Investigating the functional dynamics of membrane transport and associated proteins using advanced magnetic resonance spectroscopy

Anna Mullen

A thesis submitted for the degree of Doctor of Philosophy

Norwich Medical School & School of Chemistry

University of East Anglia

January 2021

Declaration

I certify that the work contained in this thesis submitted, by me, for the degree of Doctor of Philosophy, is my own original work, except where due reference is made to other authors or co-workers and has not previously been submitted, by me, for a degree at this or any other university.

Anna Mullen,

January 2021.

Acknowledgements

The course of a PhD is never a simple one, and my journey over the last five years proves that once again to be true. Through the years I've had some of my highest highs and lowest lows, and I would like to thank all of the people who have supported me throughout both the challenges and successes.

Firstly, to the academics and institutions which have allowed me to undertake this body of work through provision of financial support and infrastructure: the University of East Anglia (particularly the Faculty of Science, Norwich Medical School and the groups of Dr Fraser MacMillan and Prof Changjiang Dong), the Wellcome Trust, the Biochemical Society and EU-COST Action CM1306. Special thanks should be extended to the COST Action and its organisers for facilitating the multitude of conferences, workshops and seminars over the years.

Next, to the numerous collaborators who have inspired and helped to conceive or develop many of the ideas presented herein, provided material and support, and answered countless questions: Prof Bert Poolman and his group, as well as Dr Albert Guskov, at the University of Groningen; Prof Thorben Cordes and his group at the Ludwig-Maximilians-University Munich; Dr Christopher McDevitt at the Research Centre for Infectious Diseases, University of Adelaide; Dr Chris Mulligan and his group at the University of Kent; Dr Christian Teutloff at the Free University of Berlin; Dr Jess van Wonderen at the University of East Anglia.

The University of East Anglia has been my home since 2011; there are countless people who have impacted on my experiences here, and though most have since left this little corner of the country (some even prior to me embarking on my PhD studies), I would still like to express my gratitude to them. Some friends that I must name here and thank individually are Drs or soon-to-be Drs Michael O'Sullivan, Sarah Woodhouse and Alex Morritt. Your humour, solidarity and friendship have been truly invaluable to me. Additionally, I thank the many project students I have had the privilege of mentoring, supervising, or simply sharing a lab with over the years. Special recognition should go to Jenny Hall, one of my first project students, who then became a fantastic labmate and fast friend.

Finally, I thank my family, whose love and support has been steadfast through years of bemused conversations and attempts to explain what exactly it was I was doing. And yes, ***now*** you can call me 'Dr'.

Table of contents

Glossary of abbreviations, symbols and terms

Å	Ångstrom (= 0.1 nm)
ABC	ATP-Binding Cassette
APS	Ammonium persulfate
B_0	Applied magnetic field
cw	Continuous wave
DASS	Divalent Anion Sodium Symporter
DDM	n-Dodecyl β-D-maltoside
DEER	Double Electron-Electron Resonance
DMSO	Dimethyl sulfoxide
DPPH	2,2-diphenyl-1-picrylhydrazyl radical
DQC	Double Quantum Coherence
DTT	Dithiothreitol
EDNMR	ELDOR-Detected NMR
ENDOR	Electron-Nuclear DOuble Resonance
EPR	Electron Paramagnetic Resonance
ESE	Electron Spin Echo
ESEEM	Electron Spin Echo Envelope Modulation
FID	Free Induction Decay
FRET	Förster (or Fluorescence) Resonance Energy Transfer
FSE	Field-Swept Echo
Gdn.HCl	Guanidinium hydrochloride
HF	Hyperfine interaction
HFC	Hyperfine coupling
HYSCORE	HYperfine Sub-level CORrElation
IAP	3-(3-(2-iodo-acetamido)-propyl-carbamoyl)-PROXYL
IMAC	Immobilised Metal Affinity Chromatography
KP_i	Potassium phosphate
LB	Lysogeny broth
MMM	Multiscale Modelling of Macromolecular systems
MTSL	*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
PAGE	Polyacrylamide Gel Electrophoresis

PDB	Protein Data Bank
PELDOR	Pulsed Electron-eLectron DOuble Resonance
PMSF	Phenylmethylsulfonyl fluoride
PRE	Paramagnetic Relaxation Enhancement
PROXYL	2,2,5,5-tetramethylpyrrolidinyl-1-oxyl radical
RDC	Residual Dipolar Coupling
RIDME	Relaxation Induced Dipole Modulation Enhancement
RMSD	Root Mean Squared Deviation
SBD/SBP	Substrate Binding Domain/Substrate Binding Protein
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SDSL	Site-Directed Spin Labelling
SLC	Solute Carrier
sm	Single molecule
STD-NMR	Saturation Transfer Difference Nuclear Magnetic Resonance
T_1	Spin lattice (longitudinal) relaxation
T_2	Spin spin (transverse) relaxation
TEMED	Tetramethylethylenediamine
TEMPO	2,2,6,6-tetramethylpiperidinyl-1-oxyl radical
ZFS	Zero Field Splitting

List of figures

List of tables

Abstract

Protein structures were originally determined by X-ray crystallography, a technique which uses a regularly arranged lattice to measure the electron density of a protein in three-dimensional space. NMR later emerged as an alternative for solving solution-based structures but is generally limited to small proteins or complexes. Membrane transporters are often multimeric, macromolecular systems that undergo large conformational changes. To understand these ‘machineries’ and their dynamic nature we must use an approach that is capable of capturing their various conformations.

Site-directed spin labelling can be used in combination with continuous wave EPR techniques to investigate the structure, topology and chemical environments of the strategically placed probes. Pulsed Electron-Electron Double Resonance spectroscopy can precisely measure distances between two labels, thus allowing the conformations of these transporters to be mapped as they undergo their dynamic rearrangements.

Here this approach has been applied to various systems in order to gain better insight into the functional dynamics of membrane-bound or -associated proteins. This includes members of the generally well-studied substrate binding proteins from ATP-binding cassette (ABC)-type transporters, where crystal structures have been published and an accepted model of function exists. This subclass of proteins is responsible for binding a huge range of natural substrates, and the reported inter-lobe flexibilities can vary widely. Further to this, preliminary investigations of a multimeric secondary active transporter (the first of its family to have its structure resolved) are reported here, with the view to work towards answering several questions such as binding order and potential synchronisation of the monomers.

The systems studied here are of interest due to their medical relevance; substrate binding domains are vital for sequestering nutrients in many cell types that are involved in disease (e.g. pathogens and cancer), and the malfunction of many mammalian homologues of secondary active transporters have been linked to neurological disorders.

Introduction

1 Membrane Transporters

In all Kingdoms of life, efficient and selective movement of nutrients and toxins into and out of organisms is vital, regardless of whether the system in question is comprised of one or trillions of cells. At a cellular level, this involves transport across a membrane, usually facilitated by specialised 'machineries' composed of proteins and their derivatives. As a broad category of proteins, membrane transporters have been linked to a plethora of medical disorders and pathogenic infections, either as a result of their malfunction or due to their role in a pathogen's survival and virulence. Understanding how these proteins function or fail has therefore become an area of study widely invested in, as they represent enticing potential targets for novel drugs (such as antimicrobials and anticancer agents).

The Transporter Classification Database[1] (TCDB; www.tcdb.org) contains more than 18,000 transporter protein sequences, classified into 1356 families based on functional and phylogenetic information and further categorised into 72 superfamilies. These can be broadly sorted into several classes (numbering of list below not consistent with that of the TCDB system[2]):

1. Channels/pores
2. Electrochemical potential-driven transporters
3. Primary active transporters
4. Group translocators
5. Transport electron carriers
6. Accessory factors involved in transport
7. Incompletely characterised transport systems

Despite the identification of a vast number of genes and protein sequences related to membrane transport, relatively few of their three-dimensional structures have actually been resolved. In the Protein Data Bank[3] (PDB; www.rcsb.org), there are over 140,000 protein structures, as resolved by any method; only around 2800 coordinate files belong to membrane proteins (n.b. not necessarily transporters), corresponding to just over 900 unique proteins (according to blanco.biomol.uci.edu/mpstruc[4]). This is likely a result of the nature of membrane proteins themselves; for example, it is often difficult to produce and purify such molecules in the quantities needed for structural studies. These challenges, and

the merits and drawbacks of the various methodologies used to resolve these protein structures, will be discussed in a later section (3. EPR on Membrane Transporters).

To introduce the systems investigated in this thesis, only two of the classes of membrane transporter described previously need to be explored in more depth: the primary active transporters and electrochemical potential-driven transporters (or secondary active transporters).

1.1 Primary Active Transport (ABC-type)

This method of transport uses a form of metabolic energy to drive the translocation of substrates across a membrane; this is frequently in the form of the chemical potential energy but may be energy derived from redox chemistry or photons.

ATP-binding cassette (ABC) systems are found ubiquitously in nature[5-6], responsible for the selective import and export of various molecules and, in some cases, for non-transport processes such as DNA repair[7]. These systems hydrolyse adenosine triphosphate (ATP) in order to drive these processes. As such, the ABC transporters are a large superfamily of primary active transporters. The members of this superfamily are typically multimeric complexes consisting of at least two transmembrane proteins/domains (TMDs), two highly conserved membrane-associated ATPase subunits (the eponymous ABCs, also called nucleotide binding domains, NBDs) which transiently dimerise, and, in most importers, a substrate/solute binding protein (SBP, also known as a substrate/solute binding domain, SBD). Due to their prominence in nature, in terms of both universal presence and functional importance, these systems have been a focal point of much research. In microorganisms, these present as potential targets for novel treatments, and indeed as the source of newly evolved resistances[8-10]. Moreover, for decades ABC systems have been implicated in drug-resistance in cancers[11], but it has been more recently proposed that the role of these systems in tumorigenesis extends beyond drug efflux[12]. Further to this, malfunctions and mutations in these protein complexes have been linked to a large number of diseases and disorders[13].

Nucleotide-binding domains (NBDs)

High resolution structural data for many of these complexes and their components exist, particularly for the ABC (or NBD) modules. These display remarkable sequence identity across the various species from which the systems in question were isolated, and as such are considered the defining characteristic of the superfamily[6, 14-15]. These also have a consistent global fold, which can be broken into two subdomains: a larger catalytic core and a smaller α-helical subdomain. The larger segment is structurally similar to the RecA-like motor ATPases, consisting of six α-helices and two β-sheets. This subdomain contains the Walker A motif (or P-loop; GXXGXGKS/T where X can be any amino acid) and Walker B motif (ΦΦΦΦD, where Φ is a hydrophobic residue)[16]. This is where ATP hydrolysis takes place; the glycine-rich P-loop binds the nucleotide electrostatically via the triphosphate component, while the aspartate residue of the Walker B coordinates the catalytic Mg^{2+} ion via its water molecule ligand[17-18]. A conserved aromatic residue, usually a tyrosine, stacks with the adenine ring in a region known as the A-loop[19].

Homologues of this subdomain are also found in a variety of systems outside of the ABC superfamily. The smaller, helical subdomain consists of three or four α-helices and contains the 'ABC signature' motif (also referred to as the LSGGQ motif, C motif, or linker peptide)[18]. The two subdomains are linked by two loop regions; one of these contains a highly conserved glutamine residue and is referred to as the Q-loop. This is believed to be involved in the interactions between the NBDs and TMDs, particularly in relaying the hydrolysis of ATP to give a conformational change in the TMDs[5, 7, 17]. The H-loop (or switch region) contains a conserved histidine, which hydrogen bonds the γ-phosphate of the ATP and is required for hydrolysis. This residue is also involved in interactions when the two NBDs dimerise and has contacts with the Walker A motif and D-loop[7, 18]. The D-loop is a conserved sequence, preceded by the Walker B motif, containing an aspartate residue[19].

A lot of research to date has focussed on using the structural information about the NBDs to understand their mechanism of function. Some of this has been based on the crystal structures, investigating where certain interactions take place and how the assemblies are oriented and rearranged spatially. For example, it is understood that in order to hydrolyse the ATP, the two NBDs act cooperatively and

dimerise. The D-loops of the two NBDs run alongside each other in the dimer and help form the hydrolysis site[19]. The nucleotide is positioned at the interface, flanked by the Walker A motif and Q-, H- and D-loop regions of one unit and the LSGGQ motif of the other. Additionally, mutational studies have been used to confirm that certain conserved residues are necessary for the units to function (either structurally, i.e. forming hydrogen bonds, salt bridges or hydrophobic interactions to stabilise the conformations, or mechanistically, i.e. hydrogen bonds to activate the water for hydrolysis)[5, 7, 17-18].

Transmembrane Domains (TMDs)

The TMD has a largely α-helical secondary structure and is embedded in the phospholipid bilayer; this undergoes conformational change as a result of ATP hydrolysis in the NBDs in order to facilitate transport. Due to the variety of substrates that are conveyed by ABC transporters, it is not surprising to note that the structures and sequences of the TMDs can vary quite significantly[20]. There are usually 6 α-helices per TMD monomer, however there are cases in which 5-11 helices per monomer occur[21]. These monomers are arranged and oriented in such a way as to form a pore or channel that is only accessible to one side of the lipid bilayer at any one time (potentially via occluded intermediates). Though there is no notable sequence identity across this subunit, they do exhibit consistently similar three-dimensional topologies than can be divided into separate folds: importer types I, II and III (these folds can only be determined once the protein's structure has been resolved), and exporters[20]. ABC importers have only been discovered in prokaryotes, whilst the exporters exist in both pro- and eukaryotes.

The research into ABC-type transporters discussed within this thesis focusses on type I and II importer systems, and so only the details of these will be examined at length here. The importer classes were originally divided into two (types I and II) based on general their size and architectural differences[22-23]. Type III importers refer to the family of energy-coupling factor (ECF) transporters; while the existence of this group of transporters has been known since the 1970s[24], they were only recently classified as a subgroup of ABC transporters[20, 25].

The type I importers are usually structurally characterised by having fewer α-helices when compared with type II importers[22-23]. These TMDs are typically composed of 5 core TM helices per monomer (the two monomers making up the TMD are either identical, i.e. homodimeric, or structurally very similar), but in some

cases there is a sixth helix which wraps and links the two monomer units, creating a TMD of 12 helices in total[19]. These are thought to undergo domain-wide rigid-body movements during conformational change[22]. Additionally, the type I class has been found to exhibit severely limited ATPase activity in the absence of SBDs (see below for more detailed structural and functional information on these subunits), whilst ligand-bound SBDs promote activity[26].

In contrast, the type II importers are typically larger than the type I importers; their TMDs typically consist of 10 α-helices per monomer[20, 22]. The TMD monomer structures can be broken down into two subdomains: TMs 1/2 and 6-10 make up one, and TMs 3/4/5/5a make up the other (also referred to as the inner subdomain). Fewer structures of these importers have been resolved than those of type I, but the mechanism of translocation has been posited to involve two distinct rigid-body rotations of these subdomains per TMD monomer[22] (see below for more mechanistic detail).

A common feature of these importer classes (and potentially other folds of ABC-type transporters)[19, 22-23, 27] is the so called 'coupling helix' of the TMDs, which exist at their interface with their associated NBDs. These short helices dock into the surface of the NBD, at a groove between the RecA-like and helical subdomains[22-23] (containing the Q-loop), which rotate towards one another upon ATP binding[19]. These coupling helices share little or no sequence identity when aligned, but architecturally they appear to be well-conserved[23, 27]. Another key characteristic that is shared by the type I and II importer classes is that they both require SBDs to sequester and deposit substrate for translocation.

Substrate-binding Domains (SBDs)

The SBDs were the first components of ABC-type transporters to be identified[21]; the first was the L-arabinose binding protein (ABP) in 1974[28]. These subunits are water-soluble and can either be found moving freely in the periplasm of Gram-negative bacteria (generally in excess relative to the translocator complexes[29]), or tethered to the membrane (either lipid-anchored or attached to a membrane-spanning peptide) or fused to their TMDs in Gram-positive bacteria or archaea[20, 30]; TMD-fused SBDs are also found in Gram-negative bacteria, but are far less common[21]. Regardless of their arrangement and location relative to the rest of the ABC complex, they share a similar overall fold, consisting of two globular lobes connected by a linker region[30]. This is interesting considering they vary widely in

both size and sequence, which itself is likely consistent with the sheer variety of target substrates they are responsible for acquiring[19].

Though the overall fold of these subunits is ubiquitous, there are defining characteristics and differences that have been used in the extensive work that has been done to divide them into 6 distinct groups, clusters A-F[29-30]. Further research into these classifications has resulted in updated subdivisions[31]. Some of these SBDs are found outside of ABC transporters[30].

Cluster A SBDs associate only with type II importers, and are responsible for binding metal ions, either directly (subgroup A-I) or via chelators (A-II). This cluster has a single, rigid α-helix linker region. These SBPs are unique to ABC importers[29-31].

SBDs in cluster B bind carbohydrates, branched-chain amino acids, natriuretic peptides and auto-inducer 2. The structural feature that defines this group is the presence of three interconnecting strands between the lobes. The cluster can be further separated into five subdivisions, based on the native substrate and the architecture of the interconnecting strands. Some of these SBDs are associated with systems outside of ABC transport[29-31].

Cluster C SBDs have a varied pool of target substrates: di- and oligopeptides, arginine, nickel ions and cellobiose. These are typically larger than the SBDs in other clusters (around 50-70 kDa) due to their ‘extra’ domain that helps in the binding of large substrates. This cluster of SBDs is associated only with type I importers[29-31].

Cluster D can be divided into four subclusters based on the respective substrates: carbohydrates such as maltose (D-I), poly-amines (D-II), tetrahedral oxyanions (D-III-a & D-III-b) and ferrous/ferric iron (D-IV). Though some of these subclusters bear functional similarities with other SBD clusters (D-I and B both bind carbohydrates), they are all defined as cluster D as their hinge region consists of two short strands (4-5 amino acids each). The majority of these are associated with ABC transport, but some D-II SBPs have been linked to transcription factors[29-31].

Cluster E associate either with tripartite ATP-independent periplasmic (TRAP) transporters (E-I subcluster), or tripartite tricarboxylate (TT) transporters (E-II subcluster), which employ electrochemical ion gradients to drive transport. E-I

SBPs bind organic acids, amino alcohols, dipeptides and glycerol-3-phosphate, while E-II SBPs bind amino acids. The defining structural characteristic of this group is a single β-strand which is part of both lobe domains[29-31].

Cluster F SBDs bind a range of substrates, which define the subclusters: thiamine, pyrimidines, sulfonates and bicarbonate (F-I), methionine (F-II), glycine betaine, proline betaine and choline (F-III), and amino acids (F-IV). The cluster F SBPs are identified by two peptide stretches at the hinge, similar to cluster D, but 8-10 amino acids each. F-I, II & III are all associated with type I importers only. F-IV SBPs are found in type I importers, ligand-gated ion channels and G-protein coupled receptors (GPCRs)[29-31].

The general consensus is that these proteins function by a common mechanism known as the 'Venus Flytrap' model[32-34], wherein the substrate is bound in the cleft between the two lobes, which close together via a pivot, bend or rotation of the hinge/linker region. However, since it has been established that there are marked differences between, and even within, the various defined clusters of SBP, it calls into question whether a single mechanism can be applied so universally without addressing the previously mentioned structural differences. Indeed, the various crystal structures point towards there being significant disparities in the degrees of flexibility between the respective lobes.

Generally Accepted Mechanism of ABC Importers

The mode of action for these importers is generally accepted to be the alternating-access model (fig. 1.1.1). This describes the alternation of the complex between inward- and outward-facing conformations. The resting state of these importers is inward-facing, i.e. the TMDs hold the NBDs apart and block off the channel at the outer side of the membrane[22]. When the SBP is docked and loaded with substrate, two molecules of ATP bind at the NBD interface[23] and they dimerise. This brings the coupling helices of the two TMDs together, switching the conformation to outward-facing. This is domain-wide in type I importers but is only thought to involve one subdomain in type II (TMs 1/2 and 6-10)[20, 22]; the inner subdomain (TMs 3-5a) of the type II importer is thought to move separately in response to the SBD, mediating access to the translocation pathway[20].

The SBP is then able to deposit the substrate in the cleft created between the TMDs. It has been suggested that the movement of the 'outside' end of the TMDs

(where the SBD docks) that occurs during conformational change may physically separates the lobes of the SBD and encourage deposition of the substrate into a transient binding pocket in the TMD gap of type I importers[19, 23]; it is suggested that there is little or no such interaction of type II importer TMDs with their substrates and instead specificity is conferred exclusively by the SBDs[19, 23]. The ATP is then hydrolysed, and the NBDs dissociate from one another, reorienting the TMDs to the inward-facing position, thereby occluding the proposed transient binding pocket. This allows the substrate to move into the cytoplasm and for the products of ATP hydrolysis (adenosine diphosphate, or ADP, and an inorganic phosphate group, P_i) to be released[22, 27].

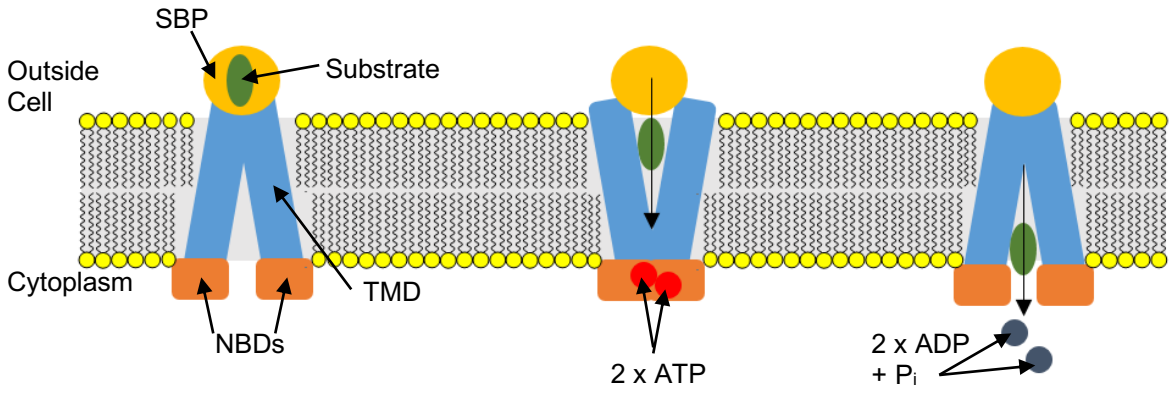


Fig. 1.1.1: representation of the alternating-access model for the mechanism of gram-positive bacterial ABC-type importers, showing the translocation of a substrate coupled with ATP hydrolysis.

1.3 Secondary Active Transport

This is also known as coupled transport or co-transport. There is no direct coupling of metabolic energy, instead electrochemical potential difference (derived from pumping ions in or out of the cell) is used to drive the movement of the coupled substrate, often against its concentration gradient. The transport of the ion and the coupled substrate can be in the same direction, or in opposite directions, to one another across the membrane (symport and antiport, respectively). Usually, the ion pumped to create this electrochemical gradient is a monovalent cation, typically H^+ or Na^+.

Solute Carrier (SLC) Family

The Solute Carrier (SLC) family of membrane transporters is comprised of around 400 secondary active transporters and facilitative transporters (i.e. those that allow solutes to be transported downhill with their electrochemical gradients). This large family is then categorised into 65 families[1, 35-37] and is responsible for the transport of a wide variety of substrates, both charged and uncharged, organic and inorganic. The criteria for a group of proteins to be classified as SLCs are functional, rather than their relation to other SLCs (either by amino acid sequence or phylogenetics); whilst the sequence identity of members within an SLC family is typically greater than about 20-25%, the homology between the separate families is negligibly low[38].

The malfunctions of most SLC families have been linked to various diseases and disorders including certain types of anaemia[39], diabetes[40], neurodegenerative diseases (including Alzheimer's, Huntington's[41] and Parkinson's[42]) and mental health issues (such as depression, anxiety and anorexia nervosa[42]). Despite their appeal as potential therapeutic targets[43] there are relatively few crystal structures resolved for this type of transporter[44]. However, from the structures of 44 SLC proteins (covered in a 2017 review[44]) that have been reported, they can be grouped into several common folds despite their vast differences in both sequence and function. The largest of these can be designated as: Major Facilitator Superfamily (MFS) fold, Leucine Transporter (LeuT) fold, 'other antiparallel' (or NhaA-like) folds. Of the remaining identified folds, the most important for this thesis is 'others with hairpin (HP) domains'.

The MFS fold is named after a large group of membrane transporters which are responsible for facilitating the relocation of small solutes across membranes as a result of chemiosmotic gradients[45]. This fold consists of 12 TM helices overall, in two 6-helix bundles referred to as N- (TMs 1-6) and C-terminal (TMs 7-12) domains; in each domain, there are two 3-TM inverted repeats. The bundles are connected by an extended cytoplasmic loop (between TMs 6 and 7). These bundles are arranged into a 'clam shell' type structure, with an aqueous cleft between the bundles; this is lined with amino acids that bind the target substrate, and thereby confer the substrate specificity. It is generally thought these proteins function as monomers. Their binding sites differ based on the chemical and structural features of the substrate (polarity, charge, size, shape) but largely occur at the interface between the two helical bundles[44]. These types of transporter are thought to use an alternating access-type mechanism sometimes referred to as 'rocker-switch'[46] (fig. 1.2.1). This means the aqueous cleft is open either to the extracellular environment or cytoplasm at any one time. Conformational change from one orientation to the other is coupled with simultaneous sealing of the opposite end (thereby preventing a continuous pathway through the transporter). The sealing of the translocation pathway is aided by 'gates' on either side of the structure. Of the SLC proteins whose structures have been resolved, this is the largest cluster (15 of 44)[44].

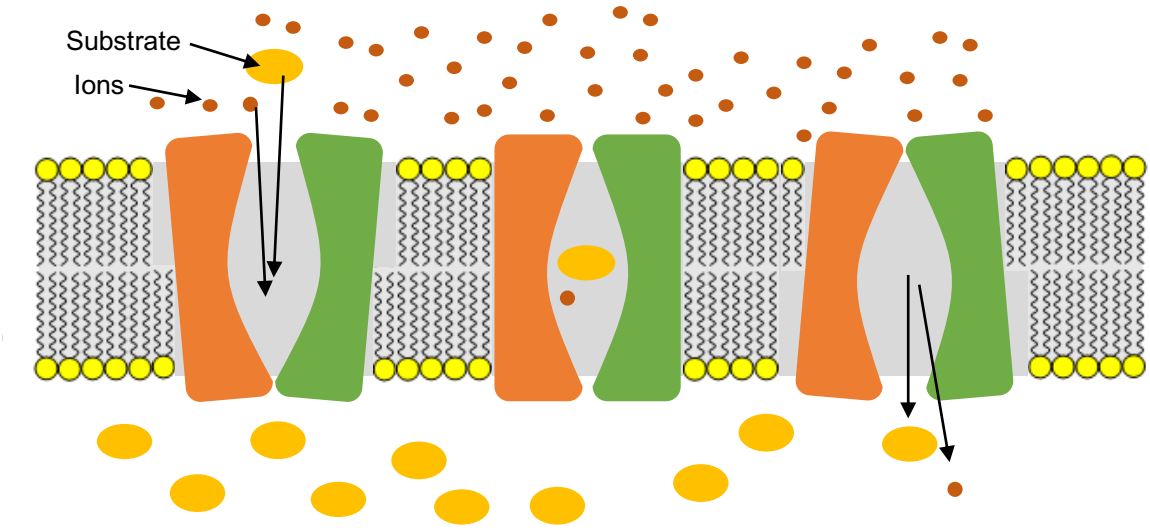


Fig. 1.2.1: representation of the alternating-access rocker-switch model for the mechanism of MFS-fold secondary active transporters, in this example showing the symport of a substrate coupled (against concentration gradient) with ions (with concentration gradient). The separate 6-helix bundles are shown in different colours (orange and green).

The LeuT fold is named after the Leucine Transporter, a dimeric bacterial homologue of the sodium-dependent neurotransmitter transporter (Neurotransmitter Sodium Symporter, NSS, family). This involves protomers of 10 TM helices arranged as a 5+5 inverted repeat (i.e. 5 TM helices in each inverted repeat)[44, 47]. These repeats are oriented anti-parallel to one another. An identifying feature of the LeuT fold is that the first TMs of the two inverted repeats (TM1 and TM6) are comprised of two short, disjointed α-helices each. These short segments are connected by a highly conserved 'unwound' stretch[47-48]. The first two helices in each repeat (TMs 1, 2, 6 & 7) form a four-helix bundle that comprises the core domain; this is surrounded by another bundle of the third and fourth helices per repeat (TMs 3, 4, 8 & 9), called the scaffold domain[44], where the protomers meet to form oligomers. Individual transporters may have unique structural features though the overall fold is the same, e.g. more helices (though 5+5 core conserved) in vSGLT[49-50].

The substrate binding site is thought to be a pocket between certain TMs; the discontinuous helices TM1 and TM6 sit approximately halfway through the lipid bilayer, and their associated unwound sections are proposed to provide local polar environments for ion/substrate binding[47]; the substrate-to-ion binding stoichiometry is within the 1:1 to 1:3 range and ion binding/release is directly coupled with substrate binding/release[44]. Like the MFS fold, these proteins are thought to operate via an alternating access mechanism. In this, the core domain rotates to transport the substrate, whilst the scaffold is a stationary support[51], and the substrate is suggested to be occluded from both the extracellular environment and the cytoplasm by 'gates'[47]; this mechanism is sometimes referred to as 'gated pore'[46, 52] (fig. 1.2.2). However far fewer of the SLC proteins whose structures have been resolved are grouped in this category (only 7 of the 44), making it difficult to build a complete picture of the functional dynamics these systems undertake[44, 47].

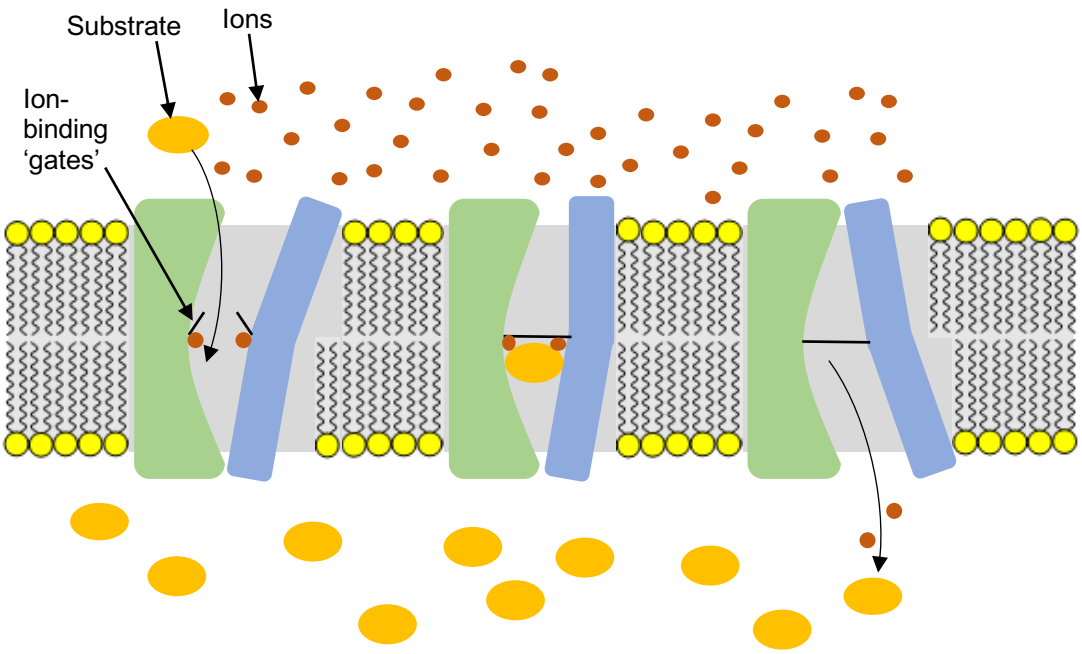


Fig. 1.2.2: representation of the alternating-access gated pore model for the mechanism of LeuT-fold secondary active transporters, in this example showing the symport of a substrate coupled (against concentration gradient) with ions (with concentration gradient). A single protomer is shown for simplicity (scaffold domain in green, transport domain in blue).

The 'other antiparallel' fold, sometimes referred to as NhaA-like (after the Na^+/H^+ antiporter NhaA), is similar to the LeuT fold in that it contains 2 sets of inverted repeats, usually with 5 TMs each[48], though this number of TMs is variable[44]. This fold also contains discontinuous helices, but they occur in a different location when compared with the LeuT-fold (TM 4 in the first repeat, and TM 9 in the second)[48]. These are also similar to LeuT-fold proteins in that their protomers are also made up of two domains: a transport domain (TMs 1, 2, 6 & 7) and an ion-coordinating scaffold domain (all remaining helices)[44, 48]. Less is known of this fold, and even fewer structures (5 out of the 44) have been resolved for these proteins[44]. It has been suggested that these proteins function via an 'elevator mechanism'[53] (fig. 1.2.3), similar to that of LeuT-fold proteins, wherein the transport domain moves relative to the scaffold (discussed in more detail below).

The remaining SLCs whose structures have been resolved (the final 17 of 44)[44] do not adopt folds that fit the above models, and therefore grouped under 'other' folds. However, within this is a very small group of oligomeric proteins with similar protomer folds containing so-called hairpin (HP) domains (3/44 SLC proteins of known structure)[44]. The overall structures have scaffold domains (where the protomers meet) and transport domains. Since so few of these structures have been resolved, it is difficult to infer a broad functional theory. However, some ideas

have been proposed, the most prevalent of which is the elevator mechanism[46, 54-55]. The prototypical system around which this mechanism has been built is $GltPh$[54-55], a homotrimeric sodium/aspartate transporter of the 'HP domain' fold group[44]. In this, the transport domain, including the HP domains, moves relative to the scaffold domain (or oligomer interface). The HPs are proposed to act as gates or flaps[52]; they open to allow substrate and ion binding/release and close to occlude the space as the transport domain moves up and down like an elevator[46, 48, 54-55]. It has been shown that the protomers of these systems work independently of one another[56-57].

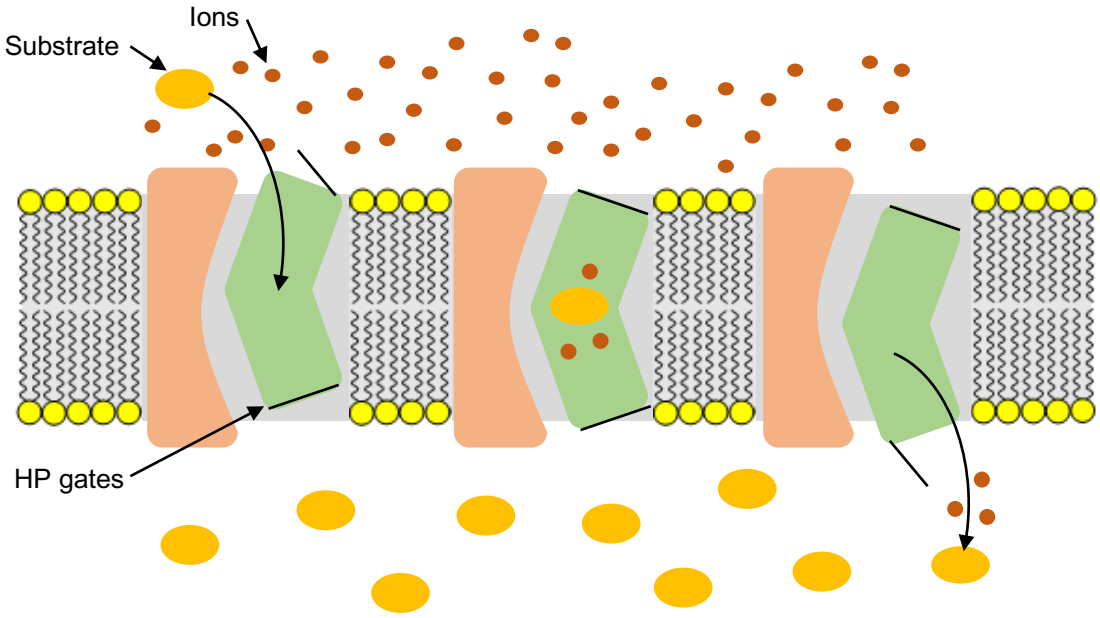


Fig. 1.2.3: representation of the alternating-access elevator model for the mechanism of some secondary active transporters, in this example showing the symport of a substrate coupled (against concentration gradient) with ions (with concentration gradient). A single protomer is shown for simplicity (scaffold domain in orange, transport domain in green).

SLC13/DASS

In order to introduce the secondary active transporter system covered in this thesis, we will focus on the specifics of the SLC13 family, also known as the Divalent Anion:Na^+ Symporter (DASS) family. These have been found in archaea, plants, bacteria and animals and can be divided into two groups based on their target substrates: Na^+-sulfate (NaS) or Na^+-carboxylate (NaC)[58]. More specifically, they are responsible for the transport of organic di- and tricarboxylates (often substrates and products of the Krebs Cycle), dicarboxylate amino acids, and inorganic sulfate and phosphate[58]. For all of these, the Na^+:anion coupling ratio is

3:1. The properties and functions (both structural and physiological) of the vertebrate, and particularly mammalian, members of the DASS family have been reviewed[58]. However, it is often beneficial to use bacterial homologues as they are typically easier to produce and purify in the quantities needed for structural and functional studies. For this particular SLC family, the structure of only one example has been successfully resolved: VcINDY (*Vibrio cholerae* I'm Not Dead Yet). The first crystal structure was published in 2012 (PDB accession code 4F35)[59], and later to higher resolution in 2017 (PDB 5ULD)[60]. This has been used to create homology models of other DASS family members (particularly the human variants)[61]. The VcINDY protein fits within the 'HP domain' fold group as described above as a homodimer[44] and so is proposed to fit with the elevator mechanism[46, 61-62]. The specific details of VcINDY will be discussed and explored more thoroughly in a later section.

2 Electron Paramagnetic Resonance Spectroscopy

2.1 Fundamental theory of EPR

Spectroscopy is the study of the interaction of matter with electromagnetic radiation. In the case of Electron Paramagnetic Resonance (EPR) spectroscopy, microwave radiation is utilised to study unpaired electrons. An electron is a subatomic particle that carries a negative elementary electric charge and has a mass approximately $\frac{1}{1836}$ that of a proton[63]. It is the charge, coupled with the angular momentum (as derived from the spin, a fundamental quantum mechanical property) of the electron contribute its intrinsic magnetic moment, μ_s.

$$\mu_S = g_e \frac{\mu_B}{\hbar} S \quad (1)$$

where g_e is the electron spin g-factor (a dimensionless quantity; 2.00232 for a free electron)[64], *S* is its spin angular momentum (½) and μ_B is the Bohr magneton.

This property allows the unpaired electron to interact with an externally applied magnetic field to separate previously degenerate energy levels. When this energy level separation is matched by that of incident electromagnetic radiation, resonance occurs. This will be explored in greater depth in a later section ('Electronic Zeeman and the resonance condition').

The spin Hamiltonian

In quantum mechanics, a 'Hamiltonian', $\hat{H}$, refers to an operator which encompasses the sum of both the kinetic and potential energies for all particles that compose a particular system. When considering a complete Hamiltonian of a molecular system, this is very complex as it expresses the spatial and spin terms for all of the electrons and nuclei. However, in the context of an EPR experiment, we can simplify the complete Hamiltonian and describe the system with the effective spin Hamiltonian, $\hat{H}_{eff}$:

$$\hat{H}_{eff} = \hat{H}_{EZI} + \hat{H}_{NZI} + \hat{H}_{HFI} + \hat{H}_{EEI} + \hat{H}_{NQI} \quad (2)$$

where each contributing Hamiltonian refers to the following interactions: electron Zeeman interaction (EZI), nuclear Zeeman interaction (NZI), hyperfine interaction (HFI), electron-electron interactions (EEI) , and nuclear quadrupolar interaction

(NQI), respectively. Many of these interactions are explored in more detail in later sections, but for now we can categorise them as follows:

- EZI – involves the interaction of the unpaired electron spin and an applied magnetic field
- NZI – involves the interaction of the nuclear spins of any magnetic nuclei and an applied magnetic field
- HFI – involves the interaction of the unpaired electron spin and the nuclear spins of any nearby magnetic nuclei
- EEI – involves the interaction if the spins of multiple unpaired electrons with one another; this encompasses multiple phenomena including those of strongly coupled electrons in zero-field splitting (ZFS) and weakly coupled electrons in dipole-dipole coupling and exchange interactions.
- NQI – involves the interaction of the electric quadrupole moment (in magnetic nuclei with $I > ½$) with the local electric field gradient at the nucleus.

Electron Zeeman and the resonance condition

In the absence of an external magnetic field, the energy levels related to the orientation of the electron's spin are degenerate. Upon application of a magnetic field and its interaction with the electron's magnetic moment (μ), an electron can align either parallel or antiparallel to the field; this separates the orientations into a higher energy +½ (μ antiparallel to the field) and lower energy –½ (μ parallel to the field) state (fig. 2.1.1). The separation on these states in energy is directly proportional to the strength of the applied magnetic field (see equation 3)[65-66]. This phenomenon is known as the electronic Zeeman Effect.

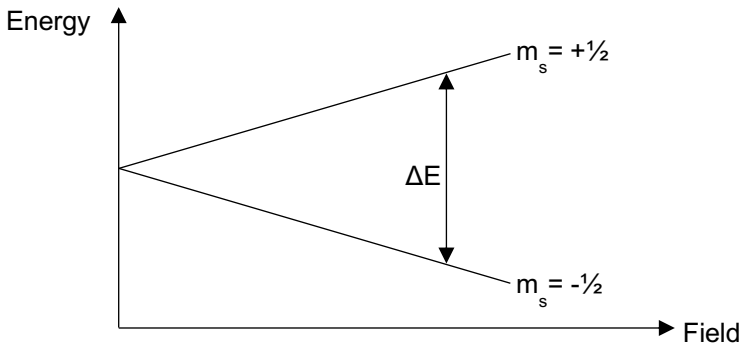


Fig. 2.1.1: representation of the separation of degenerate energy levels of the two spin states of an electron upon application of a magnetic field as described by the Zeeman effect.

This difference in energy levels creates a Boltzmann distribution between the two alignment states of the electron. When energy, in the form of electromagnetic radiation, is applied to the system the Boltzmann distribution is altered; as a photon is absorbed an electron's spin can be 'flipped' and promoted from the lower to higher energy state. The energy level difference (ΔE) is related to radiation frequency by Planck's Law:

$$\Delta E = \mathrm{h}\upsilon \tag{3}$$

Where h is Planck's constant and υ is the frequency of the electromagnetic radiation (Hz). The contribution of the electron Zeeman interaction to the spin Hamiltonian can be described as:

$$\hat{\mathrm{H}}_{EZI} = \mathrm{g_e} \cdot \mu_\mathrm{B} \cdot B_0 \cdot \boldsymbol{S} \tag{4}$$

Where g_e is the g-value of a free electron (this is explained in the later section 'g-factor'), B_0 is the applied magnetic field and $\boldsymbol{S}$ is the electron spin. This can then be related to the resonance condition (eq. 5), which is fulfilled when the energy of a photon equals the difference in energy between the +½ and –½ states at a particular magnetic field[67]:

$$\Delta E = h\upsilon = g_e \cdot B_0 \cdot \mu_B \tag{5}$$

The resonance condition for electrons is fulfilled using microwave radiation (i.e. electromagnetic radiation with frequencies in the GHz range). As the frequency of the radiation used increases, so must the applied field.

Much of the radiation in the microwave range is heavily absorbed by water molecules in the atmosphere. However, there are 'bands' at which certain frequencies of microwaves have minimal atmospheric absorption; it is at these frequencies that EPR is performed[68]. The most commonly used frequency is ~9.5 GHz (referred to as X-band).

g-factor

The g-factor is a proportionality constant; it relates the magnetic moment of a particular species to its angular momentum and is given in units of Bohr magnetons. This value provides information about the electronic structure of a molecule; an unpaired electron responds both to the externally applied magnetic field and the local magnetic fields caused by surrounding atoms (with magnetic nuclei) and other unpaired electrons.

The observed g-factor of a paramagnetic molecule differs from g_e (i.e. is larger or smaller than 2.00232)[67] when the ratio between the magnetic moment and angular momentum of the unpaired electron differs from that of a free electron. An electron's magnetic moment is constant; therefore it must be some loss or gain of angular momentum through spin-orbit coupling (i.e. an electromagnetic interaction between the spin of the unpaired electron and its orbit around the nucleus with which it is associated) that causes the g-factor to differ. The magnitude of the difference between the g-factor and g_e is indicative of the electronic character of the orbital containing the unpaired electron[68]. For example, the g-factors for organic radicals are relatively similar to g_e, ranging from about 1.99 to 2.01. The g-factors associated with transition metals can commonly vary far more, from about 1.5 to 6[69].

g-value anisotropy

Thus far the g-values referred to have been those that would be observed in fluid systems, where rapid isotropic motion would occur thereby negating the influence of directionality. In reality the g-factor is not simply a single number, but a 2nd order tensor shown in the 3x3 matrix below:

$$\begin{bmatrix} g_{xx} & g_{yx} & g_{zx} \\ g_{xy} & g_{yy} & g_{zy} \\ g_{xz} & g_{yz} & g_{zz} \end{bmatrix} \quad (6)$$

The tensor can be diagonalised (i.e. certain equivalent components, e.g. g_{xy} and g_{yx}, cancel out), leaving the three components g_{xx}, g_{yy} and g_{zz}. In a system undergoing rapid isotropic motion these values are averaged, and a single g-value will be obtained from its EPR spectrum.

$$g_{iso} = \frac{1}{3} \cdot \left(g_{xx} + g_{yy} + g_{zz}\right) \quad (7)$$

When a system is no longer in such a state to move freely (i.e. when frozen or otherwise immobilised) and unable to average out the effects of direction-dependent (anisotropic) interactions of the paramagnetic species with the field and microwave radiation, the g-value separates into the three components (fig. 2.1.2), which can be related to the x, y and z Cartesian axes; conventionally the axis with the highest symmetry is assigned as z[70-71]. A system in which all symmetries are equivalent (x=y=z) is described as isotropic and will show a single g-value in its EPR spectrum.

For many molecules, the three principle axes are not equal. A system in which the x and y axes are identical but different to the z axis (x=y≠z) is described as being 'axial'. For such a system, only two g-values are observed in its powder-type EPR spectrum: the parallel, g|| (g_{zz}), and the perpendicular, g⊥ (g_{xx} and g_{yy})[72].

In 'rhombic' systems, none of the axes are equivalent (x≠y≠z). The peak at the lowest field, corresponding with g_{xx}, is positive. The central peak, g_{yy}, is first positive, then negative (resembles a typical first derivative). The highest frequency peak, g_{zz}, is negative[70].

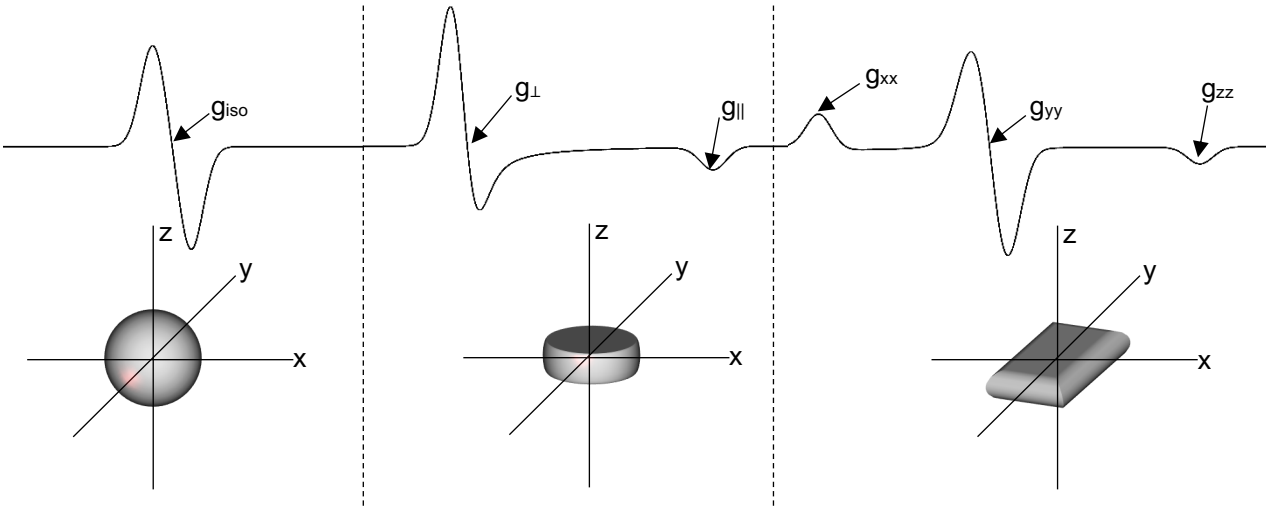


Fig. 2.1.2: characteristic frozen spectra (top) of systems with different symmetries and depictions of such molecules on Cartesian axes (bottom), from left to right: isotropic, axial and rhombic. The appropriate g-value labels are applied to their respective spectrum, obtained by simulations of rigid-limit spectra at X-band.

In the context of real-world studies, including those of biological systems, the molecules being investigated may be in a range of conformations in any given sample. As a result, the structural surroundings of the paramagnetic species will vary slightly, thereby yielding slight spreads in the g-values. This is reflected in Gaussian (inhomogeneous) lineshapes of the spectra and is known as g-strain.

Hyperfine interactions

Thus far, the interactions between an unpaired electron and the atomic nucleus it orbits have only been considered in the context of spin-orbit coupling and its influence on the g-value[68]. However, there are interactions between the electron's magnetic moment and any surrounding nuclei with non-zero nuclear spin values ($I \neq 0$); these are termed nuclear hyperfine interactions (HFI). There are two mechanisms by which these interactions occur.

