Minimum RMSD errors from the (D, σ_D) contour plots for Models 2 and 3, with and without the region about the central peak excluded in the calculation of RMSD errors.

C.10 X-band measurements and simulation with Model 3

A subset of the Gd(III) complexes studied in Chapter 3 were additionally measured at X-band. These data are plotted in Figure C.17, along with simulations using Model 3 and the ZFS parameter values given in Table 3.2. The simulations which gave quite nice results at Q-band and above give very poor fits to the X-band data. Surprisingly, the simulations produced lineshapes which are broader than the experimental EPR lineshapes. This is in contrast to our expectation that if the simulation produces an artificially narrow central peak and the X-band spectrum is dominated by the $|-1/2\rangle \rightarrow |1/2\rangle$ transition (Figure C.18), then the simulated X-band spectrum should produce a narrower lineshape. This was also unexpected given the success of Benmelouka *et al.* [21] in simulating X-band spectra using ZFS parameter values determined from 240 GHz data. The cause of this discrepancy has not yet been resolved.

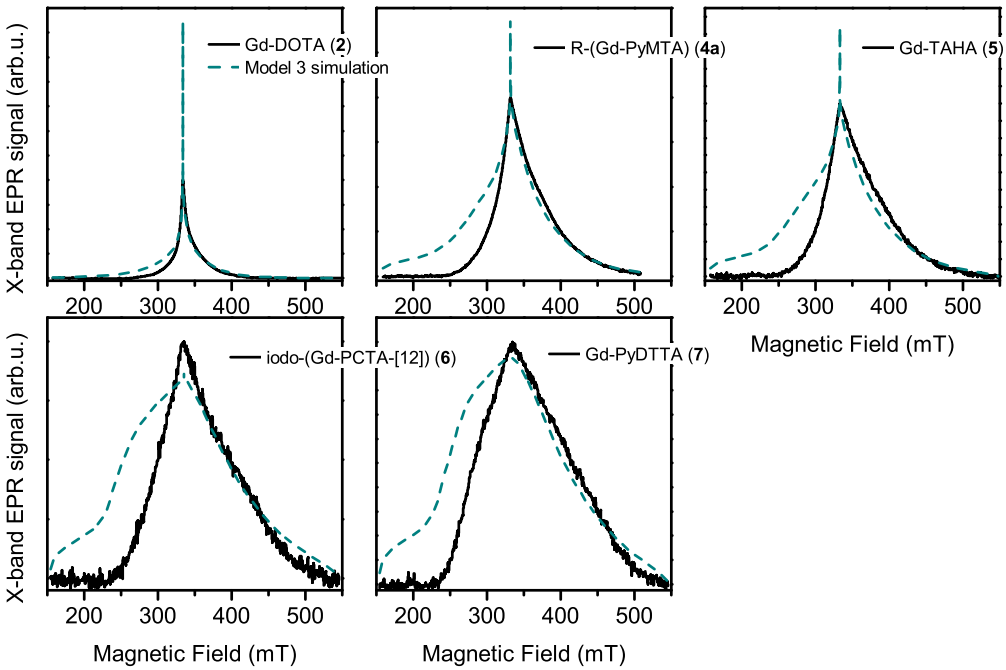

Figure C.17: X-band field-swept EDEPR spectra at 30 K measured for a subset of the Gd(III) complexes from Chapter 3. Overlaid are simulations with Model 3 using the ZFS parameter values from Table 3.2.

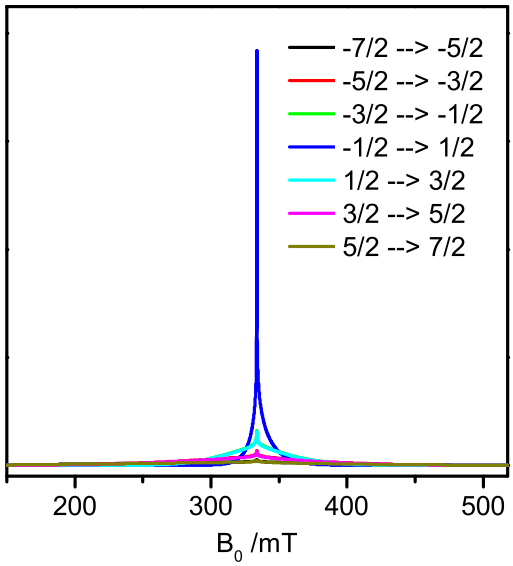

Figure C.18: Evolution of allowed EPR transitions at X-band and 30 K.

C.11 Comparison of 240 GHz spectra of the complexes R-(Gd-PyMTA) (4ab)

The two measured R-(Gd-PyMTA) complexes included in this work did not show any significant difference in the magnitude of the ZFS parameter D, despite the slight difference in structure. After the completion of the initial study, the complex R-(Gd-PyMTA) (**4b**) was measured at 240 GHz, and shows a nearly identical EPR spectrum to R-(Gd-PyMTA) (**4a**), indicating that the two complexes have nearly identical ZFS parameter values. This similarity is likely because the atoms which directly coordinate the Gd(III) are identical for both complexes, with the change in functional group being well separated from the first coordination shell.

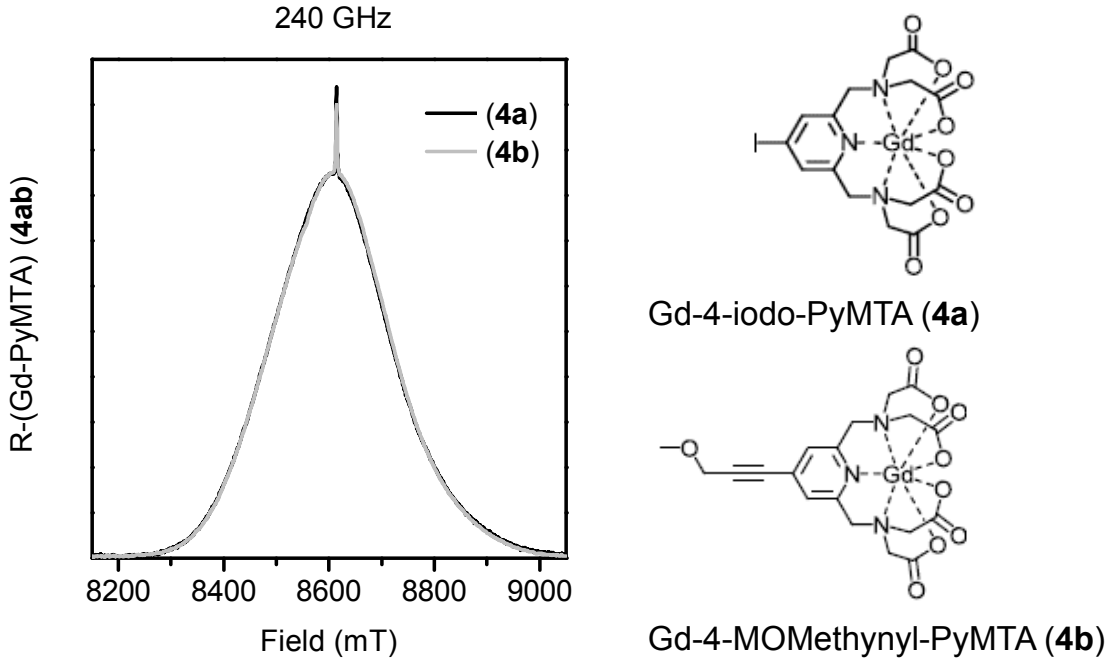

Figure C.19: Measured EPR spectra at 240 GHz and $\sim$ 5 K of R-(Gd-PyMTA) (**4a**) and R-(Gd-PyMTA) (**4b**).

Appendix D

UCSB Knot computing cluster

Knot is a computing cluster available to UCSB students and researchers. This cluster specializes in highly parallel computing jobs, and made much of the simulation work in Chapter 3 possible. Use of the Knot cluster is supported by the Center for Scientific Computing at the CNSI and MRL through the grants NSF MRSEC (DMR-1121053) and NSF CNS-0960316.

D.1 Accessing the Knot cluster

A user account may be requested through the UCSB CSC website at http://csc.cnsi.ucsb.edu/acct. On Windows machines, an SSH client such as PuTTY is required. In PuTTY, the Host Name for the Knot cluster is `knot.cnsi.ucsb.edu` (or more generally `machinename.cnsi.ucsb.edu`). Alternatively, you can log in directly to your account by using a Host Name in the format

`username@machinename.cnsi.ucsb.edu`.

D.2 Navigating within Knot

The terminal connected to Knot uses standard Linux/Unix commands for navigation. Commands generally follow the structure `command -[options] argument`. Details about commands and their options can be found using `commandname --help`. Some of the more useful commands are:

`pwd` "print working directory" displays the full path of the current directory

`ls` lists the contents of the current directory

`cd folder` changes current directory to the indicated folder

`cd ..` moves up one directory

`mkdir foldername` creates a new directory

`cp location1 location2` copies the file at location1 to location2

`mv file1 file2` renames or moves file1 into file2

`rm file` removes file

Adding the `-r` option to commands such as `cp` or `rm` will affect the directory and its contents. Use caution when removing files as the linux terminal does not have an easily recoverable equivalent of the Windows Recycle Bin if something is accidentally deleted. The Tab button on the keyboard autocompletes the names of files and directories.

D.3 Moving files to and from the server

A client such as PSCP (PuTTY Secure Copy client) is used to transfer files between your local computer and the cluster using an SSH connection. On your local computer, pscp.exe needs to be in the directory containing the files to be transferred. Files can be copied from the current directory on your local computer to your home directory on the server in the Windows Command Prompt using the command

```
pscp filename username@knot.cnsi.ucsb.edu:/home/username
```

Directories and their contents can be moved by adding the `-r` option to this command. Files are moved from the server to your local computer by reversing the order of this command

```
pscp username@knot.cnsi.ucsb.edu:/home/username/filename .
```

Instead of specifying the full path of the target directory on your local computer, the `.` may be used to specify that the target directory is the current directory.

D.4 Editing files on the server

To edit files on the server, a basic text editor such as Vim is used. To open a file on the server in Vim, type

```
vim filename
```

Vim has two modes: (1) Insert mode where you can type like a normal text editor, and (2) Command mode where you give commands to the editor to do things. Press `i` to enter insert mode and `ESC` to enter Command mode. Some of the more useful commands in command mode are

`:q` exit

`:wq` save and exit

D.5 Submitting Matlab jobs

To run jobs on Knot, they must be submitted to the queue. The server uses the TORQUE resource manager to manage and schedule jobs. To submit a job, you supply TORQUE with a batch job script (e.g. `run.job`) which contains the number of nodes and cores required and the name of your job. This script is written in BASH. An example script for submitting a MATLAB job would look like

```
#!/bin/bash
#PBS -l nodes=1:ppn=8
#PBS -l walltime=72:00:00


#PBS -m bae
#PBS -M youremail@ucsb.edu
```

```
cd $PBS_O_WORKDIR

/sw/bin/matlab -nodisplay -nodesktop < simulation.m
```

In BASH scripts, # will generally indicate a comment line. However, #PBS starts a line containing directions that will be read by the queue. The required number of nodes and cores are defined in the second line (ppn is processors per node). The walltime command tells the queue the maximum wall clock time allowed for the job. The default maximum time allowed for jobs is 72 hours. This is useful for timing out potential run-away jobs. TORQUE will also prioritize queuing for shorter jobs. The command #PBS -M youremail@ucsb.edu tells the system to email you notifications about the submission and failure or completion of the job. cd $PBS_O_WORKDIR ensures that we are in the current working directory. The final line tells the server to run the script simulation.m in MATLAB which is located on Knot at the location/sw/bin/matlab. The additional commands -nodisplay -nodesktop ensures that MATLAB does not attempt to display anything since we are not running a GUI.

The easiest way to use parallel processing in MATLAB is to parallelize for loops. The Knot cluster is currently running the R2012b distribution of MATLAB. For initiating parallel processing, this version uses the MATLAB command matlabpool, rather than the updated parpool command. matlabpool can use up to all of the cores in a given computing node in parallel, but cannot parallelize across different nodes. If the EasySpin package is used in the script simulation.m, the necessary EasySpin functions must be

saved in the same directory as `simulation.m`.

The job `run.job` is submitted to the queue using the command

```
qsub run.job
```

Short jobs that take less than one hour to run can be prioritized by the command

```
qsub -q short run.job
```

The status of jobs, in various levels of detail can be checked by the commands

`qstat` prints a detailed report of all jobs running

`qstat -u username` prints a detailed report of jobs submitted by `username`

`showq` prints a less detailed report of all jobs running

A job can be canceled with the command

```
qdel job_id
```

where `job_id` is an unique identifying code assigned to each job submitted that can be found by looking at the status of the job using one of the above commands.

Bibliography

[1] Elwy H Abdelkader, Michael D Lee, Akiva Feintuch, Marie Ramirez Cohen, James D Swarbrick, Gottfried Otting, Bim Graham, and Daniella Goldfarb. A New Gd3+ Spin Label for Gd3+-Gd3+ Distance Measurements in Proteins Produces Narrow Distance Distributions. *Journal of Physical Chemistry Letters*, 6(24):5016–5021, 2015.

[2] A. Abragam. *The Principles of Nuclear Magnetism*. Oxford University Press, 1961.

[3] A. Abragam and B. Bleaney. *Electron paramagnetic resonance of transition ions*. Clarendon Press, 1970.

[4] Sara Ahlgren, Anna Orlova, Daniel Rosik, Mattias Sandström, Anna Sjöberg, Barbro Baastrup, Olof Widmark, Gunilla Fant, Joachim Feldwisch, and Vladimir Tolmachev. Evaluation of maleimide derivative of DOTA for site-specific labeling of recombinant affibody molecules. *Bioconjugate Chemistry*, 19(1):235–243, 2008.

[5] S. Aime, P.L. Anelli, M. Botta, F. Fedeli, M. Grandi, P. Paoli, and F. Uggeri. Gd(III) Complexes of DOTA-like Ligand Derivatives. *Inorganic Chemistry*, 31(12):2422–2428, 1992.

[6] D. Akhmetzyanov, P. Schöps, A. Marko, N. C. Kunjir, S. Th. Sigurdsson, and T. F. Prisner. Pulsed EPR dipolar spectroscopy at Q- and G-band on a trityl biradical. *Phys. Chem. Chem. Phys.*, 17(37):24446–24451, 2015.

[7] Christian Altenbach. LabVIEW programs for the analysis of EPR Data: ShortDistances.

[8] Christian Altenbach, Sabine L. Flitsch, H. Gobind Khorana, and Wayne L. Hubbell. Structural studies on transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry*, 28(19):7806–7812, 1989.

[9] Christian Altenbach, Ana Karin Kusnetzow, Oliver P. Ernst, Klaus Peter Hofmann, and Wayne L. Hubbell. High-resolution distance mapping in rhodopsin reveals the

pattern of helix movement due to activation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(21):7439–7444, 2008.

[10] Christian Altenbach, Thomas Marti, H. G. Khorana, and Wayne L. Hubbell. Transmembrane protein structure: spin labeling of bacteriorhodopsin mutants. *Science*, 248(4959):1088–1092, 1990.

[11] Christian Altenbach, K. J. Oh, René J. Trabanino, Kálmán Hideg, and Wayne L. Hubbell. Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: Experimental strategies and practical limitations. *Biochemistry*, 40(51):15471–15482, 2001.