One of these mechanisms is the dipolar interaction, or the anisotropic hyperfine interaction (influenced by the orientation of the molecule in relation to the field); this is an interaction between the magnetic moments of the unpaired electron and the nucleus. It is dependent of the shape of the electron orbital and the average distance between the electron and nucleus. The symbols A_{aniso} is often used to represent the anisotropic hyperfine coupling constant[68, 73].

The other mechanism is the Fermi contact interaction, or isotropic hyperfine interaction (the orientation of the molecule to the applied field does not matter).

This only takes into account s-orbital electrons, since it has no nodes; the mechanism depends on there being unpaired electron spin density at the nucleus with which it is coupled. Hence, p-, d- and f-orbitals are not considered[68]. The symbols A_{iso}[74] is used to denote the isotropic hyperfine coupling constants. The magnitude of the isotropic coupling constant is indicative of the probability that the unpaired electron can be found at the nucleus. This gives information on the electronic structure of the molecule. The contribution of the HFI to the spin Hamiltonian of a system is given by:

$$\hat{\mathrm{H}}_{HFI} = \sum \mathbf{I} \cdot \mathbf{A} \cdot \boldsymbol{S} \qquad (8)$$

Where **I** is non-zero nuclear spin and **A** is the hyperfine interaction matrix (a combination of the anisotropic and isotropic contributions to the HFI).

An unpaired electron can couple with multiple magnetic nuclei, giving rise to complex patterns of both isotropic and anisotropic hyperfine splittings of EPR spectra. The effects of the two types of interaction are easily separated by recording the EPR spectra under different conditions, e.g. a room temperature recording of a solution, in which the isotropic interactions are prevalent.

In many cases the isotropic hyperfine splitting pattern for a paramagnetic species tumbling freely in solution can be predicted fairly easily. A radical containing n equivalent nuclei with spin I will give rise to an EPR spectrum with the number of lines 2nI+1[75]. For example, an unpaired electron coupled to a single proton (1H has a nuclear spin of ½) will yield two lines: 2(1x½)+1 = 2. In comparison an unpaired electron coupled to a nitrogen nucleus (^{14}N has I = 1) gives three lines: 2(1x1)+1 = 3.

This can be explained by re-examining the Zeeman Effect. When an unpaired electron is coupled with a magnetic nucleus, the nuclear spin alters the number of allowed transitions that the electron can undergo (fig. 2.1.3). Like the electron, a nucleus (consisting of protons and neutrons, both of which have an associated mass) with a non-zero spin value has a magnetic moment. This local field will either have an additive effect on the external magnetic field or subtract from it. The higher and lower energy states of the electron are therefore split into either integer or half-integer levels, depending on the value of I, separated in steps of 1 (i.e. if I = 3/2, the m_I levels will be: -3/2, -1/2, +1/2 and +3/2); this is the nuclear Zeeman Effect.

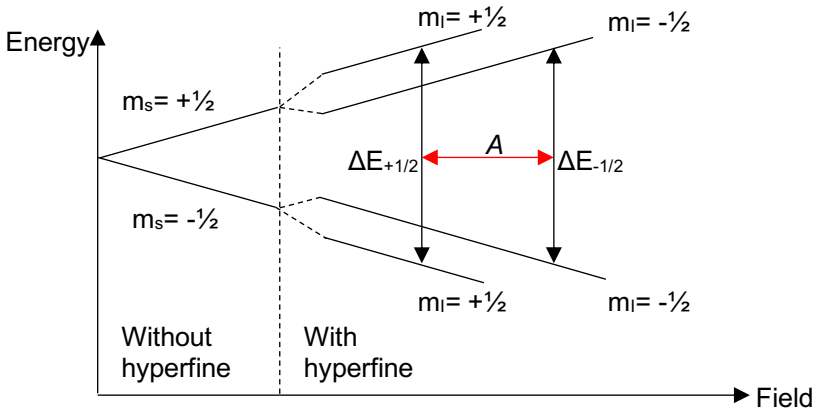


Fig. 2.1.3: representation of the electronic Zeeman Effect, before and after the influence of the nuclear Zeeman. The splitting caused by the isotropic hyperfine interaction (A) with a magnetic nucleus of I = ½.

The relative intensities of the resonance peaks split by equivalent magnetic nuclei observed on the EPR spectrum follow the pattern of Pascal's Triangle (see fig. 2.1.4)[75]. The situation is complicated when the various magnetic nuclei of inequivalent electronic environment and differing nuclear spins are coupled to the same unpaired electron. In these cases, the primary lines, caused by splitting by nuclei with larger spin density (i.e. where the electron resides to a greater extent), are split into secondary patterns, giving rise to more complex spectra.

n =	*I* = ½	*I* = 1
1	1 : 1	1 : 1 : 1
2	1 : 2 : 1	1 : 2 : 3 : 2 : 1
3	1 : 3 : 3 : 1	1 : 3 : 6 : 7 : 6 : 3 :1
4	1 : 4 : 6 : 4 : 1	1 : 4 : 10 : 16 : 19 : 16 : 10 : 4 : 1
5	1 : 5: 10 : 10 : 5 : 1	1 : 5 : 15 : 30 : 45 : 51 : 45 : 30 : 15 : 5 : 1

Fig. 2.1.4: the relative intensity ratios of peaks caused by hyperfine splitting of a resonance by different numbers of equivalent magnetic nuclei, n, with nuclear spins (I) of ½ and 1, as dictated by Pascal's Triangle.

The isotropic hyperfine coupling constant, A_{iso}, can be measured on an EPR spectrum as the difference in field strength between equivalent points on each of the resonance transitions attributed to a single nucleus; typically, the measurement is given in units of field (G, mT etc.) or frequency (MHz). Much like the g-value, the coupling constant is a tensor, which can be related to the Cartesian axes x, y and z and diagonalised as before. This gives the three

components A_{xx}, A_{yy} and A_{zz} when the rotation of the paramagnetic centre is impeded[74].

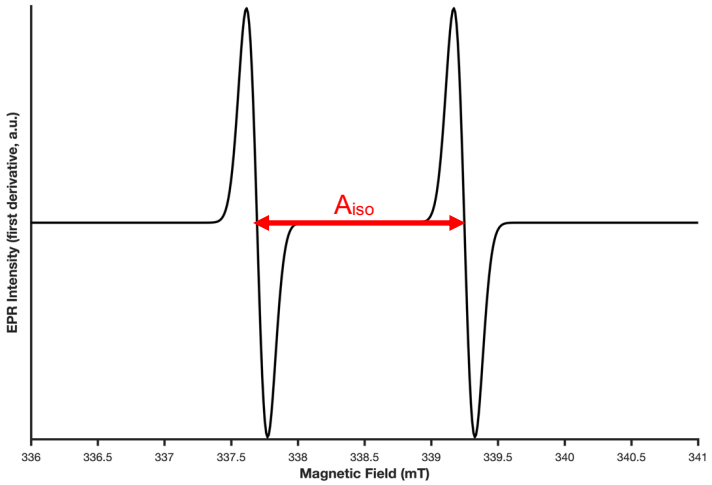


Fig. 2.1.5: simulated X-band spectrum with the unpaired electron coupled to a magnetic nucleus with I = ½, in an isotropic system, with the hyperfine coupling constant A_{iso} labelled.

Zero-field splitting

Zero field splitting (ZFS) describes the removal of spin state energy level degeneracy for systems with $S > ½$ regardless of an applied magnetic field; this is caused by the dipolar interactions and/or spin-orbit coupling between multiple unpaired electrons. The former tends to dominate in organic radical species, whilst the latter prevails in transition metal centres. For the purpose of the thesis, ZFS will largely be discussed in the context of paramagnetic metal cofactors, and the dipole-dipole interaction of two unpaired electrons (in the form of spin labels) will be explored separately.

$$\hat{\mathrm{H}}_{ZFS} = \boldsymbol{S}_A \cdot \boldsymbol{D} \cdot \boldsymbol{S}_B \tag{9}$$

The dipolar tensor, $\boldsymbol{D}$, is traceless and symmetric, diagonalisation of which yields parameters D and E:

$$\boldsymbol{D} = \begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_z \end{bmatrix} = \begin{bmatrix} -\frac{1}{3}D + E & 0 & 0 \\ 0 & -\frac{1}{3}D - E & 0 \\ 0 & 0 & \frac{2}{3}D \end{bmatrix} \tag{10}$$

We can define D and E under the following conditions:

$$D \equiv \frac{3D_z}{2} \tag{11}$$

$$E \equiv \frac{(D_x - D_y)}{2} \tag{12}$$

Then we can rewrite the ZFS Hamiltonian as follows:

$$\hat{\mathrm{H}}_{ZFS} = D\left[S_z^2 - \frac{1}{3}S(S+1)\right] + E\left(S_x^2 - S_y^2\right) \tag{13}$$

These are dependent on both the average distance between the unpaired electrons and the degree of departure of the system away from cubic symmetry.

According to the Kramers degeneracy theorem, a system with an odd number of unpaired electrons (i.e. where there is half-integer spin) will have doubly degenerate energy levels in the absence of a magnetic field. These pairs of degenerate energy levels are referred to as Kramers doublets. The separation between these doublets is dependent on parameters *D* and *E*, which in turn, as previously mentioned, is dependent on symmetry. In systems with cubic symmetry ($D = E = 0$) the sublevels remain degenerate. In axial ($D \neq 0$, $E = 0$) and rhombic ($D \neq 0$, $E \neq 0$) symmetry, the sublevels are non-degenerate and transitions may be observed (at zero field)[76-77]. The level of rhombicity is given as a ratio of E/D and lies between 0 and 1/3, in accordance with the definitions of *E* and *D* stated in equations 11 and 12 (i.e. $E \leq D/3$).

We can consider divalent manganese, Mn^{2+}, as an example, which has a partially filled 3d orbital ($3d^5$). In the high spin state there five unpaired electrons, and the total spin, *S*, is 5/2. This results in six energy levels ($2S+1$): -5/2, -3/2, -1/2, +1/2, +3/2 and +5/2. The ZFS lifts the degeneracy of these energy levels, resulting in the separated Kramers doublets, $m_s = \pm 5/2, \pm 3/2$ and $\pm 1/2$.

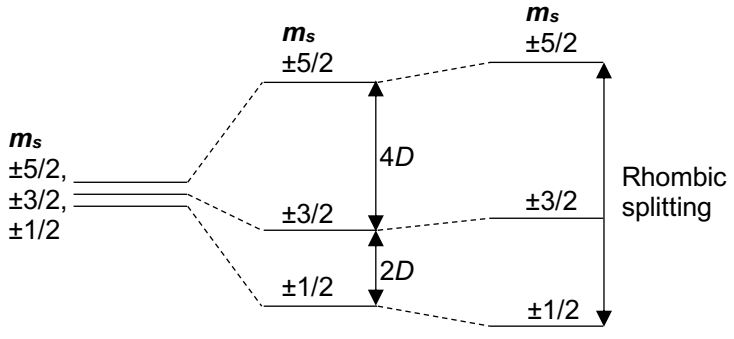


Fig. 2.1.6: representation of the loss of degeneracy in the spin state energy levels of an S = 5/2 system in the absence of an applied magnetic field caused by zero field splitting (ZFS). The effects of both the axial and rhombic symmetry are demonstrated qualitatively by the extent of the separation of the Kramers doublets.

When an external magnetic field is applied to the system, the degeneracy within the Kramers doublets is lifted in accordance with the electron Zeeman effect. In cases where the magnitude of *D* is exceeded by the energy of the microwaves applied to the system (*D* << *hv*) both inter- and intra-manifold transitions occur. Conversely, if *D* exceeds the energy of the applied radiation (*D* >> *hv*) only the intra-manifold transitions are observed[76].

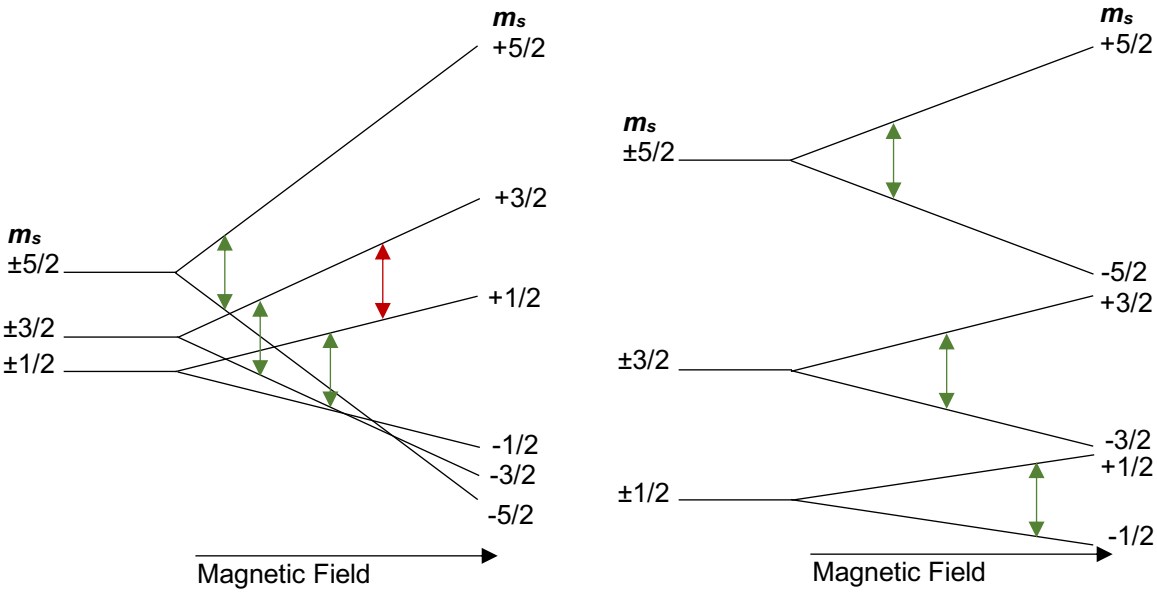


Fig. 2.1.7: representation of separation of the Kramers doublets of a S = 5/2 system upon application of a magnetic field, in accordance with the electron Zeeman Effect where: the magnitude of parameter D is less than the applied microwave radiation (left), and the magnitude of parameter D is greater than the applied microwave radiation (right). Examples of intra- and inter-manifold transitions are demonstrated by green and red arrows, respectively.

Hyperfine coupling of the unpaired electrons and non-zero nuclear spin can further split the energy levels in each manifold. In the case of Mn^{2+} here, where I = 5/2, this gives a sextet per m_s (m_I = +5/2, +3/2, +1/2, -1/2, -3/2 and -5/2); this means there are a total of 36 m_s,m_I states. According to the selection rules: $\Delta m_s = \pm 1$, $\Delta m_I = 0$, 30 of these transitions are formally allowed[78].

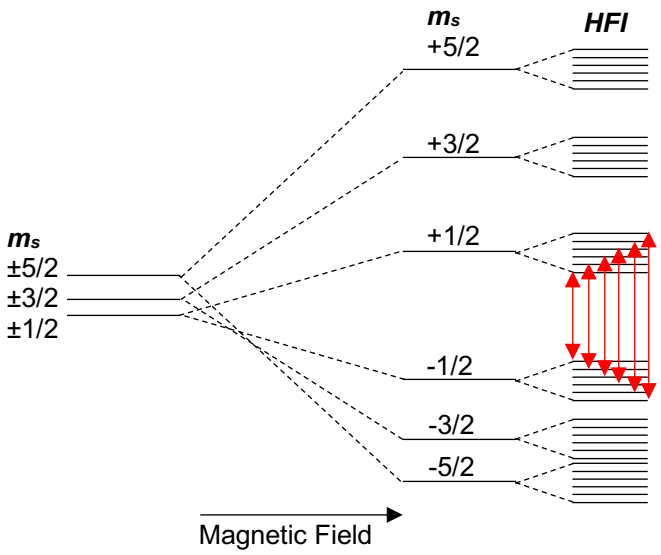


Fig. 2.1.8: representation of separation of the Kramers doublets of a S = 5/2 system upon application of a magnetic field (D < hν), in accordance with the electron Zeeman Effect. The electron spin energy levels are then split by hyperfine interactions with I = 5/2 nuclei. Examples of formally allowed transitions are demonstrated by red arrows.

In our example here of divalent manganese, the characteristic sextet, centred around g = 2, can be attributed to the intra-manifold m_s = ±1/2 transitions[78]. ZFS effects can cause mixing of the m_I levels, leading to formally forbidden (Δm_s = ±1, Δm_I = ±1) transitions. These may be visible in experimental spectra as doublets between the central transitions[77].

In the case of systems with an even number of unpaired electrons (i.e. where there is integer spin; a 'non-Kramers' system), the energy levels are still arranged into doublets but with the addition of a singlet. These doublets, however, are non-degenerate even at zero field (if E ≠ 0), and the separation of the energy levels is usually greater than the incident microwave radiation. Thus, no transitions are typically observable. Transitions between the doublets are formally forbidden as the Δm_s > ±1; if they are observed they will usually be very weak, though parallel-mode EPR can enhance them.

Dipole-dipole interactions

Dipole-dipole interactions occur as a result of two or more spins. Here, these interactions will be discussed in reference to multiple weakly coupled electron spins (these can also occur between an electron and a magnetic nucleus, though the coupling is much weaker). The interactions inform on the distance and orientation of the unpaired electrons. In the context of strongly coupled unpaired electrons and ZFS, we could consider the collection of interacting spins as a singular system. However, here it is better to consider the individual spins separately and describe the interactions between them. The Hamiltonian contribution of the dipolar coupling can be expressed similarly to the ZFS interactions

$$\hat{\mathrm{H}}_{DD} = \boldsymbol{S}_A \cdot \boldsymbol{D} \cdot \boldsymbol{S}_B \tag{14}$$

However, **D** here is the electron dipole-dipole coupling tensor, which can be described as follows:

$$\boldsymbol{D} = \frac{\mu_0}{4\pi\hbar} \cdot \frac{g_A g_B \mu_B^2}{r_{AB}^3} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -\omega_{dd} & 0 & 0 \\ 0 & -\omega_{dd} & 0 \\ 0 & 0 & 2\omega_{dd} \end{bmatrix} \tag{15}$$

Where r is the distance vector between coupled spins A and B and ω_{dd} is the dipolar coupling frequency, which can be expressed:

$$\omega_{dd} = \frac{g_A g_B \mu_B^2 \mu_0}{4\pi\hbar} \cdot \frac{1}{r_{AB}^3} \cdot (1 - 3cos^2\Theta) \tag{16}$$

Where θ is the angle between the distance vector r and the external magnetic field B_0. The dipolar coupling is a through-space interaction, which is typically measurable for distances up to approximately 80 Å (but ongoing advances are continually pushing this limit).

2.2 Continuous-wave EPR

Spectrum acquisition

There are significant engineering difficulties in configuring a microwave source that can stably provide a range of frequencies. In comparison, it is relatively simple to use an electromagnet or superconducting magnet to sweep a range of applied fields. Therefore, a cw-EPR spectrum is a plot of signal intensity against field, with the frequency of the radiation remaining constant[79]. EPR is a resonance spectroscopy; transitions of electrons between the energy levels are measured from the absorption of photons. This is detected by monitoring the disruption of a standing wave leading to reflections of microwaves from the sample cavity.

An EPR spectrum is recorded as a first derivative; the procedure allows a small signal to be distinguished from noise, thereby increasing the resolution of the recorded spectrum. During spectrum acquisition, the magnetic field is modulated at a certain frequency (in the order of kHz) across a specified amplitude or width (in units of field, G). This field oscillation causes any signal detected to oscillate at the same frequency. The greater the difference of the signal at the two ends of the modulation window, the greater the observed intensity of the detected signal. Increasing modulation amplitude will result in greater intensity of peaks on the spectrum. However, if the amplitude of modulation is greater than the linewidth of the peak, line broadening will occur, distorting the spectrum and giving unreliable information. It is therefore important to be aware of this parameter when recording spectra[65, 80]. This means a balance must be struck between maximising the signal to noise ratio, whilst avoiding line-broadening.

Saturation and relaxation

Upon application of the magnetic field, electrons can populate either the higher or lower energy state; the distribution of the magnetic moments between the parallel and antiparallel orientations in an S = ½ system is described by Boltzmann statistics (eq. 17) when the system is in thermal equilibrium[67, 71]:

$$\frac{n_{upper}}{n_{lower}} = e^{\frac{-\Delta E}{kT}} \quad (17)$$

Where n_{upper} is the population of the excited state and n_{lower} is that of the ground state, k represents the Boltzmann constant and T denotes the temperature of the system (K).

When an EPR spectrum is recorded electrons from the lower energy level are moved to the higher state. This forces the system away from equilibrium. If there was no method by which the system could return to this equilibrium, there would be a gradual loss of signal as the populations of the two states equalise; once the population difference is zero there is no EPR signal detected. This is called saturation[71].

There are two relaxation processes by which a system is returned to equilibrium following excitation. The first is spin-lattice (or longitudinal) relaxation, T_1. This method causes changes in the spin orientation by energy transfer from local field fluctuations, either caused by lattice vibration in solids, or by molecular motions and collisions in solutions[81]. The second method is spin-spin (or transverse) relaxation, T_2, which does not involve energy transfer to the lattice. This can occur through various different mechanisms, such as inhomogeneities in the field, or direct but energy-conserving spin-spin interactions[81].

The parameter $P_{½}$, or power of half-saturation, is used to characterise power saturation experiments. This value is the microwave power (mW) at which the signal is half-saturated.

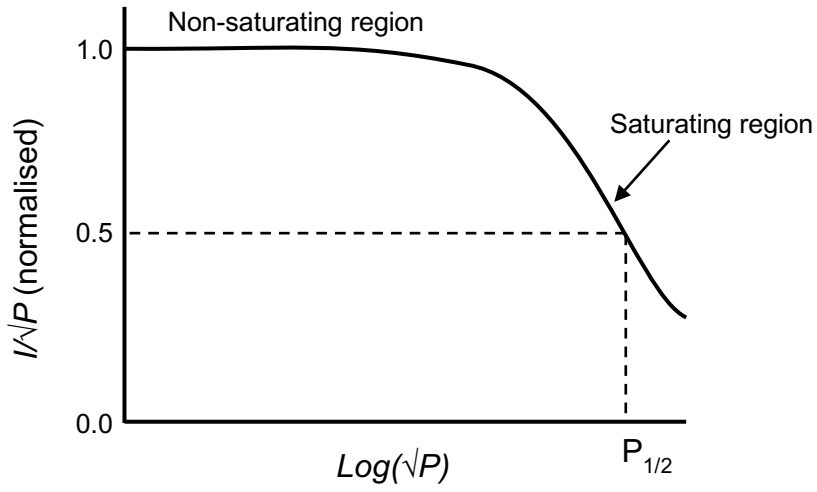


Fig. 2.2.1: a representation of a saturation profile (plot of signal amplitude as a function of the square root of microwave power), or power saturation curve, showing the measurement of the $P_{½}$ parameter.

Transition metal centres tend to have higher $P_{½}$ than radical species. This highlights the differences in metal centre and radical spin-lattice relaxation times, T_1; it might be deduced that metal centres have longer τ_1 than radicals, but this is not true[70]. The very short relaxation times of metals lead to 'lifetime broadening' (which is a common issue in acquiring the EPR spectra of transition metals). Increasing the power of the microwaves being applied to the system increases the number of electron spin transitions per unit time, and therefore decreases the lifetime. Applying Heisenberg's uncertainty principle (relating the lifetime of an excited state to its energy uncertainty) means a shorter lifetime increases the energy uncertainty, thus resulting in broadened spectral lines, making it difficult to detect a signal[82]. Measuring the EPR spectra of transition metals at low temperatures solves this problem as it slows the process of relaxation.

Lifetime broadening is not such an issue in radical species. Instead, linewidth is affected by several other factors. The unpaired electron's dipole may interact with those of surrounding nuclei and other electrons. These dipole-dipole interactions can cause slight fluctuations in the magnetic field. This results in small changes in the energy levels, thereby mixing them and increasing the energy uncertainty; this broadens the linewidth. Additionally, the flipping of one electron's spin can induce the transitions of others by affecting the local magnetic field. These processes dominate the relaxation times of radical species[83].

Site-directed spin labelling

Thus far the molecules discussed have contained some endogenous paramagnetic centre. However, many molecules of interest (i.e. proteins) do not contain any inherent source of paramagnetic character. This problem is overcome by a technique known as site-directed spin labelling (SDSL). This refers to the attachment of an exogenous paramagnetic species to a protein at a specific location. Most commonly this comes in the form of a stable organic radical with a reactive group that will interact specifically with an amino acid sidechain.

Many of the most commonly used spin labels used are nitroxide radicals; these consist of a heterocyclic ring containing a nitrogen atom, with the unpaired electron found on the –NO group protected by a number of neighbouring methyl groups. The functional groups that allow site-specific attachment often react with the thiol group of a reduced cysteine residue, via a methanethiosulfonate, maleimide or iodoacetamide group. S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol3-yl)methyl methanesulfonothioate, or MTSL, is one of the most commonly used spin labels.

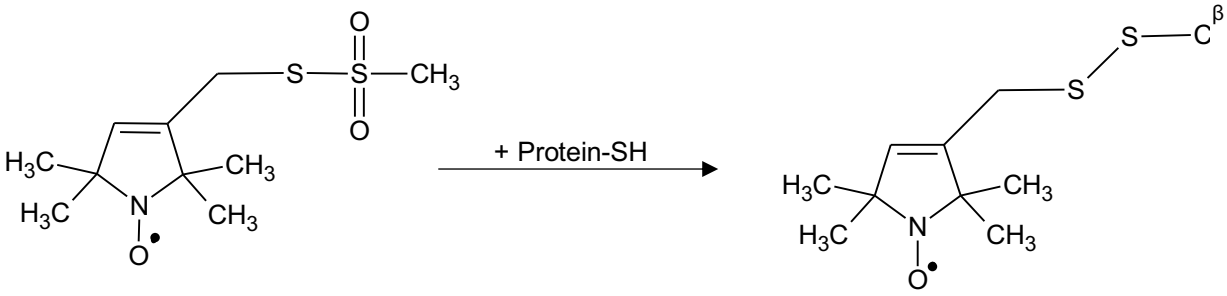


Fig. 2.2.2: the reaction of spin label MTSL with the thiol group of a reduced cysteine residue in a protein.

Cysteine residues may be removed from the native protein structure, and new single or double cys-variants produced by site-directed mutagenesis[84-85].

Correlation times

Spin labels and the analysis of their characteristic EPR spectra can provide information about the protein being studied. For example, the spectral shape of a singly-labelled variant is sensitive to motion; the rotational correlation time (τ_c) informs on the local environment. τ_c is a product of a combination of factors; nearby side chains and solvent molecules influence the rotational freedom. The spin label's chemical structure (i.e. the length of the connecting alkyl chain and bulkiness of the heteroatomic ring) and any hydrogen bonds it may form with protein side chains will also contribute to this parameter[85].

If the labelled site is solvent-exposed the rotational freedom of the radical side chain is only slightly hindered. This gives rise to a spectrum that largely resembles a nitrogen-coupled electron free in solution. That is, the hyperfine splitting is relatively small and each peak is approximately equal in intensity (1:1:1) with a narrow linewidth on the central peak (ΔH_0)[70]. This is consistent with rotational correlation times of about 0.1 ns. In contrast, a labelled site that is buried within the protein and interacting with neighbouring side chains will have much more restricted mobility. This yields a spectrum with a noticeably larger hyperfine splitting, broadened and more asymmetric peaks, and a wider linewidth on the central peak[84]. This corresponds with τ_c values of about 2.5 to 5.0 ns.

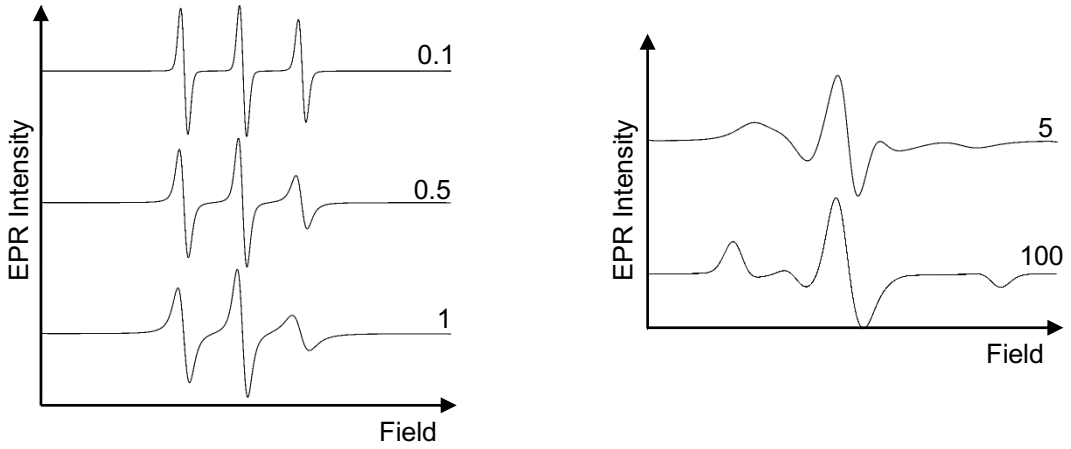


Fig. 2.2.3: the effect of increasing rotational correlation time, τ_c (ns), has on the broadening of the spectral lines of a typical spin label at X-band, measured at room temperature; 0.1 ns represents free rotational diffusion, 1 to 5 ns is expected in a spin-labelled protein, 100 ns demonstrates strong immobilisation as it begins to resemble a the anisotropy of a powder spectrum. Spectra obtained by simulation.

Knowledge of the central peak linewidth allows general assignment of the label site to a particular secondary structure[84]. In addition, the polarity of the solvent has an influence on the spin label spectrum, changing the values of the A_{zz} and g_{xx} components of the hyperfine and g tensors, respectively; a polar environment shifts the A_{zz} to higher values, while the g_{xx} moves to lower values. The former change can be readily observed in X-band spectra, but shifts in the g-tensor component require much the higher sensitivity and resolution afforded by higher frequency spectrometers (i.e. W-band or higher; see 'Multifrequency EPR')[84, 86].

Labelling at two positions allows further information to be obtained from EPR spectra; distances between the two spin labels can be measured. This can be done with low temperature cw-EPR for distances between about 10 and 20 Å[84].

For larger distances, i.e. between about 20 and 70 Å, a pulsed EPR method must be employed: Pulse Electron Double Resonance (PELDOR), also known as Double Electron-Electron Resonance (DEER)[87].

Multifrequency EPR

A typical X-band spectrum recorded at room temperature will yield very low resolution of both the g- and A-tensors; only g_{iso} and A_{iso} can be distinguished. Lowering temperature (achieved either by use of cryogenic liquids or a closed cycle refrigerator/'displex') allows further resolution, but even then, only the A_{zz} and g_{zz} components can be evaluated. In order to achieve greater resolution, higher fields and frequencies must be applied. At Q-band and above it becomes possible to better resolve the x and y components of the g-tensor. In turn, this makes it easier to distinguish hyperfine coupling constants due to decreased overlap between the components of both the A- and g-tensors[84, 86].

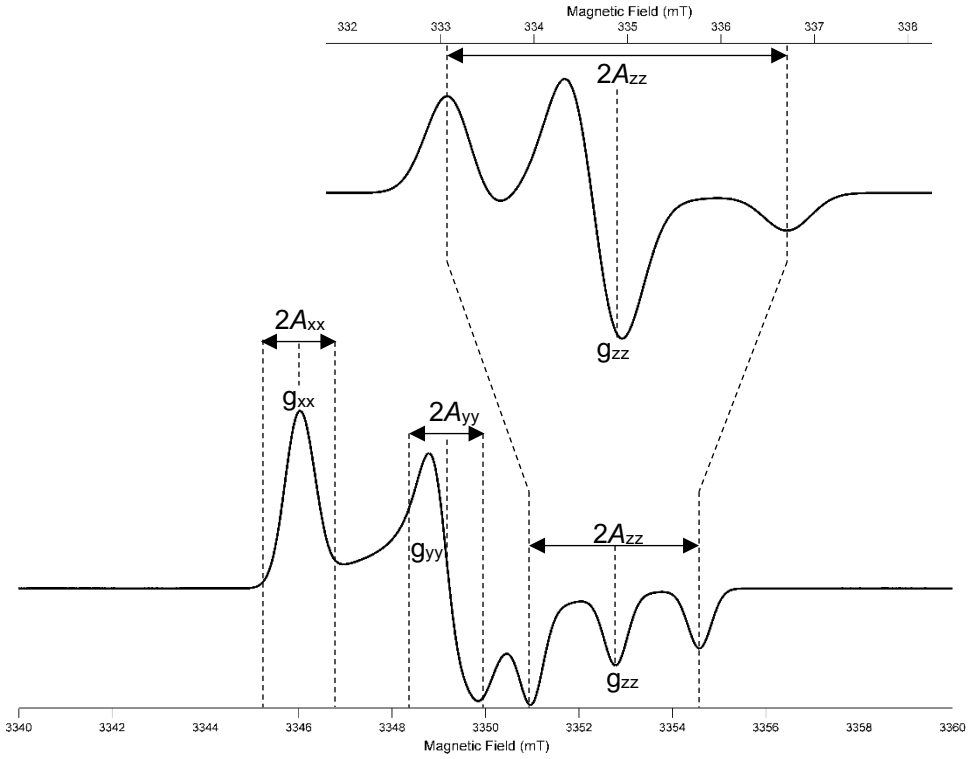


Fig. 2.2.4: illustration of the increased resolution of low temperature spectra with field and frequency; at X-band (top) only the g_{zz} and A_{zz} components are distinguishable, but at W-band (bottom) it is possible to measure A_{xx}, A_{yy}, g_{xx} and g_{yy} too. Spectra obtained by simulation of powder spectra at X- and W-band.

However, certain spectral features, such as line-broadening (g-strain) scale with the applied magnetic field (B_0) and therefore resolution does not necessarily improve with higher frequency EPR. Additionally, the positive effect of the enhanced Boltzmann distribution is eventually offset by loss in the resonator efficiency.

Dipolar broadenings

Low temperature cw-EPR spin-spin distance determination employs the concept of dipolar interactions in order to determine short distances (i.e. < ~20 Å). When 'spin A' interacts with a parallel 'spin B', the resonance frequency is decreased; the magnetic moment of a parallel electron spin decreases the effective field strength. However, when the first 'spin A' interacts with an antiparallel 'spin B' the resonance frequency is increased; an antiparallel electron magnetic moment adds to the applied field. Both interactions cause equivalent shift in frequency but are opposite in sign.

The dipolar frequency, ω_{dd}, is the difference between the original resonance frequency of spin 1 and the shifted resonance frequencies caused by interaction with the parallel and antiparallel orientations of the second spin. It is this parameter (ω_{dd}) that is inversely proportional with the distance between the spins, r. The spectra of doubly-labelled protein with different inter-spin distances can be evaluated qualitatively too; the nearer the two spins in space, the broader the powder spectrum peaks. This phenomenon is referred to as dipolar broadening[88].

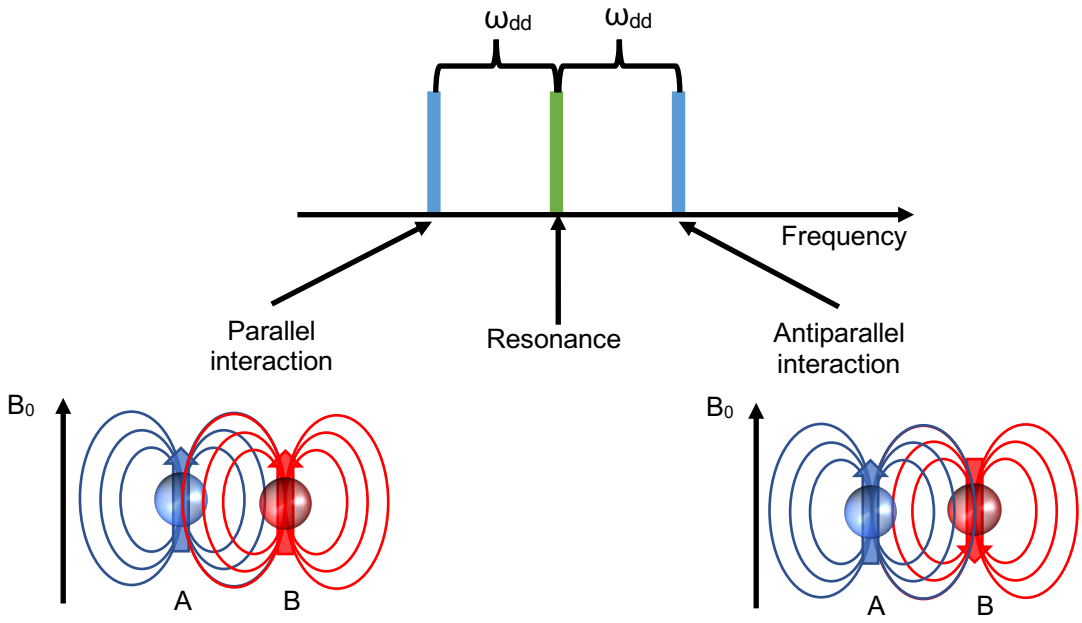


Fig. 2.2.5: the mechanism by which a resonance is broadened by dipolar interactions; parallel spins cause a decrease in resonance frequency, while antiparallel spins cause an increase in the frequency. Both interactions cause equivalent shifts.

2.3 Pulsed EPR concepts and theory

Reference frames, magnetisation and pulses

In order to discuss and describe the position and motions of an electron in the context of pulsed EPR an axis system is needed. The Cartesian axes are a well-known starting point; in EPR the laboratory frame refers to a static frame of reference in which the applied magnetic field (B_0) sitting along the z-axis with the microwave magnetic field (B_1) on the x-axis. The y-axis is of course orthogonal to both the x- and z-axes.

An electron, having angular momentum (as a result of its mass and spin), in a magnetic field experiences a torque such that it precesses about the z-axis. The frequency of this precession is known as the Larmor frequency, ω_L (eq. 18).

$$\omega_L = -\gamma B_0 \tag{18}$$

where γ is the gyromagnetic ratio, a ratio of the electron's magnetic dipole moment to its angular momentum; a free electron has a γ value of about 2.8 MHz/G. Net magnetisation will be along the z-axis; the majority of spins will be precessing about the z-axis, even though they will be randomly scattered in 2 dimensions (or across the x- and y-axes).

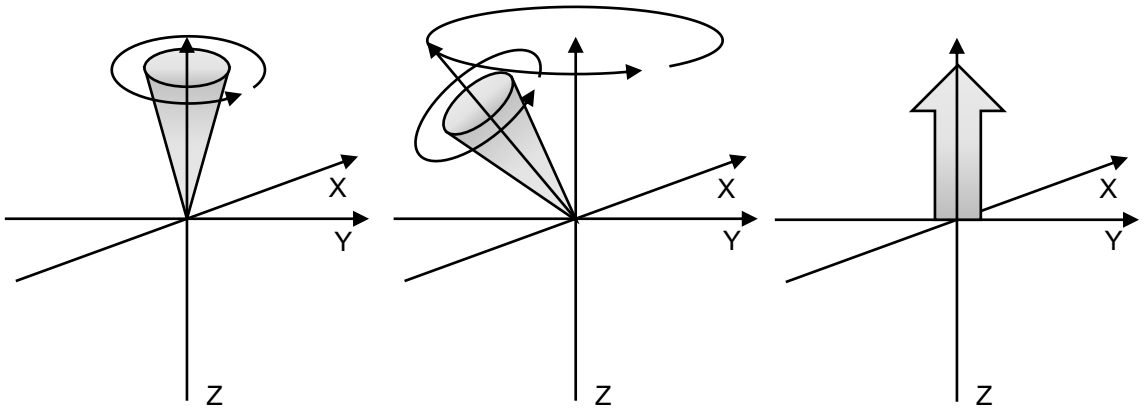


Fig. 2.3.1: diagram showing: the precession of a spin when parallel to the applied magnetic field, left; the precession of a spin at an angle to the z-axis (as a result of torque), centre; a simplified representation of the net magnetisation vector, M_0, right.

It is more convenient to move away from a static laboratory frame when considering magnetic resonance, where motions are happening in three dimensions whilst also precessing or rotating. EPR experiments typically use

resonators that produce linearly polarised microwaves, which create a magnetic field, B_1, perpendicular to the far stronger applied magnetic field, B_0. We can use a rotating frame of reference wherein the laboratory frame rotates at the same rate as the B_1 components (i.e. at the Larmor frequency).

$$\omega_L = \omega_0 \tag{19}$$

where ω_0 is the microwave frequency (rotation of B_1 components). This simplifies the explanation of a moving coordinate system by making the precessing magnetisation component/vector (M_0) appear stationary, hence B_0 can be disregarded. M_0 now interacts with the B_1 static field and precesses about it at frequency ω_1.

$$\omega_1 = -\gamma B_1 \tag{20}$$

ω_1 is known as the Rabi frequency. If B_1 is parallel to the x axis the magnetisation vector will precess in the zy plane, about the +x axis (for as long as microwaves are applied to the system). The tip angle, α, is the angle by which M_0 is displaced from the z-axis.

$$\alpha = -\gamma |B_1| t_p \tag{21}$$

where t_p is the pulse length (or duration for which B_1 is applied). The notation used to define pulses refers to the rotation of M_0 that they cause (in terms of pi) and the direction from which they originate (microwaves can be phase-shifted to be parallel to the y-axis). Hence, a $+x\ \frac{\pi}{2}$ pulse denotes an application of field B_1 along the +x axis that tips the bulk magnetisation 90° into the –y axis; signal detection is aligned with the y-axis. M_0 orientation therefore depends on the length, magnitude and direction of the B_1 pulse.

Free induction decay

Once a pulse ($+x\ \frac{\pi}{2}$) has been applied, the net magnetisation lies along the –y axis within the rotating frame. This means that in the laboratory frame it is rotating through the xy plane at ω_L. As the vector rotates it creates a signal in the resonator, which is maximised if the M_0 is completely within the xy plane and is called free induction decay (FID).

Off resonances

Up to this point, it has been assumed that the system was on resonance; i.e. microwave frequency was equal to the Larmor frequency. In reality there will be many different frequencies which will not all be on resonance, so the effects of off resonance signals must be considered.

Within the rotating frame (which turns at the frequency of the applied microwave field, B_1) a +x $\frac{\pi}{2}$ pulse leaves the net magnetisation along the –y axis. If the frequency of the microwave field does not equal that of the precessing magnetisation vector, the M_0 will rotate through the xy plane either faster or slower than frequency of B_1, ω.

$$\Delta\omega = \omega - \omega_0 \tag{22}$$

If the frequency difference, Δω, is zero both the microwave field rotation rate and magnetisation vector appear stationary to one another; this is expected for on resonance signals. If Δω>0 the magnetisation vector rotates faster than the microwave field and will turn anticlockwise. If Δω<0 the magnetisation vector rotation is slower than that of the microwave field and will turn clockwise.

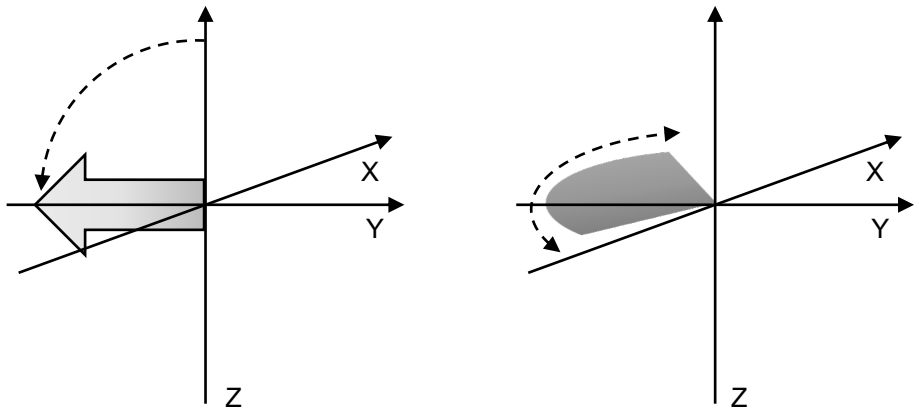


Fig. 2.3.2: representation of the magnetisation vector being tipped into the –y axis (left) and the subsequent 'fanning out' of the magnetisation vector caused by differences in rotation frequency, Δω (right), which leads to loss in magnetisation.

Since the net magnetisation precesses at Δω when off resonance, the effects of B_0 do not disappear. This means the magnetisation tips out of the z-axis differently by an effective magnetic field, B_{eff}, the vector sum of B_1 and B_0. The magnetisation is then tipped around B_{eff} at a different effective rate, ω_{eff}.

$$\omega_{eff} = \sqrt{(\omega_1)^2 + (\Delta\omega)^2} \tag{23}$$

B_{eff} does not lie in exactly the xy plane, meaning the magnetisation does not get tipped entirely into the xy plane; instead, it takes on the motion of a cone about B_{eff}. As the precession frequency ($\Delta\omega$) increases, the efficiency with which magnetisation is tipped into the xy plane decreases, with the FID signal decreasing as a result. This is especially a problem in broad EPR spectra, comprised of a wide range of frequencies ($\Delta\omega > \omega_1$). It is therefore important to maximise ω_1 to maximise the detectable FID signal. Minimising pulse length achieves the same effect.

Relaxation and broadening

Thus far, when magnetisation has been tipped into the xy plane by a microwave pulse it has been described under the assumption that it remains there. In reality this is not the case; alignment of magnetisation away from the applied field B_0 (into the xy plane) is a higher energy state compared to orienting parallel with it (along the z-axis). Since electrons are constantly interacting with their surroundings and environment, the magnetisation in the xy plane will decay and relax back to the z-axis. The process of relaxation, as previously described, can be split into two components: spin-lattice (longitudinal) relaxation time, T_1, and spin-spin (transverse) relaxation time, T_2.

T_1 describes the time needed for the magnetisation to return to M_0. This is driven by the system relaxing back to thermal equilibrium after perturbation by the microwave pulse and occurs through interactions with the surroundings (hence, spin-lattice relaxation). This parameter is important as it dictates how quickly an experiment can be repeated; magnetisation must be fully returned to the z-axis before it can be tipped into the xy plane again.

T_2 denotes the exponential decay and decoherence of magnetisation on the xy plane (i.e. M_x and M_y decaying to 0) and is related to homogeneous and inhomogeneous broadening. Inhomogeneous broadening is a result of different spins across a sample experiencing a range of different local fields; different local effective fields will cause some spins to slow down and others to speed up. This means that as the magnetisation precesses, the manifold spreads out along the xy plane, causing a decay in the signal (since many decoherent magnetisation components will cancel each other out). The shape of this decay is not generally exponential but reflects the shape of the EPR spectrum; the time constant of this decay is called T_2^*.

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{inhom}} = \frac{1}{T_2} + \gamma \Delta B_0 \quad (24)$$

where ΔB_0 is the difference in strength of the varying local field and T_{inhom} represents the relaxation rate contribution attributable to inhomogeneous field. T_2^* is always equal to or less than T_2.

The dephasing of the magnetisation vector is contributed to and sped up by the effects of homogeneous broadening; this arises from spins that experience the same magnetic field interacting and causing mutual flip-flopping of the spins.

2.4 Pulsed EPR applications

Inversion recovery (T_1)

As has been explained above electrons in a magnetic field will align parallel or antiparallel, with most in the lower energy level (i.e. parallel to the field, B_0 along the z-axis) in a Boltzmann distribution. The net magnetisation is a vector sum of the parallel and antiparallel orientations of spins. At equilibrium the majority of electrons will be aligned with the field and so the net magnetic moment will be parallel to the field.

A +x $\frac{\pi}{2}$ pulse causes the populations of the parallel and antiparallel states to equalise, resulting in no net magnetisation along the z-axis. The transfer of electrons from parallel to antiparallel (required to equalise populations) creates a magnetic moment, and therefore net magnetisation, in the xy plane. The magnetisation vector behaves as below after a $\frac{\pi}{2}$ pulse:

$$M_z(t) = M_0 \cdot \left[1 - 2 \cdot e^{-\frac{t}{T_1}}\right] \tag{25}$$

where $M_z(t)$ is the magnetisation along the z-axis at a given time, t. A +x π pulse inverts the Boltzmann distribution, putting the majority of electrons in the antiparallel orientation and net magnetisation lying along –z. The magnetisation vector behaves as below after a π pulse:

$$M_z(t) = M_0 \cdot \left[1 - e^{-\frac{t}{T_1}}\right] \tag{26}$$

A typical echo detected EPR pulse sequence ($\frac{\pi}{2} - \tau - \pi - \tau - echo$) is used in conjunction with an inversion pulse (π); the two-pulse sequence is gradually stepped through time following the inversion pulse.

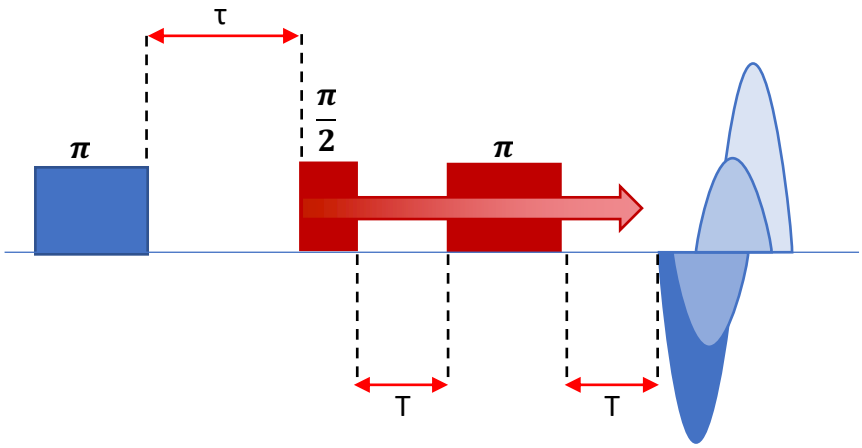


Fig. 2.4.1: schematic representation of the pulse sequence showing the timings and lengths of the pulses used to obtain an inversion recovery profile.