[12] Andrei V. Astashkin and Arnold M. Raitsimring. Electron spin echo envelope modulation theory for high electron spin systems in weak crystal field. *The Journal of Chemical Physics*, 117(13):6121–6132, 2002.

[13] F. P. Auteri, A. H. Beth, and B. H. Robinson. An automated method to find the proper modulation phase in continuous wave EPR spectroscopy. *J. Magn. Reson.*, 80(3):493–501, 1988.

[14] Michael Balcewicz. *Production of Field Standards for Use in EPR.* PhD thesis, University of California, Santa Barbara, 2017.

[15] Debamalya Banerjee, Hiromasa Yagi, Thomas Huber, Gottfried Otting, and Daniella Goldfarb. Nanometer-Range Distance Measurement in a Protein Using Mn 2+ Tags. *J. Phys. Chem. Lett.*, 3:157160, 2012.

[16] J.E. Banham, C.M. Baker, S. Ceola, I.J. Day, G.H. Grant, E.J.J. Groenen, C.T. Rodgers, G. Jeschke, and C.R. Timmel. Distance measurements in the borderline region of applicability of CW EPR and DEER: A model study on a homologous series of spin-labelled peptides. *J. Magn. Reson.*, 191(2):202–218, 2008.

[17] Alessandro Barge, Giancarlo Cravotto, Eliana Gianolio, and Franco Fedeli. How to determine free Gd and free ligand in solution of Gd chelates. A technical note. *Contrast media & molecular imaging*, 1(5):184–188, 2006.

[18] O. Béjà, E. N. Spudich, J. L. Spudich, M. Leclerc, and E. F. DeLong. Proteorhodopsin phototrophy in the ocean. *Nature*, 411(6839):786–789, 2001.

[19] Oded Béjà, L. Aravind, V. Koonin, Marcelino T. Suzuki, Andrew Hadd, Linh P. Nguyen, Steven B. Jovanovich, Christina M. Gates, Robert A. Feldman, John L. Spudich, Elena N. Spudich, and Edward F. DeLong. Bacterial Rhodopsin: Evidence for a New Type of Phototrophy in the Sea. *Science*, 289(5486):1902–1906, 2000.

[20] F Benetollo, G Bombieri, L Calabi, S Aime, and M Botta. Structural Variations Across the Lanthanide Series of Macrocyclic DOTA Complexes: Insights into the

Design of Contrast Agents for Magnetic Resonance Imaging. *Inorg. Chem.*, 19:1319, 2003.

[21] Meriem Benmelouka, Johan Van Tol, Alain Borel, Saritha Nellutla, Marc Port, Lothar Helm, Louis Claude Brunel, and André E. Merbach. Multiple-frequency and variable-temperature EPR study of gadolinium(III) complexes with polyaminocarboxylates: Analysis and comparison of the magnetically dilute powder and the frozen-solution spectra. *Helvetica Chimica Acta*, 92(11):2173–2185, 2009.

[22] Meriem Benmelouka, Johan Van Tol, Alain Borel, Marc Port, Lothar Helm, Louis Claude Brunel, and André E. Merbach. A High-Frequency EPR Study of Frozen Solutions of Gd(III) Complexes: Straightforward Determination of the Zero-Field Splitting Parameters and Simulation of the NMRD Profiles. *J. Am. Chem. Soc.*, 128:7807–7816, 2006.

[23] Antonio Bianchi, Luisella Calabi, Claudia Giorgi, Pietro Losi, Palma Mariani, Paola Paoli, Patrizia Rossi, Barbara Valtancoli, and Mario Virtuani. Thermodynamic and structural properties of Gd 3+ complexes with functionalized macrocyclic ligands based upon 1,4,7,10- tetraazacyclododecane. *J. Chem. Soc. Dalton Trans*, (5):697–705, 2000.

[24] S. W Annie Bligh, Nick Choi, Evagoras G. Evagorou, Mary Mcpartlin, W. Jon Cummins, and J. Duncan Kelly. Synthesis and crystal structure of a gadolinium(III) complex of a tetraimine schiff-base macrocycle: A potential contrast agent for magnetic resonance imaging. *Polyhedron*, 11(19):2571–2573, 1992.

[25] F. Bloch. Nuclear induction. *Physical Review*, 70(7-8):460–474, 1946.

[26] H. Blok, J. A. J. M. Disselhorst, S. B. Orlinskii, and J. Schmidt. A continuous-wave and pulsed electron spin resonance spectrometer operating at 275GHz. *Journal of Magnetic Resonance*, 166(1):92–99, 2004.

[27] G Bombieri, N Marchini, S Ciattini, A Mortillaro, and S Aime. The crystallized solvent could influence the lanthanide water bonding? *Inorganica Chimica Acta*, 359(10):3405–3411, 2006.

[28] P. P. Borbat, A. J. Costa-Filho, K. A. Earle, J. K. Moscicki, and J. H. Freed. Electron spin resonance in studies of membranes and proteins. *Science*, 291(5502):266–269, 2001.

[29] Peter P. Borbat and Jack H. Freed. Pulse Dipolar Electron Spin Resonance: Distance Measurements. In Cristiane R. Timmel and Jeffrey R. Harmer, editors, *Struct. Bond.*, volume 152, pages 1–82. Springer Berlin Heidelberg, 2014.

[30] Alain Borel, Hoon Kang, Christelle Gateau, Marinella Mazzanti, R. B. Clarkson, and R. Linn Belford. Variable Temperature and EPR Frequency Study of Two

Aqueous Gd(III) Complexes with Unprecedented Sharp Lines. *J. Phys. Chem. A*, 110:12434–12438, 2006.

[31] Alain Borel, Sabrina Laus, Andrzej Ozarowski, Christelle Gateau, Aline Nonat, Marinella Mazzanti, and Lothar Helm. Multiple-frequency EPR spectra of two aqueous Gd 3+ polyamino polypyridine carboxylate complexes: A study of high field effects. *Journal of Physical Chemistry A*, 111(25):5399–5407, 2007.

[32] Suzanne Brandon, Albert H. Beth, and Eric J. Hustedt. The global analysis of DEER data. *Journal of Magnetic Resonance*, 218:93–104, 2012.

[33] Curtis F. Brewer and John P. Riehm. Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins. *Analytical Biochemistry*, 18(2):248–255, 1967.

[34] H. A. Buckmaster and Y. H. Shing. A survey of the EPR spectra of Gd3+ in single crystals. *Physica Status Solidi (a)*, 12(2):325–361, 1972.

[35] David E Budil and Keith A. Earle. Sample Resonators for Quasioptical EPR. In Oleg Grinberg and Lawrence J. Berliner, editors, *Very High Frequency (VHF) ESR/EPR*, chapter 11. Springer US, 2004.

[36] Paul T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy*. Clarendon Press, 1991.

[37] C. Allen Chang, Lynn C. Francesconi, Mary F. Malley, Krishan Kumar, Jack Z. Gougoutas, Michael F. Tweedle, Daniel W. Lee, and Lon J. Wilson. Synthesis, Characterization, and Crystal Structures of M(DO3A) (M = Fe, Gd) and NA[M(DOTA)] (M = Fe, Y, Gd). *Inorg. Chem*, 32:3501–3508, 1993.

[38] Franklin H. Cho, Viktor Stepanov, Chathuranga Abeywardana, and Susumu Takahashi. 230/115 GHz Electron Paramagnetic Resonance/Double Electron-Electron Resonance Spectroscopy. In *Methods in Enzymology*, volume 563, chapter 5, pages 95–118. 2015.

[39] R. B. Clarkson, Alex I. Smirnov, T. I. Smirnova, H. Kang, R. L. Belford, K. Earle, and Jack H Freed. Multi-frequency EP R determination of zero field splitting of high spin species in liquids: Gd(III) chelates in water. *Molecular Physics*, 95(6):1325–1332, 1998.

[40] Jessica A Clayton, Katharina Keller, Mian Qi, Julia Wegner, Vanessa Koch, Henrik Hintz, Adelheid Godt, Songi Han, Gunnar Jeschke, Mark S. Sherwin, and Maxim Yulikov. Quantitative analysis of zero-field splittings in Gd(III) complexes. *in preparation*.

[41] Jessica A. Clayton, Mian Qi, Adelheid Godt, Daniella Goldfarb, Songi Han, and Mark S. Sherwin. Gd3+-Gd3+ distances exceeding 3 nm determined by very high frequency continuous wave electron paramagnetic resonance. *Phys. Chem. Chem. Phys.*, 19:5127–5136, 2017.

[42] Paul A S Cruickshank, David R Bolton, Duncan A Robertson, Robert I Hunter, Richard J Wylde, and Graham M Smith. A kilowatt pulsed 94 GHz electron paramagnetic resonance spectrometer with high concentration sensitivity, high instantaneous bandwidth, and low dead time. *Rev. Sci. Instrum.*, 80(68):103102–2505, 2009.

[43] Timothy F. Cunningham, Miriam R. Putterman, Astha Desai, W. Seth Horne, and Sunil Saxena. The double-histidine Cu2+-binding motif: A highly rigid, site-specific spin probe for electron spin resonance distance measurements. *Angewandte Chemie - International Edition*, 54(21):6330–6334, 2015.

[44] Timothy F. Cunningham, Matthew D. Shannon, Miriam R. Putterman, Rajith J. Arachchige, Ishita Sengupta, Min Gao, Christopher P. Jaroniec, and Sunil Saxena. Cysteine-specific Cu2+ chelating tags used as paramagnetic probes in double electron electron resonance. *Journal of Physical Chemistry B*, 119(7):2839–2843, 2015.

[45] Arina Dalaloyan, Mian Qi, Sharon Ruthstein, Shimon Vega, Adelheid Godt, Akiva Feintuch, and Daniella Goldfarb. Gd(III) - Gd(III) EPR distance measurements - the range of accessible distances and the impact of zero field splitting. *Phys. Chem. Chem. Phys.*, 17(28):18464–18476, 2015.

[46] Yuanmu Deng, Ramasamy P. Pandian, Rizwan Ahmad, Periannan Kuppusamy, and Jay L. Zweier. Application of magnetic field over-modulation for improved EPR linewidth measurements using probes with Lorentzian lineshape. *Journal of Magnetic Resonance*, 181(2):254–261, 2006.

[47] Vasyl P. Denysenkov, Thomas F. Prisner, J. Stubbe, and Marina Bennati. High-Frequency 180 GHz PELDOR. *Appl. Magn. Reson.*, 29(2):375–384, 2005.

[48] Andrin Doll, Mian Qi, Stephan Pribitzer, Nino Wili, Maxim Yulikov, Adelheid Godt, and Gunnar Jeschke. Sensitivity enhancement by population transfer in Gd(III) spin labels. *Phys. Chem. Chem. Phys.*, 17(11):7334–7344, 2015.

[49] Wolfgang Doster. The protein-solvent glass transition. *Biochim. Biophys. Acta*, 1804(1):3–14, 2010.

[50] Gareth R. Eaton, Sandra S. Eaton, David P. Barr, and Ralph T. Weber. *Quantitative EPR*. Springer, Vienna, 2010.

[51] Devin T. Edwards. *High-Field EPR for Studies of Structure in Biological Systems*. PhD thesis, University of California, Santa Barbara, 2013.

[52] Devin T. Edwards, Thomas Huber, Sunyia Hussain, Katherine M. Stone, Maia Kinnebrew, Ilia Kaminker, Erez Matalon, Mark S. Sherwin, Daniella Goldfarb, and Songi Han. Determining the oligomeric structure of proteorhodopsin by Gd3+-based pulsed dipolar spectroscopy of multiple distances. *Structure*, 22(11):1677–86, 2014.

[53] Devin T. Edwards, Zhidong Ma, Thomas J. Meade, Daniella Goldfarb, Songi Han, and Mark S. Sherwin. Extending the distance range accessed with continuous wave EPR with Gd3+ spin probes at high magnetic fields. *Phys. Chem. Chem. Phys.*, 15:11313–11326, 2013.

[54] Devin T. Edwards, Yun Zhang, Steffen J. Glaser, Songi Han, and Mark S. Sherwin. Phase cycling with a 240 GHz, free electron laser-powered electron paramagnetic resonance spectrometer. *Phys. Chem. Chem. Phys.*, 15(15):5707–5719, 2013.

[55] Harold P. Erickson. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proced. Online*, 11:32–51, 2009.

[56] Akiva Feintuch, Gottfried Otting, and Daniella Goldfarb. Gd3+ Spin Labeling for Measuring Distances in Biomacromolecules: Why and How? In *Methods in Enzymology*, volume 563, chapter 16, pages 415–457. 2015.

[57] Alistair J. Fielding, Maria Grazia Concilio, Graham Heaven, and Michael A. Hollas. New developments in spin labels for pulsed dipolar EPR. *Molecules*, 19(10):16998–17025, 2014.

[58] R.A. Fields and C.A. Hutchison Jr. The determination of hydrogen coordinates in lanthanum nicotinate dihydrate crystals by neodymium(3+)-proton double resonance. *Journal of Chemical Physics*, 82:1771–1722, 1985.

[59] Thomas Friedrich, Sven Geibel, Rolf Kalmbach, Igor Chizhov, Kenichi Ataka, Joachim Heberle, Martin Engelhard, and Ernst Bamberg. Proteorhodopsin is a light-driven proton pump with variable vectoriality. *Journal of Molecular Biology*, 321(5):821–838, aug 2002.

[60] Luca Garbuio, Enrica Bordignon, Evan K. Brooks, Wayne L. Hubbell, Gunnar Jeschke, and Maxim Yulikov. Orthogonal Spin Labeling and Gd(III)-Nitroxide Distance Measurements on Bacteriophage T4-Lysozyme. *J. Phys. Chem. B*, 117(11):3145–3153, 2013.

[61] Luca Garbuio, Kaspar Zimmermann, Daniel Häussinger, and Maxim Yulikov. Gd(III) complexes for electron-electron dipolar spectroscopy: Effects of deuteration, pH and zero field splitting. *Journal of Magnetic Resonance*, 259:163–173, 2015.

[62] Christelle Gateau, Marinella Mazzanti, Jacques Pécaut, Frank A. Dunand, and Lothar Helm. Solid-state and solution properties of the lanthanide complexes of a new nonadentate tripodal ligand derived from 1,4,7-triazacyclononane. *Dalton Trans.*, (12):2428–2433, 2003.

[63] Harishchandra Ghimire, Robert M Mccarrick, David E Budil, and Gary A Lorigan. Significantly improved sensitivity of Q-Band PELDOR/DEER experiments relative to X-Band is observed in measuring the intercoil distance of a Leucine zipper motif peptide (GCN4-LZ). *Biochemistry*, 48(25):5782–5784, 2009.

[64] Adelheid Godt, Miriam Schulte, Herbert Zimmermann, and Gunnar Jeschke. How Flexible Are Poly(para-phenyleneethynylene)s? *Angew. Chem. Int. Ed.*, 45(45):7560–7564, 2006.

[65] Daniella Goldfarb. Gd3+ spin labeling for distance measurements by pulse EPR spectroscopy. *Phys. Chem. Chem. Phys.*, 16(16):9669–10234, 2014.