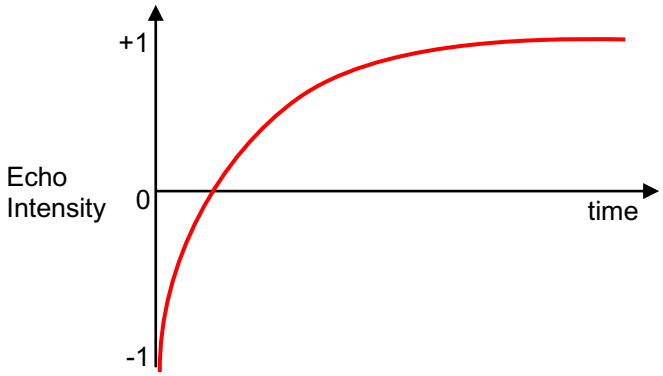


Fig. 2.4.2 a visualisation of how the echo intensity varies with time during an inversion recovery experiment.

A fit of the decaying exponential or bi-exponential is used to measure T_1 from the data recorded.

Hahn echo decay (T_2)

For T_2 the magnetisation vector behaves as below after a $\frac{\pi}{2}$ pulse:

$$M_{-y}(t) = e^{-\frac{t}{T_2}} \tag{27}$$

where $M_{-y}(t)$ is the magnetisation along the –y direction at a given time, t.

A standing Hahn echo is created using the standard echo detected EPR pulse sequence. In this, a $\frac{\pi}{2}$ pulse tips magnetisation into the xy plane; due to factors explained previously, the magnetisation undergoes decoherence. This is followed by a π pulse after a time delay,τ, which inverts the decaying signal. This causes the magnetisation to refocus after a second time delay (equal to the

aforementioned inter-pulse delay, τ) before subsequently dephasing once more. The echo that results is essentially a reversed FID followed immediately by a regular FID.

To experimentally measure T_2 a variation of the Hahn echo sequence is used; the time delay after which the second pulse (i.e. the π pulse) is applied varies. As a result, the echo intensity reduces with increased time delay due to lost decoherence (from spin relaxation) which cannot be refocused. The fit of the exponential decay of echo intensity is measured to extract T_2.

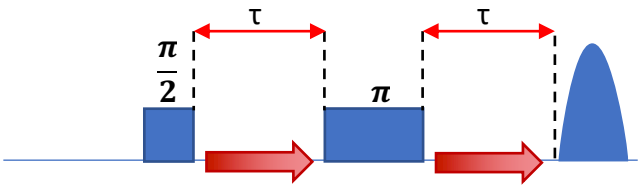


Fig. 2.4.3: schematic representation of the pulse sequence used to measure the decay of a Hahn echo.

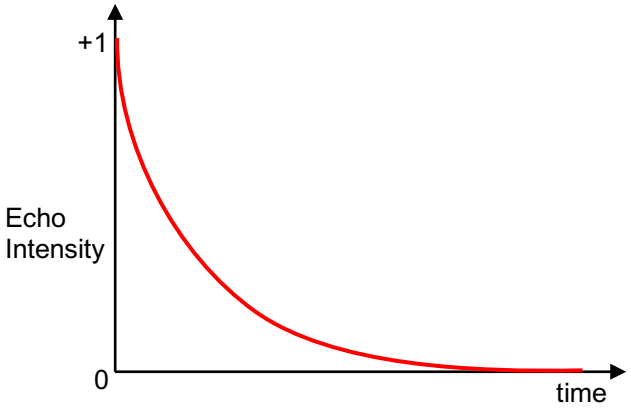


Fig. 2.4.4: a visualisation of how the echo intensity varies with time during a Hahn echo decay experiment.

Field swept echo detected EPR

Using the standard echo detected EPR pulse sequence whilst sweeping the magnetic field yields a field dependent EPR spectrum made using echo detection (field swept echo, FSE). The resulting spectrum resembles an absorbance envelope much like what would be recorded using cw-EPR (if the first derivative spectrum was integrated once).

Pulsed electron electron double resonance

Pulsed electron electron double resonance (PELDOR), also known as double electron electron resonance (DEER), is a method used for measuring the dipole-dipole interactions between two unpaired electron spins for distances around 20-80 Å.

The technique uses pulses at two different microwave frequencies to separately excite the spins; the perturbation of the inhomogeneous EPR spectrum by the 'pump' frequency is monitored at the 'observe' frequency. For there to be a change at the 'observe' frequency there must be a mechanism that relates the two regions of the spectrum (i.e. dipole-dipole interaction). For distances greater than about 20 Å this interaction is hidden in the inhomogeneous linewidth, but the PELDOR technique can resolve it.

The interaction is dependent on both the spin-spin distance (r) and the angle (θ) between the inter-spin vector and the applied magnetic field, B_0.

$$\upsilon_{dd} = \frac{\mu_0 g_1 g_2 \mu_B}{2h} \cdot \frac{1}{r^3} \cdot (3\cos\theta^2 + 1) \qquad (28)$$

where υ_{dd} is the dipolar coupling, μ_0 represents the vacuum permeability, and $g_1$and g_2 are the g-values associated with the two unpaired electron spins.

The most commonly used method for PELDOR in recent years employs a 4-pulse sequence. This involves a standard echo detected EPR sequence ($\frac{\pi}{2} - \tau_1 - \pi - \tau_1 - echo$) to generate a reporter echo followed by a second π pulse after a different time delay (τ_2) to invert the magnetisation and refocus the echo; all of this is performed (overall: $\frac{\pi}{2} - \tau_1 - \pi - \tau_1 - echo - \tau_2 - \pi - \tau_2 - echo$) at the 'observe' microwave frequency, $\nu_{observe}$. The second microwave frequency, ν_{pump}, is used for a single π pulse. This is set to begin at a time slightly prior to the first Hahn echo. The sequence is repeated with the 'pump' pulse displaced incrementally in time such that it 'sweeps through' the first reporter echo; the modulation of the first echo is reflected in oscillations of the monitored second echo's intensity. All timings for the pulses at $\nu_{observe}$ are kept constant.

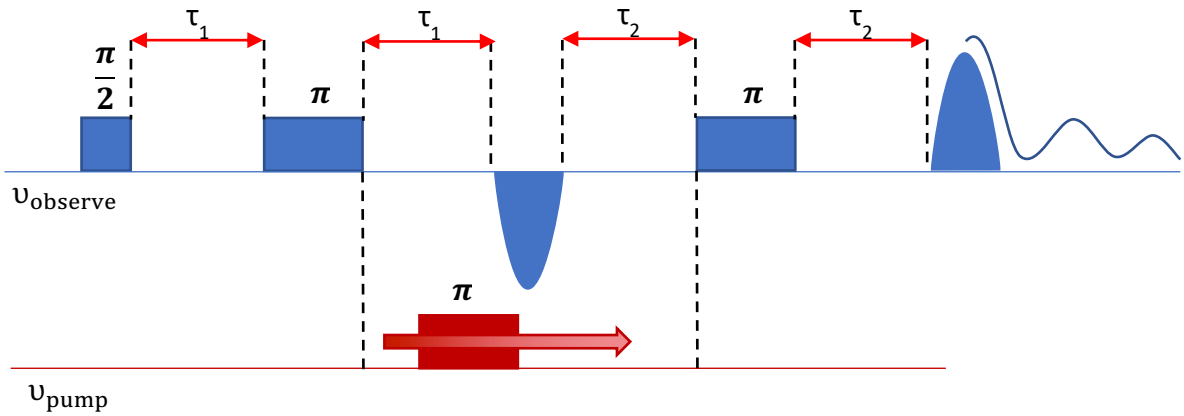


Fig. 2.4.5: schematic representation of the 4-pulse PELDOR pulse sequence showing the timings and lengths of the pulses at the different frequencies, and generation of an oscillating echo.

The first pulse at $\nu_{observe}$ ($\frac{\pi}{2}$) tips the magnetisation into the xy plane; this frequency is chosen to target the observers spins A, which precess at frequency ω_A. Inhomogeneous local fields and other factors dephase the magnetisation, where one of the factors is the contribution of the dipole-dipole interaction. This changes ω_A by $\pm\frac{1}{2}\omega_{AB}$; the sign relates to the spin state of the coupled pump spins B.

The second $\nu_{observe}$ pulse (π) after time τ_1 inverts the dephasing magnetisation and causes it to refocus such that the precession of the A spins is returned to ω_A and an undetected Hahn echo occurs after τ_1.

While the spins begin to relax again, and the magnetisation starts to dephase once more an inversion pulse (π) is applied at ν_{pump} to the B spins. This changes the resonant frequency of local A spins via the dipole-dipole interaction and causes the $+\frac{1}{2}\omega_{AB}$ contributions to the precession frequency of A spins to shift to $-\frac{1}{2}\omega_{AB}$ and vice versa.

The change in the dipole-dipole interaction causes a phase lag in the spins, altering the ability of the now-decoherent magnetisation to refocus when the final $\nu_{observe}$ pulse (π) is applied after time τ_2, thereby affecting the intensity of the detected Hahn echo (which occurs after τ_2). The extent of the phase lag is dependent on both the time of the pump pulse, t, and the dipolar frequency, ν_{dd}; by incrementally changing t the two factors can be separated. The PELDOR echo intensity variation is a function of the cosine of the phase lag, and the oscillations are proportional to ν_{dd}.

Unlike a standard FID, the refocused echo decay is controlled by more than just T_2; instead, the time constant for echo decay is called phase memory time, T_M. There are many contributors to T_M, of which transverse relaxation and spectral diffusion are likely the largest. Nuclear spins flipping and Brownian motion of molecules can cause spins to change their precession frequencies; this affects the ability of magnetisation vectors to refocus and thereby influences echo intensity.

The data recorded for this experiment is provided in the time domain. As explained above, the oscillations of the echo intensity occur at the dipolar interaction frequency; since this frequency is inversely proportional to the cube of the inter-spin distance, longer distances result in lower frequency oscillations and vice versa. This data actually encodes both the intermolecular and intramolecular interactions of spins in the sample. The intermolecular contribution is seen in the exponential decay (the 'background factor'), while the oscillations are a result of the intramolecular interactions (the 'form factor'). The background factor can be fitted with an exponential function and subtracted in order to best deconvolute the analysis of the form factor. Fourier transformation of this time-domain trace moves the data to the frequency domain, yielding the characteristic Pake pattern. Inter-spin distance can be extracted from this using the splitting of the peaks in the pattern.

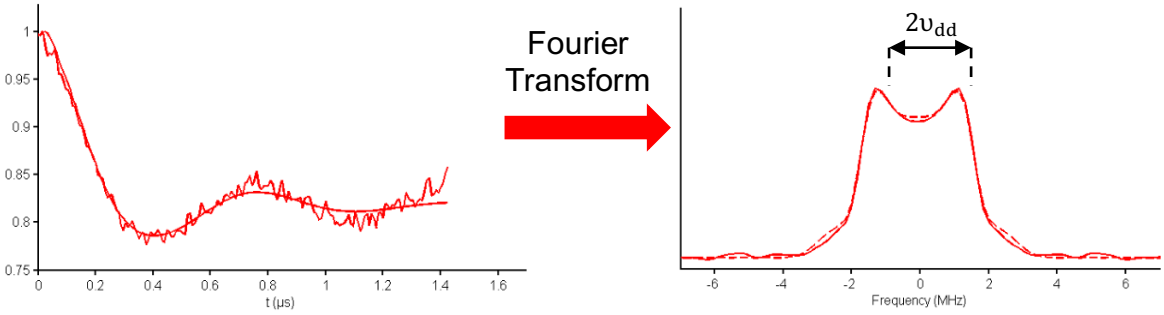


Fig. 2.4.6: examples of the time-domain trace recorded from a PELDOR experiment, and the characteristic Pake pattern that results from a Fourier transform of the data, creating a trace in the frequency domain. The distance between the peaks on the Pake pattern can be used to determine the dipolar frequency and therefore the inter-spin distance.

In reality, the distance is not calculated from the Pake pattern anymore. Small distortions in the time-domain trace are amplified in the process of Fourier transformation, and can have a large effect on the distance distributions calculated as a result. It is common practice nowadays to use software to apply mathematical

algorithms and regularisations to the data to reduce the impact of noise, anomalies and artefacts.

In order to set up a PELDOR experiment, there are a few prior experiments that must be done. First, an FSE is performed to locate the positions of the optimal pump and observe pulse frequencies. For the sake of simplicity (and to best relate to the majority of data presented in this thesis), the example of a nitroxide spin label will be used throughout, though it is possible to use other paramagnetic species such as metals. In general, pumping at the maximum in the EPR signal intensity gives peak pumping efficiency and modulation depth. The observe position can then be found at the low field edge of the nitroxide FSE, as it is the next highest point of signal intensity (after the pump position) and it is usually far enough from the pump position to avoid overlap problems. In order to calculate the frequency difference, the difference in magnetic field between the positions is taken from the FSE.

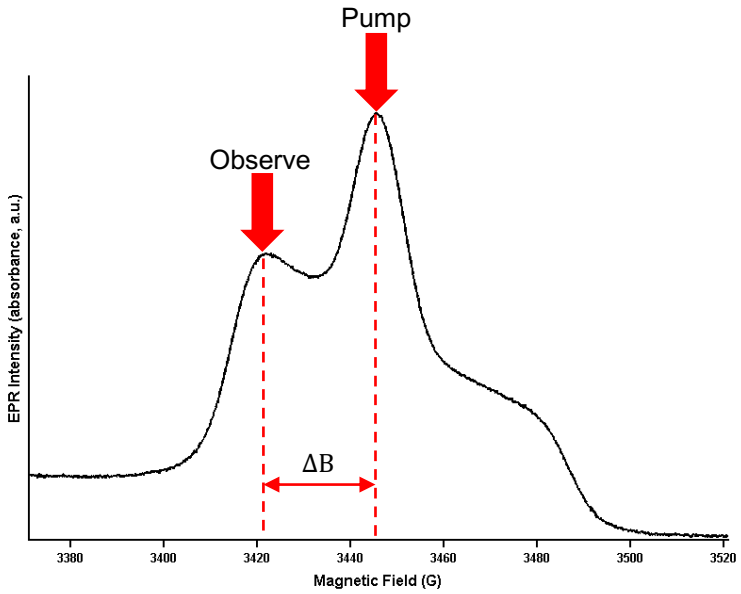


Fig. 2.4.7: example of a FSE experiment with the optimal pump and observe positions for PELDOR (and their difference in field) labelled.

$$\Delta B = B_{observe} - B_{pump} \tag{29}$$

$$\Delta\nu = \Delta B \cdot 2.83 \frac{MHz}{G} \tag{30}$$

This allows for the pump frequency to be determined from the observe frequency.

$$\nu_{pump} = \nu_{observe} + \Delta\nu \tag{31}$$

Typically for the pulse lengths used, the two frequencies must differ by more than 50 MHz to avoid overlap; the maximum and low field edge of a nitroxide FSE are usually around 65-70 MHz apart.

Additionally, relaxation parameters must be known to define the minimal time between repetitions of the pulse sequences, or Shot Repetition Time (SRT), which is limited by the spin-lattice relaxation time (T_1). Lower temperatures prolong T_1 but are necessary to produce a sufficiently long T_M; temperatures of around 50 K provide a good compromise between short T_1 and long T_M. At these temperatures spin-spin relaxation is driven by interactions with weakly coupled protons in the surrounding environment. Reducing the proton content of a sample (e.g. through exchange with deuterated solvent) can improve signal resolution.

Usually, if the power and lengths of the standard echo detected EPR sequence pulses have been optimised to produce maximum echo intensity, the π pump pulse set at 0 dB attenuation will achieve good modulation depth. If the separation between the pump and observe frequencies is greater than the bandwidth of resonator, the pump pulse power or length may need to be optimised.

It is important to note that the timing of the pump pulse must be carefully selected to avoid collision with the second and third observe frequency pulses as simultaneous or overlapping pulses will cause artefacts in the PELDOR data. A variable delay is used to time the start of the pump pulse and must be set at a greater value than the length of the second pulse in ν_{observe}. To avoid collision with the third ν_{observe} pulse during the pump pulse sweep, the number of points must be limited.

$$no.points < \frac{\tau_1+\tau_2-\tau_3-p}{\Delta t} \tag{32}$$

where τ_1 is the delay between the first two ν_{observe} ($\frac{\pi}{2}$ and π) pulses, τ_2 is the delay between the ν_{pump} and third ν_{observe} pulses, τ_3 is the delay between the start of the second ν_{observe} pulse and beginning of the ν_{pump} pulse, p represents the length of the ν_{pump} pulse, and Δt defines the increments of time by which the start of the ν_{pump} pulse is stepped (see Fig. 2.4.5 for 4-pulse PELDOR sequence).

2.5 Complementary methods

The phenomenon of dipolar interactions has been exploited by various forms of spectroscopy. The distance dependence of these interactions makes them a useful factor to study in order to extract structural information. This has already been explored in the context of the technique employed during the course of this thesis (PELDOR/DEER). In the following section, complementary and related techniques that also study dipole-dipole couplings (both magnetic and electronic) are researched, explaining their benefits, pitfalls and their experimental considerations.

Electron-electron dipolar methods

'2+1'

This is a single-frequency precursor of the modern PELDOR/DEER method, i.e. spins are pumped and probed at the same frequency[89]. The result, in theory, is an echo that oscillates at a frequency inversely proportional to the distance between the spins. Whilst the technique is relatively simple to perform, on a machine setup requiring only a single frequency, the oscillating echo is often obscured by hyperfine modulation and hindered by interference of the pump and probe spins. Hence, more recently the 4-pusle PELDOR method has typically been employed instead.

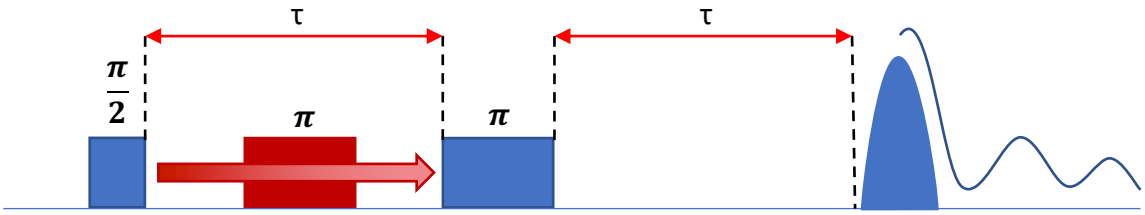


Fig. 2.5.1: schematic representation of the '2+1' pulse sequence showing the timings and lengths of the pulses and generation of an oscillating echo.

RIDME

A recently developed technique, Relaxation Induced Dipolar Modulation Enhancement (RIDME) uses the principles of DEER and applies them at a single microwave frequency[90]. During this pulse sequence, a standard Hahn echo is first set up ($\frac{\pi}{2} - \tau_1 - \pi - \tau_1 - echo$); the second spin is then inverted by longitudinal relaxation (T_1) before the final set of pulses. The resulting echo oscillates much like in a DEER experiment. This can, however, be obscured by hyperfine

modulation and transversal relaxation; operating at higher magnetic field/microwave frequency can reduce this (Q-band)[91].

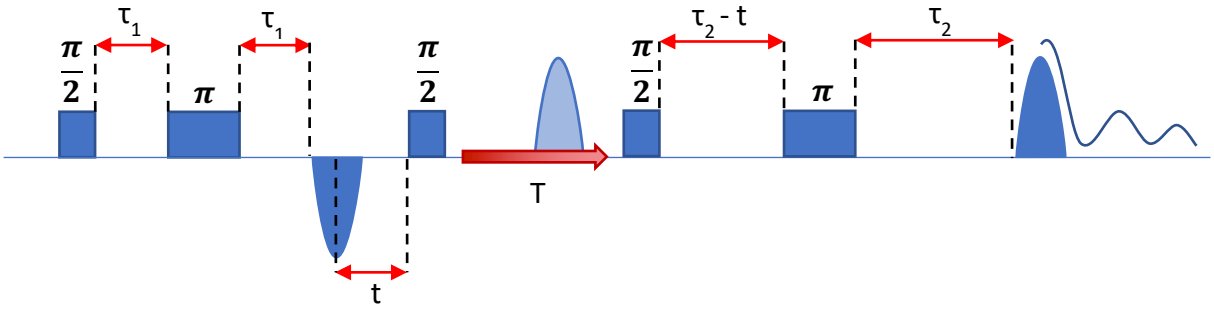


Fig. 2.5.2: schematic representation of the RIDME pulse sequence showing the timings and lengths of the pulses and generation of an oscillating echo.

DQC

Double Quantum Coherence (DQC) is a single-frequency pulsed technique that uses a set of pulses and phase cycling to suppress unwanted coherence pathways[92-93]. The method excites the entire spin system and should, in theory, have larger signal intensity and greater modulation depth relative to PELDOR (it also benefits from having no orientation selective effects) whilst measuring comparable distances. However, the technique is not routinely employed (the Freed group are the only consistent producers of this data, on a home-built K-band setup). The method is made difficult by the spectral width of nitroxide spin labels (it is hard to get the bandwidth to excite all spins).

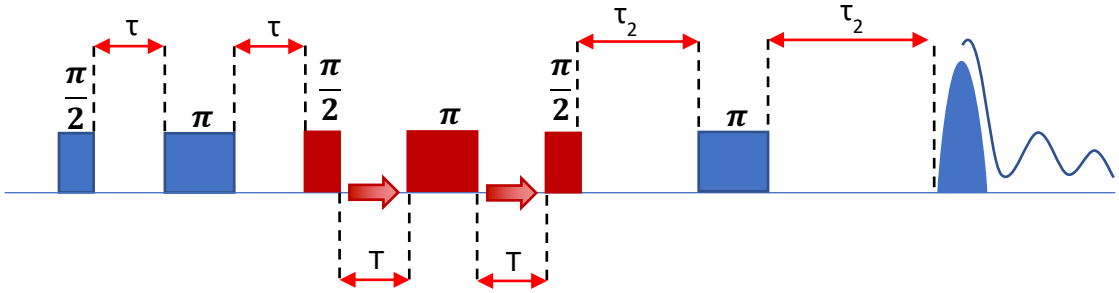


Fig. 2.5.3: schematic representation of the DQC pulse sequence showing the timings and lengths of the pulses and generation of an oscillating echo.

Electron-nuclear dipolar methods

cw-ENDOR

Continuous wave Electron-Nuclear Double Resonance (cw-ENDOR) uses microwave radiation to saturate the EPR transitions, then radio frequencies to desaturate. Hyperfine structure is observed as changes in EPR intensity as a

function or radio frequency. As such, these are small distances (<~1nm) being measured (dipolar coupling of nuclear moments around the moment of the unpaired electron) but can inform on the environment of an unpaired electron, often in the form of an intrinsic paramagnetic centre/cofactor[94].

Pulsed ENDOR

This technique, like other pulsed methods, holds the magnetic field constant, selected for where the EPR transition occurs. The microwave frequency of the EPR transition is also static, and the radio frequencies are swept through as in an NMR experiment. There are two main variants of the pulsed ENDOR method: Davies, and Mims. The Davies experiment essentially creates a 'hole' in the EPR spectrum using a MW pulse. RF radiation, corresponding with the NMR resonance, is then applied to fill that hole. Finally, more MW pulses are applied, resulting in an inverted echo, which reports on the EPR signal that is restored during the RF mixing period. Mims ENDOR is based on the 3-pulse ESEEM experiment in the MW frequency, with an RF mixing period between the second and third MW pulses (the timing between these pulses is constant, unlike the ESEEM experiment). The resulting echo intensity is measured as a function of the applied RF. The methods are fairly complementary in that they are implemented to study HFCs of differing magnitudes; Mims ENDOR is best suited for weak couplings (i.e. ~2 to 0.1 MHz), whilst Davies ENDOR is more appropriate for assigning stronger couplings. This complementarity simplifies the differentiation between strongly and weakly coupled nuclei and performing both experiments can inform on the type of nuclei around the unpaired electron, the distances between the nuclei and unpaired electron, the spin density distribution (i.e. the hyperfine structure) and on the electric field gradient at the nuclei (nuclear quadrupole; electric interaction between nonspherical nucleus and inhomogeneous electric field, only for nuclei where I >½)[79, 94-95].

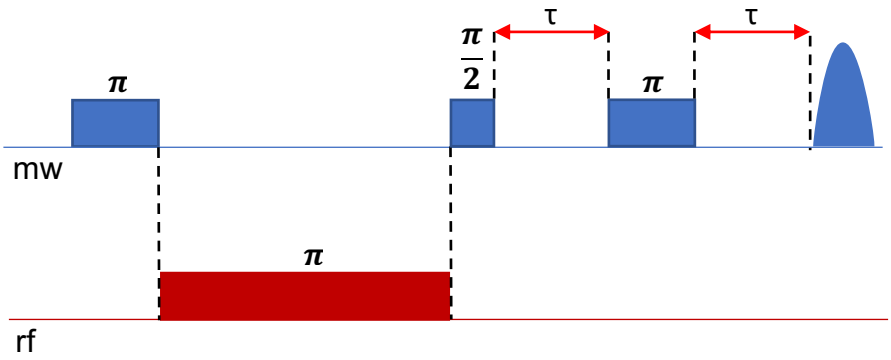


Fig. 2.5.4: schematic representation of the Davies ENDOR pulse sequence showing the timings and lengths of the pulses and generation of an echo.

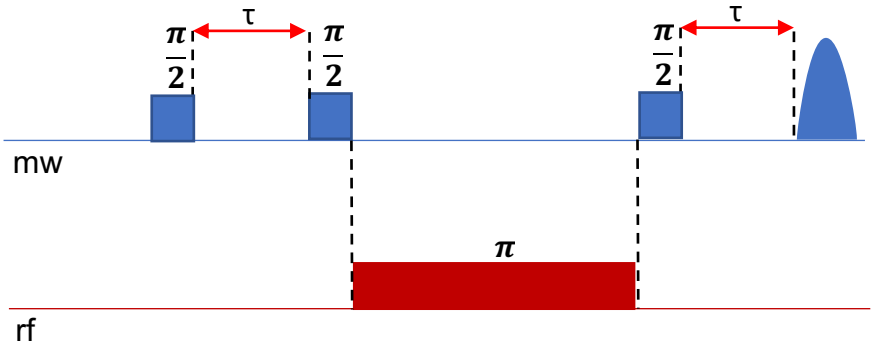


Fig. 2.5.5: schematic representation of the Mims ENDOR pulse sequence showing the timings and lengths of the pulses and generation of an echo.

ESEEM

Electron Spin Echo Envelope Modulation (ESEEM) uses a number of microwave pulses (typically 2 or 3) to create a spin echo; the time between two of the pulses is varied and the echo usually decays exponentially and depends on the spin-lattice and/or the spin-spin relaxation. Sometimes the decaying echo is modulated due to weak interactions with magnetic nuclei in the environment around the paramagnetic centre[96].

Electron Spin Echo Envelope Modulation (ESEEM) is typically done with two or three MW pulses, creating an electron spin echo. When the delay between two of the pulses is increased, the echo intensity typically decays exponentially as a result of spin-lattice and spin-spin relaxation (T_1 and T_2 respectively); in some cases, the decay is modulated as a result of weak HFCs. The decay of the echo in 2-pulse ESEEM is dominated by T_2 (usually in the microsecond time scale), which is often too short to be useful. In 3-pulse ESEEM (time is varied between the 2nd and 3rd pulse), the decay is governed by the T_1, which is usually much longer than

T_2. A Fourier transformation of the decaying echo's modulation allows the extraction of the frequencies of the HFCs. The modulation depth of these traces is related to the HFC and the nuclear quadrupole interactions (where there are nuclei with I >½)[97].

Fig. 2.5.6: schematic representation of the ESEEM pulse sequence showing the timings and lengths of the pulses and generation of an echo.

HYSCORE

HYperfine Sublevel CORrElation (HYSCORE/HYSCOR) is a 4-pulse ESEEM experiment done in two dimensions; on one axis the timing between the 2nd and 3rd pulse is varied (i.e. t_1), and along another axis the timing between the 3rd and 4th is varied (t_2). The resulting modulated decays are Fourier transformed in both dimensions, giving a 2D spectrum in the frequency domain. This appears as four quadrants but can be simplified to two as the upper right/lower left are equivalent, and the upper left/lower right are equivalent. Nuclei with HFC > Larmor frequency appear as cross-peaks in one quadrant whilst those with HFC < Larmor frequency appear in the other. HYSCORE spectra can be complicated when nuclei with I >½ are involved, due to peaks arising from the additional nuclear Zeeman levels and the influence of nuclear quadrupolar interactions. Despite this, the resulting complex spectra of systems with multiple nuclei are still manageable, where they would be potentially impossible to deconvolute from standard ESEEM data[98-99].

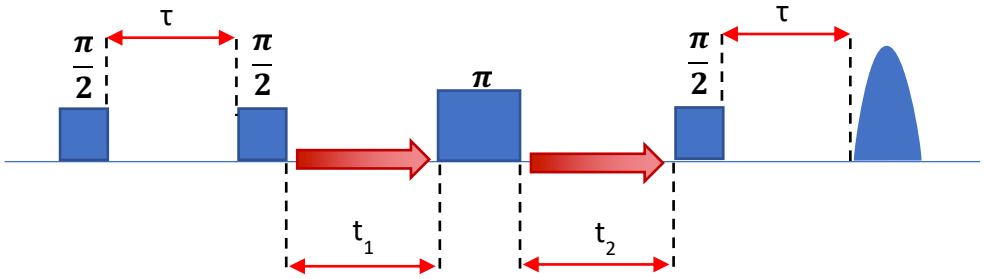


Fig. 2.5.7: schematic representation of the HYSCORE pulse sequence showing the timings and lengths of the pulses and generation of an echo. This is a two-dimensional experiment; t_1 is varied whilst t_2 is held constant in one dimension, and t_1 is held constant whilst t_2 is varied in the other dimension.

EDNMR

ELDOR-Detected NMR (EDNMR) uses selective microwave pulses to pump formally forbidden transitions (where both electron and nuclear spins change their direction; $\Delta M_S = \pm 1$, $\Delta M_I = \pm 1, \pm 2$, etc). It has higher sensitivity compared to the more conventional ENDOR methods. Additionally, it is less affected by fast T_1 relaxation. However, the EDNMR experiment typically resolves broader spectral lines than the ENDOR experiment. This method has been hindered by the 'central blind spot', where spin-forbidden transitions close in frequency (~5-10 MHz) to the allowed EPR transition are obscured due to the pumping of the allowed transition. Instrumentation restrictions such as limited resonator bandwidth cause the forbidden transitions to be suppressed when they widely off-resonance with the allowed EPR transition. However, operating at higher magnetic field has helped to mitigate these problems. The pulse sequence uses two microwave frequencies, one of which is a strong, long pulse to excite the nuclear transitions of the spin manifold (unlike conventional ENDOR methods which use radio frequency for this), known as the 'high-turning-angle' (HTA) pulse. A Hanh echo sequence is then used at the other microwave frequency for detection[100].

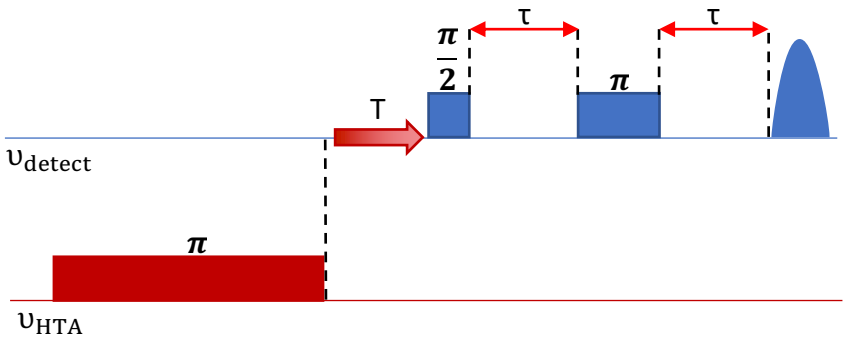


Fig. 2.5.8: schematic representation of the EDNMR pulse sequence showing the timings and lengths of the pulses at the different frequencies and generation of an echo.

PRE

Paramagnetic Relaxation Enhancement (PRE) is an NMR technique that uses the electron-nuclear dipolar coupling to obtain structural distance information[101] that can complement Nuclear Overhauser Effect (NOE) restraints (explained in section 'NOE'). The method can measure distances in the range of 15-24 Å for the lone spin of a nitroxide radical. The distances between the spin label and NMR active nuclei can be determined from the effect of the paramagnetic species on the relaxation rates of the magnetic nuclei[102]. This technique can be also be used to screen for the binding of small molecules as PRE will be transferred onto small molecules that bind macromolecules in the intermediate exchange regime (like STD-NMR, see section 'STD-NMR').

Nuclear-nuclear dipolar methods

The interaction of two nuclear spins can also occur through nuclear-nuclear dipolar coupling.

NOE

The Nuclear Overhauser Effect (NOE) can be used to obtain distance restraints and is used in structure determination. The phenomenon is defined as the transfer of nuclear spin polarisation from one population of spin-active nuclei to another via cross-relaxation. Radio frequency (RF) radiation is used to saturate or perturb one population and the change in resonance intensity of the other nucleus population is measured as a consequence of the proximity between the two[101, 103].

STD-NMR

Saturation Transfer Difference (STD) NMR similarly uses saturating RF pulses. The method allows the detection of transiently binding ligands to macromolecules. The saturation is targeted at the resonance regions of the protons of the protein or macromolecule in question (it is essential that this resonance position does not overlap with that of the ligand being investigated). Any ligand that binds to the macromolecule (with an expected K_d in the low mM to high nM range) will receive saturation via NOE and the signal of the ligand will change. A difference spectrum of the non-saturated vs saturated conditions will yield only the signals of the binding ligand(s). Often it is possible to elucidate which parts of the ligand are in contact with the macromolecule (i.e. epitope mapping) as they will exhibit the largest difference between the two conditions. Additionally, it is possible to study the K_d of a ligand with this method by measuring the STD effect as ligand is titrated in, though certain experimental parameters must be closely controlled to obtain accurate results[104-106].

RDC

Residual Dipolar Coupling (RDC) can provide relatively long-range distance information (when compared with traditional NMR or NOE techniques). This method requires alignment of the molecules being studied in order for the averaging of the dipolar couplings to be incomplete. This technique can inform on a range of things including the global fold of a protein or complex and the dynamics of relatively slow (>10^{-9} s) motions in proteins. This method is not affected by spin diffusion, which often makes NOE studies size limited (to around 25 kDa)[107-108].

Other dipolar methods

Another technique that utilises dipole-dipole interactions to measure inter-probe distances, similarly to PELDOR, is Förster Resonance Energy Transfer (FRET, sometimes referred to as Fluorescence Resonance Energy Transfer) spectroscopy. Unlike EPR and NMR methodologies, this harnesses the interactions of electric dipoles (rather than magnetic dipoles). These dipoles are induced by the separation of opposite charges; magnetic dipoles, by contrast, arise due to the movement of charges.

The excitation of a molecule from the electronic ground state to an excited state is induced by the absorption of photonic energy. The phenomenon of fluorescence then arises due to the subsequent emission of energy (in the form of photons), bringing the molecule back to its ground state.

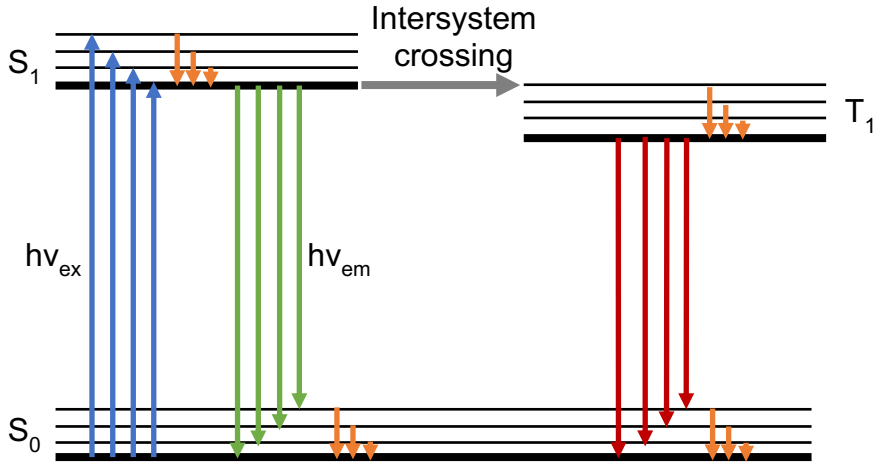


Figure 2.5.9: an adapted form of the Jablonski diagram, demonstrating the excitation of a molecule from a ground singlet state, S_0, to an excited singlet state, S_1. T_1 is the excited triplet state that arises from intersystem crossing, represented by a grey arrow, due to formally forbidden spin-exchange. Absorbance of excitation-wavelength light is shown by blue arrows, emission of light via fluorescence is shown by green arrows, vibrational relaxation is represented by orange arrows and emission of light via phosphorescence is shown by red arrows.

The probes or labels used in this methodology are light-sensitive fluorophores (or fluorochromes) with specifically reactive linkers that allow targeted labelling of the system in question, much like spin labels in SDSL. Here, two different probes are used in order for the non-radiative energy transfer to occur; the donor fluorophore is excited by a certain wavelength of incident light and transfers the excitation to a nearby acceptor fluorophore through electronic dipolar coupling, which then emits light at a different wavelength via fluorescence in order to relax back to the ground state. Typically, the labels used for studying conformational dynamics with FRET are small organic molecules (<1 nm).

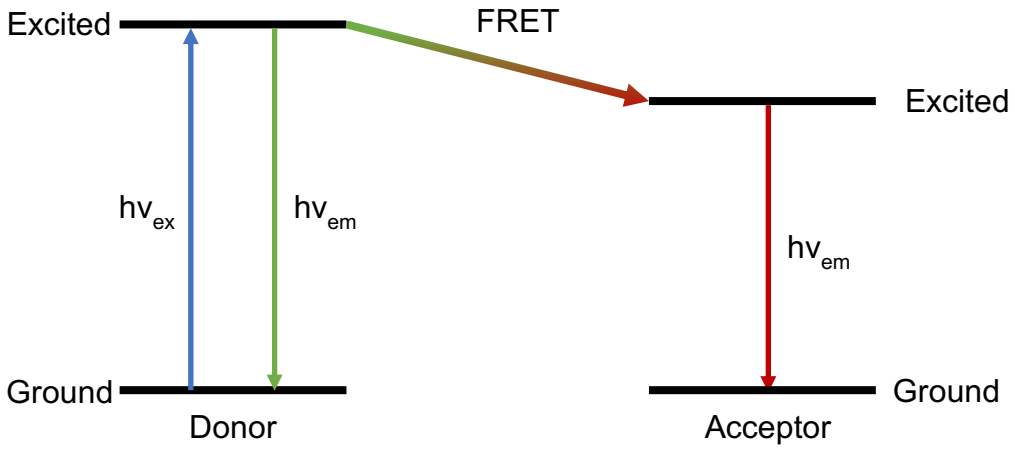


Figure 2.5.10: a simplified electronic diagram of the FRET phenomenon. The FRET donor is excited from its ground to excited state by a photon of energy (blue arrow). The donor can then either relax back to its ground state by emission of a photon of lower energy (fluorescence, green arrow), or by donating its energy to a nearby acceptor through FRET (gradient green-red arrow). The now-excited acceptor can then relax back to its ground state by emission of a photon of energy (fluorescence, red arrow).

The efficiency, E, of this energy transfer is indicative of the distance between the donor and acceptor fluorophores, r, and is analogous to the dipolar frequency in PELDOR.

$$\mathrm{E} = \frac{1}{1+(r/R_0)^6} \tag{33}$$

Where R_0 is the Förster radius. This is the distance between the donor and acceptor fluorophores at which energy transfer efficiency is 50%. This depends on the overlap integral of the donor emission and acceptor excitation spectra, J, (demonstrated in fig. 2.5.12) and their orientation.

$$R_0 = \sqrt[6]{\frac{9\cdot\ln(10)}{128\pi^5\cdot N_A}\cdot\frac{\kappa^2\cdot Q_D\cdot J}{n^4}} \tag{34}$$

Where N_A is Avogadro's number, κ^2 is dipolar coupling orientation factor, Q_D is fluorescence quantum yield of the donor in the absence of the acceptor fluorophore and n is the refractive index of the nearby medium.

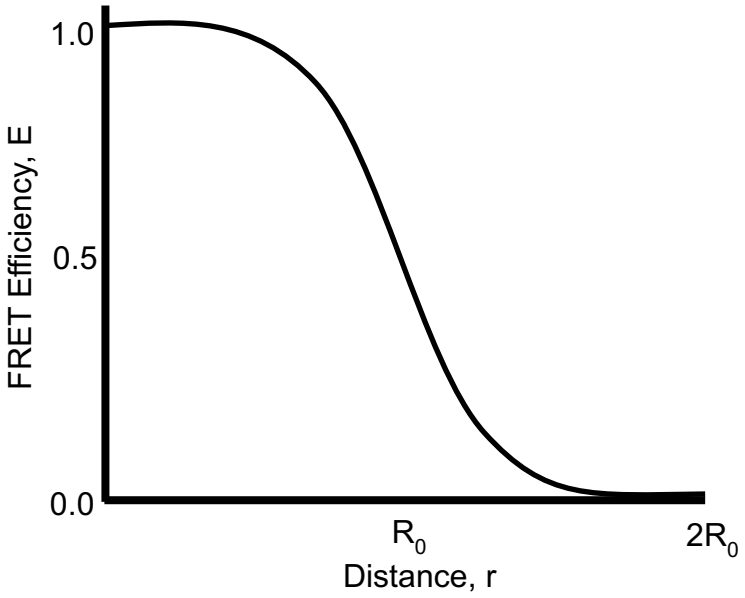


Figure 2.5.11: representation of the distance dependence of FRET; the FRET efficiency, E, is a steep function of distance when that distance nears the value of the Förster radius, R_0.

There are three important factors for FRET experiments: the donor emission spectrum should significantly overlap with the acceptor excitation spectrum; the donor-acceptor distance should be between 1-15 nm; and there should be favourable orientation of the donor and acceptor (such that transition dipole moments are properly aligned with respect to one another)[109].

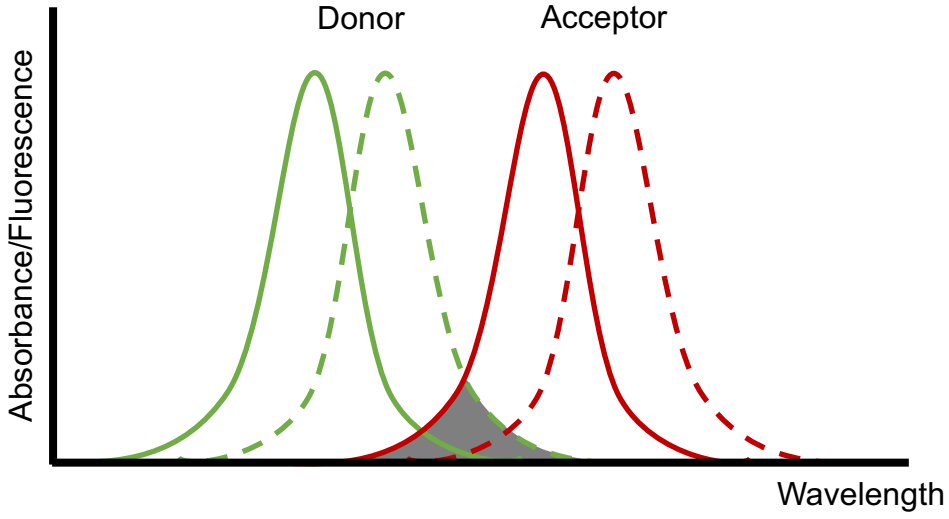


Figure 2.5.12: representation of favourable spectral overlap for donor/acceptor pairs for FRET spectroscopy, showing how donor (green) emission and acceptor (red) excitation spectra should overlap (grey) significantly. The excitation spectra are shown in solid lines and the emission spectra in dashed lines.

3 EPR on Membrane Transporters

The following is an updated adaptation of a previously published review of research using EPR spectroscopy to study membrane transporters, specifically those belonging to the classes being explored in this thesis (previously published paper Mullen *et al*., 2016[110]).

3.1 Context

Membrane transporter proteins comprise a wide variety of proteins which underpin a diverse range of functions - from the import of vital solutes to the efflux of a multitude of drugs – making this class of proteins a clear target for a plethora of therapeutic treatments in the bid to tackle a myriad of medical challenges. Resolving their three-dimensional structures at a molecular level is an integral part of understanding their functional mechanisms. However, the issues that arise when attempting to solve the structures of such dynamic proteins often render the more typical methods unsuccessful. Such challenges include the fact that it can be difficult to solubilise and stabilise membrane-associated proteins for crystallisation, and that X-ray crystallography relies on the arrangement of protein into the regular assembly of a crystal lattice, influenced by the crystal liquor around it (i.e. high salt concentration, heavy metals, stabilising agents). This often means that only single, static conformations can be sampled, leaving much of the dynamic cycle of these proteins unobserved and continually debated[111]. The issues surrounding crystallography of membrane-associated proteins and complexes is reflected in the numbers deposited in the Protein Data Bank[3] (PDB; www.rcsb.org); of the ~140,000 protein structures (the majority of these , ~127,000, are X-ray crystal structures) only around 2,700 are membrane proteins[4], the first of which was only resolved in 1985[112].

Additionally, membrane transporters may consist of large proteins, or can exist as multidomain complexes making them difficult to measure for size-limited methods such as nuclear magnetic resonance (NMR) spectroscopy (though these boundaries are being pushed)[113], which emerged as a method of resolving solution-based structures of proteins, particularly for those that are intrinsically disordered, and therefore not determinable by X-ray crystallography. Beyond the size-limitation, it requires isotopically labelled protein, which is expensive, and each measurement requires the residue sidechains of each amino acid to be assigned manually, which is both labour intensive and time-consuming. There are

numerous examples in the literature and in the PDB which demonstrate that membrane proteins are also especially susceptible to conformational changes due to environmental changes e.g. when being extracted from their native membrane[114].

Cryo-electron microscopy (cryo-EM) has received much attention of late as an alternative to X-ray crystallography and NMR; the 2017 Nobel Prize in chemistry was awarded for the use of this technique in determining high-resolution structures of biomolecules in solution[115]. Recent technological advances have developed this technique to near-atomic resolution and has been embraced as a method for studying proteins for which it is difficult to grow crystals. Since the wavelength of an electron is shorter than that of X-ray radiation (Compton wavelength of an electron is in the picometre scale, whilst the X-ray radiation used in crystallography is in the nanometre scale), better resolution may eventually be possible with cryo-EM compared to X-ray crystallography.

In this method, an electron beam is fired at a frozen protein solution and scattered electrons create a magnified image. Transmission electron microscopy (TEM) is typically performed on samples under vacuum, which would dehydrate an aqueous sample, with powerful electron beams which would heat and destroy the biomolecules. However, the method has been developed to use ‘vitreous’ samples, wherein liquid nitrogen-cooled ethane can freeze an aqueous protein sample rapidly to give a disordered glass. The water molecules are frozen so quickly they cannot form an ordered ice lattice; thus, the protein remains in its native state. The cryo-EM method also uses ‘gentler’ electron beams to prevent sample destruction. Multiple angles are used to study the samples, and computational methods can then collate this array of images of randomly oriented proteins into a three-dimensional structure of the biomolecule[116]. Currently, the lower size limit for proteins that can be imaged with cryo-EM is about 50 kDa to about 3 Å resolution[117]. Larger structures have been measured to higher (i.e. >2 Å) resolution[118].

Despite these limitations, cryo-EM is an attractive choice for elucidating structural data, particularly for membrane proteins and complexes, as they can be solubilised in detergent or reconstituted into lipid environments. Though this limits the resolution due to poor signal to noise, it also provides a unique insight into the native topology and arrangement of these systems.

However, there are few biophysical methods which can provide both structural and dynamic information on the molecular architectures, routinely offering reliable and precise distance determination over a wide range (up to ~10 nm). Electron Paramagnetic Resonance (EPR) spectroscopy is one such technique, another being the fluorescence-based technique Förster Resonance Energy Transfer (or FRET) spectroscopy. For the structural biologist, these techniques often yield similar results, and are complementary in their requirements for labelable variants of the native protein being studied (both typically use cysteine sidechain thiol chemistry for the attachment of their respective probes). As with all biophysical techniques, both have their pros and cons. EPR labels are much smaller than the fluorescent probes used in FRET, with respect to both the smaller chains connecting the labelled residue and the 'active' part of the probe. This allows EPR to study smaller distance changes and is a more accurate measure of the local dynamics being investigated. Additionally, FRET requires at least two chemically different probes to act as donor and acceptor. Conversely, EPR uses only one type of spin label, making studying oligomers less complex and the labelling process less involved. Under optimised conditions, PELDOR can measure distances up to 160 Å (an upper limit that is being pushed by ongoing research)[119]. FRET, however, works in a liquid environment (i.e. more native-like than the frozen solutions required for pulsed EPR measurements) and is more sensitive, making it more accessible for time-resolved measurement and single-molecule experiments; its use in studying ABC transporters was covered well in a review a few years ago[120].

Both are now being more widely used, especially in combination with molecular dynamics (MD) approaches as complementary structural biology techniques, being called upon to overcome such obstacles mentioned above. EPR can be harnessed in many ways to provide structural constraints that aid visualisation of the structure and dynamics of a protein in a range of states throughout its mechanistic cycle. Importantly, the method can also be applied to instances where the protein is situated within a membrane environment (either in vivo or reconstituted into proteoliposomes or even styrene maleic acid lipid particles, SMALPs[121]), allowing it to be used to obtain information about the protein in or as close to its native crowded environment as possible. The studies reviewed here aim to demonstrate the application of EPR to investigate the structural elements

and mechanistic dynamics of a range of primary (ABC-type transporters) and secondary active membrane transporters.