[66] Daniella Goldfarb, Yaakov Lipkin, Alexey Potapov, Yehoshua Gorodetsky, Boris Epel, Arnold M. Raitsimring, Marina Radoul, and Ilia Kaminker. HYSCORE and DEER with an upgraded 95 GHz pulse EPR spectrometer. *J. Magn. Reson.*, 194:8–15, 2008.

[67] J. A. T. Gonzalez, M. P. Longinotti, and H. R. Corti. The viscosity of glycerol-water mixtures including the supercooled region. *J. Chem. Eng. Data*, 56(4):1397–1406, 2011.

[68] Michal Gordon-Grossman, Ilia Kaminker, Yana Gofman, Yechiel Shai, and Daniella Goldfarb. W-Band pulse EPR distance measurements in peptides using Gd 3+-dipicolinic acid derivatives as spin labels. *Phys. Chem. Chem. Phys.*, 13(13):10771–10780, 2011.

[69] O. Hayes Griffith and A. S. Waggoner. Nitroxide free radicals: spin labels for probing biomolecular structure. *Accounts of Chemical Research*, 2(1):17–24, 1969.

[70] I. Gromov, J. Shane, J. Forrer, R. Rakhmatoullin, Yu Rozentzwaig, and A. Schweiger. A Q-band pulse EPR/ENDOR spectrometer and the implementation of advanced one- and two-dimensional pulse EPR methodology. *Journal of Magnetic Resonance*, 149(2):196–203, 2001.

[71] Chidambaram Gunanathan, Yael Diskin-Posner, and David Milstein. Lanthanide-Organic Framework of a Rigid Bis-Gd Complex: Composed by Carbonate Ions Spacers. *Cryst. Growth Des.*, 10(10):4235–4239, 2010.

[72] M. A. Hemminga and L. J. Berliner. *ESR Spectroscopy in Membrane Biophysics*, volume 27 of *Biological Magnetic Resonance*. Springer US, Boston, MA, 2007.

[73] M. M. Hertel, V. P. Denysenkov, M. Bennati, and T. F. Prisner. Pulsed 180-GHz EPR/ENDOR/PELDOR spectroscopy, 2005.

[74] Wayne L. Hubbell and Christian Altenbach. Investigation of structure and dynamics in membrane proteins using site-directed spin labeling. *Curr. Opin. Struct. Bio.*, 4(4):566–573, jan 1994.

[75] Wayne L. Hubbell, David S. Cafisco, and Christian Altenbach. Identifying conformational changes with site-directed spin labeling. *Nat. Struct. Biol.*, 7:735–739, 2000.

[76] Wayne L. Hubbell, Adrian Gross, Ralf Langen, and Michael A. Lietzow. Recent advances in site-directed spin labeling of proteins. *Curr. Opin. Struct. Bio.*, 8(5):649–656, 1998.

[77] Sunyia Hussain. *Studies of proteorhodopsin to investigate transmembrane protein function and dynamics*. PhD thesis, University of California, Santa Barbara, 2014.

[78] Sunyia Hussain, John M. Franck, and Songi Han. Transmembrane protein activation refined by site-specific hydration dynamics. *Angewandte Chemie - International Edition*, 52(7):1953–1958, 2013.

[79] Sunyia Hussain, Maia Kinnebrew, Nicole S. Schonenbach, Emily Aye, and Songi Han. Functional consequences of the oligomeric assembly of proteorhodopsin. *Journal of Molecular Biology*, 427:1278–1290, 2015.

[80] Eric J. Hustedt and Albert H. Beth. Nitroxide spin-spin interactions: applications to protein structure and dynamics. *Ann. Rev. Biophysi. Biomol. Struct.*, 28(1):129–153, 1999.

[81] Matthew N. Idso, Naomi Baxter, Sirish Narayanan, Evelyn Chang, Julia Fisher, Bradley Chmelka, and Songi Han. Tuning the function of proteorhodopsin by oligomeric assembly in different micellar solutions. *in preparation*.

[82] Gunnar Jeschke and Yevhen Polyhach. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.*, 9(16):1895, 2007.

[83] Gunnar Jeschke, Muhammad Sajid, Miriam Schulte, and Adelheid Godt. Three-spin correlations in double electronelectron resonance. *Physical Chemistry Chemical Physics*, 11(31):6580, 2009.

[84] Gunnar Jeschke, Muhammad Sajid, Miriam Schulte, Navid Ramezanian, Aleksei Volkov, Herbert Zimmermann, and Adelheid Godt. Flexibility of shape-persistent molecular building blocks composed of p-Phenylene and Ethynylene units. *J. Am. Chem. Soc.*, 132:10107–10117, 2010.

[85] Matthias J. N. Junk, Hans W. Spiess, and Dariush Hinderberger. DEER in biological multispin-systems: A case study on the fatty acid binding to human serum albumin. *Journal of Magnetic Resonance*, 210(2):210–217, jun 2011.

[86] Katharina Keller, Michal Zalibera, Mian Qi, Vanessa Koch, Julia Wegner, Henrik Hintz, Adelheid Godt, Gunnar Jeschke, Anton Savitsky, and Maxim Yulikov. EPR characterization of Mn(II) complexes for distance determination with pulsed dipolar spectroscopy. *Phys. Chem. Chem. Phys.*, 16, 2016.

[87] Younggyu Kim, Sam O. Ho, Natalie R. Gassman, You Korlann, Elizabeth V. Landorf, Frank R. Collart, and Shimon Weiss. Efficient site-specific labeling of proteins via cysteines. *Bioconjugate Chemistry*, 19(3):786–791, 2008.

[88] Adriana L. Klyszejko, Sarika Shastri, Stefania A. Mari, Helmut Grubmüller, Daniel J. Muller, and Clemens Glaubitz. Folding and assembly of proteorhodopsin. *Journal of Molecular Biology*, 376(1):35–41, 2008.

[89] Mark S. Konings, William C. Dow, David B. Love, Kenneth N. Raymond, Steven C. Quay, and Scott M. Rocklage. Gadolinium complexation by a new DTPA-amide ligand. Amide oxygen coordination. *Inorganic Chemistry*, 29(8):1488–1491, 1990.

[90] J Krzystek, Andrew Ozarowski, and Joshua Telser. Multi-frequency, high-field EPR as a powerful tool to accurately determine zero-field splitting in high-spin transition metal coordination complexes. *Coordination Chemistry Reviews*, 250(17-18):2308–2324, 2006.

[91] Nitin C. Kunjir, Gunnar W. Reginsson, Olav Schiemann, and Snorri Th. Sigurdsson. Measurements of short distances between trityl spin labels with CW EPR, DQC and PELDOR. *Phys. Chem. Chem. Phys.*, 15(15):19673–19685, 2013.

[92] Jens O. Lagerstedt, Jitka Petrlova, Silvia Hilt, Antonin Marek, Youngran Chung, Renuka Sriram, Madhu S. Budamagunta, Jean F. Desreux, David Thonon, Thomas Jue, Alex I. Smirnov, and John C. Voss. EPR assessment of protein sites for incorporation of Gd(III) MRI contrast labels. *Contrast Media and Molecular Imaging*, 8(3):252–264, 2013.

[93] Aurélie Lasoroski, Rodolphe Vuilleumier, and Rodolphe Pollet. Vibrational dynamics of zero-field-splitting hamiltonian in gadolinium-based MRI contrast agents from ab initio molecular dynamics. *The Journal of Chemical Physics*, 141(1):014201, 2014.

[94] Mariane Le Fur, Maryline Beyler, Nicolas Lepareur, Olivier Fouge, Carlos Platas-Iglesias, Olivier Rousseaux, and Raphae L Tripier. Pyclen Trinbutylphosphonate ester as potential chelator for targeted radiotherapy: from Yttrium(III) complexation to 90Y radiolabeling. *Inorganic Chemistry*, 55(16):8003–8012, 2016.

[95] Malcolm H. Levitt. *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. Wiley, second edition, 2008.

[96] D.R. Lide, editor. *CRC Handbook of Chemistry and Physics*. CRC Press, 85 edition, 2004.

[97] Petra Lueders, Heidrun Jager, Marcus A Hemminga, Gunnar Jeschke, and Maxim Yulikov. Distance measurements on orthogonally spin-labeled membrane spanning WALP23 polypeptides. *Journal of Physical Chemistry B*, 117(7):2061–2068, 2013.

[98] Jakob Maciejko, Michaela Mehler, Jagdeep Kaur, Tobias Lieblein, Nina Morgner, Olivier Ouari, Paul Tordo, Johanna Becker-Baldus, and Clemens Glaubitz. Visualizing Specific Cross-Protomer Interactions in the Homo-Oligomeric Membrane Protein Proteorhodopsin by Dynamic-Nuclear-Polarization-Enhanced Solid-State NMR. *Journal of the American Chemical Society*, 137(28):9032–9043, 2015.

[99] M. Malferrari, A. Nalepa, G. Venturoli, F. Francia, W. Lubitz, K. Möbius, and A. Savitsky. Structural and dynamical characteristics of trehalose and sucrose matrices at different hydration levels as probed by FTIR and high-field EPR. *Physical Chemistry Chemical Physics*, 16(16):9831–9848, 2014.

[100] Dikla Man-Aharonovich, Gazalah Sabehi, Oleg A. Sineshchekov, Elena N. Spudich, John L. Spudich, and Oded Béjà. Characterization of RS29, a blue-green proteorhodopsin variant from the Red Sea. *Photochem. Photobiol. Sci.*, 3(5):459–462, 2004.

[101] Nurit Manukovsky, Akiva Feintuch, Ilya Kuprov, and Daniella Goldfarb. Time domain simulation of Gd3+-Gd3+ distance measurements by EPR. *The Journal of Chemical Physics*, 147(4):044201, 2017.

[102] Nurit Manukovsky, Veronica Frydman, and Daniella Goldfarb. Gd3+ Spin Labels Report the Conformation and Solvent Accessibility of Solution and Vesicle-Bound Melittin. *J. Phys. Chem. B*, 119:137732–13741, 2015.

[103] Andrea Martorana, Giuliano Bellapadrona, Akiva Feintuch, Enza Di Gregorio, Silvio Aime, and Daniella Goldfarb. Probing protein conformation in cells by EPR distance measurements using Gd3+ spin labeling. *J. Am. Chem. Soc.*, 136(38):13458–65, 2014.

[104] Andrea Martorana, Yin Yan, Yu Zhao, Qing-Feng Li, Xun-Cheng Su, and Daniella Goldfarb. Mn(II) tags for DEER distance measurements in proteins via C-S attachment. *Dalton Trans.*, 44:20812–20816, 2015.

[105] Florencia C. Mascali, H. Y. Vincent Ching, Rodolfo M. Rasia, Sun Un, and Leandro C. Tabares. Using genetically encodable self-assembling Gd(III) spin labels to make in-cell nanometric distance measurements. *Angew. Chem.*, 55:11041–11043, 2016.

[106] Erez Matalon, Thomas Huber, Gregor Hagelueken, Bim Graham, Veronica Frydman, Akiva Feintuch, Gottfried Otting, and Daniella Goldfarb. Gadolinium(III) spin labels for high-sensitivity distance measurements in transmembrane helices. *Angew. Chem. Int. Ed.*, 52(45):11831–11834, 2013.

[107] Michaela Mehler, Frank Scholz, Sandra J. Ullrich, Jiafei Mao, Markus Braun, Lynda J. Brown, Richard C. D. Brown, Sarah A. Fiedler, Johanna Becker-Baldus, Josef Wachtveitl, and Clemens Glaubitz. The EF loop in green proteorhodopsin affects conformation and photocycle dynamics. *Biophysical Journal*, 105(2):385–397, 2013.

[108] F. Mentink-Vigier, A. Collauto, A. Feintuch, I. Kaminker, V. Tarle, and D. Goldfarb. Increasing sensitivity of pulse EPR experiments using echo train detection schemes. *Journal of Magnetic Resonance*, 236:117–125, 2013.

[109] Virginia Meyer, Michael A. Swanson, Laura J. Clouston, Przemysław J. Boratyński, Richard A. Stein, Hassane S. Mchaourab, Andrzej Rajca, Sandra S. Eaton, and Gareth R. Eaton. Room-temperature distance measurements of immobilized spin-labeled protein by DEER/PELDOR. *Biophys. J.*, 108(5):1213–1219, 2015.

[110] Louise S. Natrajan, Ntai M. Khoabane, Benjamin L. Dadds, Christopher A. Muryn, Robin G. Pritchard, Sarah L. Heath, Alan M. Kenwright, Ilya Kuprov, and Stephen Faulkner. Probing the structure, conformation, and stereochemical exchange in a family of lanthanide complexes derived from tetrapyridyl-appended cyclen. *Inorganic Chemistry*, 49(17):7700–7709, 2010.

[111] D.J. Newman and W. Urban. Interpretation of S-state ion E.P.R. spectra. *Advances in Physics*, 24(6):793–844, 1975.

[112] Gerardo Palazzo, Antonia Mallardi, Alejandro Hochkoeppler, Lorenzo Cordone, and Giovanni Venturoli. Electron transfer kinetics in photosynthetic reaction centers embedded in trehalose glasses: trapping of conformational substates at room temperature. *Biophysical Journal*, 82:558–568, 2002.

[113] Eric F. Pettersen, Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13):1605–1612, oct 2004.

[114] Yevhen Polyhach, Enrica Bordignon, Sandhya Gandra, Adelheid Godt, and Gunnar Jeschke. High sensitivity and versatility of the DEER experiment on nitroxide

radical pairs at Q-band frequencies. *Phys. Chem. Chem. Phys.*, 14:10762–10773, 2012.

[115] Charles P Poole. *Electron Spin Resonance - A Comprehensive Treatise on Experimental Techniques*. Dover Publications, Inc., second edition, 1983.

[116] A.M. Portis. Rapid passage effects in electron spin resonance. *Physical Review*, 100(4):1219–1221, 1955.

[117] Alexey Potapov, Hiromasa Yagi, Thomas Huber, Slobodan Jergic, Nicholas E Dixon, Gottfried Otting, and Daniella Goldfarb. Nanometer-scale distance measurements in proteins using Gd3+ spin labeling. *J. Am. Chem. Soc.*, 132:9040–9048, 2010.

[118] Mian Qi, Andreas Groß, Gunnar Jeschke, Adelheid Godt, and Malte Drescher. Gd(III)-PyMTA label is suitable for in-cell EPR. *J. Am. Chem. Soc.*, 136:15366–15378, 2014.

[119] Mian Qi, Miriam Hulsmann, and Adelheid Godt. Spacers for geometrically well-defined water-soluble molecular rulers and their application. *J. Org. Chem.*, 81(6):2549–2571, 2016.

[120] Peter Z. Qin and Kurt Warncke, editors. *Methods in Enzymology: Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Part A, Volume 563*. Elsevier, 2015.

[121] Peter Z. Qin and Kurt Warncke, editors. *Methods in Enzymology: Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions Part B, Volume 564*. Elsevier, 2015.

[122] Mark D Rabenstein and Yeon-Kyun Shin. Determination of the distance between two spin labels attached to a macromolecule. *Biophysics*, 92:8239–8243, 1995.

[123] Nicole Radzwill, Klaus Gerwert, and H.-J. Steinhoff. Time-resolved detection of transient movement of helices F and G in doubly spin-labeled bacteriorhodopsin. *Biophys. J.*, 80(6):2856–66, 2001.