Initially, EPR spectroscopy was developed as a physical magnetic resonance technique following its discovery in the USSR in 1944, though it soon developed into a key biophysical method for studying biological systems containing intrinsic paramagnetic species, especially those involved in bioenergetic electron transfer and metalloproteins. The development of this method over the past seven decades has resulted in the emergence of a broad range of techniques that can be employed to study paramagnetic species and their surrounding environments[122-124]. These methods have been reviewed in several excellent monographs and review articles over the past 10-20 years and the reader is referred to the following review articles and references therein[87, 125-127].

Arguably one of the major turning points in the development of EPR methodology from being a rather niche technique into a key structural biology technique was the introduction of site-directed spin labelling (SDSL). As described previously, this is a technique that allows attachment of a stable paramagnetic probe to a specific site (often made possible using site-directed mutagenesis) within proteins or nucleic acids without any intrinsic paramagnetic centres, thus making previously EPR-inactive diamagnetic proteins now accessible to this technique. One of the leading figures in this development is Prof. Wayne Hubbell (UCLA, USA) who has reviewed recent developments in the field[128-130]. Much of these recent advances have resulted from essential improvements in instrumentation, especially in sensitivity[131-132]. Compared with NMR or X-ray studies the SDSL methodology does not yet [131-133] provide global structural models but rather contributes sparse structural constraints which require essential supporting analysis packages[134-135]. Clearly questions can be asked with regards to the site-directed modification of proteins: does a labelled protein still function correctly; how physiologically relevant is measuring a distance (or distances) at cryogenic temperatures; are purified, detergent-solubilised membrane proteins true reflections of a functional protein in a membrane. However, with that in mind here we aim to focus on the key information that has been extracted from the application of EPR spectroscopy to the study of a range of primary (ABC-type transporters) and secondary active membrane transporters, especially determination of longer distances[136], distance editing[137], resolving not only distances but also relative orientations[138-140], measuring distances under physiological conditions[93] or in native environments

(whole cells or oocytes)[141-142]. To summarise, EPR – especially in combination with SDSL – is able to provide a wealth of detailed information both for the mechanistic structural biologist and the computational chemist to contribute to advancing our understanding of membrane transporters (fig. 3.1.1).

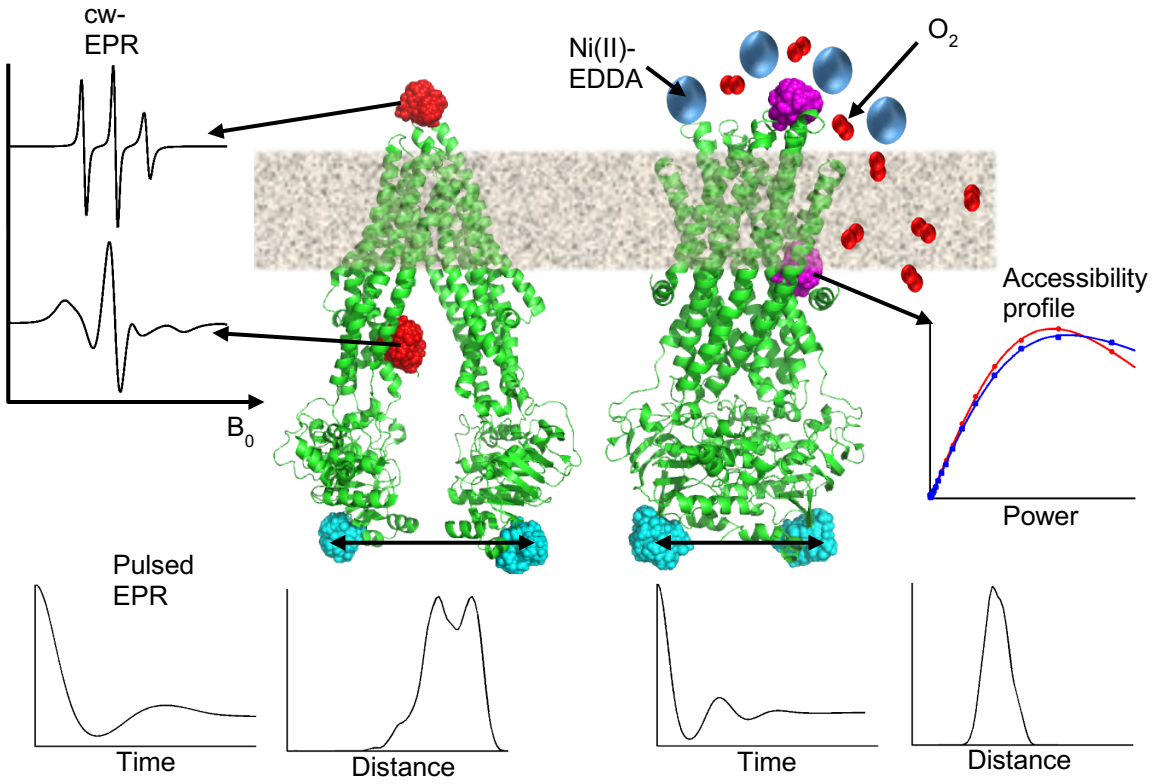


Fig. 3.1.1: a schematic view of how EPR can be used to study membrane transporters using in silico attachment of spin label rotamer libraries to the crystal structures of P-glycoprotein (Mus musculus, PDB 3G5U, left) and MsbA (Salmonella enterica, PDB 3B60, right). Clusters in red demonstrate the use of cw-EPR mobility studies to investigate local structure and topology; pink shows the use of accessibility studies and paramagnetic quenchers (blue spheres are membrane-impermeable Ni(II)-EDDA, and red are membrane-permeable O_2) to elucidate local environment; blue presents the use of pulsed EPR for distance measurements to study conformational dynamics.

The aim here is not to cover all the published EPR data on the large family of membrane transporters since this is a vast area. Instead it is to focus on recent developments on a few key areas including primary active transport and ATP-binding cassettes (ABC) as well as facilitated diffusion/secondary active transport systems. Figure 3.1.2 shows some examples of the families of structures of the various transporters for which spectroscopic data is being reported here.

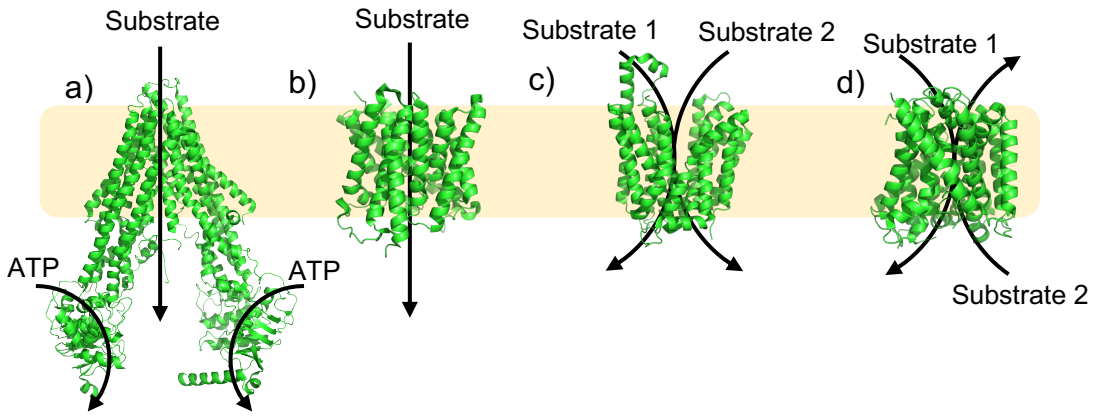


Fig. 3.1.2: a comparison of different types of transmembrane transporter: (a) The multidrug transporter P-glycoprotein (Caenorhabditis elegans, PDB 4F4C), an ATP-powered pump; (b) the uniporter cyclodextrin glycosyltransferase (Bacillus circulans, PDB 1XQE); (c) fucose/proton symporter FucP (E. coli, PDB 3O7P) and (d) glutamate/GABA antiporter GadC (E. coli, PDB 4DJK).

3.2 Primary Active Transport

Substrate Binding Proteins

Substrate-binding proteins (SBPs) are a class of proteins that form complexes with membrane proteins. Commonly used for transport or signal transduction, SBPs can be found to be associated with ATP-binding cassette (ABC) transporters[30] and more recently in other membrane protein complexes[143-144].

Despite the wealth of structural data that exists for this subclass of proteins on the PDB, EPR studies of SBPs have been mainly limited to the maltose transporter $MalEFGK_2$. The maltose-binding protein (MBP, or MalE) is associated with the ATP-binding $MalFGK_2$. Hall *et al*. used both room temperature (RT) and low temperature (LT) continuous wave (cw)-EPR techniques to demonstrate that there are two modes of ligand binding in the MBP; one being active and the other inactive, depending on the ligand orientation[145]. They argue that the inactive mode hinders the closure of MalE, which prevents the complex from interacting with the inner membrane domain. They found that ligand binding did not affect the mobility of the spin labels attached to the protein and deduced that the spectral broadening which occurred upon the addition of maltose was a result of an increase in spin–spin interaction in the double spin-labelled proteins. Using low temperature EPR techniques they found that upon maltose binding the spin distance between the two domains changed from 16.5 to 10.5 Å (1 Å = 0.1nm).

Further to this, Austermuhle *et al*. studied the interaction of MBP with the transmembrane domain of the transporter[146]. ATP hydrolysis during the closure of the MalK dimer interface coincides with the opening of MBP. They deduced this from the spin–spin interaction between two spin labels: one on MBP and the other on the transmembrane domain. In a vanadate-trapped transition state intermediate, all free MBP became tightly bound to $MalFGK_2$ and spin labels in both lobes were completely immobilised. In addition, spin–spin interactions were lost, suggesting that MBP was in an open conformation.

Building on previous work, Orelle *et al*. discovered that both MBP and ATP are required for the closure of the nucleotide binding domain (NBD) MalK[147]; SDSL and cw-EPR were used to study the opening and closing of the NBD. They found that after ATP hydrolysis the NBD is in a semi-open configuration, which is distinctly different from the open state. They propose that the release of inorganic phosphate (Pi) happens concurrently with the reorientation of the transmembrane domain to an inward-facing conformation.

Grote *et al*. used doubly spin-labelled mutants of the maltose transporter $MalFGK_2$-E to further investigate the mechanism of the transport cycle mediated by transmembrane signalling[148]. The EPR data revealed that MBP is bound throughout the transport cycle. The reciprocal communication across the membrane gives information on the effect that MalE has on ATPase activity, as they regard it to be an important mechanistic feature of receptor-coupled ABC-transporters. They suggested that further characterisation of the inter-domain relationships during substrate transport is needed to elucidate details of the conformational changes brought about by the SBP. They posited that studies of spin-labelled MalE variants should be undertaken to investigate distances between positions within both the SBP and the transporter. It was noted that although their model is specific to the maltose transporter, similarities may be found in the BtuCD-F complex.

More recently, double quantum coherence (DQC) EPR – an alternative pulsed EPR method for determining inter-label distances – was used to examine ligand binding in both the native/folded and intermediate molten-globule (MG) states of MalE. The studies showed that both the open and closed states existed in the absence of ligand maltose, in both the native/folded and MG states; the MG state is achieved here using an environment at pH 3.2. Additionally, despite the loss of

tertiary structure (as verified by methods such as circular dichroism spectroscopy), the MG state exhibits nativelike structure, particularly around the binding pocket[149].

Although it is often used as a model for type 1 ABC transporters, $MalFGK_2$ has some distinct differences from the vast majority of other transporters of this kind. As such, other transporters have been studied using EPR. One such system is the histidine transporter, $HisQMP_2$. Sippach et al. used DEER (or PELDOR) spectroscopy to study the conformational changes of $HisQMP_2$[150]. The system works with an SBP similar to MalE, HisJ, that has a high affinity for histidine. The results showed that the distances measured resemble those of the maltose transporter throughout the cycle. In the presence of HisJ, the closed conformation of the NBDs is found. These observed conformational changes led to the proposal that there are three different conformations of NBD; open, semi-open and closed and that these conformations are regulated by SBP binding.

Pneumococcal surface adhesin A (PsaA) plays a vital role in the virulence and colonisation of the *Streptococcus pneumoniae* pathogen[151] and is a vital component of the Mn^{2+}-specific ABC-type transporter, PsaBCA[152]. The Mn^{2+} ion serves to negate the effects of oxidative stress on the bacteria by neutralising reactive oxygen species (ROS) due to its role in superoxide dismutase[153]. A combination of SDSL and EPR approaches have been taken to investigate the protein environment. In order to determine the conformational flexibility of the PsaA protein, five variants (L56C, S58C, S266C, I125C and I236C) were labelled with the nitroxide spin label MTSL and characterised using cw-EPR. The combination of MD simulations and cw-EPR spectra allowed for elucidation of the flexibility of the PsaA protein lobes, hypothesising various interactions with other proteins comprising the PsaBCA complex[154]. Since the intrinsic metal (Mn^{2+}) is also paramagnetic, these single variants can be used for distance determination using multi- or rather high frequency (34 and 94 GHz) cw-EPR as has previous been demonstrated[155].

ABC-type Import

The MolBC type II importer transports molybdate ions in bacteria, vital for carbon and nitrogen cycling[156]. cw-EPR spectroscopy was applied to MolBC reconstituted into proteoliposomes to elucidate the transport mechanism; MTSL was used to label the S180C and D173C residues on gates of MolB, which are responsible for the transfer of molybdate ions through the transporter. It was deduced that in the

presence of ATP the conformation of the periplasmic gate converts to an open position allowing passage of the substrate. This was seen through an increase in mobility at the D173C label position. Once ATP was hydrolysed a shift back to the original conformation was noted. The MolBC-S180C in complex with substrate-free MolA was seen to decrease spin label mobility; however, on addition of ATP, the S180C label demonstrated increased mobility[157].

PELDOR and cw-EPR have been applied to the rest of the maltose ABC transporter $MalEFGK_2$[148]. MalF-P2 domain double mutants were spin labelled, along with a triple mutant with sites present in the MalK and P2 domains. For the MalF-P2 double mutant, inter-spin distances were determined using both cw-EPR and pulsed EPR methods, allowing for determination of rearrangements of the protein during ion transport. Furthermore, the labelled MalK domain showed communication between protein domains, and analysis of the changes of conformation in the cytoplasmic and periplasmic domains of the ABC-transporter highlighted the importance of communication between these two domains for transport. Shifts in spectra were evident on binding, showing the dependence of ATP and MalE for conformational change of the protein[148].

The conformations undertaken by the BtuCD-F ABC-transporter have also been elucidated through EPR analysis of spin-labelled mutants reconstituted in proteoliposomes. This complex is responsible for the translocation of vitamin B12 in Escherichia coli[158-159]. Resulting spectra of BtuCD compared with BtuCD-F highlighted the differences in conformation adopted. After the addition of BtuF-B12, coupling between spin labels at residues 141 and 168 within BtuC was seen, suggesting that the distance between the two labels had decreased[160].

ABC-type Efflux

The multidrug exporter MsbA shares sequence similarity with a class of ABC transporters, which are linked to multidrug resistance and cancer development[11, 161]. PELDOR studies have revealed large-scale movement in opposite directions in the periplasmic and cytoplasmic parts of the transporter upon ATP hydrolysis. A 33 Å change in distance was measured upon formation of the trapped post-ATP hydrolysis intermediate, and inhibited structural changes caused by lipopolysaccharide (LPS) binding. Results also indicate that ATP hydrolysis powers transport of LPS into an open cytoplasmic chamber before its translocation by alternating access, involving conformational changes of 10–20 Å[162].

A recent study utilising a plethora of complementary biophysical techniques (complex solid-state NMR, pulsed EPR and MD simulations) identified a novel nucleotide binding site in the MsbA NBDs, at a conserved area called the Q-loop, wherein a reverse adenylate kinase (rAK)-like reaction takes place under low-ATP conditions. This is a phosphorylation reaction which converts 2 ADP molecules (populating the NBD after the ATP-hydrolysis step) to one molecule of ATP and one of AMP. Additionally, the Mg^{2+} was replaced by Mn^{2+} for pulsed EPR studies, measuring Mn^{2+}- Mn^{2+} dipolar couplings, which showed the metal cofactor binding site (necessary for catalysing ATP hydrolysis) was not disrupted by this novel nucleotide binding site. This investigation gives some evidence of how the ATP-hydrolysis and rAK cycles of MsbA may work in a coupled manner to afford the cell continued survival under stress conditions[163].

A range of intricate EPR techniques were applied to directly characterise the coordination of Mn^{2+} in the NBDs of MsbA, including W-band ENDOR and EDNMR (these methods have been briefly explained in an earlier section of this thesis) in addition to PELDOR. This also allowed the detection of bound nucleotide throughout the hydrolysis cycle and calculation of some kinetic parameters. The resulting data from these experiments supports the two-state model of NBD dimerisation and dissociation in MsbA[164].

LmrA is a multidrug ABC transporter isolated from *Lactococcus lactis* that extrudes hydrophobic drugs from the membrane[165]. Initial EPR studies involved labelling the TMDs in order to analyse the relationship of drug recognition, transport and coupling with the hydrolysis cycle[166]. PELDOR later showed that LmrA samples far fewer conformational states upon nucleotide binding when compared with its apo state; ATP binding alone, rather than hydrolysis, is sufficient to trigger this change. It is only when the protein cycles back to its apo state that this relatively fixed conformation is lost. Results strongly suggest that alternating between two states, with distinct differences in dynamics and structure, is necessary for substrate translocation[167].

Another ABC-type efflux pump of interest is ABCB1 (also referred to as P-glycoprotein, or P-gp), which confers anticancer therapy resistance. In one study mechanistic details were elucidated from changes in the mobility and accessibility of spin-labelled transport substrate verapamil with ABCB1 reconstituted into liposomes[168]. Following this, a study used spin-labelled ATP to specifically

investigate the structure and dynamics of the NBDs[169]; results of this support a two-state model of a resting open conformation with readily accessible NBDs and an ATP-trapped transition- like state where the nucleotide is buried in the protein. Doubly labelled variants of P-gp in proteoliposomes were used with PELDOR to obtain distance constraints[170]. Measurements revealed disagreement with crystal structure data[171]; this, combined with the broad distance distributions indicated high protein mobility. Shorter NBD interdomain distances were measured following ATP addition. cw-EPR accessibility experiments were performed on a range of spin-labelled variants in the resting, nucleotide-bound and post-hydrolytic states[172]. Results were rationalised using MD simulations and supported models of the protein with a central cavity involved in an alternating access mechanism.

TM287/288 of *Thermotoga maritima* is a heterodimeric ABC exporter analogous to systems such as ABCB1 (P-gp) and MsbA and is a homologue of the eukaryotic CFTR (the malfunction of which is responsible for cystic fibrosis). Initial PELDOR studies of this system investigated the separation of the NBDs throughout the transport cycle and the role of a non-hydrolytic ATP binding site. It was found that nucleotide binding at the non-catalytic site prevented full disengagement of the NBDs from one another; this is unlike what is exhibited by homodimeric ABC efflux pumps[173].

Further work focussed on the large-scale conformational changes of the system, from inward- to outward facing, using MD simulations and validated by Q-band PELDOR distance constraints (performed on both detergent-solubilised protein and when reconstituted into proteoliposomes). The mechanism is described as a closing and twisting of the NBDs in the inward-facing state; a sharp distance distribution upon dimerising suggest a tight locking of the NBDs. Meanwhile, the cytoplasmic gate of the TMDs close to give an occluded intermediate (~38 to ~25 Å). This is followed by the periplasmic gates opening to yield the outward-facing state; the distance distributions here go from relatively narrow and short to multimodal and broad[174]. It is important to note that these investigations were performed in the absence of substrate.

The CusCFBA efflux system within E. coli is a copper-regulating system which helps maintain cellular concentrations[175-176]. Combinations of cw-EPR and PELDOR have revealed the importance of residues M36 and M38 of CusB in both Cu(I) coordination to the CusBNT (N-terminal) domain, and interaction with CusF.

Additionally, it was found that K32 is essential for interaction with CusF – mutation removed the exchange interaction, suggesting differences in protein folding and separation of CusBNT monomers. It is thought that mutations of lysine residues might affect the conformational structure of CusBNT, thereby interfering with Cu(I) coordination[177]. Later work investigated the conformational changes CusB undergoes upon binding Cu(I). It was shown in this study that CusB forms a dimer in solution, and the orientation of the monomers relative to one another changes drastically upon Cu(I)-binding, creating a more compact dimer. This was posited to alter the CusB-CusC channel pore and possibly aid in the transition of the metal ions from one subunit to the other[178].

3.3 Secondary Active Transport

Secondary active transport describes the movement of substrates using the electropotential difference of a concentration gradient across the membrane; the movement of ions along their concentration gradient (i.e. facilitated diffusion) allows the second solute to be transported against its own electrochemical gradient (i.e. active transport). The simultaneous transport of the ion and its coupled substrate can occur either together in the same direction (symport) or in opposite directions (antiport). Multidrug and toxic compound extrusion or multidrug antimicrobial extrusion (MATE) proteins, small multidrug resistance (SMR) proteins and the major facilitator superfamily (MFS) represent three of the five major classes of bacterial multidrug efflux transporters[179-180]; specifically, these proteins are responsible for the efflux of harmful or toxic compounds via coupled proton or sodium cation antiport.

Symport

One of the best-known groups of cotransporters is the ubiquitous sodium-solute symport superfamily (SSS, comprises 11 subfamilies), which utilise the energy harnessed from sodium motive force to drive the transport of the second solute against its concentration gradient[181].

One of these, PutP (Na^+/proline symporter) has been studied using a range of different EPR techniques. The earliest study exclusively used RT cw-EPR on a range of spin-labelled variants to probe label mobility and accessibility[182]. This study supported the then-recently proposed 13-helix model[183] and focused on several transmembrane domains and loops; spin labels predicted to be buried in

the TMs were more restricted and less accessible than those attached to residues proposed to sit exposed on the surface with the least restricted and most accessible labels located towards the middle of the loop regions. Only 2 of the 17 variants indicated any spectral change upon substrate binding, pointing towards the possibility that the associated areas (TM II and loop 2) undergo binding-induced conformational change.

A later study used a combination of cw-EPR and PELDOR to measure interspin distances and how they change upon substrate binding[184]. None of the double Cys variants demonstrated the dipolar interactions associated with interspin distances <1.8 nm; PELDOR was required to measure distances of ~2–8 nm. However, the data obtained for one pair were contradictory to the cw-EPR data, due to loop regions being flexible; the broad nature of the associated distance distributions supports this explanation. The L4/L7 distance was measured to be 4.8 nm; this agrees with the previous model, indicating that these loops are on opposite sides of the membrane. Subsequent addition of sodium and proline caused minimal changes.

Another study focused on the structure of the backbone of TM IX, using PELDOR to measure 16 double Cys variants[185]. This TM domain had recently been suggested to be involved in the translocation pathway[186]. The results of the measurements determined that this transmembrane domain was kinked; the bend was suggested to act as a hinge, sealing the inward-facing cavity upon substrate binding and opening for substrate release into the cell.

A more recent study specifically targeted extracellular loop 4 (EL4), proposed to act as a gate to ligand binding sites[187]; the study used cw-EPR to investigate mobility, polarity and accessibility profiles, and PELDOR to measure interspin distances of PutP reconstituted into liposomes. The results of these studies show that EL4 consists of two α- helices connected by a loop region (similar to LeuT and vSGLT). One of the residues within EL4 is demonstrated to have hydrophobic contact with the cTM1 (ten of the TMs of the 13-helix model are ‘core’, or cTMs), sealing the extracellular gate upon substrate binding. This was further consolidated in later work[188].

Another member of the SSS is Mhp1, a Na^+/benzyl-hydantoin symporter, believed to operate via the alternating access model[189-190]. A single EPR study on this system has been published, using PELDOR to determine interspin distances to

investigate the movement of transmembrane domains under different ligand conditions[191]. The results of the study indicate that sodium binding does not induce a conformational change; instead it is the second substrate (in the presence of sodium) that shifts the conformational equilibrium. This is contrary to previously published crystal structures and suggests that LeuT-fold proteins may have at least two different coupling mechanisms.

The small hydrophobic amino acid transporter LeuT is a bacterial homologue of the neurotransmitter/sodium symporter (NSS) family, one of the subgroups that make up the SSS. It is responsible for the cotransport of leucine and sodium cations and is a homodimer[114] although this is disputed. Only a single EPR group has studied this protein, using mobility and accessibility experiments alongside measurements of distance constraints[192]. These studies reported results that suggested that the crystal structures at the time represented inhibited conformations.

BetP, a Na^+/glycine betaine symporter, is one of the most studied osmo-regulated uptake systems; it is another member of the LeuT-fold group of proteins, physiologically existing as an asymmetric homotrimer[193-194]. An EPR study using cw- and pulsed EPR methods[195] resulted in determination that the C-terminal domain weakly interacts either with corresponding domains in adjacent BetP monomers, or with the lipid bilayer (the latter agreeing with a proposed functional model)[196]. It was also suggested that the packing within the trimer determined from experimental distance constraints differs significantly from that seen in the crystal structures.

The galactose permease GalP system is involved in monosaccharide transport using chemiosmosis. It is a member of the MFS; it has a 12 transmembrane α-helix fold with both the N-terminus and C-terminus located in the cytoplasm[197]. Crystallography of the protein reconstituted into liposomes shows that it exists functionally in trimers[198]. Conventional cw-EPR (i.e. label mobility studies) and saturation transfer EPR techniques showed that the labelling site was oriented into a densely packed interhelical region and that the packing of the helices is less tight than the proposed models[199]. Prior to the publication of any crystal data, EPR was used to investigate the stoichiometry and selectivity of the lipids around the protein in a membrane environment[200].

Another member of the MFS is LacY, a H^+/β-galactoside cotransporter; the crystal structure was first solved in 2003[201], with various structures published since; it is purported to function via the alternating access model[202]. Distance measurements performed using PELDOR appear to support this assertion, with ligand-induced conformational changes resulting in opposite movements of the cytoplasmic and periplasmic ends of the transmembrane bundles[203].

Glt_{Ph} is found within chemical synapses and couples aspartate transport to the symport of three sodium ions within synapses. PELDOR measurements on two trimerisation-domain mutants taken for both the apoprotein, and the protein in the presence of coupling ions and substrate indicated that the trimerisation-domain forms the stable core. Measurements performed on mutants of transporting domains showed sampling of multiple conformations in all states to similar extents; this is consistent with large- scale movement during the transport cycle. Conformations favoured in the membrane environment are different from those favoured in detergent micelles[204]. A simultaneous study by Georgieva et al. demonstrated that the domain motions involved in the outward-to-inward transition occur both in detergent and in membranes, and energies of the outward-facing and inward-facing states were similar both in the presence and absence of substrate[205].

Antiport

Among the most studied and best characterised antiport systems is NhaA of E. coli, a Na^+/H^+ cotransporter thought to regulate pH and use proton motive force to expel sodium[206-207]. cw- and pulsed EPR techniques were used to confirm the assertion that this protein can exist as a dimer[208-209]. Further experiments suggested that the dimerisation equilibrium is moderately pH-dependent. Building on this, NhaA was reconstituted into liposomes in conjunction with PELDOR to determine the physiological dimer structure[209]. Two points of contact within the homodimer were found, contradicting preceding high-resolution crystal and cryo-EM data (though it is stated that further investigation with varying pH is required to confirm or refute these claims). NHE1, or Na^+/H^+ exchanger isoform 1, is one of nine mammalian isoforms similar in structure and function to the bacterial NhaA protein; studies of this protein have been based on the structural model of NhaA, with EPR being used to map and measure distance constraints to support this[210]. OxlT, a member of the MFS, is responsible for the exchange of oxalate for formate

in *Oxalobacter formigenes*[211]. Much of the work done on this system has been modelled on other members of the MFS, many of which are mentioned in this review (e.g. LacY, GalP etc.). PELDOR was used to measure long-range distances and it was found that OxlT adopts the inward-open and outward-open states in the presence of substrate[212]; this is contradictory to previous homology-modelled predictions, which suggested the conformation in which the cytoplasmic side is sealed and periplasmic side is open would not be significantly populated.

Pho84 is an MFS phosphate/H^+ antiporter of *Saccharomyces cerevisiae*. The crystal structure of the protein has not been solved but has been modelled on other MFS members; a cw-EPR label mobility study has been used to investigate the accuracy of a homology-modelled structure using GltPh[213]. Results were promising, suggesting the model was accurate, though further studies would be required to confirm this as not many sites on the protein were sampled.

LmrP is an MFS multidrug transporter from *L. lactis* which couples proton translocation to the extrusion of cytotoxic molecules[214]. Initial EPR studies undertaken gave evidence of distinct structural changes upon ligand binding or protonation of specific residues; – most profound was the rearrangement of helix XIII upon substrate binding, strongly indicating that the C- and N-terminal interface plays an important role in controlled drug access[215]. Subsequent cw-EPR and PELDOR studies suggest alternation between outward-open and outward-closed conformations, caused by protonation of specific residues, allowing a transmembrane protonation relay. A model was proposed involving the initiation of transport via substrate binding and opening of the extracellular side, after which specific residue protonation causes substrate release on the extracellular side. This also causes a number of conformational changes leading to proton transfer and release to the intracellular side[216].

NorM of *Vibrio parahaemolyticus* was the first multidrug and toxic compound extrusion (MATE) protein to be classified and is one of the best-studied so far; several conserved acidic residues in membrane-embedded regions have been identified that take part in Na^+-coupled transport[217] but the mechanism of substrate binding could not be elucidated from crystal structures.

MATE proteins were first categorised for bacteria but have also been found in mammalian and plant cells and are believed to be universally present in all living organisms[179, 218]. It has been demonstrated that MATE proteins play an important

role in antimicrobial resistance of bacteria, making them an attractive potential target for novel antimicrobial and anticancer drugs[219], whereas in plants they contribute to homoeostasis by secretion of waste products and detoxification of metals, and in mammalian cells it is likely that they mediate final excretion of toxic organic cations[218].

Many MATE proteins are known to be Na^+/drug antiporters although some transporters have been discovered that use protons instead of sodium cations, like PmpM of Pseudomonas aeruginosa or hMATE1 in human liver, kidney and skeletal muscle cells[179]. MATE proteins show a wide range of substrate specificity among which fluoroquinolones like norfloxacin are substrates for almost all of the transporters. Each transporter can pump out several of a growing list of identified compounds with very different chemical structures, from cationic dyes and intercalators (e.g. ethidium bromide or doxorubicin) to aminoglycosides (e.g. kanamycin and streptomycin) or β-lactam antibiotics (e.g. ampicillin)[179].

The structure of MATE family proteins shows 12-13 α-helical transmembrane segments (TMS)[220]. Steed et al. published an EPR study on Vc-NorM, using the spin-labelled substrate homologue Ruboxyl as a paramagnetic probe to gather information on location and mechanism of substrate binding and expected translocation of substrate in Na^+ presence. Quantitative results for substrate binding (both in the presence and absence of competitors) and Na^+-dependence were gained from cw-EPR lineshape experiments. Multiple modes of substrate binding were detected from the data, with one high affinity binding site and non-specific binding in case of substrate excess.

DEER experiments were used to locate the high-affinity binding site. Vc-NorM was spin-labelled at six periplasmic sites and one cytoplasmic site in order to identify the substrate position in crystal structure scaffolds of Vc-NorM[221] and Ng-NorM[222]. Results point to one membrane-embedded high-affinity substrate-binding site at TMS 7, close to the loop at TMS 7/8 that probably shifts upon substrate/ion binding. cw-EPR lineshapes also showed that addition of known substrates as competitors for Ruboxyl binding led to reduction in Ruboxyl binding to the high-affinity binding site to approximately 50%. Highly interesting was the discovery that Na^+ concentration, varied from 0 to 100 mM, did not have any effect on substrate binding nor did it trigger substrate translocation. The lack of Na^+-dependence for Ruboxyl binding supports the theory that MATE transporters (or at least NorM) do

not operate according to the classical mechanism of antiport but via a transport cycle comprising multiple equilibrium states that depend on a Na^+ gradient rather than Na^+ concentration.

More recent work was done on Vc-NorM, exploring ligand-dependent conformational changes using PELDOR. On the periplasmic side of the complex, distances between spin labels suggested unique intermediate conformations as a result of Na^+, H^+ or doxorubicin binding. In particular, the arrangement of TM1 was found to deviate from the prediction (from the crystal structures). The effect of H^+ vs Na^+ concentration on a range of singly labelled variants along a small section of TM1 was studied. It was found that the lineshapes of certain variants were distinctly different in the presence of protons relative to sodium ions suggesting differing structural arrangements of TM1 in these conditions; it may be that TM1 has a role in ion sensing. Additionally, the doxorubicin-bound form of Vc-NorM was found to be stable in the presence of Na^+ but not in the presence of H^+, suggesting protons may play a role in drug dissociation. A series of residues were also identified that are posited to be important for modulating conformational dynamics in the NTD ion binding site[223].

SMR proteins are specific for a vast variety of substrates, including quaternary ammonium compounds (QAC), other lipophilic cations and a multitude of detergents, antiseptics, cationic dyes and antibiotics[224-226].

The main substrates of EmrE are tertiary and quaternary bulky aromatic cations and other positively charged hydrophobic compounds. Substrates bind at Glu-14 embedded in TMS 1 and the extrusion mechanism using the proton motive force is well characterised, although it is not yet fully understood on a structural basis. EPR and SDSL were used to further elucidate the structure and its conformations involved in the transport cycle.

The oligomerisation state was investigated using cw-EPR and SDSL, concluding in agreement with other studies that it forms a homodimer[227-228]. The question of topology of the dimers is still discussed as arguments for both parallel and antiparallel topology can be found[229-231]. Both cw-EPR[227] and pulsed EPR[232] of spin-labelled EmrE pointed towards a parallel topology which would also be the conformation with least bias within the membrane. More recent cw-EPR results though[233] support antiparallel topology. In this study, spin labels were introduced along the axes of the TM helices to gain more insight in topology and

conformational changes upon substrate binding. Accessibility assays with O_2 and NiEDDA (Ni(II)-ethylenediamine-N,N′-diacetic acid) with and without TPP^+ as substrate were performed. Both NiEDDA and O_2 function as paramagnetic relaxants and their collision frequency with spin label, resulting in characteristic changes in EPR spectra, allows conclusions on the environment of the particular spin label. NiEDDA is only soluble in aqueous phases whereas O_2 enters only the membranes and thus, membrane boundaries and lipid- facing or water-exposed residues can be identified. EPR results showed, that a putative symmetric interchange from apo to substrate-bound intermediate could not be verified by the data. Instead, a permeation pathway of the substrate through the asymmetric dimer was suggested. Both for further elucidation of substrate transport by EmrE and for the still unresolved question of topology, Amadi *et al*. suggest long range distance measurements.

PELDOR was later used to probe the previously unstudied conformational state of EmrE subsequent to protonation. Protonation of residue E14 was found to give rise to rotation and tilting of TMs 1-3, which form the proton/substrate binding site. It is presumed that the reorientation of side chains that result from this movement affects substrate affinity, and also alters access to the binding cavity. It is suggested that the transporter rests in a novel protonated-occluded state. The substrate binding releases the protons and prompts the alternating access conformational changes between the inward- and outward-facing states, allowing antiport without disrupting the proton gradient[234].

EmrD is a member of the MFS. MFS transporters are ubiquitous; in bacteria they are mainly used for nutrition uptake and extrusion of harmful compounds. They consist of 12 TM helices, have two pseudosymmetrical halves and likely transport substrates via alternating accessibility of a central cavity.

EmrD from *E. coli* is the only structurally characterised multidrug/H^+ antiporter (DHA) of the MFS so far[111, 235-236]. The crystal structure revealed a doubly occluded conformation with unexpected features and an EPR study[236] followed to assess the structure in a more native-like environment in unilamellar liposomes. In this study, 76 EmrD mutants with spin labels introduced along their TM helical axes were investigated using cw-EPR. Information gathered from EPR lineshape and from accessibility assays with O_2 and NiEDDA were mostly consistent with the

crystal structure, although major deviations were found for the orientation of TMS 5 and the topology of TMS 10 and 11.

Furthermore, EPR results showed pH-dependent conformational changes that led to opening of a cytoplasmic cleft at TMS 2–4 in the N-terminal half of EmrD, a region containing the MFS signature GxxxD(R/K)xG and conserved amino acid residues Asp-68 and Asp-123 that could play a crucial role in H^+ transport, and alongside mediated movement of TMS 6 that reduced accessibility at the periplasmic side. Thus, Steed et al. could confirm a pH- dependent conformational switch in EmrD as expected for a multidrug/H^+-antiporter. Structural data on substrate binding and transport by EmrD in vitro and in cell-based assays could not be gained because their studies could not reproduce multidrug resistance leaving the authors questioning the suitability of EmrD as a model for other MFS DAH transporters.

Conclusions

Knowledge of protein structures at atomic resolution is essential to understand function. Although crystallography remains the mainstream method to obtain structural information, crystal structures of dynamic membrane transport proteins are difficult to derive, and often crystallography only provide static snapshots. Indeed, very few membrane proteins have been crystallized in more than one conformation. Knowledge of such structural and conformational changes is a key to understand how membrane transporters translocate substrates across the membrane EPR spectroscopy can, in principle, provide such complex dynamic information over a large range of distances, allowing for measurement of distances, environment and protein dynamics. Taken together with the structural information of protein snapshots from crystals, EPR has the power to enhance our understanding of the complex functional dynamics at play in important macromolecular protein complexes such as membrane transporters.

References

1. Saier, M. H.; Reddy, V. S.; Tsu, B. V.; Ahmed, M. S.; Li, C.; Moreno-Hagelsieb, G., The Transporter Classification Database (TCDB): Recent Advances. *Nucleic Acids Res.* **2016,** *44* (D1), D372-D379.

2. Busch, W.; Saier, M. H., The Transporter Classification (TC) System, 2002. *Crit. Rev. Biochem. Mol. Biol.* **2002,** *37* (5), 287-337.

3. Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C., The Protein Data Bank. *Acta Crystallographica Section D-Structural Biology* **2002,** *58*, 899-907.

4. White, S. Membrane Proteins of Known 3D Structure. https://blanco.biomol.uci.edu/mpstruc/ (accessed 27th May 2019).

5. Davidson, A. L.; Maloney, P. C., ABC Transporters: How Small Machines Do a Big Job. *Trends Microbiol.* **2007,** *15* (10), 448-455.

6. Higgins, C. F., ABC Transporters - from Microorganisms to Man. *Annu. Rev. Cell Biol.* **1992,** *8*, 67-113.

7. Davidson, A. L.; Dassa, E.; Orelle, C.; Chen, J., Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol Mol Biol Rev* **2008,** *72* (2), 317-64, table of contents.

8. Mendez, C.; Salas, J. A., The Role of ABC Transporters in Antibiotic-Producing Organisms: Drug Secretion and Resistance Mechanisms. *Res. Microbiol.* **2001,** *152* (3-4), 341-350.

9. Blair, J. M. A.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. V., Molecular Mechanisms of Antibiotic Resistance. *Nature Reviews Microbiology* **2015,** *13* (1), 42-51.

10. Choi, C. H., ABC Transporters as Multidrug Resistance Mechanisms and the Development of Chemosensitizers for Their Reversal. *Cancer Cell Int.* **2005,** *5*, 13.

11. Gottesman, M. M.; Fojo, T.; Bates, S. E., Multidrug Resistance in Cancer: Role of ATP-Dependent Transporters. *Nat. Rev. Cancer* **2002,** *2* (1), 48-58.

12. Fletcher, J. I.; Haber, M.; Henderson, M. J.; Norris, M. D., ABC Transporters in Cancer: More Than Just Drug Efflux Pumps. *Nat. Rev. Cancer* **2010,** *10* (2), 147-156.

13. Pedersen, P. L., Transport Atpases into the Year 2008: A Brief Overview Related to Types, Structures, Functions and Roles in Health and Disease. *J. Bioenerg. Biomembr.* **2007,** *39* (5-6), 349-355.

14. Higgins, C. F.; Hiles, I. D.; Salmond, G. P. C.; Gill, D. R.; Downie, J. A.; Evans, I. J.; Holland, I. B.; Gray, L.; Buckel, S. D.; Bell, A. W.; Hermodson, M. A.,

A Family of Related ATP-Binding Subunits Coupled to Many Distinct Biological Processes in Bacteria. *Nature* **1986,** *323* (6087), 448-450.

15. Hyde, S. C.; Emsley, P.; Hartshorn, M. J.; Mimmack, M. M.; Gileadi, U.; Pearce, S. R.; Gallagher, M. P.; Gill, D. R.; Hubbard, R. E.; Higgins, C. F., Structural Model of ATP-Binding Proteins Associated with Cystic-Fibrosis, Multidrug Resistance and Bacterial Transport. *Nature* **1990,** *346* (6282), 362-365.

16. Walker, J. E.; Saraste, M.; Runswick, M. J.; Gay, N. J., Distantly Related Sequences in the Alpha-Subunits and Beta-Subunits of ATP Synthase, Myosin, Kinases and Other ATP-Requiring Enzymes and a Common Nucleotide Binding Fold. *Embo Journal* **1982,** *1* (8), 945-951.

17. Davidson, A. L.; Chen, J., ATP-Binding Cassette Transporters in Bacteria. *Annu Rev Biochem* **2004,** *73*, 241-68.

18. Jones, P. M.; George, A. M., The ABC Transporter Structure and Mechanism: Perspectives on Recent Research. *Cellular and Molecular Life Sciences* **2004,** *61* (6), 682-699.

19. Ter Beek, J.; Guskov, A.; Slotboom, D. J., Structural Diversity of ABC Transporters. *J. Gen. Physiol.* **2014,** *143* (4), 419-435.

20. Rice, A. J.; Park, A.; Pinkett, H. W., Diversity in ABC Transporters: Type I, II and III Importers. *Crit. Rev. Biochem. Mol. Biol.* **2014,** *49* (5), 426-437.

21. Biemans-Oldehinkel, E.; Doeven, M. K.; Poolman, B., ABC Transporter Architecture and Regulatory Roles of Accessory Domains. *Febs Letters* **2006,** *580* (4), 1023-1035.

22. Oldham, M. L.; Davidson, A. L.; Chen, J., Structural Insights into ABC Transporter Mechanism. *Curr Opin Struct Biol* **2008,** *18* (6), 726-33.

23. Locher, K. P., Structure and Mechanism of ATP-Binding Cassette Transporters. *Philos. Trans. R. Soc. B-Biol. Sci.* **2009,** *364* (1514), 239-245.

24. Henderson, G. B.; Zevely, E. M.; Huennekens, F. M., Mechanism of Folate Transport in Lactobacillus-Casei - Evidence for a Component Shared with the Thiamine and Biotin Transport-Systems. *Journal of Bacteriology* **1979,** *137* (3), 1308-1314.

25. Rodionov, D. A.; Hebbeln, P.; Eudes, A.; Ter Beek, J.; Rodionova, I. A.; Erkens, G. B.; Slotboom, D. J.; Gelfand, M. S.; Osterman, A. L.; Hanson, A. D.; Eitinger, T., A Novel Class of Modular Transporters for Vitamins in Prokaryotes. *Journal of Bacteriology* **2009,** *191* (1), 42-51.

26. Davidson, A. L.; Shuman, H. A.; Nikaido, H., Mechanism of Maltose Transport in Escherichia-Coli - Transmembrane Signaling by Periplasmic Binding-Proteins. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** *89* (6), 2360-2364.

27. Hollenstein, K.; Dawson, R. J.; Locher, K. P., Structure and Mechanism of ABC Transporter Proteins. *Curr Opin Struct Biol* **2007,** *17* (4), 412-8.

28. Quiocho, F. A.; Phillips, G. N.; Parsons, R. G.; Hogg, R. W., Crystallographic Data of an L-Arabinose-Binding Protein from Escherichia-Coli. *Journal of Molecular Biology* **1974,** *86* (2), 491-493.

29. Poolman, B.; Slotboom, D.-J., Substrate Capture by ABC Transporters. In *Encyclopedia of Biophysics*, Roberts, G. C. K., Ed. Springer: Berlin, Heidelberg, 2013; pp 2503-2509.

30. Berntsson, R. P.; Smits, S. H.; Schmitt, L.; Slotboom, D. J.; Poolman, B., A Structural Classification of Substrate-Binding Proteins. *FEBS Lett* **2010,** *584* (12), 2606-17.

31. Scheepers, G. H.; Lycklama, J. A.; Poolman, B., An Updated Structural Classification of Substrate-Binding Proteins. *Febs Letters* **2016,** *590* (23), 4393-4401.

32. Mao, B.; Pear, M. R.; Mccammon, J. A.; Quiocho, F. A., Hinge-Bending in L-Arabinose-Binding Protein - the Venus-Flytrap Model. *Journal of Biological Chemistry* **1982,** *257* (3), 1131-1133.

33. Quiocho, F. A.; Ledvina, P. S., Atomic Structure and Specificity of Bacterial Periplasmic Receptors for Active Transport and Chemotaxis: Variation of Common Themes. *Mol. Microbiol.* **1996,** *20* (1), 17-25.

34. Felder, C. B.; Graul, R. C.; Lee, A. Y.; Merkle, H. P.; Sadee, W., The Venus Flytrap of Periplasmic Binding Proteins: An Ancient Protein Module Present in Multiple Drug Receptors. *AAPS Pharmsci* **1999,** *1* (2), 28.

35. Hediger, M. A.; Romero, M. F.; Peng, J. B.; Rolfs, A.; Takanaga, H.; Bruford, E. A., The ABCs of Solute Carriers: Physiological, Pathological and Therapeutic Implications of Human Membrane Transport Proteins - Introduction. *Pflugers Arch.* **2004,** *447* (5), 465-468.

36. Perland, E.; Fredriksson, R., Classification Systems of Secondary Active Transporters. *Trends Pharmacol. Sci.* **2017,** *38* (3), 305-315.

37. Hediger, M. A.; Clemencon, B.; Burrier, R. E.; Bruford, E. A., The ABCs of Membrane Transporters in Health and Disease (SLC Series): Introduction. *Mol. Asp. Med.* **2013,** *34* (2-3), 95-107.

38. Hoglund, P. J.; Nordstrom, K. J. V.; Schioth, H. B.; Fredriksson, R., The Solute Carrier Families Have a Remarkably Long Evolutionary History with the Majority of the Human Families Present before Divergence of Bilaterian Species. *Mol. Biol. Evol.* **2011,** *28* (4), 1531-1541.

39. Romero, M. F.; Fulton, C. M.; Boron, W. F., The SLC4 Family of HCO3-Transporters. *Pflugers Arch.* **2004,** *447* (5), 495-509.

40. Palmieri, F., The Mitochondrial Transporter Family Slc25: Identification, Properties and Physiopathology. *Mol. Asp. Med.* **2013,** *34* (2-3), 465-484.

41. Kanai, Y.; Clemencon, B.; Simonin, A.; Leuenberger, M.; Lochner, M.; Weisstanner, M.; Hediger, M. A., The Slc1 High-Affinity Glutamate and Neutral Amino Acid Transporter Family. *Mol. Asp. Med.* **2013,** *34* (2-3), 108-120.

42. Pramod, A. B.; Foster, J.; Carvelli, L.; Henry, L. K., SLC6 Transporters: Structure, Function, Regulation, Disease Association and Therapeutics. *Mol. Asp. Med.* **2013,** *34* (2-3), 197-219.

43. Lin, L.; Yee, S. W.; Kim, R. B.; Giacomini, K. M., SLC Transporters as Therapeutic Targets: Emerging Opportunities. *Nat. Rev. Drug Discov.* **2015,** *14* (8), 543-560.

44. Bai, X. Y.; Moraes, T. F.; Reithmeier, R. a. F., Structural Biology of Solute Carrier (SLC) Membrane Transport Proteins. *Mol. Membr. Biol.* **2017,** *34* (1-2), 1-32.

45. Pao, S. S.; Paulsen, I. T.; Saier, M. H., Major Facilitator Superfamily. *Microbiol. Mol. Biol. Rev.* **1998,** *62* (1), 1-+.

46. Colas, C.; Ung, P. M. U.; Schlessinger, A., SLC Transporters: Structure, Function, and Drug Discovery. *MedChemComm* **2016,** *7* (6), 1069-1081.

47. Abramson, J.; Wright, E. M., Structure and Function of Na+-Symporters with Inverted Repeats. *Curr. Opin. Struct. Biol.* **2009,** *19* (4), 425-432.

48. Shi, Y. G., Common Folds and Transport Mechanisms of Secondary Active Transporters. In *Annual Review of Biophysics, Vol 42*, Dill, K. A., Ed. Annual Reviews: Palo Alto, 2013; Vol. 42, pp 51-72.

49. Faham, S.; Watanabe, A.; Besserer, G. M.; Cascio, D.; Specht, A.; Hirayama, B. A.; Wright, E. M.; Abramson, J., The Crystal Structure of a Sodium Galactose Transporter Reveals Mechanistic Insights into Na(+)/Sugar Symport. *Science* **2008,** *321* (5890), 810-814.

50. Watanabe, A.; Choe, S.; Chaptal, V.; Rosenberg, J. M.; Wright, E. M.; Grabe, M.; Abramson, J., The Mechanism of Sodium and Substrate Release from the Binding Pocket of Vsglt. *Nature* **2010,** *468* (7326), 988-U162.

51. Forrest, L. R.; Zhang, Y. W.; Jacobs, M. T.; Gesmonde, J.; Xie, L.; Honig, B. H.; Rudnick, G., Mechanism for Alternating Access in Neurotransmitter Transporters. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (30), 10338-10343.

52. Levine, M. V.; Cuendet, M. A.; Khelashvili, G.; Weinstein, H., Allosteric Mechanisms of Molecular Machines at the Membrane: Transport by Sodium-Coupled Symporters. *Chemical Reviews* **2016,** *116* (11), 6552-6587.

53. Lee, C.; Kang, H. J.; Von Ballmoos, C.; Newstead, S.; Uzdavinys, P.; Dotson, D. L.; Iwata, S.; Beckstein, O.; Cameron, A. D.; Drew, D., A Two-Domain Elevator Mechanism for Sodium/Proton Antiport. *Nature* **2013,** *501* (7468), 573-+.