[124] A Raitsimring, A Dalaloyan, A Collauto, A Feintuch, T Meade, and D Goldfarb. Zero field splitting fluctuations induced phase relaxation of Gd3+ in frozen solutions at cryogenic temperatures. *Journal of Magnetic Resonance*, 248:71–80, 2014.

[125] A. M. Raitsimring, A. V. Astashkin, O. G. Poluektov, and P. Caravan. High-Field Pulsed EPR and ENDOR of Gd3+ Complexes in Glassy Solutions. *Appl. Magn. Reson.*, 28:281–295, 2005.

[126] Arnold M. Raitsimring, Andrei V. Astashkin, and Peter Caravan. High-frequency EPR and ENDOR characterization of MRI contrast agents. In Lawrence Berliner and Graeme Hanson, editors, *Biological Magnetic Resonance*, chapter 14, pages 581–621. Springer, New York, 2009.

[127] Arnold M. Raitsimring, Chidambaram Gunanathan, Alexey Potapov, Irena Efremenko, Jan M. L. Martin, David Milstein, and Daniella Goldfarb. Gd3+ complexes as potential spin labels for high field pulsed EPR distance measurements. *Journal of the American Chemical Society*, 129(46):14138–14139, 2007.

[128] Marie Ramirez-Cohen, Veronica Frydman, Petr Milko, Mark A. Iron, Elwy H. Abdelkader, Michael D. Lee, James D. Swarbrick, Arnold Raitsimring, Gottfried Otting, Bim Graham, Akiva Feintuch, and Daniella Goldfarb. Overcoming artificial broadening in Gd3+-Gd3+ distance distributions arising from dipolar pseudosecular terms in DEER experiments. *Phys. Chem. Chem. Phys.*, 18(18):12847–12859, 2016.

[129] Tingting Ran, Gabriel Ozorowski, Yanyan Gao, Oleg A. Sineshchekov, Weiwu Wang, John L. Spudich, and Hartmut Luecke. Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. *Acta Crystallographica Section D: Biological Crystallography*, 69(10):1965–1980, 2013.

[130] S. Rast, P.H. Fries, E. Belorizky, A. Borel, L. Helm, and A. E. Merbach. A general approach to the electronic spin relaxation of Gd(III) complexes in solutions. Monte Carlo simulations beyond the Redfield limit. *The Journal of Chemical Physics*, 115(16):7554–7563, 2001.

[131] Sahand Razzaghi, Mian Qi, Anna I. Nalepa, Adelheid Godt, Gunnar Jeschke, Anton Savitsky, and Maxim Yulikov. RIDME Spectroscopy with Gd(III) Centers. *J. Phys. Chem. Lett.*, 5:3970–3975, 2014.

[132] Sina Reckel, Daniel Gottstein, Jochen Stehle, Frank Löhr, Mirka Kristin Verhoefen, Mitsuhiro Takeda, Robert Silvers, Masatsune Kainosho, Clemens Glaubitz, Josef Wachtveitl, Frank Bernhard, Harald Schwalbe, Peter Güntert, and Volker Dötsch. Solution NMR structure of proteorhodopsin. *Angew. Chem. Int. Ed.*, 50(50):11942–11946, 2011.

[133] Gunnar W. Reginsson, Nitin C. Kunjir, Snorri Th. Sigurdsson, and Olav Schiemann. Trityl radicals: spin labels for nanometer-distance measurements. *Chemistry - A European Journal*, 18(43):13580–13584, 2012.

[134] M. Rohrer, O. Brügmann, B. Kinzer, and T. F. Prisner. High-field/high-frequency EPR spectrometer operating in pulsed and continuous-wave mode at 180 GHz. *Applied Magnetic Resonance*, 21:257–274, 2001.

[135] P. Roser, M. J. Schmidt, M. Drescher, and D. Summerer. Site-directed spin labeling of proteins for distance measurements in vitro and in cells. *Org. Biomol. Chem.*, 14(24):5468–5476, 2016.

[136] Ajit Kumar Saha and Tara Prasad Das. *Nuclear Induction*. Saha Institute of Nuclear Physics, Calcutta, 1957.

[137] Olav Schiemann and Thomas F. Prisner. Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Q. Rev. Biophys.*, 40(01):1, 2007.

[138] Arthur Schweiger and Gunnar Jeschke. *Principles of pulse electron paramagnetic resonance*. Oxford University Press, 2001.

[139] K. Ilker Sen, Timothy M. Logan, and Piotr G. Fajer. Protein dynamics and monomer-monomer interactions in AntR activation by electron paramagnetic resonance and double electron-electron resonance. *Biochemistry*, 46(41):11639–11649, 2007.

[140] Sandip A. Shelke and Snorri Th. Sigurdsson. Site-Directed Nitroxide Spin Labeling of Biopolymers. In Christiane R Timmel and Jeffrey R Harmer, editors, *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centres in the Biosciences*, pages 121–162. Springer Berlin Heidelberg, 2011.

[141] Georgiy Yu. Shevelev, Olesya A. Krumkacheva, Alexander A. Lomzov, Andrey A. Kuzhelev, Olga Yu. Rogozhnikova, Dmitry V. Trukhin, Tatiana I. Troitskaya, Victor M. Tormyshev, Matvey V. Fedin, Dmitrii V. Pyshnyi, and Elena G. Bagryanskaya. Physiological-Temperature Distance Measurement in Nucleic Acid using Triarylmethyl-Based Spin Labels and Pulsed Dipolar EPR Spectroscopy. *J. Am. Chem. Soc.*, 136:9874–9877, 2014.

[142] Georgiy Yu. Shevelev, Olesya A. Krumkacheva, Alexander A. Lomzov, Andrey A. Kuzhelev, Dmitry V. Trukhin, Olga Yu. Rogozhnikova, Victor M. Tormyshev, Dmitrii V. Pyshnyi, Matvey V. Fedin, and Elena G. Bagryanskaya. Triarylmethyl Labels: Toward Improving the Accuracy of EPR Nanoscale Distance Measurements in DNAs. *J. Phys. Chem. B*, 8:13641–13648, 2015.

[143] Lichi Shi, Mumdooh A M Ahmed, Wurong Zhang, Gregg Whited, Leonid S Brown, and Vladimir Ladizhansky. Three-Dimensional Solid-State NMR Study of a Seven-Helical Integral Membrane Proton Pump-Structural Insights. *Journal of Molecular Biology*, 386(4):1078–1093, 2009.

[144] Lichi Shi, Evelyn M R Lake, Mumdooh A M Ahmed, Leonid S Brown, and Vladimir Ladizhansky. Solid-state NMR study of proteorhodopsin in the lipid environment: Secondary structure and dynamics. *Biochimica et Biophysica Acta - Biomembranes*, 1788(12):2563–2574, 2009.

[145] Alexandra Simperler, Andreas Kornherr, Reenu Chopra, William Jones, W D Samuel Motherwell, and Gerhard Zifferer. The glass transition temperatures of amorphous trehalosewater mixtures and the mobility of water: an experimental and in silico study. *Carbohydr. Res.*, 342:1470–1479, 2007.

[146] Charles P. Slichter. *Principles of Magnetic Resonance*. Harper & Row Publishers, New York, 1963.

[147] G. M. Smith, J. C. G. Lesurf, R. H. Mitchell, and P. C. Riedi. Quasi-optical cw mm-wave electron spin resonance spectrometer. *Rev. Sci. Instrum.*, 69(11):3924, 1998.

[148] Likai Song, Zhanglong Liu, Pavanjeet Kaur, Jackie M Esquiaqui, Robert I Hunter, Stephen Hill, Graham M Smith, and Gail E Fanucci. Toward increased concentration sensitivity for continuous wave EPR investigations of spin-labeled biological macromolecules at high fields. *J. Magn. Reson.*, 265:188–196, 2016.