54. Crisman, T. J.; Qu, S. G.; Kanner, B. I.; Forrest, L. R., Inward-Facing Conformation of Glutamate Transporters as Revealed by Their Inverted-Topology Structural Repeats. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (49), 20752-20757.

55. Reyes, N.; Ginter, C.; Boudker, O., Transport Mechanism of a Bacterial Homologue of Glutamate Transporters. *Nature* **2009,** *462* (7275), 880-885.

56. Grewer, C.; Balani, P.; Weidenfeller, C.; Bartusel, T.; Tao, Z.; Rauen, T., Individual Subunits of the Glutamate Transporter EAAC1 Homotrimer Function Independently of Each Other. *Biochemistry* **2005,** *44* (35), 11913-11923.

57. Verdon, G.; Boudker, O., Crystal Structure of an Asymmetric Trimer of a Bacterial Glutamate Transporter Homolog. *Nat. Struct. Mol. Biol.* **2012,** *19* (3), 355-357.

58. Markovich, D., Sodium-Sulfate/Carboxylate Cotransporters (SLC13). In *Co-Transport Systems*, Bevensee, M. O., Ed. Elsevier Academic Press Inc: San Diego, 2012; Vol. 70, pp 239-256.

59. Mancusso, R.; Gregorio, G. G.; Liu, Q.; Wang, D. N., Structure and Mechanism of a Bacterial Sodium-Dependent Dicarboxylate Transporter. *Nature* **2012,** *491* (7425), 622-+.

60. Nie, R. X.; Stark, S.; Symersky, J.; Kaplan, R. S.; Lu, M., Structure and Function of the Divalent Anion/Na+ Symporter from Vibrio cholerae and a Humanized Variant. *Nat. Commun.* **2017,** *8*, 10.

61. Lu, M., Structure and Mechanism of the Divalent Anion/Na+ Symporter. *Int. J. Mol. Sci.* **2019,** *20* (2), 12.

62. Mulligan, C.; Fenollar-Ferrer, C.; Fitzgerald, G. A.; Vergara-Jaque, A.; Kaufmann, D.; Li, Y.; Forrest, L. R.; Mindell, J. A., The Bacterial Dicarboxylate Transporter VcINDY Uses a Two-Domain Elevator-Type Mechanism. *Nat. Struct. Mol. Biol.* **2016,** *23* (3), 256-263.

63. Cornish, F., Classical Radiation Theory and Point Charges. *Proceedings of the Physical Society of London* **1965,** *86* (551P), 427-&.

64. Odom, B.; Hanneke, D.; D'urso, B.; Gabrielse, G., New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron. *Phys Rev Lett* **2006,** *97* (3), 030801.

65. Eaton, G. R.; Eaton, S. S.; Barr, D. P.; Weber, R. T., *Quantitative EPR*. Springer Vienna: 2010.

66. Junk, M. J. N., *Assessing the Functional Structure of Molecular Transporters by EPR Spectroscopy*. Springer Berlin Heidelberg: 2012.

67. Hagen, W. R., *Biomolecular EPR Spectroscopy*. CRC Press: 2008.

68. Jackson, S. D.; Hargreaves, J. S. J., *Metal Oxide Catalysis*. Wiley: 2009.

69. Haddy, A.; Smith, G., Transition Metal and Organic Radical Components of Carp Liver Tissue Observed by Electron Paramagnetic Resonance Spectroscopy. *Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology* **1999,** *123* (4), 8.

70. Goodman, B.; Raynor, J., Electron Spin Resonance of Transition Metal Complexes. In *Advances in Inorganic Chemistry and Radiochemistry*, Sharpe, H. J. E. a. A. G., Ed. 1970; Vol. 13, pp 135-362.

71. Scott, R. A.; Lukehart, C. M., *Applications of Physical Methods to Inorganic and Bioinorganic Chemistry*. Wiley: 2007.

72. Beltrán-López, V.; Castro-Tello, J., Powder Pattern of Systems with Axially Anisotropic G and a Tensors. The EPR Spectrum of Copper Phthalocyanine. *Journal of Magnetic Resonance (1969)* **1982,** *47* (1), 19-27.

73. Soliverez, C., The Contact Hyperfine Interaction - an Ill-Defined Problem. *Journal of Physics C-Solid State Physics* **1980,** *13* (34), 1017-1019.

74. Grinberg, O. Y.; Dadali, A. A.; Dubinskii, A. A.; Vasserman, A. M.; Buchachenko, A. L.; Lebedev, Y. S., Determination of G and a Tensor Components and Rotational Mobility of Nitroxide Radicals by 2-mm EPR Spectroscopy. *Theoretical and Experimental Chemistry* **1979,** *15* (5), 457-461.

75. Bunce, N., Introduction to the Interpretation of Electron-Spin Resonance-Spectra of Organic Radicals. *Journal of Chemical Education* **1987,** *64* (11), 907-914.

76. Roessler, M. M.; Salvadori, E., Principles and Applications of EPR Spectroscopy in the Chemical Sciences. *Chemical Society Reviews* **2018,** *47* (8), 2534-2553.

77. Stich, T. A.; Lahiri, S.; Yeagle, G.; Dicus, M.; Brynda, M.; Gunn, A.; Aznar, C.; Derose, V. J.; Britt, R. D., Multifrequency Pulsed EPR Studies of Biologically Relevant Manganese(II) Complexes. *Appl. Magn. Reson.* **2007,** *31* (1-2), 321-341.

78. Chang, C. H.; Svedruzic, D.; Ozarowski, A.; Walker, L.; Yeagle, G.; Britt, R. D.; Angerhofer, A.; Richards, N. G. J., EPR Spectroscopic Characterization of the Manganese Center and a Free Radical in the Oxalate Decarboxylase Reaction - Identification of a Tyrosyl Radical During Turnover. *Journal of Biological Chemistry* **2004,** *279* (51), 52840-52849.

79. Weil, J. A.; Bolton, J. R., *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*. Wiley: 2007.

80. Bender, C.; Berliner, L., *EPR: Instrumental Methods*. Springer US: 2012.

81. Atkins, P.; De Paula, J., *Atkins' Physical Chemistry*. OUP Oxford: 2010.

82. Mabbs, F. E.; Collison, D., *Electron Paramagnetic Resonance of D Transition Metal Compounds*. Elsevier Science: 2013.

83. Allewell, N. M.; Narhi, L. O.; Rayment, I., *Molecular Biophysics for the Life Sciences*. Springer New York: 2013.

84. Klare, J.; Steinhoff, H., Spin Labeling EPR. *Photosynthesis Research* **2009,** *102* (2-3), 377-390.

85. Altenbach, C.; Flitsch, S.; Khorana, H.; Hubbell, W., Structural Studies on Transmembrane Proteins .2. Spin Labeling of Bacteriorhodopsin Mutants at Unique Cysteines. *Biochemistry* **1989,** *28* (19), 7806-7812.

86. Owenius, R.; Engstrom, M.; Lindgren, M.; Huber, M., Influence of Solvent Polarity and Hydrogen Bonding on the EPR Parameters of a Nitroxide Spin Label Studied by 9-GHz and 95-GHz EPR Spectroscopy and Dft Calculations. *Journal of Physical Chemistry a* **2001,** *105* (49), 10967-10977.

87. Jeschke, G., DEER Distance Measurements on Proteins. In *Annual Review of Physical Chemistry, Vol 63*, Johnson, M. A.; Martinez, T. J., Eds. 2012; Vol. 63, pp 419-446.

88. Steinhoff, H.; Radzwill, N.; Thevis, W.; Lenz, V.; Brandenburg, D.; Antson, A.; Dodson, G.; Wollmer, A., Determination of Interspin Distances between Spin Labels Attached to Insulin: Comparison of Electron Paramagnetic Resonance Data with the X-Ray Structure. *Biophysical Journal* **1997,** *73* (6), 3287-3298.

89. Raitsimring, A., "2+1" Pulse Sequence as Apllied for Distance and Spatial Distribution Measurements of Paramagnetic Centers. In *Distance Measurements in Biological Systems by EPR*, Berliner, L. J.; Eaton, G. R.; Eaton, S. S., Eds. Springer US: Boston, MA, 2000; pp 461-491.

90. Kulik, L. V.; Dzuba, S. A.; Grigoryev, I. A.; Tsvetkov, Y. D., Electron Dipole-Dipole Interaction in ESEEM of Nitroxide Biradicals. *Chem Phys Lett* **2001,** *343* (3-4), 315-324.

91. Keller, K.; Zalibera, M.; Qi, M.; Koch, V.; Wegner, J.; Hintz, H.; Godt, A.; Jeschke, G.; Savitsky, A.; Yulikov, M., EPR Characterization of Mn(II) Complexes for Distance Determination with Pulsed Dipolar Spectroscopy. *Physical Chemistry Chemical Physics* **2016,** *18* (36), 25120-25135.

92. Borbat, P. P.; Freed, J. H., Multiple-Quantum ESR and Distance Measurements. *Chem Phys Lett* **1999,** *313* (1-2), 145-154.

93. Yang, Z.; Liu, Y.; Borbat, P.; Zweier, J. L.; Freed, J. H.; Hubbell, W. L., Pulsed ESR Dipolar Spectroscopy for Distance Measurements in Immobilized Spin Labeled Proteins in Liquid Solution. *Journal of the American Chemical Society* **2012,** *134* (24), 9950-9952.

94. Kulik, L.; Lubitz, W., Electron-Nuclear Double Resonance. *Photosynthesis Research* **2009,** *102* (2-3), 391-401.

95. Schweiger, A.; Jeschke, G., *Principles of Pulse Electron Paramagnetic Resonance*. Oxford University Press: 2001.

96. Rowan, L. G.; Hahn, E. L.; Mims, W. B., Electron-Spin-Echo Envelope Modulation. *Physical Review* **1965,** *137* (1A), A61-&.

97. Mims, W. B., Envelope Modulation in Spin-Echo Experiments. *Physical Review B-Solid State* **1972,** *5* (7), 2409-&.

98. Ponti, A.; Schweiger, A., Nuclear Coherence-Transfer Echoes in Pulsed EPR. *J Chem Phys* **1995,** *102* (13), 5207-5219.

99. Jeschke, G.; Schweiger, A., Hyperfine Decoupling in Electron Spin Resonance. *J Chem Phys* **1997,** *106* (24), 9979-9991.

100. Cox, N.; Nalepa, A.; Lubitz, W.; Savitsky, A., ELDOR-Detected NMR: A General and Robust Method for Electron-Nuclear Hyperfine Spectroscopy? *Journal of Magnetic Resonance* **2017,** *280*, 63-78.

101. Solomon, I., Relaxation Processes in a System of Two Spins. *Physical Review* **1955,** *99* (2), 559-565.

102. Jahnke, W., Spin Labels as a Tool to Identify and Characterize Protein-Ligand Interactions by NMR Spectroscopy. *Chembiochem* **2002,** *3* (2-3), 167-173.

103. Overhauser, A. W., Polarization of Nuclei in Metals. *Physical Review* **1953,** *92* (2), 411-415.

104. Huang, R.; Bonnichon, A.; Claridge, T. D. W.; Leung, I. K. H., Protein-Ligand Binding Affinity Determination by the Waterlogsy Method: An Optimised Approach Considering Ligand Rebinding. *Sci Rep* **2017,** *7*, 6.

105. Mayer, M.; Meyer, B., Group Epitope Mapping by Saturation Transfer Difference NMR to Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *Journal of the American Chemical Society* **2001,** *123* (25), 6108-6117.

106. Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J., Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *Journal of Chemical Education* **2011,** *88* (7), 990-994.

107. Palmer, A. G., NMR Characterization of the Dynamics of Biomacromolecules. *Chemical Reviews* **2004,** *104* (8), 3623-3640.

108. Lange, O. F.; Lakomek, N. A.; Fares, C.; Schroder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; De Groot, B. L., Recognition Dynamics up to Microseconds Revealed from an RDC-Derived Ubiquitin Ensemble in Solution. *Science* **2008,** *320* (5882), 1471-1475.

109. Vogel, S. S.; Van Der Meer, B. W.; Blank, P. S., Estimating the Distance Separating Fluorescent Protein FRET Pairs. *Methods* **2014,** *66* (2), 131-138.

110. Mullen, A.; Hall, J.; Diegel, J.; Hassan, I.; Fey, A.; Macmillan, F., Membrane Transporters Studied by EPR Spectroscopy: Structure Determination and Elucidation of Functional Dynamics. *Biochem. Soc. Trans.* **2016,** *44*, 905-915.

111. Quistgaard, E. M.; Low, C.; Guettou, F.; Nordlund, P., Understanding Transport by the Major Facilitator Superfamily (MFS): Structures Pave the Way. *Nature Reviews Molecular Cell Biology* **2016,** *17* (2), 123-132.

112. Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H., Structure of the Protein Subunits in the Photosynthetic Reaction Center of Rhodopseudomonas-Viridis at 3A Resolution. *Nature* **1985,** *318* (6047), 618-624.

113. Comellas, G.; Rienstra, C. M., Protein Structure Determination by Magic-Angle Spinning Solid-State NMR, and Insights into the Formation, Structure, and Stability of Amyloid Fibrils. In *Annual Review of Biophysics, Vol 42*, Dill, K. A., Ed. 2013; Vol. 42, pp 515-536.

114. Krishnamurthy, H.; Gouaux, E., X-Ray Structures of LeuT in Substrate-Free Outward-Open and Apo Inward-Open States. *Nature* **2012,** *481* (7382), 469-U80.

115. Callaway, E., Molecular-Imaging Pioneers Scoop Nobel. *Nature* **2017,** *550* (7675), 167-167.

116. Sigworth, F. J., Principles of Cryo-Em Single-Particle Image Processing. *Microscopy* **2016,** *65* (1), 57-67.

117. Fan, X.; Wang, J.; Zhang, X.; Yang, Z.; Zhang, J. C.; Zhao, L. Y.; Peng, H. L.; Lei, J. L.; Wang, H. W., Single Particle Cryo-Em Reconstruction of 52 kDa Streptavidin at 3.2 Angstrom Resolution. *Nat. Commun.* **2019,** *10*, 11.

118. Danev, R.; Yanagisawa, H.; Kikkawa, M., Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. *Trends Biochem.Sci.* **2019,** *44* (10), 837-848.

119. Schmidt, T.; Walti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M., Long Distance Measurements up to 160 Angstrom in the Groel Tetradecamer Using Q-Band DEER EPR Spectroscopy. *Angewandte Chemie-International Edition* **2016,** *55* (51), 15905-15909.

120. Husada, F.; Gouridis, G.; Vietrov, R.; Schuurman-Wolters, G. K.; Ploetz, E.; De Boer, M.; Poolman, B.; Cordes, T., Watching Conformational Dynamics of ABC Transporters with Single-Molecule Tools. *Biochem. Soc. Trans.* **2015,** *43*, 1041-1047.

121. Dorr, J. M.; Koorengevel, M. C.; Schafer, M.; Prokofyev, A. V.; Scheidelaar, S.; Van Der Cruijsen, E. a. W.; Dafforn, T. R.; Baldus, M.; Killian, J. A., Detergent-Free Isolation, Characterization, and Functional Reconstitution of a Tetrameric K+ Channel: The Power of Native Nanodiscs. *Proceedings of the National Academy of Sciences of the United States of America* **2014,** *111* (52), 18607-18612.

122. Milov, A. D.; Salikhov, K. M.; Shirov, M. D., Application of ELDOR in Electron-Spin Echo for Paramagnetic Center Space Distribution in Solids. *Fizika Tverdogo Tela* **1981,** *23* (4), 975-982.

123. Martin, R. E.; Pannier, M.; Diederich, F.; Gramlich, V.; Hubrich, M.; Spiess, H. W., Determination of End-to-End Distances in a Series of Tempo Diradicals of up to 2.8 nm Length with a New Four-Pulse Double Electron Electron Resonance Experiment. *Angewandte Chemie-International Edition* **1998,** *37* (20), 2834-2837.

124. Borbat, P. P.; Mchaourab, H. S.; Freed, J. H., Protein Structure Determination Using Long-Distance Constraints from Double-Quantum Coherence ESR: Study of T4 Lysozyme. *Journal of the American Chemical Society* **2002,** *124* (19), 5304-5314.

125. Prisner, T.; Rohrer, M.; Macmillan, F., Pulsed EPR Spectroscopy: Biological Applications. *Annual Review of Physical Chemistry* **2001,** *52*, 279-313.

126. Klare, J. P.; Steinhoff, H.-J., Spin Labeling Studies of Transmembrane Signaling and Transport: Applications to Phototaxis, ABC Transporters and Symporters. In *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt B*, Qin, P. Z.; Warncke, K., Eds. 2015; Vol. 564, pp 315-347.

127. Borbat, P. P.; Freed, J. H., Pulse Dipolar Electron Spin Resonance: Distance Measurements. In *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centres in the Biosciences*, Timmel, C. R.; Harmer, J. R., Eds. 2014; Vol. 152, pp 1-82.

128. Yang, Z.; Bridges, M.; Lerch, M. T.; Altenbach, C.; Hubbell, W. L., Saturation Recovery EPR and Nitroxide Spin Labeling for Exploring Structure and Dynamics in Proteins. In *Electron Paramagnetic Resonance Investigations of*

Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt B, Qin, P. Z.; Warncke, K., Eds. 2015; Vol. 564, pp 3-27.

129. Lerch, M. T.; Yang, Z.; Altenbach, C.; Hubbell, W. L., High-Pressure EPR and Site-Directed Spin Labeling for Mapping Molecular Flexibility in Proteins. In *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt B*, Qin, P. Z.; Warncke, K., Eds. 2015; Vol. 564, pp 29-57.

130. Altenbach, C.; Lopez, C. J.; Hideg, K.; Hubbell, W. L., Exploring Structure, Dynamics, and Topology of Nitroxide Spin-Labeled Proteins Using Continuous-Wave Electron Paramagnetic Resonance Spectroscopy. In *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt B*, Qin, P. Z.; Warncke, K., Eds. 2015; Vol. 564, pp 59-100.

131. Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G., High Sensitivity and Versatility of the DEER Experiment on Nitroxide Radical Pairs at Q-Band Frequencies. *Physical Chemistry Chemical Physics* **2012,** *14* (30), 10762-10773.

132. Spindler, P. E.; Glaser, S. J.; Skinner, T. E.; Prisner, T. F., Broadband Inversion PELDOR Spectroscopy with Partially Adiabatic Shaped Pulses. *Angewandte Chemie-International Edition* **2013,** *52* (12), 3425-3429.

133. Hirst, S. J.; Alexander, N.; Mchaourab, H. S.; Meiler, J., RosettaEPR: An Integrated Tool for Protein Structure Determination from Sparse EPR Data. *Journal of Structural Biology* **2011,** *173* (3), 506-514.

134. Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H., DeerAnalysis2006 - a Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006,** *30* (3-4), 473-498.

135. Polyhach, Y.; Bordignon, E.; Jeschke, G., Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. *Physical Chemistry Chemical Physics* **2011,** *13* (6), 2356-2366.

136. Ward, R.; Bowman, A.; Sozudogru, E.; El-Mkami, H.; Owen-Hughes, T.; Norman, D. G., EPR Distance Measurements in Deuterated Proteins. *Journal of Magnetic Resonance* **2010,** *207* (1), 164-167.

137. Van Wonderen, J. H.; Kostrz, D. N.; Dennison, C.; Macmillan, F., Refined Distances between Paramagnetic Centers of a Multi-Copper Nitrite Reductase Determined by Pulsed EPR (iDEER) Spectroscopy. *Angewandte Chemie-International Edition* **2013,** *52* (7), 1990-1993.

138. Abe, C.; Klose, D.; Dietrich, F.; Ziegler, W. H.; Polyhach, Y.; Jeschke, G.; Steinhoff, H.-J., Orientation Selective DEER Measurements on Vinculin Tail at X-Band Frequencies Reveal Spin Label Orientations. *Journal of Magnetic Resonance* **2012,** *216*, 53-61.

139. Reginsson, G. W.; Hunter, R. I.; Cruickshank, P. a. S.; Bolton, D. R.; Sigurdsson, S. T.; Smith, G. M.; Schiemann, O., W-Band PELDOR with 1 kW

Microwave Power: Molecular Geometry, Flexibility and Exchange Coupling. *Journal of Magnetic Resonance* **2012,** *216*, 175-182.

140. Tkach, I.; Pornsuwan, S.; Hoebartner, C.; Wachowius, F.; Sigurdsson, S. T.; Baranova, T. Y.; Diederichsen, U.; Sicoli, G.; Bennati, M., Orientation Selection in Distance Measurements between Nitroxide Spin Labels at 94 GHz EPR with Variable Dual Frequency Irradiation. *Physical Chemistry Chemical Physics* **2013,** *15* (10), 3433-3437.

141. Krstic, I.; Haensel, R.; Romainczyk, O.; Engels, J. W.; Doetsch, V.; Prisner, T. F., Long-Range Distance Measurements on Nucleic Acids in Cells by Pulsed EPR Spectroscopy. *Angewandte Chemie-International Edition* **2011,** *50* (22), 5070-5074.

142. Azarkh, M.; Okle, O.; Singh, V.; Seemann, I. T.; Hartig, J. S.; Dietrich, D. R.; Drescher, M., Long-Range Distance Determination in a DNA Model System inside Xenopus laevis Oocytes by in-Cell Spin-Label EPR. *Chembiochem* **2011,** *12* (13), 1992-1995.

143. Gonin, S.; Arnoux, P.; Pierru, B.; Lavergne, J.; Alonso, B.; Sabaty, M.; Pignol, D., Crystal Structures of an Extracytoplasmic Solute Receptor from a Trap Transporter in Its Open and Closed Forms Reveal a Helix-Swapped Dimer Requiring a Cation for Alpha-Keto Acid Binding. *Bmc Structural Biology* **2007,** *7*.

144. Neiditch, M. B.; Federle, M. J.; Pompeani, A. J.; Kelly, R. C.; Swem, D. L.; Jeffrey, P. D.; Bassler, B. L.; Hughson, F. M., Ligand-Induced Asymmetry in Histidine Sensor Kinase Complex Regulates Quorum Sensing. *Cell* **2006,** *126* (6), 1095-1108.

145. Hall, J. A.; Ganesan, A. K.; Chen, J.; Nikaido, H., Two Modes of Ligand Binding in Maltose-Binding Protein of Escherichia coli - Functional Significance in Active Transport. *Journal of Biological Chemistry* **1997,** *272* (28), 17615-17622.

146. Austermuhle, M. I.; Hall, J. A.; Klug, C. S.; Davidson, A. L., Maltose-Binding Protein Is Open in the Catalytic Transition State for ATP Hydrolysis During Maltose Transport. *Journal of Biological Chemistry* **2004,** *279* (27), 28243-28250.

147. Orelle, C.; Ayvaz, T.; Everly, R. M.; Klug, C. S.; Davidson, A. L., Both Maltose-Binding Protein and ATP Are Required for Nucleotide-Binding Domain Closure in the Intact Maltose ABC Transporter. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (35), 12837-12842.

148. Grote, M.; Polyhach, Y.; Jeschke, G.; Steinhoff, H.-J.; Schneider, E.; Bordignon, E., Transmembrane Signaling in the Maltose ABC Transporter MalFGK(2)-E Periplasmic MalF-P2 Loop Communicates Substrate Availability to the ATP-Bound Malk Dimer. *Journal of Biological Chemistry* **2009,** *284* (26), 17521-17526.

149. Selmke, B.; Borbat, P. P.; Nickolaus, C.; Varadarajan, R.; Freed, J. H.; Trommer, W. E., Open and Closed Form of Maltose Binding Protein in Its Native and Molten Globule State as Studied by Electron Paramagnetic Resonance Spectroscopy. *Biochemistry* **2018,** *57* (38), 5507-5512.

150. Sippach, M.; Weidlich, D.; Klose, D.; Abe, C.; Klare, J.; Schneider, E.; Steinhoff, H.-J., Conformational Changes of the Histidine ATP-Binding Cassette Transporter Studied by Double Electron-Electron Resonance Spectroscopy. *Biochim. Biophys. Acta-Biomembr.* **2014,** *1838* (7), 1760-1768.

151. Clatworthy, A. E.; Pierson, E.; Hung, D. T., Targeting Virulence: A New Paradigm for Antimicrobial Therapy. *Nature Chemical Biology* **2007,** *3* (9), 541-548.

152. Bajaj, M.; Mamidyala, S. K.; Zuegg, J.; Begg, S. L.; Ween, M. P.; Luo, Z.; Huang, J. X.; Mcewan, A. G.; Kobe, B.; Paton, J. C.; Mcdevitt, C. A.; Cooper, M. A., Discovery of Novel Pneumococcal Surface Antigen a (PsaA) Inhibitors Using a Fragment-Based Drug Design Approach. *ACS Chem. Biol.* **2015,** *10* (6), 1511-1520.

153. Mcdevitt, C. A.; Ogunniyi, A. D.; Valkov, E.; Lawrence, M. C.; Kobe, B.; Mcewan, A. G.; Paton, J. C., A Molecular Mechanism for Bacterial Susceptibility to Zinc. *Plos Pathogens* **2011,** *7* (11).

154. Deplazes, E.; Begg, S. L.; Van Wonderen, J. H.; Campbell, R.; Kobe, B.; Paton, J. C.; Macmillan, F.; Mcdevitt, C. A.; O'mara, M. L., Characterizing the Conformational Dynamics of Metal-Free PsaA Using Molecular Dynamics Simulations and Electron Paramagnetic Resonance Spectroscopy. *Biophysical Chemistry* **2015,** *207*, 51-60.

155. Kass, H.; Macmillan, F.; Ludwig, B.; Prisner, T. F., Investigation of the Mn Binding Site in Cytochrome C Oxidase from Paracoccus denitrificans by High-Frequency EPR. *Journal of Physical Chemistry B* **2000,** *104* (22), 5362-5371.

156. Hille, R., The Mononuclear Molybdenum Enzymes. *Chemical Reviews* **1996,** *96* (7), 2757-2816.

157. Rice, A. J.; Harrison, A.; Alvarez, F. J. D.; Davidson, A. L.; Pinkett, H. W., Small Substrate Transport and Mechanism of a Molybdate ATP Binding Cassette Transporter in a Lipid Environment. *Journal of Biological Chemistry* **2014,** *289* (21), 15005-15013.

158. Reynolds, P. R.; Mottur, G. P.; Bradbeer, C., Transport of Vitamin-B12 in Escherichia-Coli - Some Observations on the Roles of the Gene-Products of Btuc and Tonb. *Journal of Biological Chemistry* **1980,** *255* (9), 4313-4319.

159. Deveaux, L. C.; Clevenson, D. S.; Bradbeer, C.; Kadner, R. J., Identification of the Btuced Polypeptides and Evidence for Their Role in Vitamin-B12 Transport in Escherichia-Coli. *Journal of Bacteriology* **1986,** *167* (3), 920-927.

160. Hvorup, R. N.; Goetz, B. A.; Niederer, M.; Hollenstein, K.; Perozo, E.; Locher, K. P., Asymmetry in the Structure of the ABC Transporter-Binding Protein Complex BtuCD-BtuF. *Science* **2007,** *317* (5843), 1387-1390.

161. Van Veen, H. W.; Callaghan, R.; Soceneantu, L.; Sardini, A.; Konings, W. N.; Higgins, C. F., A Bacterial Antibiotic-Resistance Gene That Complements the Human Multidrug-Resistance P-Glycoprotein Gene. *Nature* **1998,** *391* (6664), 291-295.

162. Borbat, P. P.; Surendhran, K.; Bortolus, M.; Zou, P.; Freed, J. H.; Mchaourab, H. S., Conformational Motion of the ABC Transporter MsbA Induced by ATP Hydrolysis. *Plos Biology* **2007,** *5* (10), 2211-2219.

163. Kaur, H.; Abreu, B.; Akhmetzyanov, D.; Lakatos-Karoly, A.; Soares, C. M.; Prisner, T.; Glaubitz, C., Unexplored Nucleotide Binding Modes for the ABC Exporter MsbA. *Journal of the American Chemical Society* **2018,** *140* (43), 14112-14125.

164. Collauto, A.; Mishra, S.; Litvinov, A.; Mchaourab, H. S.; Goldfarb, D., Direct Spectroscopic Detection of ATP Turnover Reveals Mechanistic Divergence of ABC Exporters. *Structure* **2017,** *25* (8), 1264-+.

165. Vanveen, H. W.; Venema, K.; Bolhuis, H.; Oussenko, I.; Kok, J.; Poolman, B.; Driessen, A. J. M.; Konings, W. N., Multidrug Resistance Mediated by a Bacterial Homolog of the Human Multidrug Transporter Mdr1. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** *93* (20), 10668-10672.

166. Hellmich, U. A.; Lyubenova, S.; Moenkemeyer, L.; Kaltenborn, E.; Van Veen, H. W.; Prisner, T.; Glaubitz, C., Investigation of the Multidrug ABC-Transporter Lmra by Multinuclear Mas-NMR and EPR. *Biophysical Journal* **2009,** *96* (3), 594A-594A.

167. Hellmich, U. A.; Lyubenova, S.; Kaltenborn, E.; Doshi, R.; Van Veen, H. W.; Prisner, T. F.; Glaubitz, C., Probing the ATP Hydrolysis Cycle of the ABC Multidrug Transporter Lmra by Pulsed EPR Spectroscopy. *Journal of the American Chemical Society* **2012,** *134* (13), 5857-5862.

168. Omote, H.; Al-Shawi, M. K., A Novel Electron Paramagnetic Resonance Approach to Determine the Mechanism of Drug Transport by P-Glycoprotein. *Journal of Biological Chemistry* **2002,** *277* (47), 45688-45694.

169. Delannoy, S.; Urbatsch, I. L.; Tombline, G.; Senior, A. E.; Vogel, P. D., Nucleotide Binding to the Multidrug Resistance P-Glycoprotein as Studied by ESR Spectroscopy. *Biochemistry* **2005,** *44* (42), 14010-14019.

170. Wen, P.-C.; Verhalen, B.; Wilkens, S.; Mchaourab, H. S.; Tajkhorshid, E., On the Origin of Large Flexibility of P-Glycoprotein in the Inward-Facing State. *Journal of Biological Chemistry* **2013,** *288* (26), 19211-19220.

171. Aller, S. G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P. M.; Trinh, Y. T.; Zhang, Q.; Urbatsch, I. L.; Chang, G., Structure of P-Glycoprotein Reveals a Molecular Basis for Poly-Specific Drug Binding. *Science* **2009,** *323* (5922), 1718-1722.

172. Van Wonderen, J. H.; Mcmahon, R. M.; O'mara, M. L.; Mcdevitt, C. A.; Thomson, A. J.; Kerr, I. D.; Macmillan, F.; Callaghan, R., The Central Cavity of ABCB1 Undergoes Alternating Access During ATP Hydrolysis. *Febs Journal* **2014,** *281* (9), 2190-2201.

173. Timachi, M. H.; Hutter, C. a. J.; Hohl, M.; Assafa, T.; Bohm, S.; Mittal, A.; Seeger, M. A.; Bordignon, E., Exploring Conformational Equilibria of a Heterodimeric ABC Transporter. *eLife* **2017,** *6*, 28.

174. Goddeke, H.; Timachi, M. H.; Hutter, C. a. J.; Galazzo, L.; Seeger, M. A.; Karttunen, M.; Bordignon, E.; Schafer, L. V., Atomistic Mechanism of Large-Scale Conformational Transition in a Heterodimeric ABC Exporter. *Journal of the American Chemical Society* **2018,** *140* (13), 4543-4551.

175. Delmar, J. A.; Su, C.-C.; Yu, E. W., Structural Mechanisms of Heavy-Metal Extrusion by the Cus Efflux System. *Biometals* **2013,** *26* (4), 593-607.

176. Franke, S.; Grass, G.; Rensing, C.; Nies, D. H., Molecular Analysis of the Copper-Transporting Efflux System CusCFBA of Escherichia coli. *Journal of Bacteriology* **2003,** *185* (13), 3804-3812.

177. Meir, A.; Natan, A.; Moskovitz, Y.; Ruthstein, S., EPR Spectroscopy Identifies Met and Lys Residues That Are Essential for the Interaction between the CusB N-Terminal Domain and Metallochaperone CusF. *Metallomics* **2015,** *7* (7), 1163-1172.

178. Meir, A.; Abdelhai, A.; Moskovitz, Y.; Ruthstein, S., EPR Spectroscopy Targets Structural Changes in the E. coli Membrane Fusion CusB Upon Cu(I) Binding. *Biophysical Journal* **2017,** *112* (12), 2494-2502.

179. Kuroda, T.; Tsuchiya, T., Multidrug Efflux Transporters in the MATE Family. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2009,** *1794* (5), 763-768.

180. Bay, D. C.; Rommens, K. L.; Turner, R. J., Small Multidrug Resistance Proteins: A Multidrug Transporter Family That Continues to Grow. *Biochim. Biophys. Acta-Biomembr.* **2008,** *1778* (9), 1814-1838.

181. Reizer, J.; Reizer, A.; Saier, M. H., A Functional Superfamily of Sodium/Solute Symporters. *Biochimica Et Biophysica Acta-Reviews on Biomembranes* **1994,** *1197* (2), 133-166.

182. Wegener, C.; Tebbe, S.; Steinhoff, H. J.; Jung, H. R., Spin Labeling Analysis of Structure and Dynamics of the Na+/Proline Transporter of Escherichia coli. *Biochemistry* **2000,** *39* (16), 4831-4837.

183. Jung, H.; Rubenhagen, R.; Tebbe, S.; Leifker, K.; Tholema, N.; Quick, M.; Schmid, R., Topology of the Na+/Proline Transporter of Escherichia coli. *Journal of Biological Chemistry* **1998,** *273* (41), 26400-26407.

184. Jeschke, G.; Wegener, C.; Nietschke, M.; Jung, H.; Steinhoff, H. J., Interresidual Distance Determination by Four-Pulse Double Electron-Electron Resonance in an Integral Membrane Protein: The Na+/Proline Transporter PutP of Escherichia coli. *Biophysical Journal* **2004,** *86* (4), 2551-2557.

185. Hilger, D.; Polyhach, Y.; Jung, H.; Jeschke, G., Backbone Structure of Transmembrane Domain IX of the Na+/Proline Transporter PutP of Escherichia coli. *Biophysical Journal* **2009,** *96* (1), 217-225.

186. Hilger, D.; Bohm, M.; Hackmann, A.; Jung, H., Role of Ser-340 and Thr-341 in Transmembrane Domain IX of the Na+/Proline Transporter PutP of Escherichia coli in Ligand Binding and Transport. *Journal of Biological Chemistry* **2008,** *283* (8), 4921-4929.

187. Raba, M.; Dunkel, S.; Hilger, D.; Lipiszko, K.; Polyhach, Y.; Jeschke, G.; Bracher, S.; Klare, J. P.; Quick, M.; Jung, H.; Steinhoff, H.-J., Extracellular Loop 4 of the Proline Transporter PutP Controls the Periplasmic Entrance to Ligand Binding Sites. *Structure* **2014,** *22* (5), 769-780.

188. Bracher, S.; Guerin, K.; Polyhach, Y.; Jeschke, G.; Dittmer, S.; Frey, S.; Bohm, M.; Jung, H., Glu-311 in External Loop 4 of the Sodium/Proline Transporter PutP Is Crucial for External Gate Closure. *Journal of Biological Chemistry* **2016,** *291* (10), 4998-5008.

189. Weyand, S.; Shimamura, T.; Yajima, S.; Suzuki, S. I.; Mirza, O.; Krusong, K.; Carpenter, E. P.; Rutherford, N. G.; Hadden, J. M.; O'reilly, J.; Ma, P.; Saidijam, M.; Patching, S. G.; Hope, R. J.; Norbertczak, H. T.; Roach, P. C. J.; Iwata, S.; Henderson, P. J. F.; Cameron, A. D., Structure and Molecular Mechanism of a Nucleobase-Cation-Symport-1 Family Transporter. *Science* **2008,** *322* (5902), 709-713.

190. Shimamura, T.; Weyand, S.; Beckstein, O.; Rutherford, N. G.; Hadden, J. M.; Sharples, D.; Sansom, M. S. P.; Iwata, S.; Henderson, P. J. F.; Cameron, A. D., Molecular Basis of Alternating Access Membrane Transport by the Sodium-Hydantoin Transporter Mhp1. *Science* **2010,** *328* (5977), 470-473.

191. Kazmier, K.; Sharma, S.; Islam, S. M.; Roux, B.; Mchaourab, H. S., Conformational Cycle and Ion-Coupling Mechanism of the Na+/Hydantoin Transporter Mhp1. *Proceedings of the National Academy of Sciences of the United States of America* **2014,** *111* (41), 14752-14757.

192. Claxton, D. P.; Quick, M.; Shi, L.; De Carvalho, F. D.; Weinstein, H.; Javitch, J. A.; Mchaourab, H. S., Ion/Substrate-Dependent Conformational Dynamics of a Bacterial Homolog of Neurotransmitter: Sodium Symporters. *Nat. Struct. Mol. Biol.* **2010,** *17* (7), 822-U68.

193. Ressl, S.; Van Scheltinga, A. C. T.; Vonrhein, C.; Ott, V.; Ziegler, C., Molecular Basis of Transport and Regulation in the Na+/Betaine Symporter BetP. *Nature* **2009,** *458* (7234), 47-U1.

194. Tsai, C.-J.; Khafizov, K.; Hakulinen, J.; Forrest, L. R.; Kraemer, R.; Kuehlbrandt, W.; Ziegler, C., Structural Asymmetry in a Trimeric Na+/Betaine Symporter, BetP, from Corynebacterium Glutamicum. *Journal of Molecular Biology* **2011,** *407* (3), 368-381.

195. Nicklisch, S. C. T.; Wunnicke, D.; Borovykh, I. V.; Morbach, S.; Klare, J. P.; Steinhoff, H. J.; Kraemer, R., Conformational Changes of the Betaine Transporter BetP from Corynebacterium Glutamicum Studied by Pulse EPR Spectroscopy. *Biochim. Biophys. Acta-Biomembr.* **2012,** *1818* (3), 359-366.

196. Ott, V.; Koch, J.; Spaete, K.; Morbach, S.; Kraemer, R., Regulatory Properties and Interaction of the C- and N-Terminal Domains of BetP, an Osmoregulated Betaine Transporter from Corynebacterium Glutamicum. *Biochemistry* **2008,** *47* (46), 12208-12218.

197. Saier, M. H., Jr.; Beatty, J. T.; Goffeau, A.; Harley, K. T.; Heijne, W. H.; Huang, S. C.; Jack, D. L.; Jahn, P. S.; Lew, K.; Liu, J.; Pao, S. S.; Paulsen, I. T.; Tseng, T. T.; Virk, P. S., The Major Facilitator Superfamily. *Journal of molecular microbiology and biotechnology* **1999,** *1* (2), 257-79.

198. Zheng, H.; Taraska, J.; Merz, A. J.; Gonen, T., The Prototypical H+/Galactose Symporter GalP Assembles into Functional Trimers. *Journal of Molecular Biology* **2010,** *396* (3), 593-601.

199. Marsh, D.; Henderson, P. J. F., Specific Spin Labelling of the Sugar-H+ Symporter, GalP, in Cell Membranes of Escherichia coli: Site Mobility and Overall Rotational Diffusion of the Protein. *Biochim. Biophys. Acta-Biomembr.* **2001,** *1510* (1-2), 464-473.

200. Hubert, A.; Henderson, P. J. F.; Marsh, D., Lipid-Protein Interactions in Escherichia coli Membranes Overexpressing the Sugar-H+ Symporter, GalP EPR of Spin-Labelled Lipids. *Biochim. Biophys. Acta-Biomembr.* **2003,** *1611* (1-2), 243-248.

201. Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H. R.; Iwata, S., Structure and Mechanism of the Lactose Permease of Escherichia coli. *Science* **2003,** *301* (5633), 610-615.

202. Smirnova, I.; Kasho, V.; Kaback, H. R., Lactose Permease and the Alternating Access Mechanism. *Biochemistry* **2011,** *50* (45), 9684-9693.

203. Smirnova, I.; Kasho, V.; Choe, J.-Y.; Altenbach, C.; Hubbell, W. L.; Kaback, H. R., Sugar Binding Induces an Outward-Facing Conformation of LacY. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (42), 16504-16509.

204. Hanelt, I.; Wunnicke, D.; Bordignon, E.; Steinhoff, H.-J.; Slotboom, D. J., Conformational Heterogeneity of the Aspartate Transporter Glt(Ph). *Nat. Struct. Mol. Biol.* **2013,** *20* (2), 210-214.

205. Georgieva, E. R.; Borbat, P. P.; Ginter, C.; Freed, J. H.; Boudker, O., Conformational Ensemble of the Sodium-Coupled Aspartate Transporter. *Nat. Struct. Mol. Biol.* **2013,** *20* (2), 215-221.

206. Padan, E.; Venturi, M.; Gerchman, Y.; Dover, N., Na+/H+ Antiporters. *Biochimica Et Biophysica Acta-Bioenergetics* **2001,** *1505* (1), 144-157.

207. Hunte, C.; Screpanti, E.; Venturi, M.; Rimon, A.; Padan, E.; Michel, H., Structure of a Na+/H+ Antiporter and Insights into Mechanism of Action and Regulation by pH. *Nature* **2005,** *435* (7046), 1197-1202.

208. Hilger, D.; Jung, H.; Padan, E.; Wegener, C.; Vogel, K. P.; Steinhoff, H. J.; Jeschke, G., Assessing Oligomerization of Membrane Proteins by Four-Pulse DEER: pH-Dependent Dimerization of NhaA Na+/H+ Antiporter of E-coli. *Biophysical Journal* **2005,** *89* (2), 1328-1338.

209. Hilger, D.; Polyhach, Y.; Padan, E.; Jung, H.; Jeschke, G., High-Resolution Structure of a Na+/H+ Antiporter Dimer Obtained by Pulsed Election Paramagnetic Resonance Distance Measurements. *Biophysical Journal* **2007,** *93* (10), 3675-3683.

210. Nygaard, E. B.; Lagerstedt, J. O.; Bjerre, G.; Shi, B.; Budamagunta, M.; Poulsen, K. A.; Meinild, S.; Rigor, R. R.; Voss, J. C.; Cala, P. M.; Pedersen, S. F., Structural Modeling and Electron Paramagnetic Resonance Spectroscopy of the

Human Na+/H+ Exchanger Isoform 1, Nhe1. *Journal of Biological Chemistry* **2011,** *286* (1), 634-648.

211. Anantharam, V.; Allison, M. J.; Maloney, P. C., Oxalate - Formate Exchange - the Basis for Energy Coupling in Oxalobacter. *Journal of Biological Chemistry* **1989,** *264* (13), 7244-7250.

212. Iyalomhe, O.; Herrick, D. Z.; Cafiso, D. S.; Maloney, P. C., Closure of the Cytoplasmic Gate Formed by Tm5 and Tm11 During Transport in the Oxalate/Formate Exchanger from Oxalobacter formigenes. *Biochemistry* **2014,** *53* (49), 7735-7744.

213. Lagerstedt, J. O.; Voss, J. C.; Wieslander, A.; Persson, B. L., Structural Modeling of Dual-Affinity Purified Pho84 Phosphate Transporter. *Febs Letters* **2004,** *578* (3), 262-268.

214. Putman, M.; Van Veen, H. W.; Degener, J. E.; Konings, W. N., The Lactococcal Secondary Multidrug Transporter LmrP Confers Resistance to Lincosamides, Macrolides, Streptogramins and Tetracyclines. *Microbiology-Sgm* **2001,** *147*, 2873-2880.

215. Masureel, M.; Smriti, S.; Martens, C.; Zou, P.; Ruysschaert, J.-M.; Mchaourab, H. S.; Govaerts, C., Studying the Conformational Cycle of the Secondary Multidrug Transporter LmrP by EPR Spectroscopy. *Biophysical Journal* **2012,** *102* (3), 660A-660A.

216. Masureel, M.; Martens, C.; Stein, R. A.; Mishra, S.; Ruysschaert, J.-M.; Mchaourab, H. S.; Govaerts, C., Protonation Drives the Conformational Switch in the Multidrug Transporter LmrP. *Nature Chemical Biology* **2014,** *10* (2), 149-155.

217. Otsuka, M.; Yasuda, M.; Morita, Y.; Otsuka, C.; Tsuchiya, T.; Omote, H.; Moriyama, Y., Identification of Essential Amino Acid Residues of the NorM Na+/Multidrug Antiporter in Vibrio parahaemolyticus. *Journal of Bacteriology* **2005,** *187* (5), 1552-1558.

218. Moriyama, Y.; Hiasa, M.; Matsumoto, T.; Omote, H., Multidrug and Toxic Compound Extrusion (MATE)-Type Proteins as Anchor Transporters for the Excretion of Metabolic Waste Products and Xenobiotics. *Xenobiotica* **2008,** *38* (7-8), 1107-1118.

219. Van Veen, H. W., Structural Biology Last of the Multidrug Transporters. *Nature* **2010,** *467* (7318), 926-927.

220. Steed, P. R.; Stein, R. A.; Mishra, S.; Goodman, M. C.; Mchaourab, H. S., Na+-Substrate Coupling in the Multidrug Antiporter NorM Probed with a Spin-Labeled Substrate. *Biochemistry* **2013,** *52* (34), 5790-5799.

221. He, X.; Szewczyk, P.; Karyakin, A.; Evin, M.; Hong, W.-X.; Zhang, Q.; Chang, G., Structure of a Cation-Bound Multidrug and Toxic Compound Extrusion Transporter. *Nature* **2010,** *467* (7318), 991-U139.

222. Lu, M.; Symersky, J.; Radchenko, M.; Koide, A.; Guo, Y.; Nie, R.; Koide, S., Structures of a Na+-Coupled, Substrate-Bound MATE Multidrug Transporter. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (6), 2099-2104.

223. Claxton, D. P.; Jagessar, K. L.; Steed, P. R.; Stein, R. A.; Mchaourab, H. S., Sodium and Proton Coupling in the Conformational Cycle of a MATE Antiporter from Vibrio cholerae. *Proceedings of the National Academy of Sciences of the United States of America* **2018,** *115* (27), E6182-E6190.

224. Littlejohn, T. G.; Diberardino, D.; Messerotti, L. J.; Spiers, S. J.; Skurray, R. A., Structure and Evolution of a Family of Genes Encoding Antiseptic and Disinfectant Resistance in Staphylococcus-Aureus. *Gene* **1991,** *101* (1), 59-66.

225. Paulsen, I. T.; Skurray, R. A.; Tam, R.; Saler, M. H.; Turner, R. J.; Weiner, J. H.; Goldberg, E. B.; Grinius, L. L., The Smr Family: A Novel Family of Multidrug Efflux Proteins Involved with the Efflux of Lipophilic Drugs. *Mol. Microbiol.* **1996,** *19* (6), 1167-1175.

226. Jack, D. L.; Storms, M. L.; Tchieu, J. H.; Paulsen, I. T.; Saier, M. H., A Broad-Specificity Multidrug Efflux Pump Requiring a Pair of Homologous Smr-Type Proteins. *Journal of Bacteriology* **2000,** *182* (8), 2311-2313.

227. Koteiche, H. A.; Reeves, M. D.; Mchaourab, H. S., Structure of the Substrate Binding Pocket of the Multidrug Transporter EmrE: Site-Directed Spin Labeling of Transmembrane Segment 1. *Biochemistry* **2003,** *42* (20), 6099-6105.

228. Schuldiner, S., EmrE, a Model for Studying Evolution and Mechanism of Ion-Coupled Transporters. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2009,** *1794* (5), 748-762.

229. Korkhov, V. M.; Tate, C. G., An Emerging Consensus for the Structure of EmrE. *Acta Crystallographica Section D-Structural Biology* **2009,** *65*, 186-192.

230. Steiner-Mordoch, S.; Soskine, M.; Solomon, D.; Rotem, D.; Gold, A.; Yechieli, M.; Adam, Y.; Schuldiner, S., Parallel Topology of Genetically Fused EmrE Homodimers. *Embo Journal* **2008,** *27* (1), 17-26.

231. Hellmich, U. A.; Glaubitz, C., NMR and EPR Studies of Membrane Transporters. *Biological Chemistry* **2009,** *390* (8), 815-834.

232. Mchaourab, H. S.; Mishra, S.; Koteiche, H. A.; Amadi, S. H., Role of Sequence Bias in the Topology of the Multidrug Transporter EmrE. *Biochemistry* **2008,** *47* (31), 7980-7982.

233. Amadi, S. T.; Koteiche, H. A.; Mishra, S.; Mchaourab, H. S., Structure, Dynamics, and Substrate-Induced Conformational Changes of the Multidrug Transporter EmrE in Liposomes. *Journal of Biological Chemistry* **2010,** *285* (34), 26710-26718.

234. Dastvan, R.; Fischer, A. W.; Mishra, S.; Meiler, J.; Mchaourab, H. S., Protonation-Dependent Conformational Dynamics of the Multidrug Transporter EmrE. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (5), 1220-1225.

235. Yan, N., Structural Biology of the Major Facilitator Superfamily Transporters. In *Annual Review of Biophysics, Vol 44*, Dill, K. A., Ed. 2015; Vol. 44, pp 257-283.

236. Steed, P. R.; Zou, P.; Trone, K. E.; Mchaourab, H. S., Structure and pH-Induced Structural Rearrangements of the Putative Multidrug Efflux Pump EmrD in Liposomes Probed by Site-Directed Spin Labeling. *Biochemistry* **2013,** *52* (45), 7964-7974.

Results & Discussion

During the course of this thesis, various different systems were investigated using the methods and techniques outlined in the introductory sections. Here, these systems are broken down and categorised into ABC-type transporter SBDs (MalE, GlnPQ SBDs 1 & 2, and PsaA) and secondary active transporters (VcINDY). Though the underlying biological and mechanistic principles of these systems has been broadly explained in the preceding sections, the specifics will be explored in more detail here.

ABC-type Transporter SBDs

4 Maltose Binding Protein

4.1 Background

The ABC-type transporter responsible for importing malto-oligosaccharides and maltodextrins in *Escherichia coli*, MalEFGK$_2$, is one of the most widely studied ABC system to date[1] and is often used as a model complex for this class of transporter. The SBD subunit of this system, MalE, is sometimes referred to as maltose binding protein (MBP). As a member of the cluster B subgroup of SBPs the general architecture of the protein is two globular lobes connected by a hinge region of three interconnecting strands[2]. Crystal structures exist of this SBD in both the open, ligand-free (PDB 1OMP[3]; fig. 4.1.1 and fig. 4.1.2, green) and closed, ligand-bound forms (1ANF with maltose; fig. 4.1.1 and fig. 4.1.2, blue. Also, 3MBP and 4MBP– bound substrates maltotriose and maltotetraose respectively[4]).

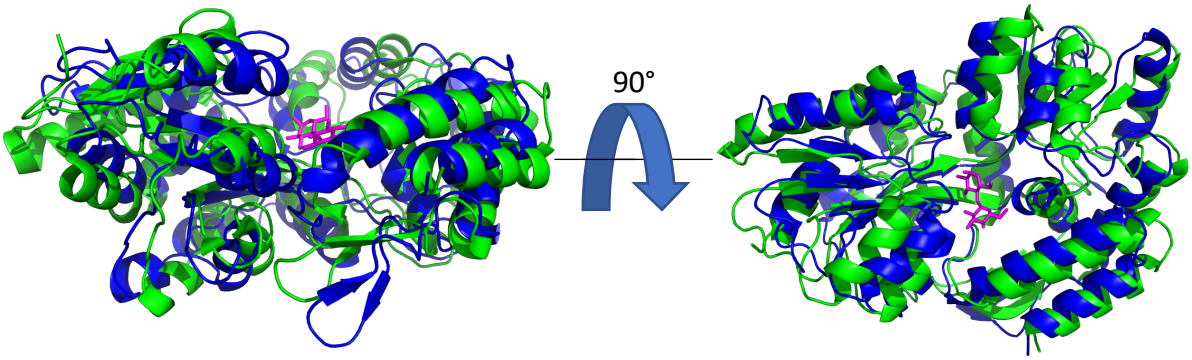


Fig. 4.1.1: overlaid crystal structures of MalE in both the apoprotein and maltose-bound states from different aspects showing significant conformational differences: green is apoprotein and blue is maltose-bound (PDB 1OMP and 1ANF respectively). Maltose is shown in magenta.

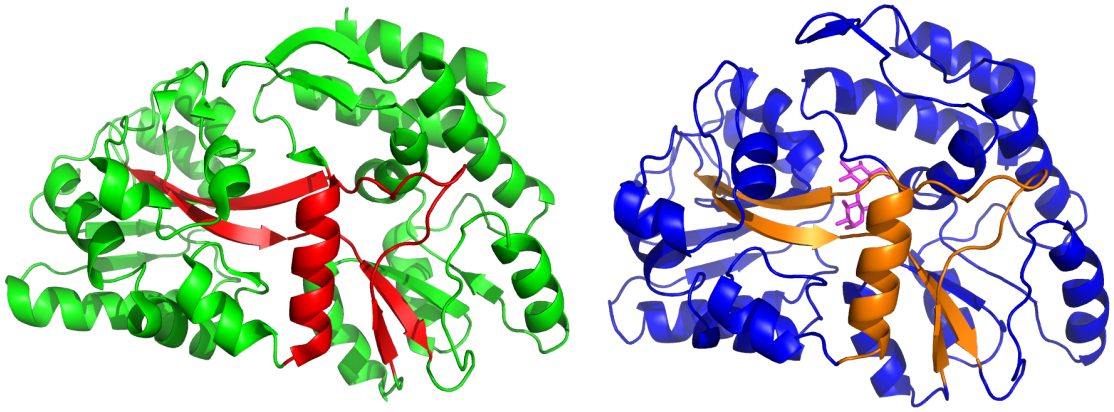


Fig. 4.1.2: crystal structures of MalE in both the apoprotein and maltose-bound states with the hinge regions highlighted: green is apoprotein (hinge highlighted in red) and blue is maltose-bound (hinge highlighted in orange; PDB 1OMP and 1ANF respectively). Maltose is shown in magenta.

The crystal structures of the apo and maltose-bound forms of MBP clearly shows the conformational changes of the lobes opening and closing around the substrate binding cleft, about the hinge region as a classic example of the posited 'Venus flytrap' mechanism. However, there is still some dispute over the overall mechanism of maltose import as it relates to the entire complex of $MalEFGK_2$.

In one theorised model the maltose-bound (demonstrated below, fig. 4.1.3), the closed form of MalE binds to the TMDs, MalF and MalG, in the inward-facing conformation; substrate binding to free MalE has a dissociation constant (K_d) of ~2 μM, and the binding of the substrate-bound MalE to the rest of the complex has a K_d greater than ~45 μM[5]. This then switches to the outward-facing state as ATP binds at the NBDs (MalK); the separation of the TMDs facilitates the opening of the MalE to deposit the substrate into the cavity for translocation. ATP hydrolysis then occurs, ADP and P_i are released, and the TMDs rearrange to the inward-

facing state, allowing the release of substrate into the cytoplasm[6-7]. A structure of the complex in a stabilised intermediate of the mechanism has been reported, in which the ATP-bound NBDs are locked in their dimerised state and the TMDs in the outward-facing conformation. In this, the MalE sits atop the TMDs in the open state and electron density for maltose was found in the bottom of the TMD cleft (when maltose was modelled in to match the electron density, this fit well with the interactions of surrounding resides)[6].

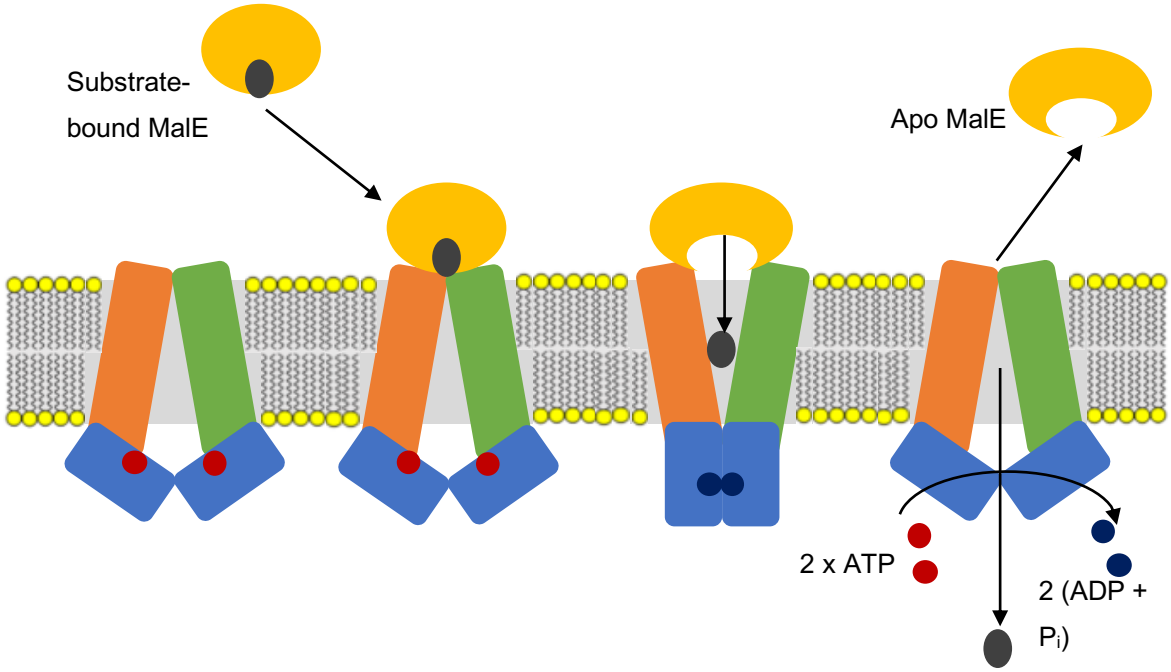


Fig. 4.1.3: representation of one of the proposed transport mechanisms of MalEFGK$_2$, in which the loaded SBD, MalE, docks onto the TMDs in the inward-facing conformation, before ATP hydrolysis causes conformational change to facilitate the deposition of substrate from translocation.

An opposing model theorises that the ligand-free, open state of MalE interacts with the TMDs in the outward-facing conformation (K_d ~ 50-80 nM), which can then allow substrate to bind (K_d ~120 μM)[5] (this is shown in fig. 4.1.4). Hydrolysis of ATP drives the conformational change of the TMDs to the inward-facing state, thereby releasing the substrate into the cytoplasm. This model states that MalE and maltose are responsible for the return of the system to the inward-facing state, but if ATP hydrolysis does not occur immediately upon substrate binding, the ligand-bound, closed MalE will dissociate from the complex. This may also occur as an autoregulatory mechanism in the presence of excess maltose[5, 8].

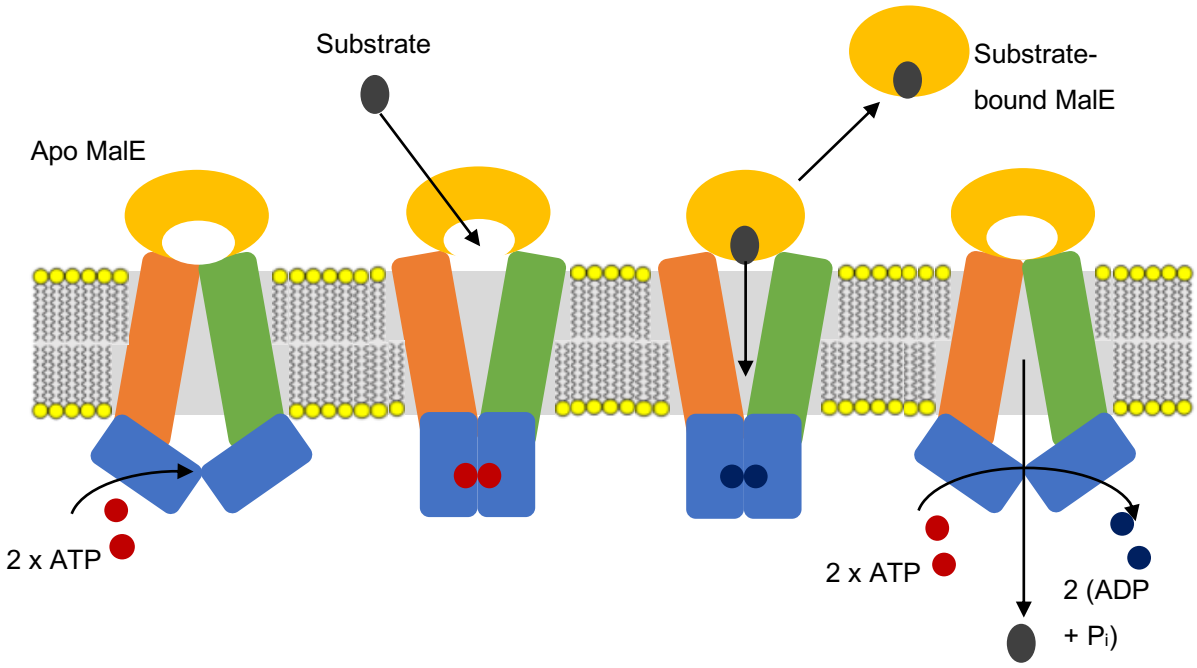


Fig. 4.1.4: representation of one of the opposing proposed transport mechanisms of MalEFGK$_2$, in which the unloaded SBD, MalE, is docked onto the TMDs in the inward-facing conformation. ATP binding causes the conformational change that allows substrate binding and deposition. Hydrolysis then drives the conformational change back to the inward-facing state, or if this is not immediate, the loaded MalE dissociates from the complex.

The functional dynamics of the overall transporter have been well studied using EPR spectroscopic techniques[9], but the focus on the SBP itself is relatively lacking[1]. However, some FRET investigations have previously been undertaken to elucidate some mechanistic detail of the binding of MalE to various malto-oligosaccharides and maltodextrins[10-11]. This provides an opportunity to compare FRET and EPR methods for distance measurements in relation to conformational change; here, this thesis takes MalE as a well-studied, prototypical model in studying SBD functional dynamics as a result of substrate binding.

4.2 Results and Discussion

in silico spin labelling

The cysteine-containing variant of MalE, T36C/S352C, was originally designed to investigate the binding of the substrate using sm-FRET. This method is very complementary with EPR spectroscopy, and so the same construct can be used. The cysteine mutation can be performed computationally, and the spin label attached to it *in silico* using the crystal structure. A library of rotamers of the spin label, based on the space around the site of attachment and its rotational freedom, can be produced. Here, where we have two spin labels being computationally attached, the predicted distance between them is calculated, and the distance distribution (i.e. weighted probabilities of distances populated between the spin label rotamers) can be predicted. The rotamer libraries were calculated for both of the labelling positions in MalE (figure 4.2.1 and table 4.2.1) under both the apo and maltose-bound from the crystal structures; these were calculated at ambient temperature (298 K) using Multiscale Modelling of Macromolecular Systems (MMM) software package that runs via MATLAB[12].

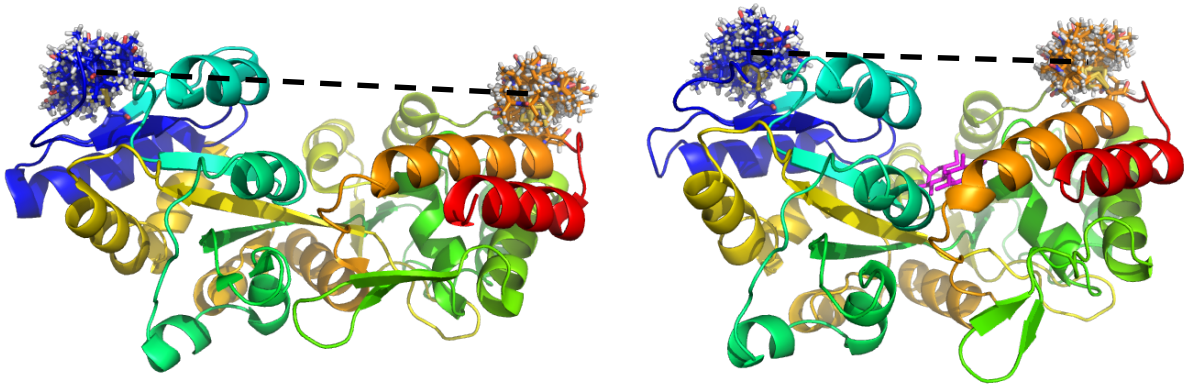


Fig. 4.2.1: the crystal structures of the apo (left) and maltose-bound (right) forms of MalE (PDB 1OMP and 1ANF respectively) with MMM-generated rotamer libraries highlighted, at the residue positions where cysteines will be introduced; the colouring indicates the sequence of the protein from N- to C-terminus, blue to red, with maltose highlighted in pink. The distances to be measured are indicated by dashed lines connecting the pairs of rotamer libraries.

Table 4.2.1: rotamer libraries of the MalE variant cysteine sites in the apo and maltose-bound forms, calculated from the crystal structures at 298 K using spin label MTSL.

Variant	Calculated Rotamer Library (298 K)	
	Apo	**+ Maltose**
T36C	72	59
S352C	47	34

The rotamer libraries calculated here show relatively small differences between the apoprotein and maltose-bound states for the individual mutation. This may indicate that the area surrounding the mutation site won't change dramatically between the different binding states and so it is unlikely that any noticeable differences will be observed in the lineshapes of the room temperature X-band cw-EPR spectra.

The predicted distances and distance distributions for the double label pairs were then calculated (table 4.2.2, below).

Table 4.2.2: the predicted distances between the pair of spin labels in the double-labelled variant of MalE as extracted from the predicted distance distributions, calculated at 298 K using spin label MTSL.

Variant	Predicted inter-label distance (Å) (MMM @ 298 K)	
	Apo	**+ Maltose**
T36C/S352C	52.5	40.0

cw-EPR & simulations

Room temperature X-band

The aforementioned variant of MalE was spin labelled with MTSL (details in section 13.1 ‘Site directed spin labelling’). Labelling efficiencies of approximately 30-50% were routinely observed; these were calculated using a calibration curve of spin label at known concentrations (double integration of the first derivative signal yields area under the curve, which is related to the concentration of the radical species in the sample), the spin label signal of the labelled protein and then comparing with the measured concentration of protein, accounting for the presence of two labelable cysteines per protein. The room temperature cw-EPR spectrum of the double-labelled variant of MalE was recorded at X-band.

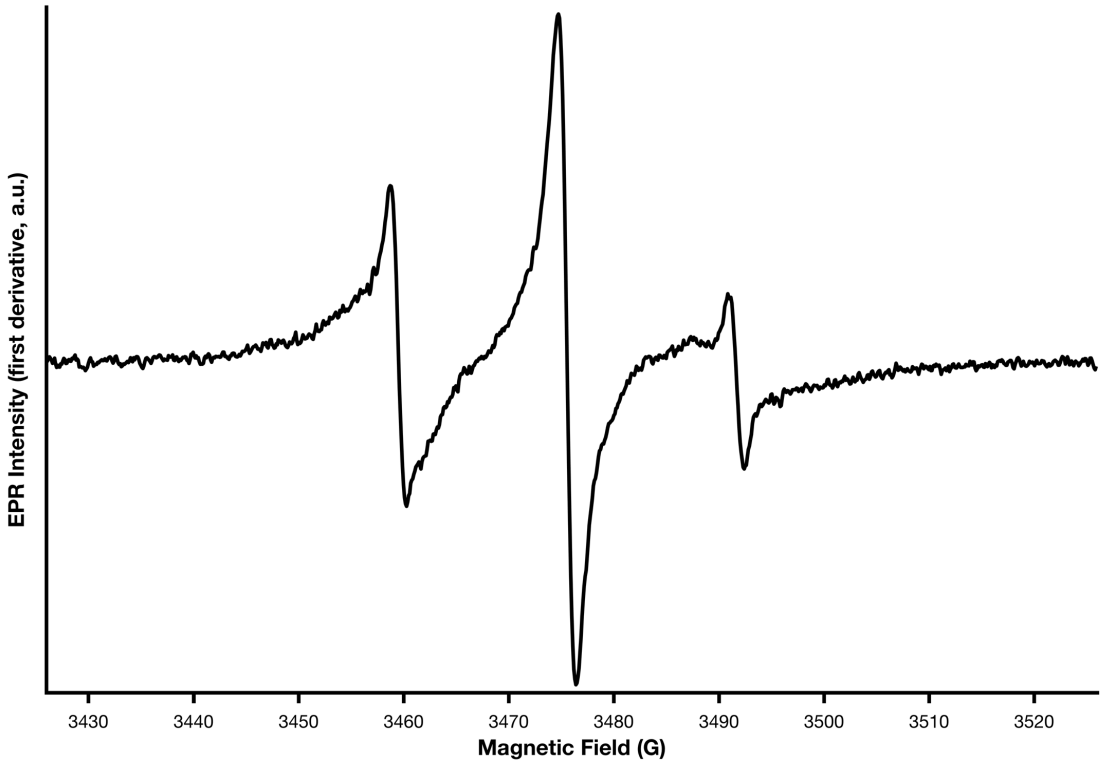


Figure 4.2.2: the room temperature cw-EPR spectrum of the apoprotein, MTSL-labelled T36C/S352C variant of MalE at X-band.

Initially, this spectrum of the apoprotein form of MalE shows that the attached spin labels have a fairly high degree of motional freedom, though not to the extent of spin label tumbling freely in solution; the spectrum exhibits relatively sharp and symmetric peaks, though the relative intensity ratio of these indicates that the label is indeed attached to a larger molecule. Looking more carefully at the spectral features indicates there may be a more restrictive environment surrounding the

spin label; this can be observed in the asymmetric features around the higher- and lower-field peaks, and in the broadness of the central peak. This may arise due to a number of different reasons, or a combination thereof. For example, through distinct conformations being sampled, different orientations of attached spin labels, or a difference in the environment around each of the cysteine mutation sites (this will be discussed further below).

Simulation was performed using the EasySpin package[13] for MATLAB. Two separate spin label species were defined to account for the spectral features described above: one with shorter τ_c to describe the more rotationally-free environment, and the other with longer τ_c to define the more restricted environment (see table 4.2.3). The relative weightings of these species were varied. The simulation was iteratively fitted to the experimental data in order to extract values for certain parameters that could then be quantitatively assessed. Approximate ranges of values of the g- and A-tensors for MTSL and MTSL-labelled proteins are well-documented[14-15] and have been used to input starting values for the simulations and fittings performed here.

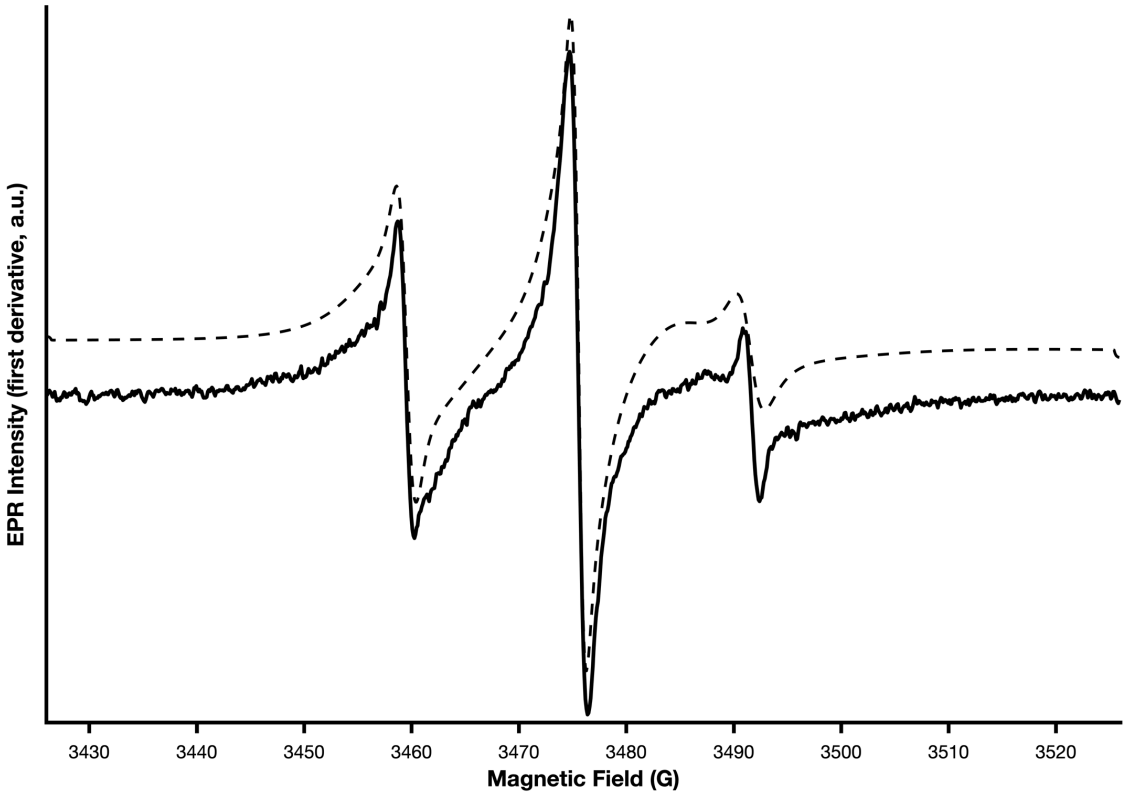


Figure 4.2.3: the fitted simulations of the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled T36C/S352C variant of MalE at X-band; experimental data is shown as a solid line, fitted simulation is shown as a dotted line.

Table 4.2.3: best fit simulation values for the x, y and z components of each of the g- and A-tensors, as well as linewidth and averaged correlation times for each spin label ' species' contributing to the room temperature X-band experimental data of the MTSL-labelled apoprotein MalE, with their respective, relative weightings. Values were obtained from simulation and fitting of the experimental data. The RMSD of the fitting is given

Component	**g-tensor**			**A-tensor (MHz)**			**Linewidth (peak-to-peak, mT)**	**τ_c (ns)**	**Relative weighting**	**Simulation RMSD**
	x	y	z	x	y	z				
1	2.0085	2.006	2.0022	14.1	16	104.3	0.01	2.7	0.21	0.028589
2	2.0085	2.006	2.002	14.5	14.2	102.3	0.01	14.4	0.79	

The simulation required two separate spin label species with different τ_c values in order to achieve sufficiently low RMSD (a measure of the difference between the fitted simulation and experimental input data), indicating that there are indeed spin labels in very different environments present in the sample. As previously mentioned, this probably arises due to the chemical and physical environmental differences between the two labelling sites, T36C and S352C, where one site has more rotational freedom around it than the other.

Free label in aqueous solution would typically have a τ_c value of about 0.1 ns (see section 2.2 'Correlation times'); the τ_c value result of the more 'mobile' population (2.7 ns) does not meet this criterion and is therefore not attributable to unbound MTSL. However, the possibility of free label cannot be completely disregarded as there may be some contribution to the sharpest features which were difficult to fit (i.e. a very small population of a third spin label 'species'), obscuring the broader features of the less mobile species. This is observed most prominently the high and low field peaks, where the simulation fits the experimental data the least.

The presence of these sharp peaks may have distorted the relative populations of the two simulated spin label species. As can be deduced from the broadness and asymmetry of the spectral lines underlying the sharper features, it would be expected that the τ_c values for the spin label species/environments would range from about 1 to 5 ns (see section 2.2 'Correlation times'). However, a weighted average of the fitted simulation results yields a τ_c of 11.9 ns, approximately one order of magnitude slower than would be visually estimated.

Low temperature X-band

The X-band cw-EPR spectrum of the T36C/S352C variant of MalE was also recorded at low temperature (50 K).

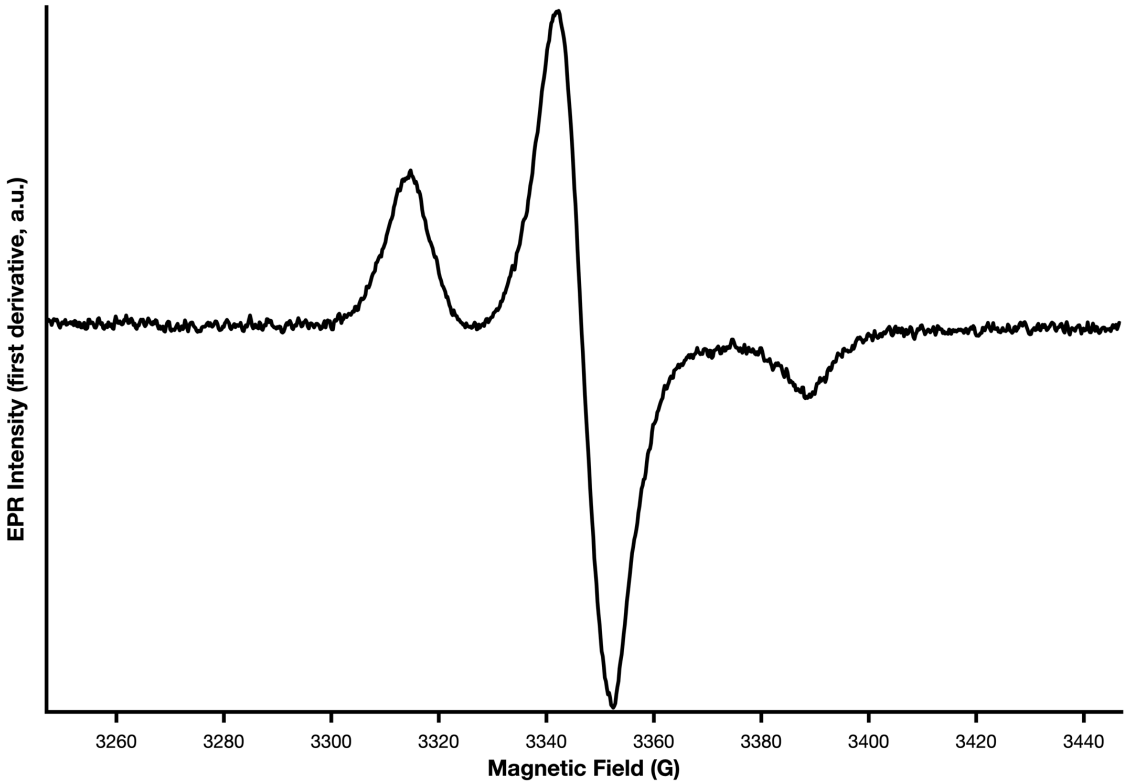


Figure 4.2.4: the low temperature cw-EPR spectra of the apoprotein, MTSL-labelled T36C/S352C variant of MalE at X-band, recorded at 50 K.

The spectrum does not appear to exhibit any broadening as a result of dipolar coupling of two spin labels at a distance of less than about 20 Å; i.e. the areas of baseline between each of the peaks do not deviate significantly from that on either side of the spectral envelope, and the central peak does not appear to be particularly broad. This supports the prediction made in the *in silico* labelling experiment, which calculated an inter-label distance of 52.5 Å in the apoprotein form. This spectrum was then simulated and fitted to the experimental data using the EasySpin package[13] for MATLAB.

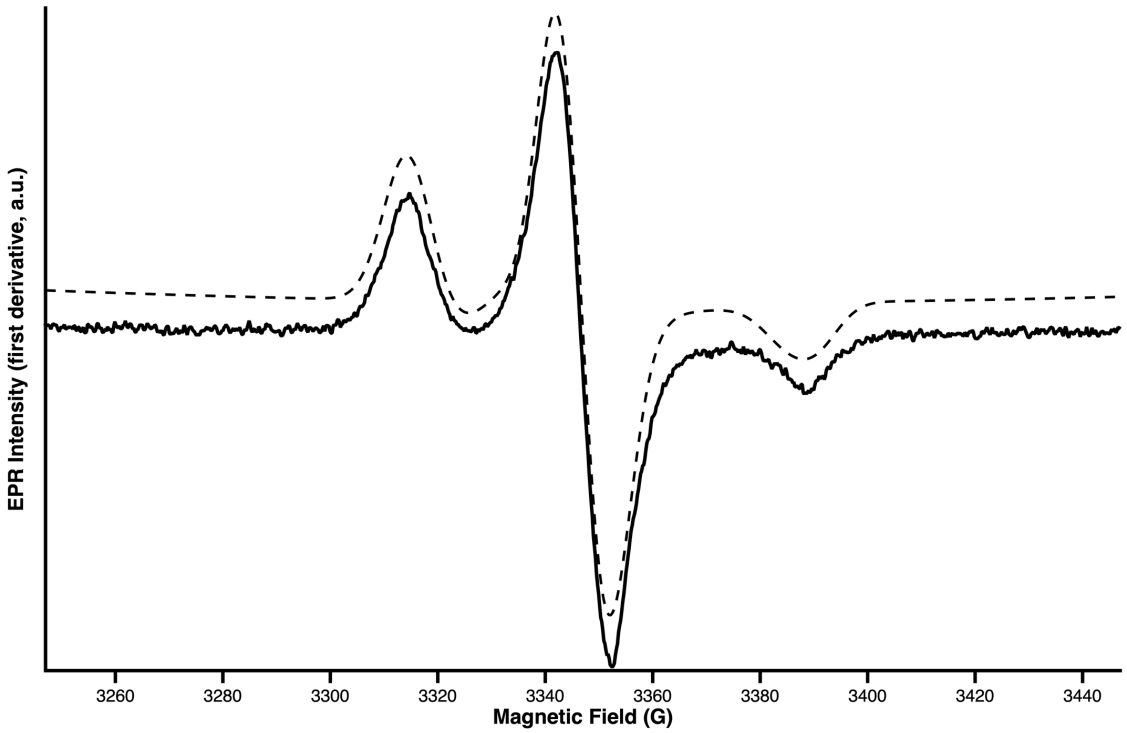


Figure 4.2.5: the fitted simulations of the low temperature cw-EPR spectrum of the apoprotein, MTSL-labelled T36C/S352C variant of MalE at X-band, recorded at 50 K; experimental data are shown as solid lines, fitted simulations are shown as dotted lines. All spectra have been normalised to same scale.

The simulation was performed by taking the results of the best fit of the simulation of the room temperature X-band cw-EPR spectrum as a starting point. The simulation obtained better visual convergence when two separate spin species were accounted for, much like the room temperature data.

Table 4.2.4: best fit simulation values for the x, y and z components of each of the g- and A-tensors and linewidth for the low temperature X-band experimental data of the MTSL-labelled MalE. Two spin label ‘ species’ are simulated for the fitting of a single spectrum; the respective, relative weightings are provided for each component. The RMSD of the fitting is given.

Component	**g-tensor**			**A-tensor (MHz)**			**Linewidth (peak-to-peak, mT)**	**Relative weighting**	**Simulation RMSD**
	x	y	z	x	y	z			
1	2.0079	2.0064	2.0030	14.5	15.7	95.1	0.85	0.35	0.023040453
2	2.0085	2.0052	2.0019	13.1	14.3	107.2	0.87	0.65	

In general, there is good agreement between the values of the x, y and z components of the g-tensors of both spin label 'species' extracted from the fitted simulations of both the room and low temperature spectra. However, this is most likely a product of the limited user-allowed variation of the fitting algorithm around sensible estimates that are well-known from literature studies[14-15]. Indeed, at X-band the values of the x and y components for both the g- and A-tensors cannot be considered reliable. As explained previously (section 2.2 'Multifrequency EPR'), the relatively small g-value anisotropy of organic radicals such as MTSL causes the individual x, y and z contributions to overlap[16-17]. As A_{zz} is significantly larger than A_{xx} or A_{yy} for nitroxide radicals, this dominates the spectrum and masks the features attributed to the remaining g- and A-tensor components. As g-values scale with the applied magnetic field, and A-values do not, W-band EPR (~95 GHz, ~3.5 T applied field for spin label spectra) or higher would provide sufficient resolution to extract experimental values of g_{xx}, g_{yy}, A_{xx} and A_{yy}.

Simulation of a room temperature spectrum in the 'intermediate' regime (i.e. for species not freely tumbling in solution with Brownian motion) requires the input of separate x, y and z components for the g- and A-tensors. However, the sample is actually still under largely isotropic conditions, so the g_{zz} and A_{zz} values extracted here should not be considered accurate; instead, averaging of the g- and A-tensors can be used to report g_{iso} and A_{iso}. Furthermore, the τ_c value (which is also computed through its x, y and z components, and averaged for reporting purposes) is often the main parameter being elucidated in these simulation fittings. The room and low temperature spectra, and by extension the values extracted by fitting of simulated spectra, are not directly comparable and are used largely independently to characterise the same systems in different ways. Nevertheless, it can be helpful to take best-fit values from one set of data to begin the simulation fitting of the other.

Pulsed EPR

The MTSL-labelled T36C/S352C variant of MalE was studied using PELDOR spectroscopy in both the apoprotein and maltose-bound forms. The raw time traces were processed using the DeerAnalysis package[18] which runs on MATLAB. This involves fitting the background factor to account for the exponential decay of the signal. An automated Tikhonov regularisation (which gives accurate approximate solutions to ill-posed problems, such as that posed by extracting distance from PELDOR data) is then applied to find a compromise between suppressing artefacts (introduced by noise) and the resolution of the distance distribution. The resulting distance distributions are yielded with colour-coded confidence bands (primarily a result of the measured dipolar evolution time).

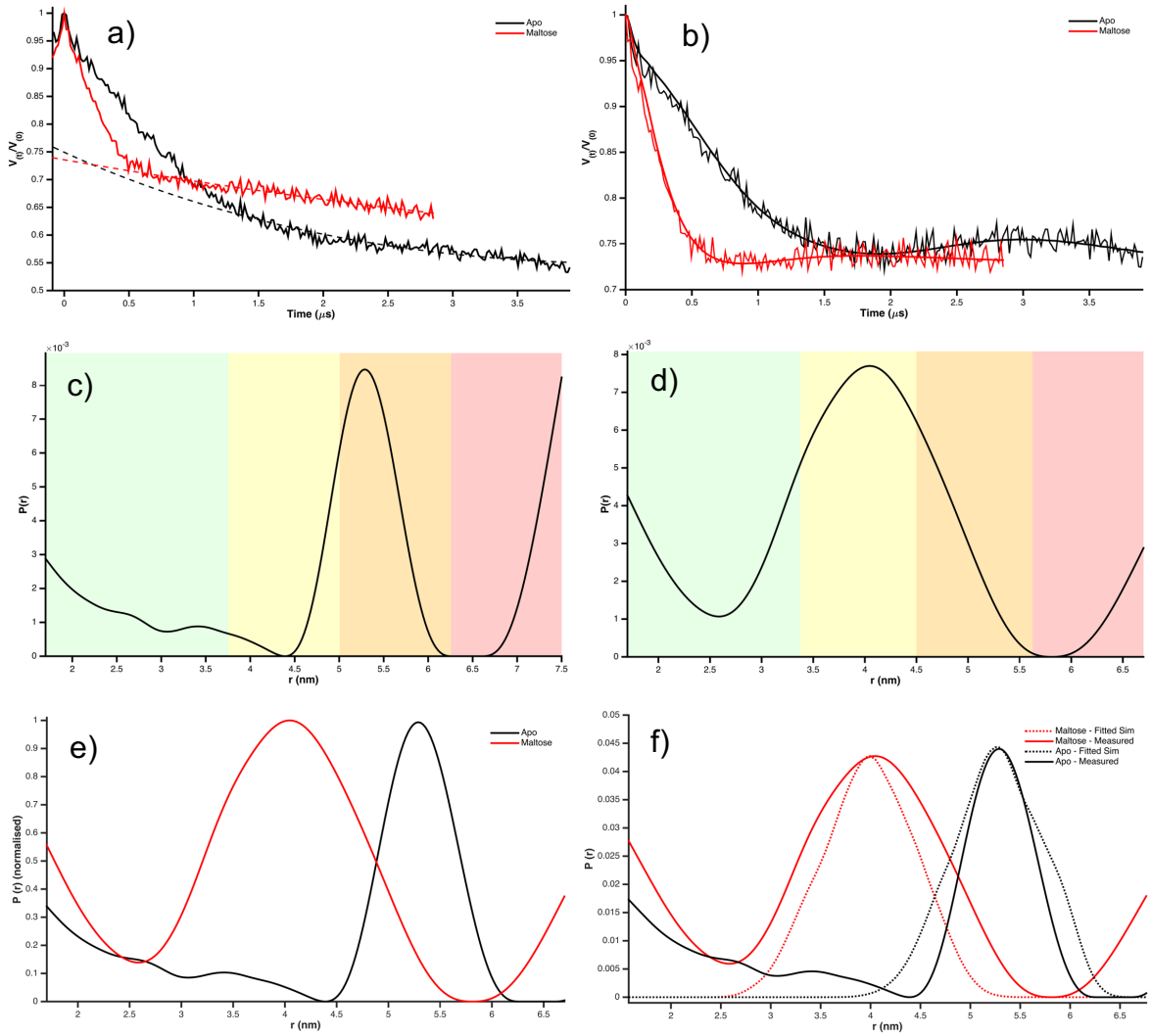


Fig. 4.2.6: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled T36C/S352C MalE: a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and maltose-bound form, red, b) background factor subtracted time trace (apoprotein in black, maltose-bound in red), c) apoprotein distance distribution output, d) maltose-bound distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black, maltose-bound in red), f) scale-normalised overlay of the distance distributions (apoprotein in black solid line, maltose-bound in red solid line) with fitted simulations (from MMM; apoprotein in black dotted line, maltose-bound in red dotted line).

The distance distribution of the apoprotein form MalE exhibits a single mode peaking at 52.9 Å, in good agreement with the prediction of 52.5 Å. The confidence in the value is fair, but the accuracy of the distribution (i.e. how sharp or broad the population is) is low. This can be related back to the raw time trace and background correction; though the dipolar evolution time is reasonably long (~4 μs) the oscillation appears shallow, enhancing the obscuring effects of noise on the subsequent data analysis process. For this reason, it is likely we can also

disregard the features in the distribution at short distances. This is reinforced by the low temperature cw-EPR results (see fig. 4.2.4 and fig. 4.2.5), which would demonstrate some broadened spectral features if any inter-label distances less than about 20 Å are present.

Such short distances (i.e. <20 Å) would also likely manifest in the time trace as a notably sharp peak and steep drop-off in signal in the first 0.5-1 µs. This is potentially the case in the time trace data for the maltose-bound form, which is the probable cause for the presence of distances <20 Å in the distribution. Measuring the samples at low temperature (20-80 K) with X-band cw-EPR would be necessary to conclude whether this is the case. The majority of the inter-label distances were measured at around 40.5 Å, however. This feature is broad, indicating either that there is a fairly wide range of conformations being sampled, or is a by-product of the processing method. The former argument is supported by the raw and background-fitted time domain data, which shows a shallow and broad oscillation, and the confidence bands suggest that the accuracy of the distribution is good. The distance also corresponds well with the *in silico* labelling experiment prediction of 40.0 Å.

Comparison with sm-FRET

The variant of MalE studied here was originally designed for sm-FRET studies, and this data was recently published[11] (see table 4.2.5 for a summary of relevant results).

Table 4.2.5: comparison of the distance-based data, both predicted and measured, with the reported FRET efficiencies[a] of the fluorescent-labelled MalE T36C/S352C in both the absence and presence of maltose. The distance measurements consist of: the distance between the C_α atoms of the mutation positions as calculated from the crystal structures, the distances between the spin labels in the double-labelled as predicted from the simulated distance distributions calculated at 298 K using spin label MTSL, and the measured distances between the spin labels in the double-labelled variants as extracted from the distance distributions obtained from the PELDOR experiments.

[a] – data reported in publication by de Boer et al., 2019[11]

Conditions	**Distance (Å)**			**FRET Efficiency, E* (ALEX sm-FRET)[a]**
	C_α-C_α	**MMM**	**PELDOR**	
MalE Apo	51.0	52.5	52.9	0.646 ± 0.003
MalE + Maltose	40.8	40.0	40.5	0.824 ± 0.002

The distance measurements obtained using PELDOR spectroscopy agree very well with the data obtained from the crystal structures. The T36-S352 C_α-C_α distance showed a change from 51.0 to 40.8 Å upon maltose binding (Δ=10.2 Å), and the MMM-predicted distance distribution of the MTSL-labelled T36C/S352C variant projected a ~12 Å change (52.5 to 40.0 Å upon substrate-binding). This value was confirmed by the PELDOR results (52.9 to 40.5 Å; Δ=12.4 Å), demonstrating how powerful a tool the prediction software can be.

A very similar variant, with cysteine mutations introduced at the lobes (two residues away either side of the variant used here; K34C/R354C) has been studied previously using sm-FRET[10], and again in the more recent publication. However, the binding of the K34C/R354C variant in the latest publication was not studied with maltose[11].

Table 4.2.6: comparison of the predicted distance-based data, with both the reported sm-FRET distances[b] and FRET efficiencies[c] of the fluorescent-labelled MalE K34C/R354C in both the absence and presence of maltose. The predicted distance measurements consist of: the distance between the C_α atoms of the mutation positions as calculated from the crystal structures, and the distances between the spin labels in the double-labelled as predicted from the simulated distance distributions calculated at 298 K using spin label MTSL. Dashes indicate where there is no published data.

[b] – data reported in publication by Kim et al., 2013[10]
[c] – data reported in publication by de Boer et al., 2019[11]

Conditions	**Distance (Å)**			**FRET Efficiency, E* (ALEX sm-FRET)[c]**
	C_α-C_α	**MMM**	**sm-FRET[b]**	
MalE Apo	59.2	65.0	58.3	0.524 ± 0.010
MalE + Maltose	50.3	51.4	49.0	-

Here, the distance measurements obtained using sm-FRET are shown to agree with the measurements made using the crystal structures. The K34-R354 C_α-C_α measurements showed a change from 59.2 to 50.3 Å (Δ=8.9 Å) upon maltose binding, and the distance distribution of the corresponding MTSL-labelled K34C/R354C variant predicted *in silico* yielded a change from 65.0 to 51.4 Å (Δ=13.6 Å). The former agrees well with the reported sm-FRET distance change of 9.3 Å (58.3 to 49.0 Å)[10].

It is difficult to quantitatively compare all of these results with the data reported from FRET studies. Whilst FRET efficiency related to the distance between the fluorescent labels (higher FRET efficiency occurs when the labels are closer together), it cannot be converted into a distance measurement without knowledge of certain experimental details[19]. However, qualitative assessment shows the FRET efficiencies correspond with a longer inter-label distance in the apoprotein state, and a shorter inter-label distance in the maltose-bound state, and the source publication of the FRET data shown here states a 10 Å closure of fluorescent dye-labelled T36C/S352C MalE upon maltose binding[11]. In comparison, the PELDOR distance distributions of the same variant (spin-labelled) report a difference of 12.4 Å.

When both variants and their respective datasets are considered together, the values describe a consistent trend, especially bearing in mind that the two double variants (T36C/S352C and K34C/R354C) are labelled in very similar positions, only offset by two residues either side of one another. The similarities between the results for these variants across both measurement methods highlights the complementarity of the two techniques. These also align with the predictions taken from the crystal structures, validating the measurements using a relatively simple and well-known protein. When applied to more complex systems they may provide insight into conformations that are not so easily accessed through crystallography (or other more typical methods).

Comparison with models

The distance constraint data recorded here are not extensive enough to inform much on the overall transport model of the $MalEFGK_2$ system. Nevertheless, the remarkable agreement in the data obtained through all three of these methods (crystallography, EPR and FRET) provide a unified picture of the conformations that MalE exists in, in both the absence and presence of maltose. This view of the two 'extremes' of binding (i.e. fully unbound and fully bound) which demonstrates relatively large conformational change between the two lobes fits well with the prevalent Venus flytrap model of SBP binding[20-22].

However, to verify that there are only two conformations accessible by MalE a more extensive investigation should be performed, in which substrate is progressively titrated to access the conformational dynamics exhibited in sub-stoichiometric situations. Should the hypothesis hold true PELDOR would allow a discernible change to be followed in the relative populations of each in the distance distributions. Conversely, alternative or intermediate conformations would result in populations at different distances being observed.

One of the potential weaknesses of FRET is that if the binding-induced changes in conformational dynamics are faster than the millisecond timescale the histograms will not show discrete populations, and the distributions will be unimodal (as an average of the populations present in the sample). In contrast, PELDOR uses frozen solutions which can provide a snapshot of the ensemble of protein molecules in their various stages of binding (though multiple similar conformations may cause broadened distributions).

In the recent FRET study different conformations of MalE were found using different malto-oligosaccharides and maltodextrins as substrates[11]. The findings demonstrated that the inter-lobe closure of MalE differs depending on the size of the substrate. Smaller substrates (maltose, maltotriose and maltotetraose) caused ~10 Å closure (i.e. difference between open and closed conformations. Meanwhile, mid-sized malto-oligosaccharides (maltopentaose and maltohexaose) gave a ~7 Å closure, and the large substrate maltoheptaose yielded a closure of ~5 Å. Non-cognate ligands also bound to MalE, but they did not close the protein to the same extent as the cognate ligands: maltooctaose and maltodecaose caused a ~5 Å closure; β-cyclodextrin, maltotriitol and maltotetraitol closed the protein <5 Å.

5 Amino Acid Transporter GlnPQ

5.1 Background

The amino acid transporter GlnPQ from *Lactococcus lactis* is a slightly unusual ABC-type transporter in that it uses multiple SBPs per complex. This consists of two separate pairs of SBP, tethered to the rest of the complex in tandem, for a total of four SBPs[23]; on each tandem pair there is SBD1, which binds asparagine and glutamine, and SBD2, which binds glutamine.

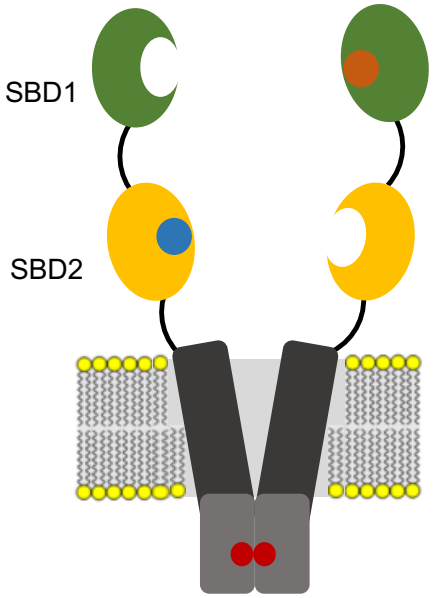


Fig. 5.1.1: representation of the overall architecture of the GlnPQ complex, highlighting the unusual arrangement of multiple SBDs tethered in tandem to the rest of the transporter complex.

At present, there are four published crystal structures of the SBDs of this system: two of SBD1 (in the apoprotein state, PDB 4LA9[24], and in the asparagine-bound form, PDB 6FXG[25]) and two of SBD2 (in the apoprotein form, PDB 4KR5[24], and in the glutamine-bound state, PDB 4KQP[24]). Though their overall sequence identity is very low (<50%), the fold of these two different proteins is very similar[26].

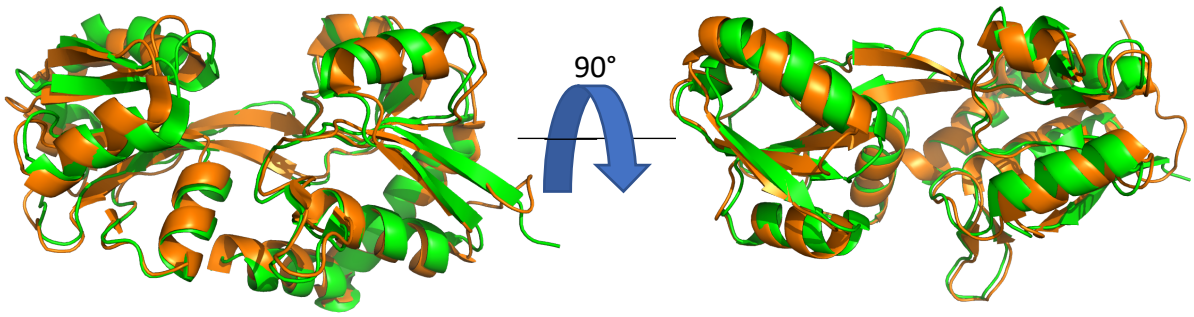


Fig. 5.1.2: overlaid crystal structures of the GlnPQ SBDs in their apoprotein states from different aspects showing the similarity of their folds: green is SBD1 and orange is SBD2 (PDB 4LA9 and 4KR5 respectively).