[149] Y. Song, T.J. Meade, A.V. Astashkin, E.L. Klein, J.H. Enemark, and A. Raitsimring. Pulsed dipolar spectroscopy distance measurements in biomacromolecules labeled with Gd(III) markers. *J. Magn. Reson.*, 210(1):59–68, 2011.

[150] Richard A Stein, Albert H Beth, and Eric J Hustedt. A straightforward approach to the analysis of double electron-electron resonance data. In *Methods in Enzymology*, volume 563, pages 531–567. 2015.

[151] H-J Steinhoff. Inter-and intra-molecular distances determined by EPR spectroscopy and site-directed spin labeling reveal protein- protein and protein-oligonucleotide interaction. *Biol. Chem.*, 385:913–920, 2004.

[152] Heinz-Jürgen Steinhoff. Methods for study of protein dynamics and protein-protein interaction in protein-ubiquitination by electron paramagnetic resonance spectroscopy. *Frontiers in Bioscience*, 7:97–110, 2002.

[153] Heinz-Jurgen Steinhoff, Nicole Radzwill, Wilhelm Thevis, Volker Lenz, Dietrich Brandenburg, Alfred Antson, Guy Dodson, and Axel Wollmeri. Determination of Interspin Distances between Spin Labels Attached to Insulin: Comparison of Electron Paramagnetic Resonance Data with the X-Ray Structure. *Biophys. J.*, 73:3287–3298, 1997.

[154] Viktor Stepanov and Susumu Takahashi. Determination of nitrogen spin concentration in diamond using double electron-electron resonance. *Physical Review B*, 94(2):024421, jul 2016.

[155] Stefan Stoll and Arthur Schweiger. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.*, 178(1):42–55, 2006.

[156] Katherine M Stone, Jeda Voska, Maia Kinnebrew, Anna Pavlova, Matthias J N Junk, and Songi Han. Structural insight into proteorhodopsin oligomers. *Biophysical Journal*, 104(2):472–481, 2013.

[157] T J Stone, T Buckman, P L Nordiot, and H M Mcconnell. Spin-Labeled Biomolecules. *Proc. Nat. Acad. Sci. U.S.A.*, 54(4):1010–1017, 1965.

[158] James W. Stoner, Dennis Szymanski, Sandra S. Eaton, Richard W. Quine, George A. Rinard, and Gareth R. Eaton. Direct-detected rapid-scan EPR at 250 MHz. *Journal of Magnetic Resonance*, 170(1):127–135, sep 2004.

[159] S Takahashi, L.-C Brunel, D T Edwards, J Van Tol, G Ramian, S Han, and M S Sherwin. Pulsed electron paramagnetic resonance spectroscopy powered by a free-electron laser. *Nature*, 489:409–413, 2012.

[160] John R. Taylor. *An Introduction to Error Analysis: the study of uncertainties in physical measurements*. University Science Books, 1997.

[161] L. K. Templeton, D. H. Templeton, A. Zalkin, and H. W. Ruben. Anomalous scattering by praseodymium, samarium and gadolinium and structures of their ethylenediaminetetraacetate (edta) salts. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 38(8):2155–2159, aug 1982.

[162] Francois-Xavier Theillet, Andres Binolfi, Beata Bekei, Andrea Martorana, Honor May Rose, Marchel Stuiver, Silvia Verzini, Dorothea Lorenz, Marleen van Rossum, Daniella Goldfarb, and Philipp Selenko. Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature*, 530(7588):45–50, feb 2016.

[163] Igor Tkach, Karin Halbmair, Claudia Höbartner, and Marina Bennati. High-frequency 263 GHz PELDOR. *Appl. Magn. Reson.*, 45(10):969–979, oct 2014.

[164] René Tschaggelar, Besnik Kasumaj, Maria Grazia Santangelo, Jörg Forrer, Patrik Leger, Henry Dube, François Diederich, Jeffrey Harmer, Rolf Schuhmann, Inés García-Rubio, and Gunnar Jeschke. Cryogenic 35 GHz pulse ENDOR probehead accommodating large sample sizes: Performance and applications. *Journal of Magnetic Resonance*, 200:81–87, 2009.

[165] Yury D. Tsvetkov, A.D. Milov, and A.G. Maryasov. Pulsed electron-electron double resonance (PELDOR) as EPR spectroscopy in nanometre range. *Russ. Chem. Rev.*, 77(6):487, 2008.

[166] Uptima. TCEP Product Description, Technical, and Scientific Information. Technical report.

[167] J Van Tol, L.-C Brunel, and R J Wylde. A quasioptical transient electron spin resonance spectrometer operating at 120 and 240 GHz. *Rev. Sci. Instrum.*, 76(68), 2005.

[168] J H Van Vleck. The dipolar broadening of magnetic resonance lines in crystals. *Physical Review*, 74(9):1168–1183, 1948.

[169] Wei Wu Wang, Oleg A. Sineshchekov, Elena N. Spudich, and John L. Spudich. Spectroscopic and Photochemical Characterization of a Deep Ocean Proteorhodopsin. *Journal of Biological Chemistry*, 278(36):33985–33991, 2003.

[170] Meaghan E Ward, Lichi Shi, Evelyn Lake, Sridevi Krishnamurthy, Howard Hutchins, Leonid S Brown, and Vladimir Ladizhansky. Proton-detected solid-state NMR reveals intramembrane polar networks in a seven-helical transmembrane protein proteorhodopsin. *Journal of the American Chemical Society*, 133(43):17434–17443, 2011.

[171] M. Weger. Passage Effects in Paramagnetic Resonance Experiments. *Bell System Technical Journal*, 39(4):1013–1112, jul 1960.

[172] John A. Weil and James R. Bolton. *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*. Wiley, second edition, 2007.

[173] Adarshi Welegedara, Yin Yang, Michael Lee, James Swarbrick, Thomas Huber, Bim Graham, Daniella Goldfarb, and Gottfried Otting. Double-Arm Lanthanide Tags Deliver Narrow Gd3+-Gd3+ Distance Distributions in DEER Measurements. *Chemistry - A European Journal*, jul 2017.

[174] Jan Wilhelm and Erwin Frey. Radial Distribution Function of Semiflexible Polymers. *Physical Review Letters*, 77(12):2581–2584, 1996.

[175] Hiromasa Yagi, Debamalya Banerjee, Bim Graham, Thomas Huber, Daniella Goldfarb, and Gottfried Otting. Gadolinium Tagging for High-Precision Measurements of 6 nm Distances in Protein Assemblies by EPR. *J. Am. Chem. Soc.*, 133:10418–10421, 2011.

[176] Yin Yang, Feng Yang, Yan Jun Gong, Jia Liang Chen, Daniella Goldfarb, and Xun Cheng Su. A Reactive, Rigid GdIII Labeling Tag for In-Cell EPR Distance Measurements in Proteins. *Angewandte Chemie - International Edition*, 56(11):2914–2918, feb 2017.

[177] Zhongyu Yang, Ming Ji, Timothy F Cunningham, and Sunil Saxena. Cu2 + as an ESR Probe of Protein Structure and Function. In *Methods in Enzymology*, volume 563, pages 459–481. 2015.

[178] Maxim Yulikov. Spectroscopically orthogonal spin labels and distance measurements in biomolecules. In B.C. Gilbert, V. Chechik, and D.M. Murphy, editors, *Electron Paramagnetic Resonance: Volume 24*, chapter 1, pages 1–31. 2015.

[179] Maxim Yulikov, Petra Lueders, Muhammad Farooq Warsi, Victor Chechik, and Gunnar Jeschke. Distance measurements in Au nanoparticles functionalized with nitroxide radicals and Gd3+DTPA chelate complexes. *Physical Chemistry Chemical Physics*, 14(30):10732, 2012.