From these crystal structures, it can be observed that the proteins fit the conventional arrangement expected of SBPs; two globular lobes, connected by a hinge region. In this case, since they belong to the cluster F-IV group of SBPs, that hinge region consists of two peptide stretches of about 8-10 amino acids each[2].

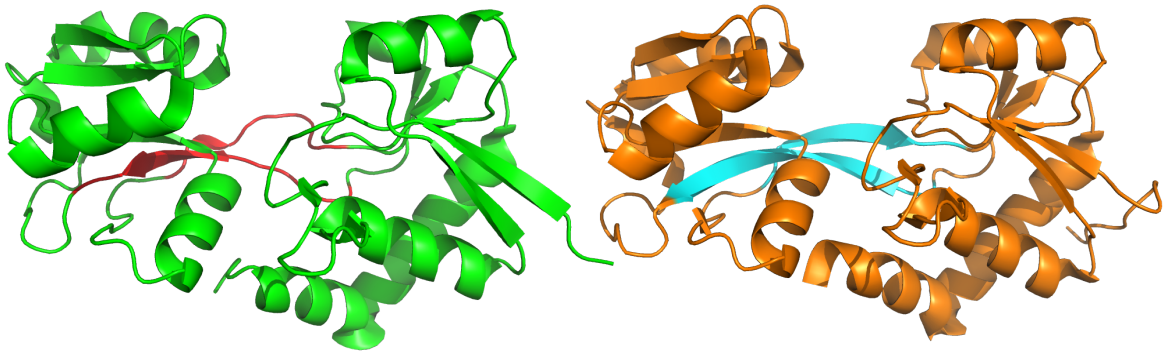


Fig. 5.1.3: crystal structures of the GlnPQ SBDs in their apoprotein states with the hinge regions highlighted: green is SBD1 (hinge highlighted in red) and orange is SBD2 (hinge highlighted in cyan; PDB 4LA9 and 4KR5 respectively).

The aforementioned deviation in sequence identity between these two proteins is likely related to the marked difference in their substrate specificities and dissociation constants. SBD1 is known to bind both glutamine and asparagine, with K_d values of 92 μM and 0.2 μM respectively. Meanwhile, SBD2 binds glutamine with high affinity (K_d 0.9 μM)[26]; it is also known to bind glutamate, though no binding affinity data has yet been obtained. It is thought that these proteins bind their respective substrates via an induced fit mechanism[11, 26], which would fit with the conformational dynamics of the Venus flytrap model. Recent studies using FRET showed that the GlnPQ SBDs sample several different conformations depending on the substrate that is being bound and, in some cases,

the SBDs will bind non-cognate and non-transported ligands, which interfere with the uptake of cognate ligands, but do not trigger protein closure[11].

Here, the majority of the studies have been directed towards SBD1 and its binding of glutamine and asparagine and considering how the results obtained using various EPR techniques compare with the K_d information and FRET data. SBD2 studies have also been undertaken to provide insight into the binding of glutamine and glutamate.

5.2 Results and Discussion

in silico spin labelling

The cysteine-containing variants of the GlnPQ SBDs were designed for studies using sm-FRET, which is complementary with EPR spectroscopy (both use the thiol sidechain chemistry of cysteines to specifically label positions within a protein). The cysteine mutations (Q87C/T159C for SBD1, and T369C/S451C for SBD2) and spin label attachments can be simulated *in silico* from the crystal structures, with rotamer libraries of the spin label produced based on the sampling of space around the site of mutation. Here, where we have two spin labels per protein being modelled, the distances between each spin label pair can be calculated, and their distance distributions predicted (i.e. weighted probabilities of distances populated between the spin label rotamers).

The rotamer libraries were calculated for the labelling positions in both SBD1 and SBD2, in the apo and substrate-bound forms for each, from the published crystal structures; these were calculated at ambient temperature (298 K) using Multiscale Modelling of Macromolecular Systems (MMM) software package that runs via MATLAB[12].

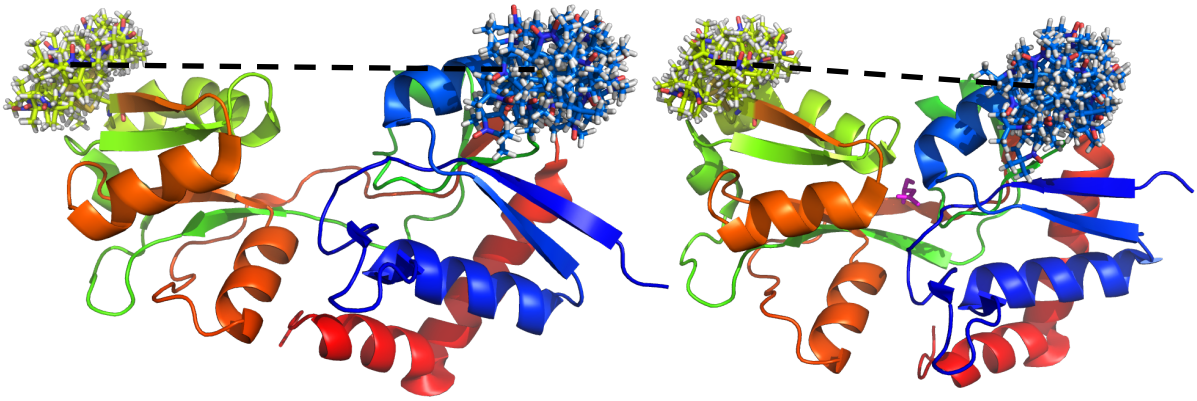


Fig. 5.2.1: the crystal structures of the apo (left) and asparagine-bound (right) forms of GlnPQ SBD1 (PDB 4LA9 and 6FXG respectively) with MMM-generated rotamer libraries highlighted, at the residue positions where cysteines will be introduced; the colouring indicates the sequence of the protein from N- to C-terminus, blue to red, with asparagine highlighted in pink. The distances to be measured are indicated by dashed lines connecting the pairs of rotamer libraries.

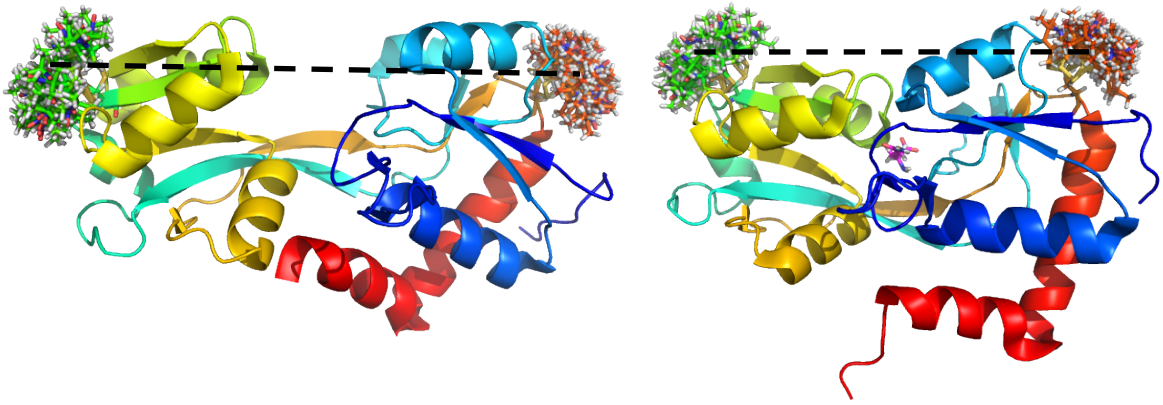


Fig. 5.2.2 the crystal structures of the apo (left) and glutamine-bound (right) forms of GlnPQ SBD2 (PDB 4KR5 and 4KQP respectively) with MMM-generated rotamer libraries highlighted, at the residue positions where cysteines will be introduced; the colouring indicates the sequence of the protein from N- to C-terminus, blue to red, with glutamine highlighted in pink. The distances to be measured are indicated by dashed lines connecting the pairs of rotamer libraries.

Table 5.2.1: rotamer libraries of the GlnPQ SBD variants in the apo, Gln-bound and Asn-bound forms, calculated from the crystal structures at 298 K. Dashes indicate where there is no published crystal structure.

Variant	**Calculated Rotamer Library (298 K)**			
	Apo	**+ Gln**	**+ Glu**	**+ Asn**
SBD1 Q87C	85	-	-	86
SBD1 T159C	115	-	-	118
SBD2 T369C	69	56	-	-
SBD2 S451	47	33	-	-

These calculated rotamer libraries mostly indicate only small differences between the different substrate-binding states for the individual mutation sites in each protein. This can be interpreted as a prediction that the space around the site of mutation won't change much between the different conformations (i.e. apoprotein, or substrate-bound), and so it is unlikely that any noticeable differences will be observed in the lineshapes of the room temperature X-band cw-EPR spectra.

The predicted distances and distance distributions for the double label pairs were then calculated.

Table 5.2.2: the predicted distances between the pairs of spin labels in the double-labelled variants of the GlnPQ SBDs as extracted from the predicted distance distributions, calculated at 298 K. In cases where there are multiple modes in the distance distributions, the predominant value is shown in bold. Dashes indicate where there is no published crystal structure.

Variant	**Predicted inter-label distance (Å) (MMM @ 298 K)**			
	Apo	**+ Gln**	**+ Glu**	**+ Asn**
SBD1 Q87C/T159C	50.2	-	-	37.5
SBD2 T369C/S451	57.6	47.1	-	-

From these predictions it can be seen that reasonably large changes in the modes of the distance distributions, of 10 Å or more, can be expected upon substrate binding.

cw-EPR & simulations of SBD1

Room temperature X-band

The Q87C/T159C variant of SBD1 was spin labelled with MTSL (13.1 'Site directed spin labelling'); the best labelling efficiency of ~94% was achieved (issues with other protein batches will be discussed later). As before, this was calculated using the measured signals of a calibration curve of spin label at known concentrations and that of the labelled protein sample, comparing with the measured concentration of the protein after the labelling process. The room temperature cw-EPR spectra of the double-labelled SBD1 were recorded at X-band in three states: apoprotein, Gln-bound and Asn-bound (details of substrate binding in section 13.1 'Substrate additions').

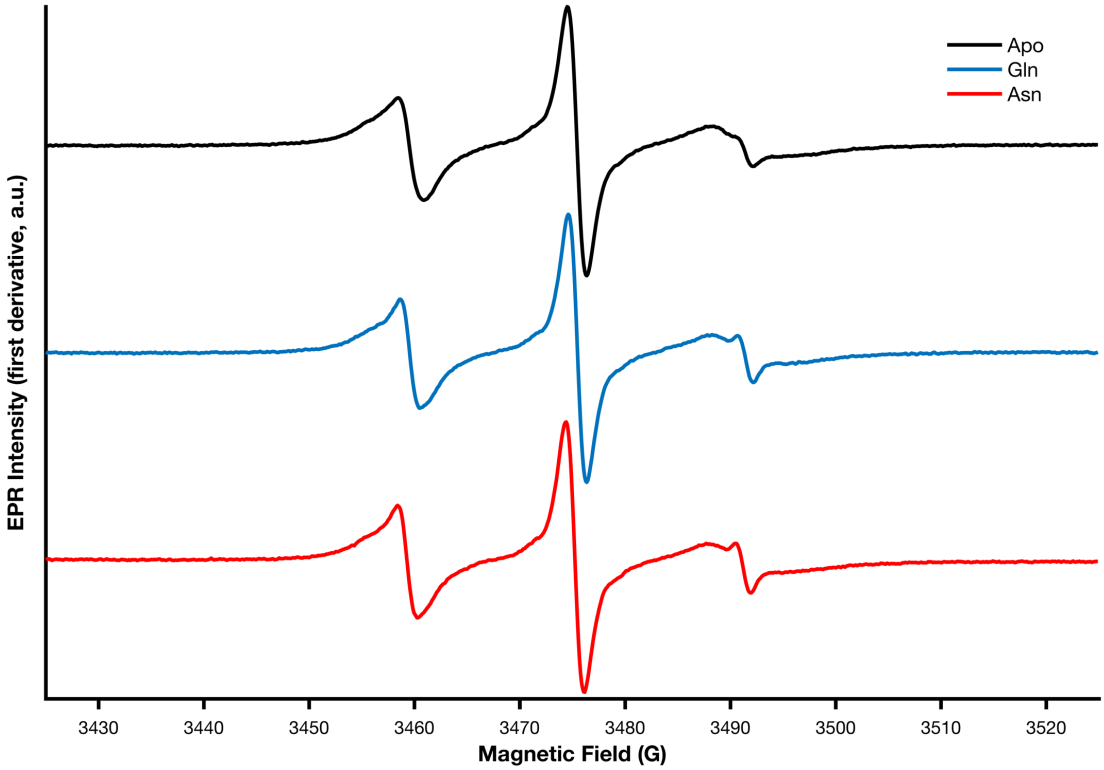


Fig. 5.2.3: the room temperature cw-EPR spectra of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in various binding states: apoprotein in black, Gln-bound in blue and Asn-bound in red. All spectra have been normalised to same scale.

In all three cases the spectra indicate that the attached MTSL has a high degree of motional freedom, as would be expected from the rotamer library predictions from the *in silico* label attachment calculations (85 and 115 for Q87C and T159C respectively in the apoprotein form), though those predictions should always be considered with a degree of scepticism due to the assumption that the crystal

structure accurately reflects a physiologically-relevant state and is taken as a rigid body. A qualitative assessment of the differences between the three spectra above indicates small variations mainly seen in the outer peaks; in the substrate-bound forms a slightly sharper feature on the high-field peak can be observed, as well as a somewhat more defined shoulder on the low-field peak.

As stated above, a labelling efficiency of ~94% was achieved. This result is taken from early labelling experiments of SBD1. Various problems were encountered with later samples, both in those that were produced and purified in-house and those provided by collaborators. These difficulties included lower yields, loss of protein during the labelling process, lower labelling efficiency (from ~30% to ~50%), and inactive protein (i.e. samples that would not bind any substrate). The reasons for this are as yet unknown and have hindered efforts in performing more complex experiments (which are outlined later for potential future work).

In order to extract accurate values for various parameters, to then be able to quantitatively assess and compare them, these spectra must be simulated and fitted. This was performed using the EasySpin package[13] for MATLAB. As previously mentioned, approximate starting values for the g- and A-tensors of MTSL[14-15] are used as input for the simulations and fittings performed here.

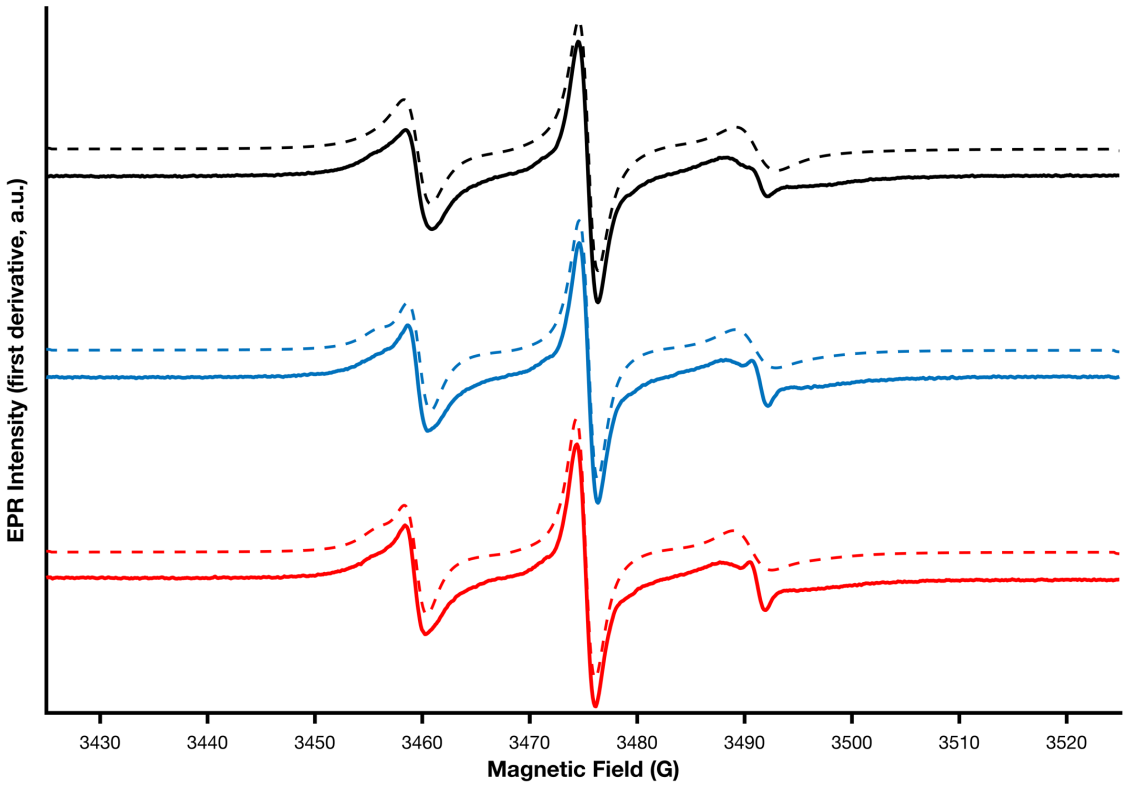


Fig. 5.2.4: the fitted simulations of the room temperature cw-EPR spectra of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in various binding states: apoprotein in black, Gln-bound in blue and Asn-bound in red; experimental data are shown as solid lines, fitted simulations are shown as dotted lines. All spectra have been normalised to same scale.

For each of the simulations it was found that two spin label species were needed to achieve sufficiently low fitting RMSDs. These account for one species with shorter τ_c and the other with longer τ_c; the distinction between these environments is minimal in all parameters other than τ_c, where there was a difference of 1 order of magnitude (table 5.2.3) within each fitted simulation result.

However, as previously discussed at greater length (section 4.2. 'cw-EPR & simulations'), the limitations of room temperature X-band cw-EPR means that the the g- and A-tensors should only be considered estimates (though, provided the input values and allowed variation parameters are sensible based on well-reported literature values, they should be 'good' approximations). Instead it is perhaps more appropriate to report averages of the tensors (i.e. g_{iso} and A_{iso}) if comparisons or differences are being drawn between the three binding states. Here, the extracted τ_c values are the main output being studied (the g- and A-tensors are more accurately reported by the fitted simulations of the low temperature data below).

There is little difference between the three binding states; they all consist of two different components of differing τ_c, wherein the shorter τ_c (1.8 to 2.7 ns) dominates the relative weightings. Across the three fittings, the relative weightings of the separate species within each are reasonably consistent at 0.67:0.33 ±0.06, which may represent the relative ratio of labelling at the individual cysteine sites. From the MMM *in silico* labelling experiments, the apoprotein and Asn-bound forms of SBD1 both have very similar rotamer library populations of 85/115 (43:57) and 86/118 (42:58) at the Q87C/T159C sites respectively. Whilst it is important to remember that these predictions are not necessarily indicative of behaviour of protein in aqueous environment, there is correlation between the dominance of one species with a shorter τ_c in the experimental data and a more 'accessible' labelling site as predicted from computational calculations.

The weighted averages of the τ_c of all three species are slightly different: 6.3 ns for the apoprotein, 8.1 ns in the Gln-bound state and 7.8 ns in the Asn-bound state. As there is no crystal structure data for the Gln-bound form, it is difficult to draw any conclusions from this. However, an initial comparison of the apoprotein and Asn-bound states contradicts the expectation based on the MMM predictions; a longer τ_c is extracted from the form with the higher number of rotamers at both of the cysteine sites.

Though the differences between the spectra of SBD1 in its various binding states are minimal, the discussion of the other extracted parameters (i.e. g- and A-tensors, particularly those of the apoprotein form) will be continued in more detail as comparison to those of SBD2 (section 5.2 'cw-EPR & simulations of SBD2').

Table 5.2.3: best fit simulation values for the x, y and z components of each of the g- and A-tensors, as well as linewidth and averaged correlation times for each of the room temperature X-band experimental data of the MTSL-labelled SBD1 in different binding states. Where multiple spin label ' species' are simulated for fitting a single spectrum the respective, relative weightings are provided for each component. The RMSD of each fitting is given.

Condition	Component	g-tensor			A-tensor (MHz)			Linewidth (peak-to-peak, mT)	τ_c (ns)	Relative weighting	Simulation RMSD
		x	y	z	x	y	z				
Apo	1	2.0083	2.0058	2.0027	15	16	101	0.01	2.7	0.65	0.017545
	2	2.0082	2.0059	2.0021	14	14	101	0.01	13	0.35	
+ Gln	1	2.0081	2.0057	2.0028	15	16	102	0.01	1.8	0.73	0.019826
	2	2.0085	2.0059	2.002	14	13	102	0.01	25.3	0.27	
+ Asn	1	2.008	2.0058	2.003	15	17	101	0.01	2.1	0.62	0.021707
	2	2.0081	2.0061	2.002	14	14	103	0.01	17	0.38	

Low temperature X-band

The X-band cw-EPR spectra of the Q87C/T159C variant of SBD1 were also recorded at low temperature (20 K) for all three binding states: apoprotein, Gln-bound and Asn-bound.

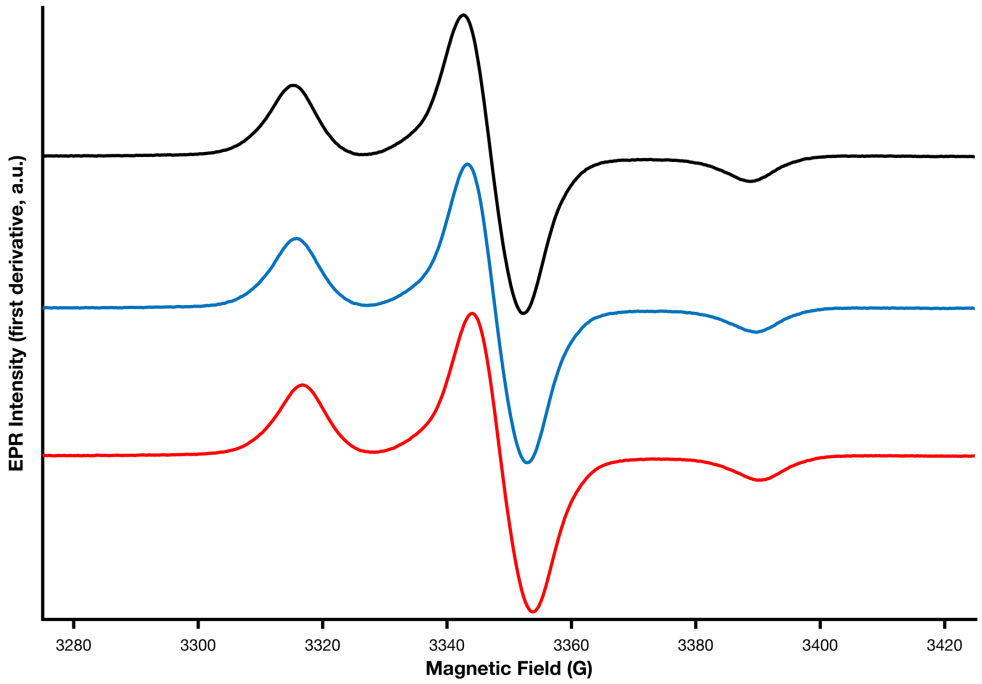


Fig. 5.2.5: the low temperature (20K) cw-EPR spectra of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in various binding states: apoprotein in black, Gln-bound in blue and Asn-bound in red. All spectra have been normalised to same scale.

The spectral characteristics of the three binding states at low temperature do not appear to differ much; when overlaid (see fig. 5.2.6), it is clear that there is no observable dipolar broadening in either of the substrate-bound states when compared with the apoprotein spectrum, meaning there is unlikely to be any changes in inter-label distances to less than 20 Å. This agrees with the results of the *in silico* labelling models, which calculated an inter-label distance of 37.5 Å in the asparagine-bound form of SBD1 Q87C/T159C. As this binding state is posited to demonstrate the greatest lobe closure (based on FRET studies[11, 26]), it can be inferred that the glutamine-bound form, by virtue of its spectral similarities to that of the Asn-bound SBD1, also does not experience inter-label distances of below 20 Å.

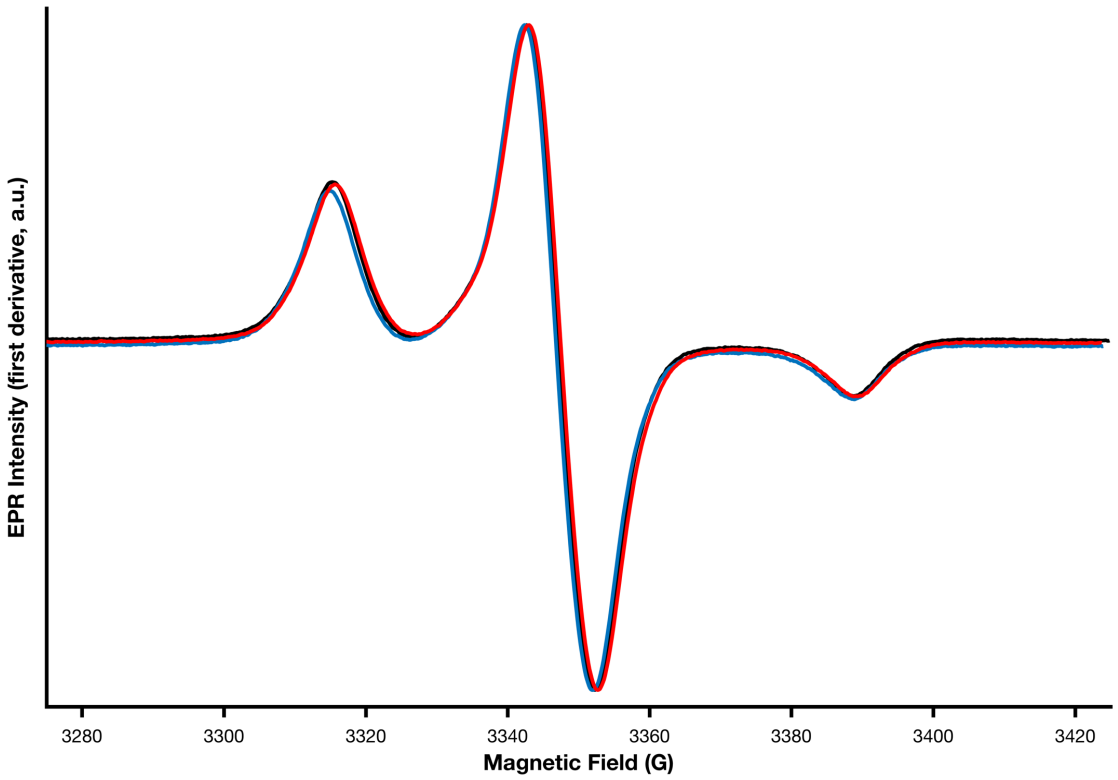


Fig. 5.2.6: the overlaid low temperature (20K) cw-EPR spectra of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in various binding states (apoprotein in black, Gln-bound in blue and Asn-bound in red) highlighting the minimal differences between them. All spectra have been normalised to same scale and centred upon one another by minor frequency adjustments.

These spectra were then simulated and fitted to the experimental data using the EasySpin package[13] for MATLAB. The best-fit results of the fitted simulations of the room temperature cw-EPR spectra were taken as starting parameters for this analysis. Two spin label species were used per simulation.

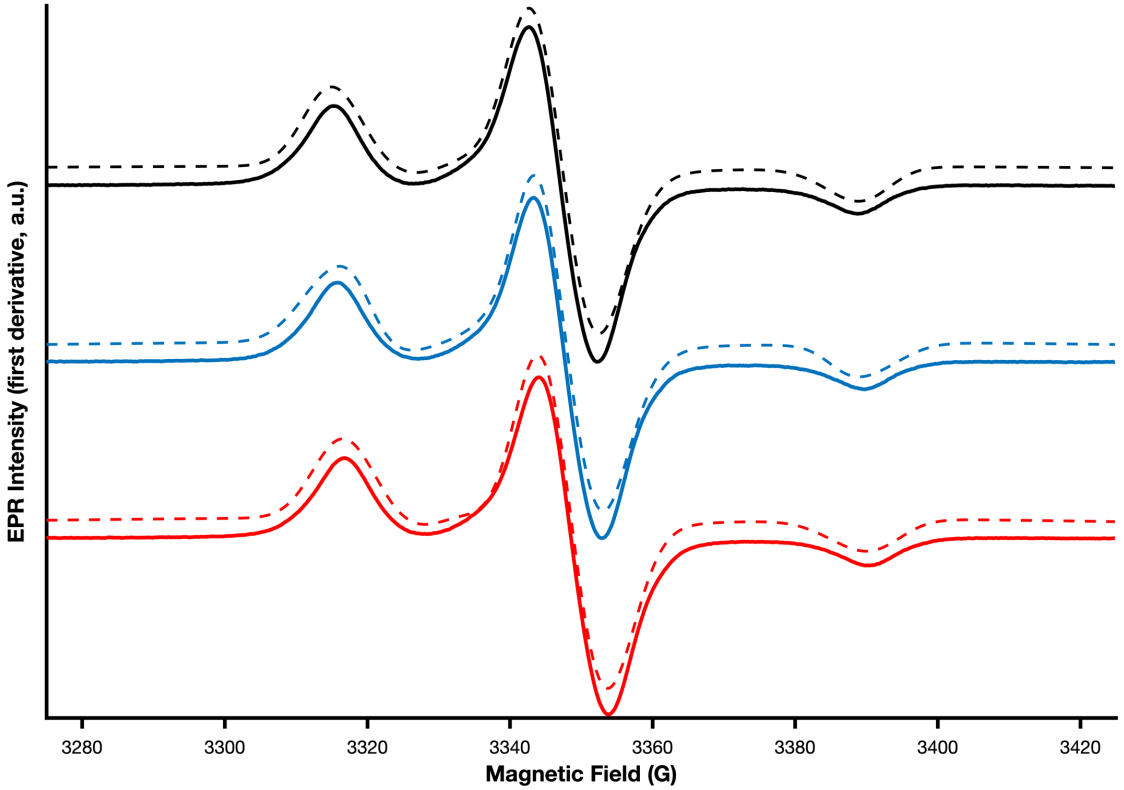


Fig. 5.2.7: the fitted simulations of the low temperature cw-EPR spectra of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in various binding states: apoprotein in black, Gln-bound in blue and Asn-bound in red; experimental data are shown as solid lines, fitted simulations are shown as dotted lines. All spectra have been normalised to same scale.

The results of the fitted simulations confirm that the three binding states do not yield noticeably different low temperature cw-EPR spectra. There is clear agreement in all parameters extracted from each of the fittings (see table 5.2.4). However, one should bear in mind the limited capability of low temperature X-band cw-EPR to resolve individual components of the g- and A-tensors. The similarities between the parameters extracted from these fitted simulations across the three different states are not necessarily accurate; it may be the case that there are variances across the binding states that are not distinguishable at this resolution. If the experiments were repeated at higher frequency (i.e. W-band or above), the potential subtle differences between the chemical and physical environments of the spin labels across the binding states may be observable[16-17]. These results will be discussed in greater detail, and compared with the equivalent data for SBD2, in a later section (5.2 ‘cw-EPR & simulations of SBD2’).

Table 5.2.4: best fit simulation values for the x, y and z components of each of the g- and A-tensors and linewidth for each of the low temperature X-band experimental data of the MTSL-labelled SBD1 in different binding states. Where multiple spin label ' species' are simulated for fitting a single spectrum the respective, relative weightings are provided for each component. The RMSD of each fitting is given.

Condition	Component	g-tensor			A-tensor (MHz)			Linewidth (peak-to-peak, mT)	Relative weighting	Simulation RMSD
		x	y	z	x	y	z			
Apo	1	2.0082	2.0065	2.0025	14.2	13.7	99.7	0.90	0.69	0.025963
	2	2.0083	2.0061	2.0024	13.5	15.1	107.2	0.70	0.31	
+ Gln	1	2.0079	2.0062	2.0021	14.8	14.7	105.9	0.90	0.82	0.025293
	2	2.0076	2.0060	2.0021	15.1	13.7	96.5	0.57	0.18	
+ Asn	1	2.0073	2.0052	2.0015	15.0	14.6	106.5	0.78	0.65	0.024734
	2	2.0079	2.006	2.0018	15.2	14.8	94.4	0.83	0.35	

Power saturation

The room temperature cw-EPR spectra of the Q87C/T159C variant of SBD1 at X-band were also measured as a function of applied microwave powers of for all three binding states: apoprotein, Gln-bound and Asn-bound. The attenuation of the microwave power was increased in 3 dB steps (on this attenuation scale a 3 dB increase halves the applied microwave power in mW). Intensity was taken as the first derivative peak-to-trough height of the central resonance line (m_I = 0) of the nitroxide spin label and is related to $P_{1/2}$ as shown in equation 35[27].

$$I = A\sqrt{P}\left[1 + \frac{\left(2^{1/\varepsilon} - 1\right)P}{P_{1/2}}\right]^{-\varepsilon} \quad (35)$$

Where I is signal intensity (a.u.), A is a scaling factor related to the slope of the increase in I with √P in the non-saturating regime, P is the applied microwave power (mW) and ε is a measure of the homogeneity of the saturation of the resonance line (where the homogeneous limit is 3 and inhomogeneous limit is 1)[27].

The data were processed initially for an estimate of the $P_{1/2}$ value as shown below (fig. 5.2.8 – 5.2.10).

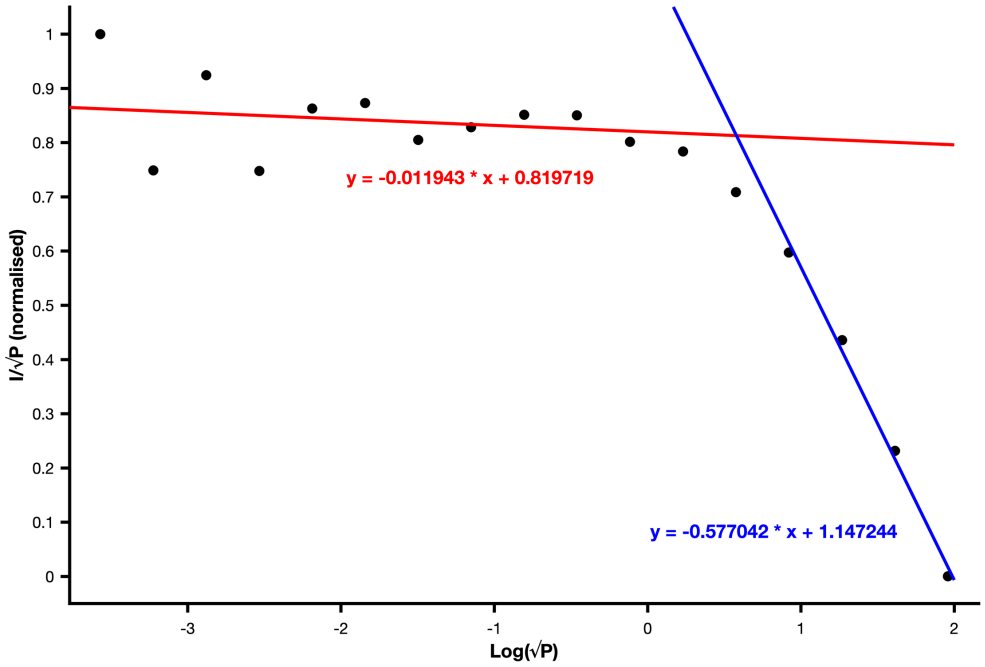


Fig. 5.2.8: the processed room temperature cw-EPR power saturation data of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in the apoprotein binding state. Two sets of data points were selected to fit linear trendlines (these lines are shown in red and blue, with their respective equations).

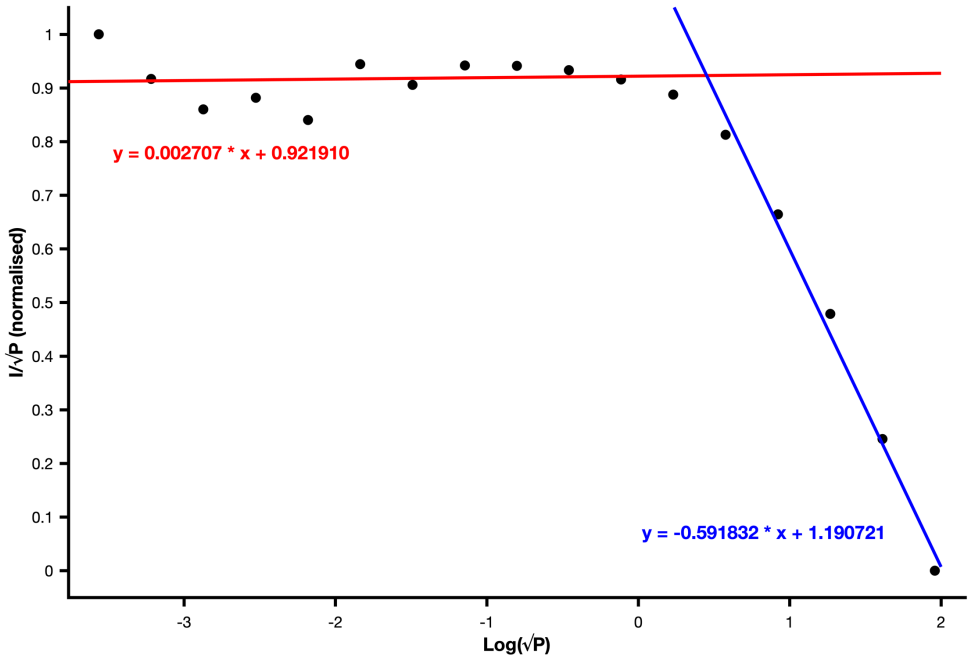


Fig. 5.2.9: the processed room temperature cw-EPR power saturation data of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in the Gln-bound state. Two sets of data points were selected to fit linear trendlines (these lines are shown in red and blue, with their respective equations).

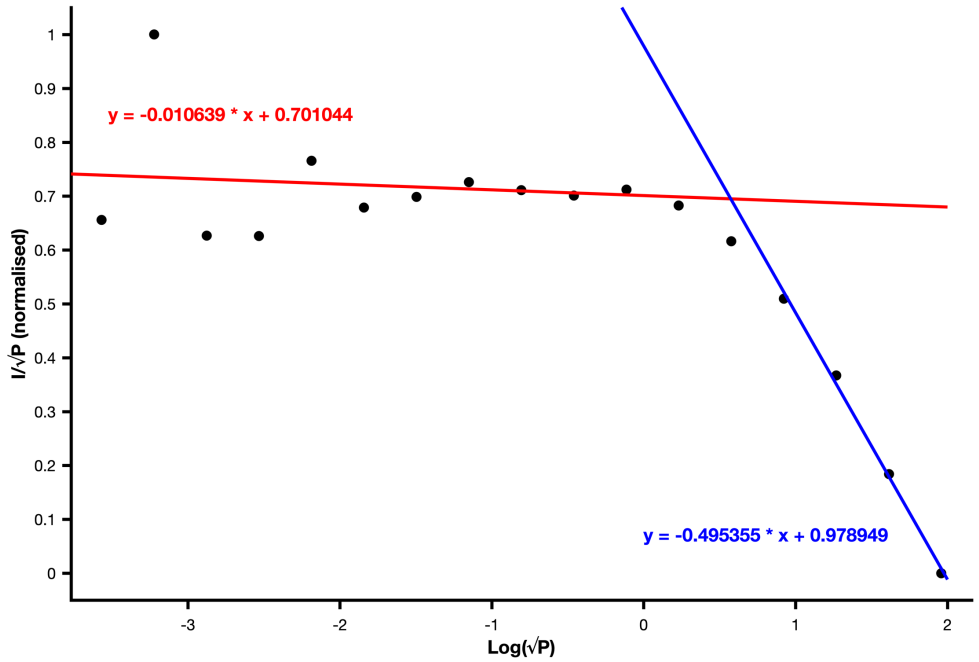


Fig. 5.2.10: the processed room temperature cw-EPR power saturation data of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in the Asn-bound state. Two sets of data points were selected to fit linear trendlines (these lines are shown in red and blue, with their respective equations).

The equations of the trendlines can be used to find the x value at which they cross, and then solved to give an estimate for the $P_{1/2}$ value in mW.

Table 5.2.5: the estimated $P_{1/2}$ values (mW) for the MTSL-labelled Q87C/T159C variant of SBD1 under various binding conditions, as calculated from the processed room temperature cw-EPR power saturation data. Two trendlines were applied to distinct sections of each of the plots and solved for the x-value at which they intercepted.

Condition	Estimated $P_{1/2}$ (mW)
SBD1 Apo	0.42
SBD1 + Gln	0.62
SBD1 + Asn	0.47

The estimate $P_{1/2}$ value across the binding states does not differ significantly. There is a slight increase when substrate is bound (particularly in the Gln-bound state), but this is probably within the level of error of this method. The trendlines, particularly that of the 'flat' section of the plot, are heavily influenced by the signal to noise level of the data from which the intensity is extracted. This is why at lower powers some of the data points may not be used for the trendline fitting; this is especially noticeable in the data presented for the Asn-bound state (fig. 5.2.10). The accuracy here may be improved by increasing the number of scans at each step. This, however, could significantly increase the time of the experiment and should thus be considered in balance with the potential to decrease noise in the spectra.

As can be seen in from the processed data plots, the 'sloped' trendline is applied to only a handful of data points. In order to increase the accuracy the increments of attenuated microwave power at which spectra are recorded could be decreased (i.e. every 1 dB rather than every 3 dB). Since this regime gives greater signal intensity, the experimental run-time will not increase excessively.

For a more accurate value for $P_{1/2}$ equation 35 was fit to the data itself (figure 5.2.11).

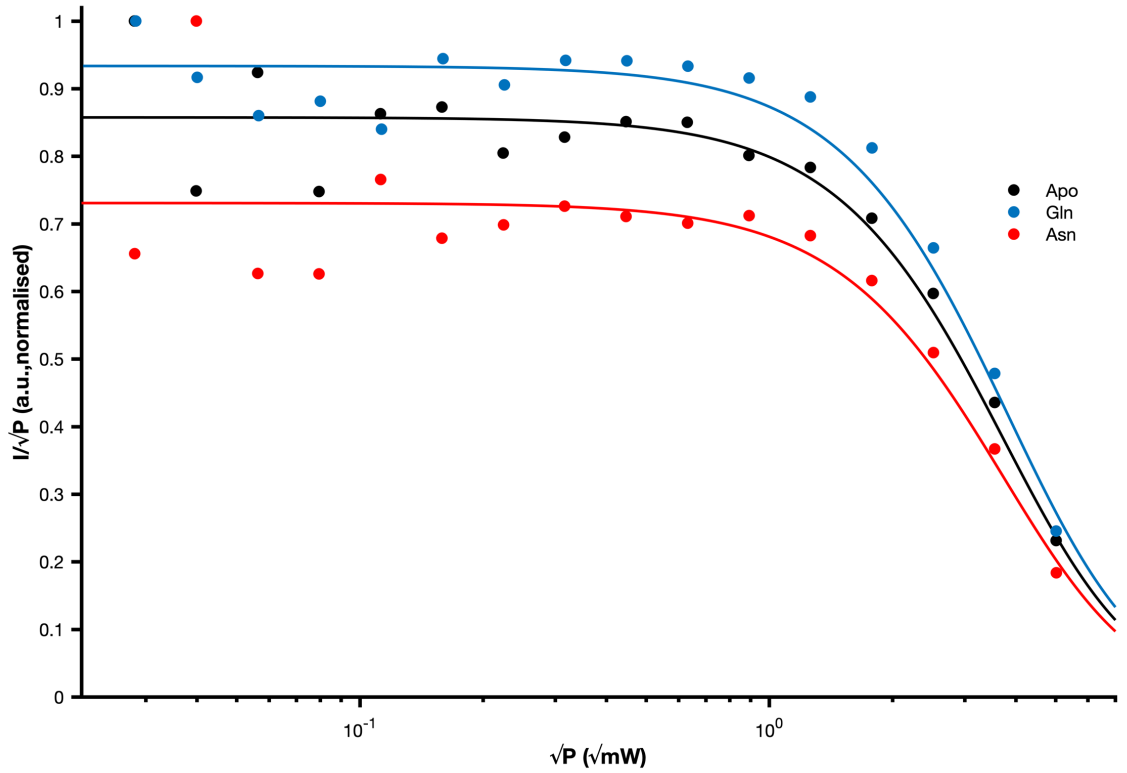


Fig. 5.2.11: the room temperature cw-EPR power saturation data of the MTSL-labelled Q87C/T159C variant of SBD1 at X-band in various binding states: apoprotein (black), Gln-bound (blue) and Asn-bound state (red). These data were each fit with equation 35, varying parameters A, $P_{1/2}$ and ε. The data points are marked by circles and the fits are shown by the lines in the respective colours of the binding state.

Table 5.2.6: the calculated $P_{1/2}$ values (mW) for the MTSL-labelled Q87C/T159C variant of SBD1 under various binding conditions, as determined from the fitting of the room temperature cw-EPR power saturation data with equation 35. The scaling factor, A, and saturation homogeneity measure, ε are also stated. R^2 values of the fits are given.

Condition	$P_{1/2}$ (mW)	A	ε	R^2
SBD1 Apo	0.5003	0.8575	1.964	0.9384
SBD1 + Gln	0.5004	0.9336	1.966	0.9612
SBD1 + Asn	0.5008	0.7803	1.964	0.8642

This method demonstrates that there is a good fit of the equation to the experimental data (the R^2 values provide a measure of the match to the data, i.e. the closer to 1 the better the fit), though the fit is poorer for the Asn-bound state data. As previously stated, this is likely due to the poor signal to noise of the data collected at lower powers. The $P_{1/2}$ values calculated from this fitting method generally agree with those estimated by fitting two straight trendlines, though there is no real variation between the values for each of the binding states. Furthermore, it should be noted that in order to obtain 'sensible' $P_{1/2}$ values, the equation had to be fitted iteratively. Firstly, the resulting best-fit value for the A factor was taken from the fit that allowed A, ε and $P_{1/2}$ to vary. This value was then placed into the equation without allowing it to vary, and only ε and $P_{1/2}$ were then fit. The resulting best-fit value for ε was then taken and placed into the equation and only $P_{1/2}$ was then allowed to vary for the final fitting.

The data would likely be best measured again, perhaps using smaller attenuation increments at high power and certainly more scans per slice at low powers to account for the signal to noise (as the value of A is heavily influenced by the data obtained at lower powers). Additionally, integrating each spectrum twice (to find the area under the curve) may yield more accurate signal intensity, though this also poses a problem in the data processing in correcting any baseline abnormalities in a consistent manner.

cw-EPR of SBD2

Room temperature X-band

The T369C/S451C variant of SBD2 was spin labelled with MTSL (details in section 13.1 'Site directed spin labelling'), attaining a best labelling efficiency of ~94%. This was calculated as previously detailed with a calibration curve of known concentrations of spin label in solution. The room temperature cw-EPR spectrum of the apoprotein MTSL-labelled SBD2 was recorded at X-band. As with SBD1, the best labelling efficiencies were achieved with the earlier preparations of protein; later batches either did not produce similar yields or labelled less well (from ~10% to ~55%) or significant losses were suffered during the labelling process.

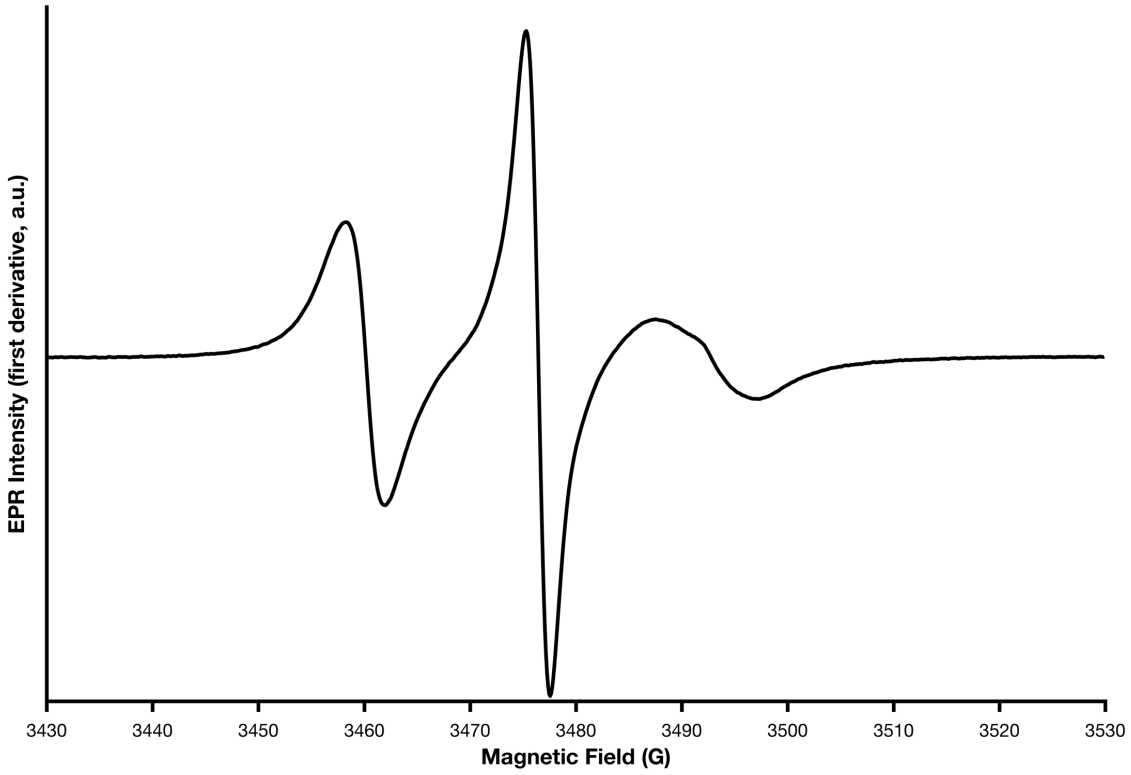


Fig. 5.2.12: the room temperature cw-EPR spectrum of the apoprotein, MTSL-labelled T369C/S451C variant of SBD2 at X-band.

Here the spectrum suggests that the MTSL is fairly mobile whilst attached to SBD2. This agrees well with the *in silico* label attachment calculation's rotamer library prediction (69 and 47 for T369C and S451C respectively). As discussed previously these projections should be viewed with a degree of doubt considering the crystal structure is assumed to be a rigid body and does not account for the actual physiological conditions of the protein in solution.

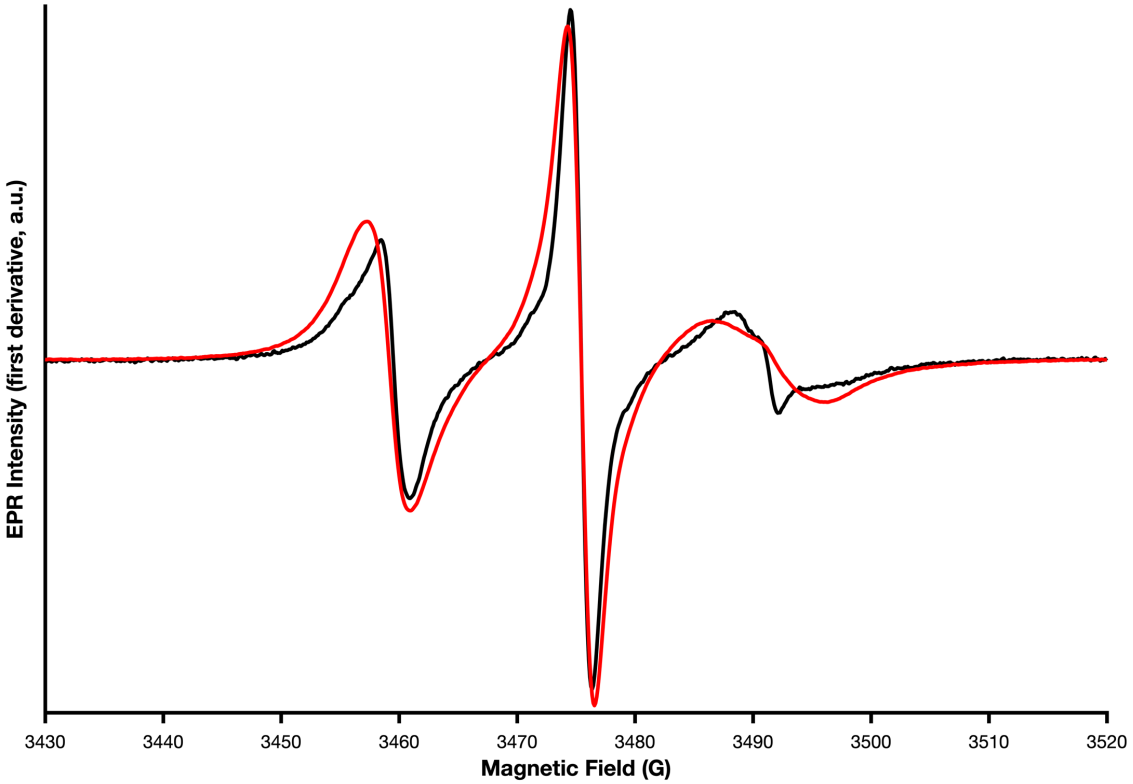


Fig. 5.2.13: the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled Q87C/T159C variant of SBD1 (black) and T369C/S451C variant of SBD2 (red) at X-band. The spectra have been normalised to the same scale and centred over one another.

An initial, qualitative assessment of the spectral features compared with those of the equivalent spectrum of SBD1 (apoprotein form) would predict a slightly longer rotational correlation time, τ_c, for apo-SBD2; all of the peaks exhibit greater broadness and the high-field peak is much more asymmetric in the spectrum of SBD2. These differences in characteristic features may support the idea that the spectrum of the SBD1 sample has significant influence from unbound label in the solution.

This correlates well with the comparison of rotamer library numbers for each cysteine mutation site in each of the proteins as calculated from the *in silico* labelling experiments. SBD1 sites Q87C and T159C have libraries of 85 and 115 rotamers respectively, whilst SBD2 sites T369C and S451C have 69 and 47 respectively. Lower populated rotamer libraries imply less rotational space around the proposed mutation site and therefore a slower rotational correlation time.

As before, the spectrum was simulated and fitted against the experimental data in order to extract accurate values for quantitative analysis.

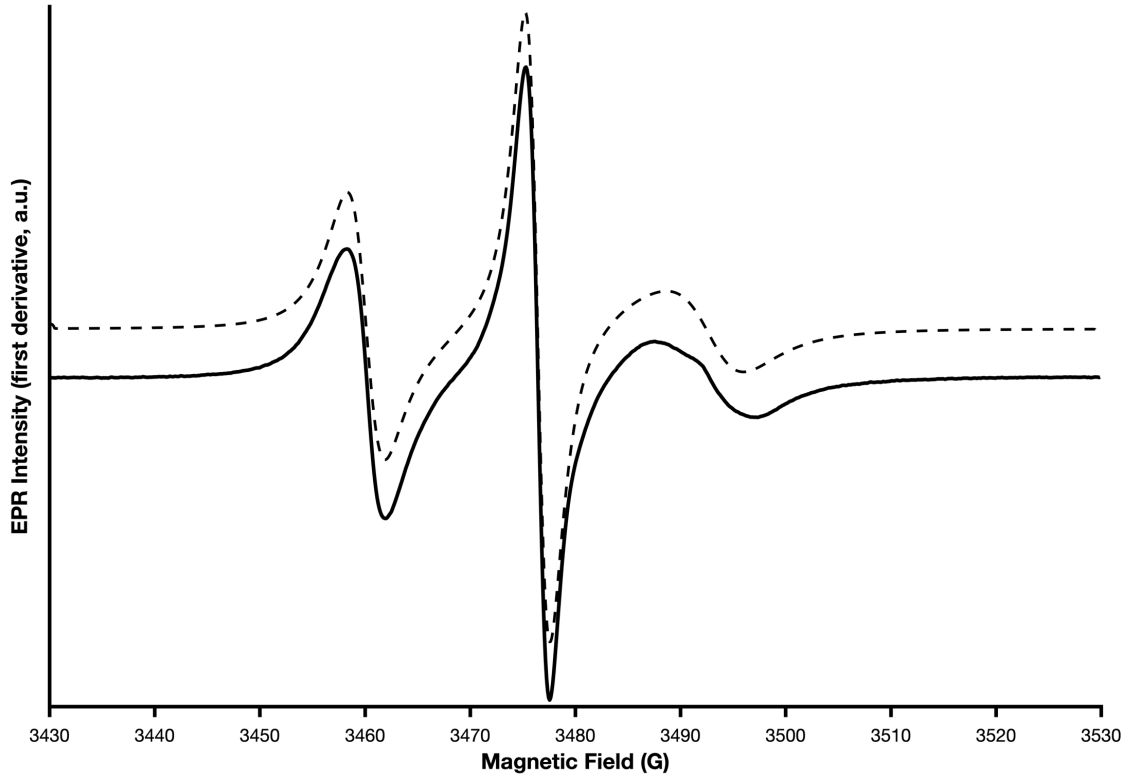


Fig. 5.2.14: the fitted simulation of the room temperature cw-EPR spectrum of the apoprotein, MTSL-labelled T369C/S451C variant of SBD2 at X-band; experimental data is shown as a solid line, fitted simulation is shown as a dotted line.

Table 5.2.7: best fit simulation values for the x, y and z components of each of the g- and A-tensors, as well as averaged correlation time and linewidth for the room temperature X-band experimental data of the apoprotein, MTSL-labelled SBD2. The RMSD of the fitting is given.

g-tensor			**A-tensor (MHz)**			**Linewidth (peak-to-peak, mT)**	**τ_c (ns)**	**Relative weighting**	**Simulation RMSD**
x	y	z	x	y	z				
2.0075	2.0056	2.0024	16.9	16.1	106.4	0.01	9.8	0.41	0.014308
2.0075	2.0062	2.0021	16.5	15.8	96.3	0.01	5.0	0.59	

The fitting of the simulated spectrum converged well with the experimental spectrum of SBD2 in the apoprotein state with two spin label species defined, yielding a τ_c of 9.8 ns for one component and 5.0 ns for the other (0.41:0.59 relative weighting, respectively). As previously discussed, the other parameters (i.e. the individual x, y and z components of the g- and A-tensors) should be considered good estimates at best.

Whilst the rotational correlation time, τ_c, is arguably the most 'important' output of the simulation and fitting of the room temperature X-band cw-EPR spectrum, there are other characterising parameters of the isotropic system: g_{iso} and A_{iso}. As previously discussed, the limitations of room temperature cw-EPR at this frequency/applied field range means the individual x, y and z components of these tensors are averaged.

Discrepancies between the g_{iso} values of the two SBDs (2.00553 and 2.00523 weighted average of SBD1 and SBD2 respectively) likely arise from equipment/setup differences over time affecting things like the homogeneity of the applied magnetic field at the sample cavity. Meanwhile, the differences in A_{iso} between the two apoprotein SBDs are negligible (within 1 MHz of each other).

By comparison of the extracted τ_c values of the apoprotein forms of both SBD1 and SBD2, there is good agreement with what would be predicted from the rotamer library outputs of the *in silico* labelling experiments. The fitted simulation of the apoprotein SBD1 is dominated by the short τ_c (2.7 ns) species, with a weighted average of 6.3 ns between the two simulated label species. In contrast, the equivalent fitting for apoprotein SBD2 gives a relatively even weighting of species with τ_c values of 5.0 ns and 9.8 ns. However, this yields a weighted average of 7.0 ns, which is not too different to the corresponding SBD1 value. The SBD1 Q87C/T159C variant was predicted to have 85 and 115 rotamers at each respective labelling site, whilst SBD2 T369C/S451C was predicted to have 69 and 47 respectively.

It therefore fits that the doubly labelled SBD2 variant would have a longer τ_c, though the weighted averages of the extracted values are not very dissimilar between the two SBDs. However, qualitative assessment of the spectral characteristics shows the apoprotein SBD2 to be significantly broader and more asymmetric than SBD1. This may be because of the intrinsic difficulties in simulating and fitting experimental data recorded under a single set of parameters;

the breadth of the spectrum (i.e. how the high and low field nitroxide peaks are separated from the central peak) can be influenced by both τ_c and the A-tensor, and finding a best fit of both simultaneously could lead to errors.

Low temperature X-band

The low temperature (50 K) cw-EPR spectrum of the apoprotein MTSL-labelled SBD2 was recorded at X-band.

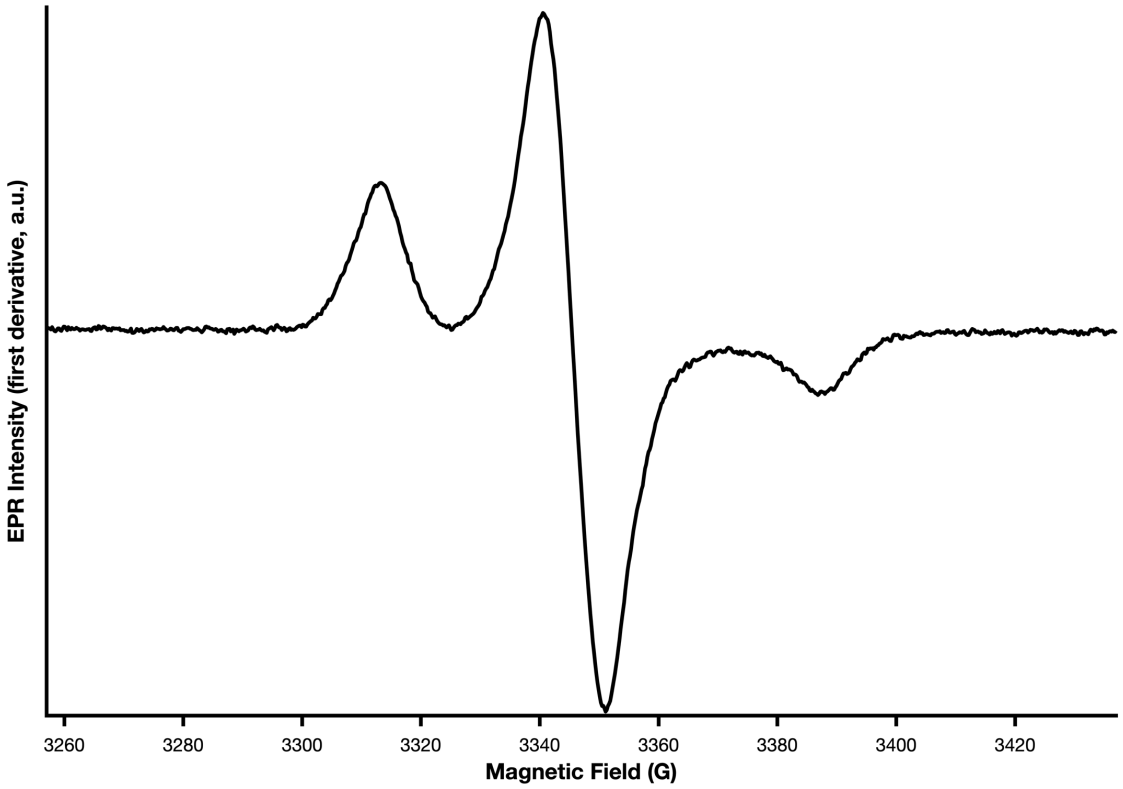


Fig. 5.2.15: the low temperature (50K) cw-EPR spectra of the MTSL-labelled T369C/S451C variant of SBD2 at X-band in the apoprotein binding state.

From initial assessment of this spectrum it appears there is no apparent broadening as a result of dipolar coupling of two spin labels closer than about 20 Å. This would agree with the *in silico* prediction of an inter-label distance of 57.6 Å for the apoprotein form of SBD2.

A simulated spectrum was then fitted to the experimental data (figure 5.2.16) using the EasySpin package[13] for MATLAB. The best-fit results of the fitted simulation of the room temperature cw-EPR spectrum was used to provide starting parameters for this analysis.

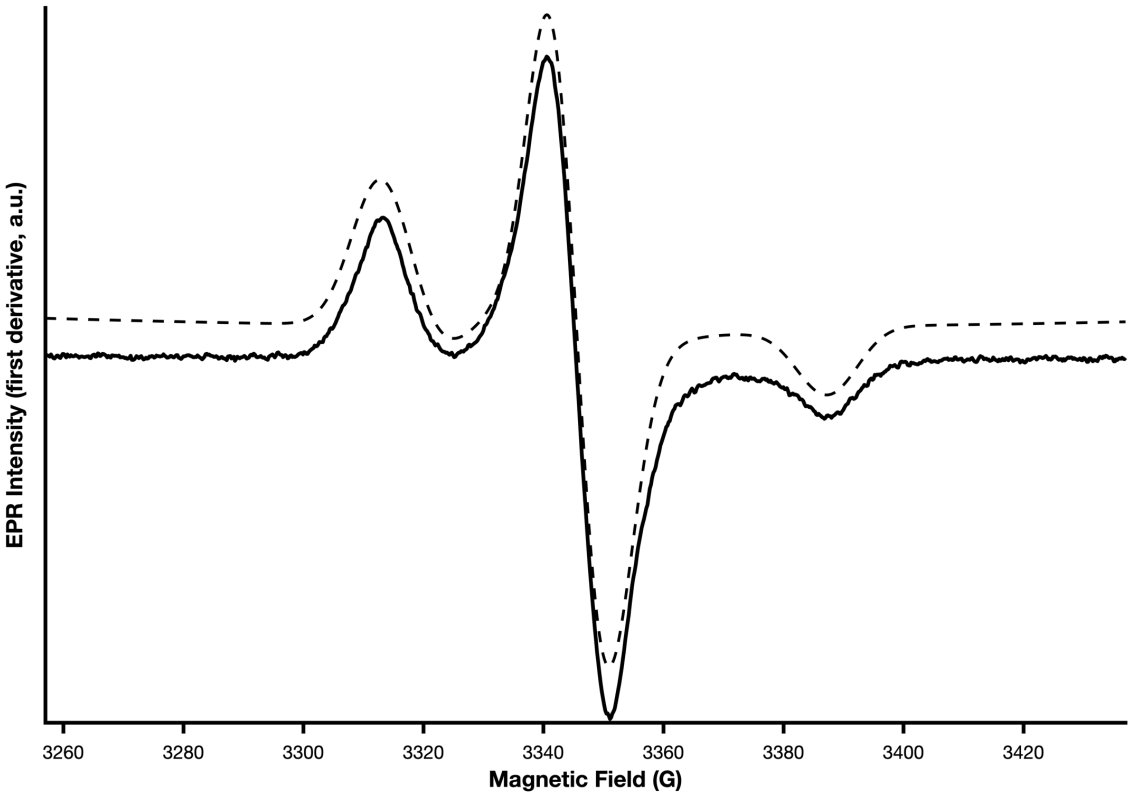


Fig. 5.2.16: the fitted simulations of the low temperature cw-EPR spectra of the MTSL-labelled T369C/S451C variant of SBD2 at X-band in the apoprotein binding state; experimental data are shown as solid lines, fitted simulations are shown as dotted lines.

As with the fitting of the simulated room temperature X-band data, two spin label species were used to achieve a low RMSD. These two different components are very similar across the parameters except for the linewidth and A_{zz} values.

Table 5.2.8: best fit simulation values for the x, y and z components of each of the g- and A-tensors and linewidth for the low temperature X-band experimental data of the apoprotein, MTSL-labelled SBD2. The respective, relative weightings are provided for each component spin label ‘ species’ . The RMSD of the fitting is given.

Component	**g-tensor**			**A-tensor (MHz)**			**Linewidth (peak-to-peak, mT)**	**Relative weighting**	**Simulation RMSD**
	x	y	z	x	y	z			
1	2.0075	2.006	2.002	14.1	13.3	106.2	0.92	0.83	0.025475
2	2.0077	2.0057	2.002	14.2	14.1	93.2	0.79	0.17	

Generally, the values for the different parameters extracted here align with those of apoprotein SBD1 under the same conditions. However, the A_{zz} values of each of the two spin label species and the ratio of these species in each SBD differ significantly. Whilst in both SBD1 and SBD2 there is one low A_{zz} (<100 MHz) and one high A_{zz} (>100 MHz) species, in SBD1 the low A_{zz} has the higher relative weighting (99.7 MHz and 107.2 MHz at 0.69:0.31 respectively). Meanwhile, the opposite is true of the fitted data for the SBD2 spectrum (106.2 MHz and 93.2 MHz at 0.83:0.17 respectively). However, taking a weighted average of the values yields similar results: 102.03 MHz for SBD1 and 103.99 MHz for SBD2.

It is important to bear in mind the limitations of spectroscopy at this frequency; it is not capable of resolving all of the individual components of the g- and A-tensors. For maximal reasonable accuracy the spectra should be recorded at several frequencies (X-, Q- and W-band) under low temperature conditions (to maximise anisotropy of the g- and A-tensors) and the data simultaneously fitted to gather consistent extracted parameters.

Pulsed EPR

The MTSL-labelled Q87C/T159C variant of SBD1 was studied using PELDOR spectroscopy across its binding states: apoprotein, Gln-bound and Asn-bound. The resulting time traces were processed using the DeerAnalysis package[18] for MATLAB. The background factor is fitted to account for the exponential decay of the signal and an automated Tikhonov regularisation is then applied the ill-posed problem of extracting a distance from the dipolar coupling of two electron spins. The resulting distance distributions are yielded with colour-coded confidence bands (fig. 5.2.17).

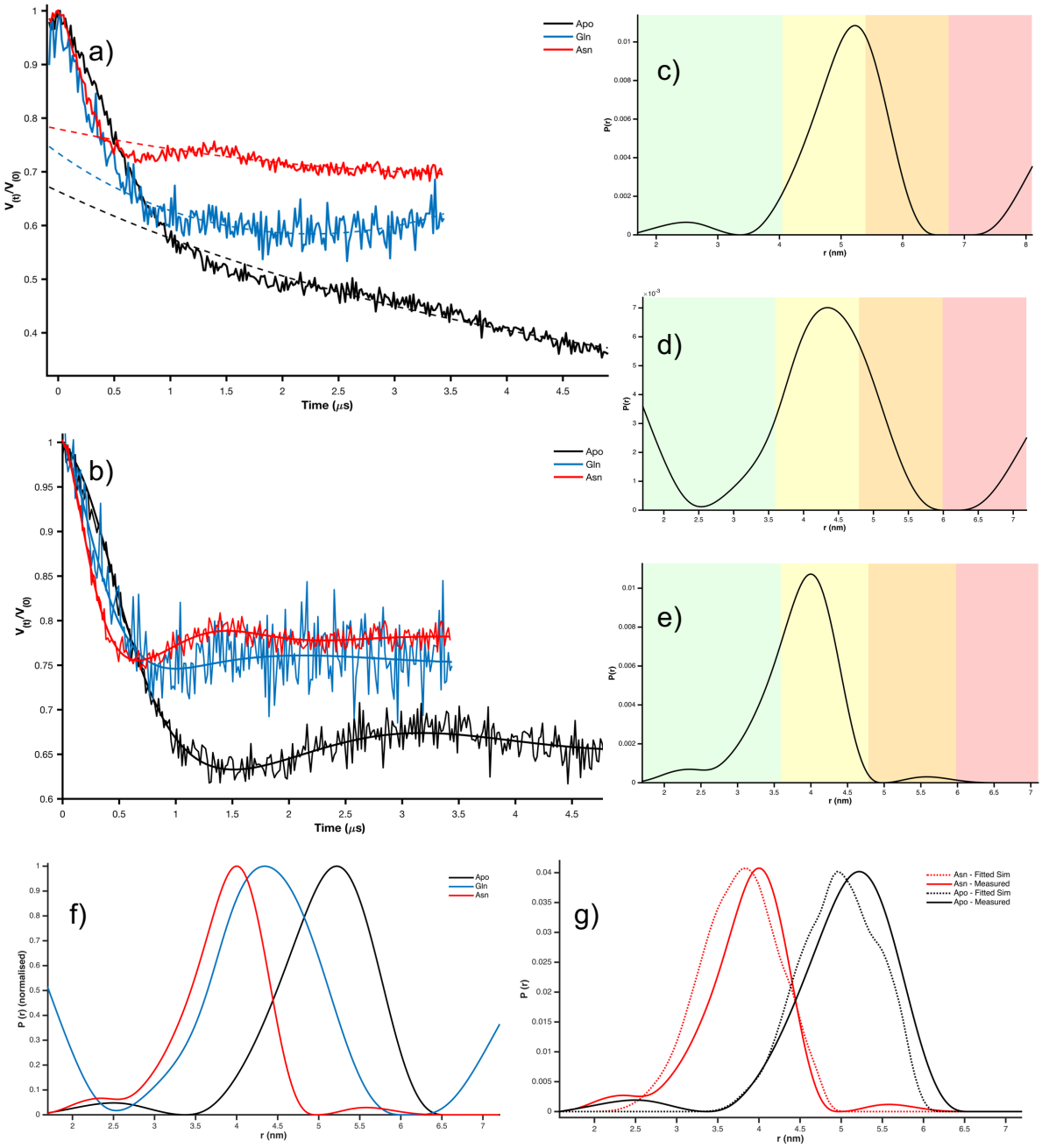


Fig. 5.2.17: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled Q87C/T159C variant of SBD1, a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, Gln-bound, blue, and Asn-bound, red, b) background factor subtracted time trace (apoprotein in black, Gln-bound in blue, Asn-bound in red), c) apoprotein distance distribution output, d) Gln-bound distance distribution output, e) Asn-bound distance distribution output, f) scale-normalised overlay of the distance distributions (apoprotein in black, Gln-bound in blue and Asn-bound in red), f) scale-normalised overlay of the distance distributions (apoprotein in black solid line, Asn-bound in red solid line) with fitted simulations (from MMM; apoprotein in black dotted line, Asn-bound in red dotted line).

The distance distributions shown here are all unimodal and show a clear change in inter-label distance between the three binding states of SBD1. In the apoprotein form the population peaks at 52.3 Å. Upon glutamine binding, the distance decreases to 43.5 Å, whilst asparagine binding further shortens the distance to 39.9 Å. In all cases, the distribution peaks are within the higher confidence bands. However, the oscillation in the unprocessed time trace of the Gln-bound SBD1 appears to be incomplete; i.e. remeasuring this PELDOR experiment with a longer dipolar evolution time (in order to observe a full oscillation) would likely be beneficial to the accuracy of this data. This, and the level of noise in the trace, is probably the reason for the broad distribution exhibited above.

As mentioned above (section 5.2 ‘cw-EPR & simulations of SBD1’), issues with later preparations of the SBD1 protein were experienced. The apoprotein data displayed previously (fig. 5.2.17) is a result of one such sample. Initially, the apoprotein PELDOR trace yielded a bimodal distance distribution (shown below, fig. 5.2.18).

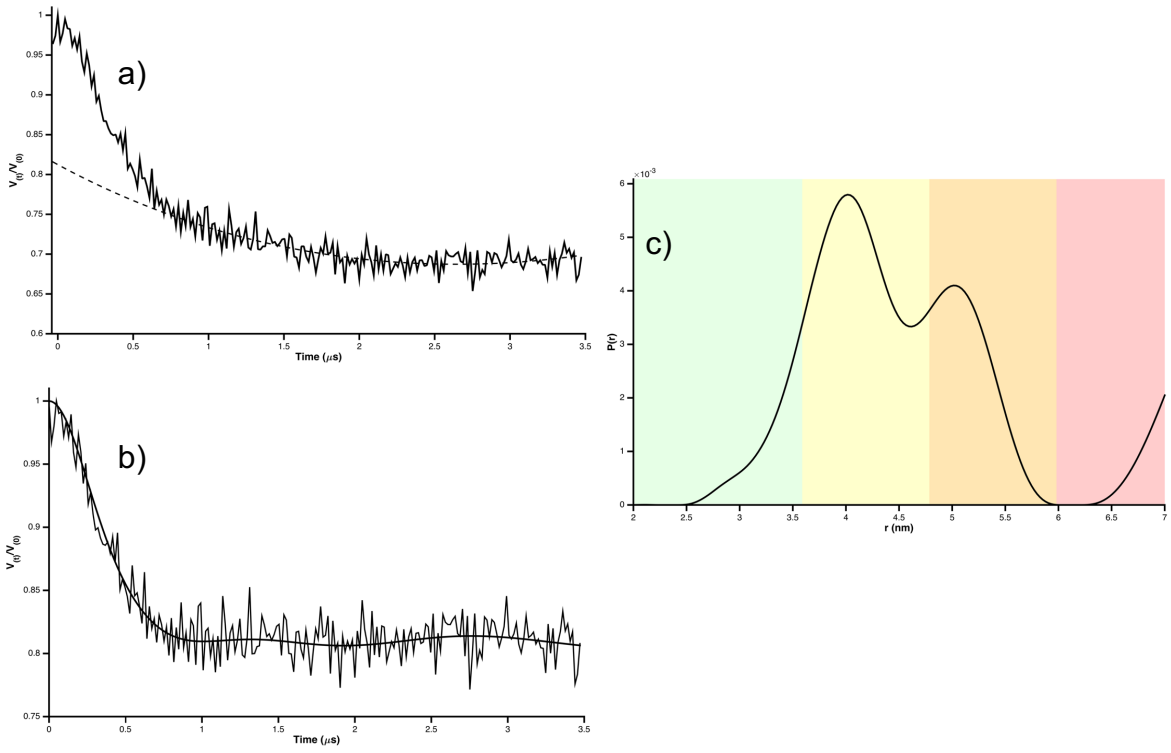


Fig. 5.2.18: the 4P-PELDOR data and resulting distance distribution of MTSL-labelled ‘apoprotein’ Q87C/T159C variant of SBD1, a) ‘raw’ data (trace of the echo decay and oscillation in the time domain), b) background factor subtracted time trace, c) apoprotein distance distribution output.

The modes populated here indicate two conformations of the protein, one of which was suspected to be a result of contamination (i.e. substrate-bound). This was presumed due to the prediction that the protein would not occupy a ‘closed’ state unless substrate was present, based on the FRET measurements previously acquired[26]. The longer distance in this data matches well with that seen in the unimodal distribution of the apoprotein data displayed above (fig. 5.2.17). This would indicate that the issue of contamination had indeed been solved with the more recent purifications. However, these samples failed to show any indication of substrate binding in the PELDOR traces, even upon incubation with large excess amounts of glutamine (see below, fig. 5.2.19).

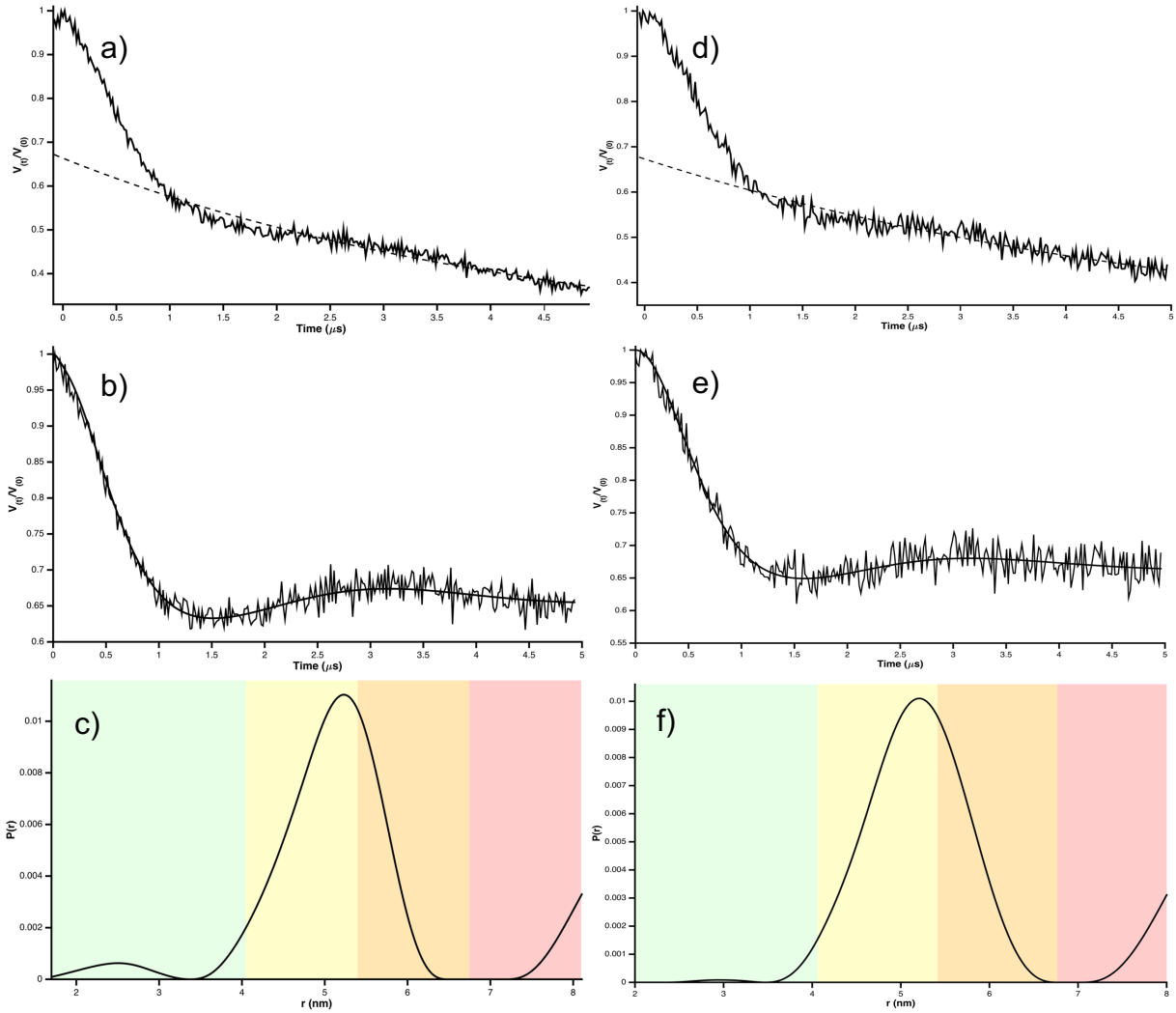


Fig. 5.2.19: the 4P-PELDOR data and resulting distance distribution of MTSL-labelled Q87C/T159C variant of SBD1, a) ‘raw’ data (trace of the echo decay and oscillation in the time domain) of apoprotein form, b) background factor subtracted time trace of apoprotein form, c) apoprotein distance distribution output, d) ‘raw’ data (trace of the echo decay and oscillation in the time domain) of excess Gln form, e) background factor subtracted time trace of excess Gln form, f) excess Gln distance distribution output.

This suggests that the data presented above for the apoprotein form of this SBD1 variant may be inaccurate, as the protein appears to be inactive. However, the agreement of the longer distance extracted from the ‘old’ apoprotein data (fig. 5.2.18) and the distance of the single population seen in the ‘new’ data (fig. 5.2.17 and fig. 5.2.19) is notable.

The initial PELDOR results for two of the binding states of SBD2 (below, fig. 5.2.20) were measured using glutamate as the substrate. Both states exhibit bimodal distance distributions, with peaks at very similar distances. However, the relative populations of these modes differ between the two states.

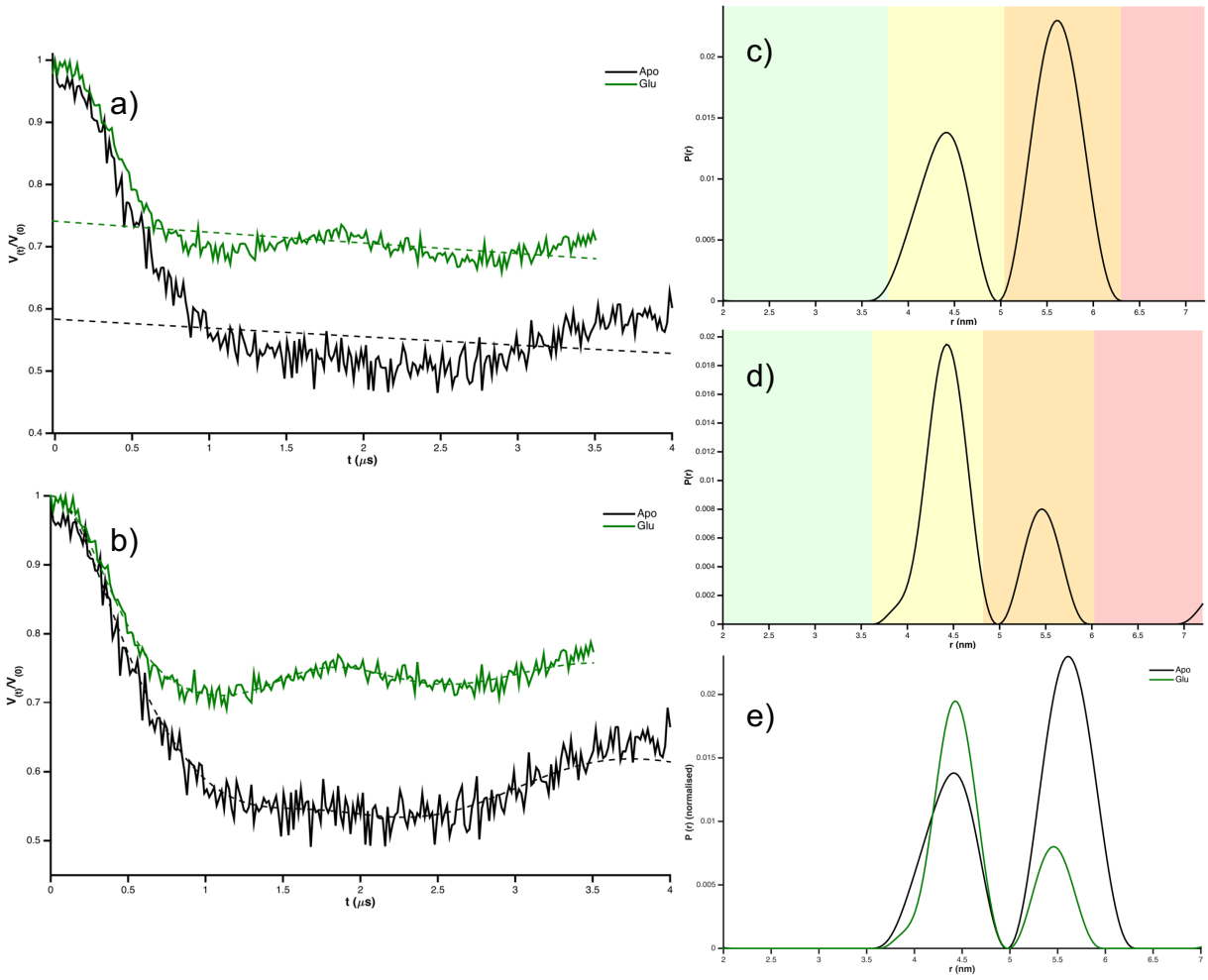


Fig. 5.2.20: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled T369C/S451C variant of SBD2, a) ‘raw’ data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and Glu-bound, green, b) background factor subtracted time trace (apoprotein in black and Glu-bound in green), c) apoprotein distance distribution output, d) Glu-bound distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black and Glu-bound in green).

In the apoprotein form, the longer distance mode at 56.3 Å is greater in population than the shorter distance (44.3 Å). In the glutamate-bound state, the opposite is true; the longer distance (54.6 Å) is less populated than the shorter distance (44.2 Å). In both cases, the longer distance population is within a lower confidence band, as a result of experimental constraints, i.e. the echo did not persist long enough to measure across a greater dipolar evolution time. As such the data should therefore be remeasured across a longer dipolar evolution time for better accuracy; this may be achieved with greater protein concentration, higher labelling efficiency and/or more extensive deuteration (of the buffer, or higher v/v percentage d_8-glycerol).

Much like the 'old' data of apoprotein SBD1 (see above, fig. 5.2.18), the bimodal distribution of the apoprotein SBD2 here may be interpreted as being a result of contamination with substrate. In contrast, the bimodal distribution of the glutamate-bound form may be a result of being incubated with too little substrate to force full closure of the SBD, or that the SBD does indeed undergo some detectable open-closed conformational changes.

As previously mentioned, the difficulties in producing newer samples of the SBDs meant further experiments were hindered. The bimodal distribution seen in the above experiments was replicated in further preparations of the apoprotein SBD2 (see fig. 5.2.21). The population distribution between the modes is different here (with the shorter distance having a higher population) but the distances they occur at match fairly well with the previous data at 44.3 Å and 55.4 Å. However, the raw time trace data does not show pronounced oscillations above the noise level and the relative ratio of the two distances may be an artefact of the fitting of the background function.

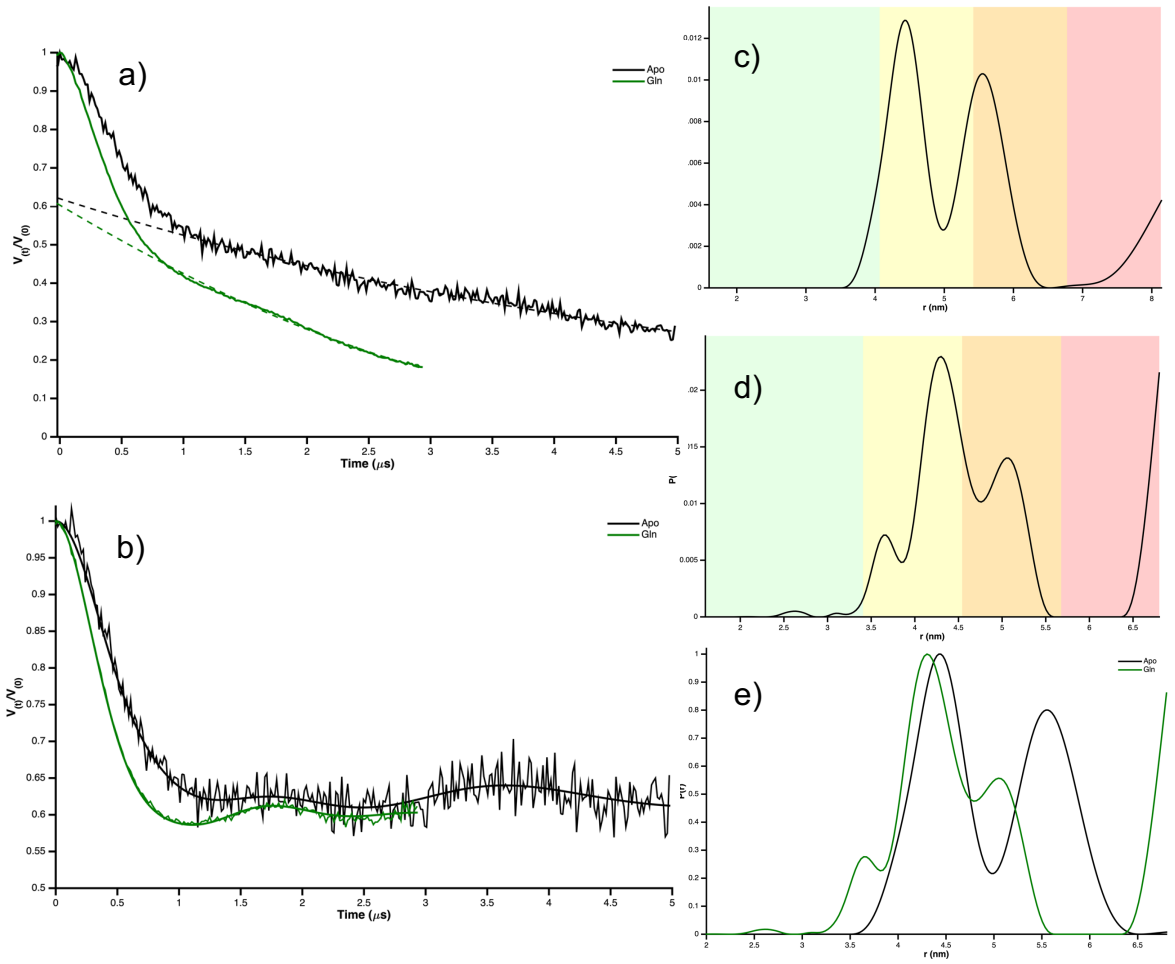


Figure 5.2.21: the 4P- PELDOR data and resulting distance distributions of MTSL-labelled T369C/S451C variant of SBD2, a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and excess Gln, green, b) background factor subtracted time trace (apoprotein in black and excess Gln in green), c) apoprotein distance distribution output, d) excess Gln distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black and excess Gln in green).

Addition of excess glutamine shift the distance distribution to a broad feature at a shorter distance of 43.0 Å. Though there are shoulders to this feature it is difficult to assess their accuracy, considering the dipolar evolution time was only measured to 3 μs. As can be seen in the raw time trace data (figure 5.2.21a) the echo decayed significantly by this time. Attempts to measure a longer dipolar evolution time were made difficult by very low signal to noise and the further-decayed echo prevented processing by normal means.

Comparison with sm-FRET

These GlnPQ SBD variants (SBD1 Q87C/T159C and SBD2 T369C/S451C) were initially designed for FRET studies[11, 26]. The complementarity of this technique and the equivalent EPR experiments allowed the same variants to be used here and permits an almost direct comparison of the results of the respective methods.

Table 5.2.9: comparison of the distance-based data, both predicted and measured, with the reported FRET efficiencies of the fluorescent-labelled GlnPQ SBDs in a range of conditions. The distance measurements consist of: the distance between the C_α atoms of the mutation positions as calculated from the crystal structures, the distances between the spin labels in the double-labelled as predicted from the simulated distance distributions calculated at 298 K using spin label MTSL, and the measured distances between the spin labels in the double-labelled variants as extracted from the distance distributions obtained from the PELDOR experiments. In cases where there are multiple modes in the distance distributions, the predominant value is shown in bold. Dashes indicate where there are missing data points.

[a] – data reported in publication by de Boer et al., 2019[11]

Conditions	Distance (Å)			FRET Efficiency, E* (ALEX sm-FRET)[a]
	C_α-C_α	MMM	PELDOR	
SBD1 Apo	44.3	50.2	52.3	0.612 ± 0.003
SBD1 + Gln	-	-	43.5	0.710 ± 0.003
SBD1 + Asn	33.6	37.5	39.9	0.805 ± 0.002
SBD2 Apo	48.6	57.6	44.3, **56.3**	0.492 ± 0.003
SBD2 + Glu	-	-	**44.2**, 54.6	0.633 ± 0.004
SBD2 + Gln	39.3	47.1	43.0	0.677 ± 0.002

The lack of a complete crystal dataset (i.e. crystal structures of both SBDs in all binding states) makes it difficult to draw a comparison between distance measurements taken from theoretical methods with those obtained through experimental means. However, it can be seen that the SBD1 apoprotein and asparagine-bound distance changes (Δ=10.7 Å, 9.4 Å, 12.7 Å and 12.4 Å for the

crystal C_α-C_α, FRET inter-dye[11], MMM-predicted and PELDOR-measured inter-label measurements respectively) all agree well. When the differences between the C_α position, fluorescent dye linkers and spin label linkers and their relative orientations are considered, this likely explains the discrepancies seen in these values.

Whilst there is good agreement between the trends in the PELDOR distances and FRET efficiencies, it is difficult to compare the results quantitatively; calculation of inter-dye distance requires knowledge of certain experimental details[28]]. However, the reported FRET efficiency values can be used for a semi-quantitative assessment of SBD closure. Consider SBD1, where the apoprotein state has a reported FRET efficiency of 0.612 and inter-label distance of 52.3 Å as measured by PELDOR. Upon glutamine binding the FRET efficiency improves to 0.710 (reported distance change of 4.6 Å[11]) and the distance between spin labels is shorter, at 43.5 Å (Δ=8.8 Å). Upon asparagine binding the FRET efficiency is even higher (0.805, Δ=9.4 Å[11]) and the inter-label distance even shorter (39.9 Å, Δ=12.4 Å vs apo). This pattern also corresponds with the K_d values; the shorter PELDOR distance, higher FRET efficiency and greater inter-dye distance change is seen for the 'tighter binding' substrate, asparagine, which has a reported K_d of 0.2 μM. In contrast, glutamine has a hundredfold higher K_d of 92 μM, and exhibits a longer inter-spin distance, smaller inter-dye difference and lower FRET efficiency.

The equivalent data for SBD2, however, is not complete; there is no K_d yet reported for glutamate binding. The comparison of the PELDOR results with the other data is complicated in that the apoprotein and glutamate-bound samples show bimodal distributions. However, the dominant population of 56.3 Å in the apoprotein is consistent with the predicted distance yielded by MMM (57.6 Å). The predominant distance of 44.2 Å in the glutamate-bound also agrees fairly well with the prediction for glutamine-bound SBD2 (47.1 Å). The distance distribution obtained for Gln-bound SBD2 was very broad and the time trace data was only recorded to 3 μs but had a peak at 43.0 Å`

These data show some tentative agreement with the reported FRET data, wherein the apoprotein to Glu-bound distance change is reported as 7.2 Å. The equivalent PELDOR data showed a change of 12.1 Å when looking at the change in dominant populations of the distance distributions. Meanwhile, for the apoprotein to Gln-bound conformational change, a larger distance change of 9.4 Å is reported

from the FRET experiments. Though the PELDOR data has some uncertainties associated with it, the preliminary results also show a greater lobe closure of 13.3 Å. These measurements show agreement that glutamine binding likely causes greater SBD2 closure and indicates that the K_d of glutamate is probably higher than that of glutamine (K_d 0.9 μM), in the same way that the closure of SBD1 by asparagine (K_d 0.2 μM) is demonstrated to be greater than that of glutamine (K_d 92 μM) through both FRET and PELDOR distance measurements.

6 Pneumococcal Surface Adhesin A

6.1 Background

Streptococcus pneumoniae, or pneumococcus, is the most common cause of bacterial pneumonia in children (approximately 5 million children under 5 years old die each year from pneumonia worldwide) and is also involved in several other invasive diseases (e.g. bacterial meningitis)[29-30]. The treatment for such diseases is complicated by the ever-growing resistance to existing antibiotics.

Pneumococcal Surface Adhesin A (PsaA) is the metal ion-binding subunit of the manganese-specific ABC-type transporter in S. pneumoniae, PsaBCA. The complex is known to be important for the survival and virulence of this bacterium, as manganese is involved in several essential metabolic processes, such as the regulation of oxidative stress[31]. However, the mechanism of metal ion binding and transport are not yet well understood. It should be noted that, despite its name, there has been substantial debate as to whether PsaA is actually involved in adhesion; it is doubtful that it would protrude through the cell wall if it is attached to the rest of the protein complex of the cell membrane (described more fully below), as would be expected in gram-positive ABC transporters[32-34]. However, a significant reduction in cell adhesion is witnessed in PsaA mutants[35].

Whilst the structures of many SBPs from gram-negative bacteria have been resolved, only a few of these proteins from gram-positive organisms have been successfully isolated and characterised[36]. The SBPs of gram-negative bacteria tend to adopt a two-domain fold where both domains are globular, connected by one or more chains. It is between these domains that the substrate binds. It has been found that these proteins generally contain a ‘hinge’ region about which conformational change occurs upon substrate binding[36]. However, in the case of PsaA – a cluster A-I type SBD – the ‘hinge’ consists of a single, rigid α-helix linker[2], but still must undergo some bending or twisting conformational change upon substrate binding[37].

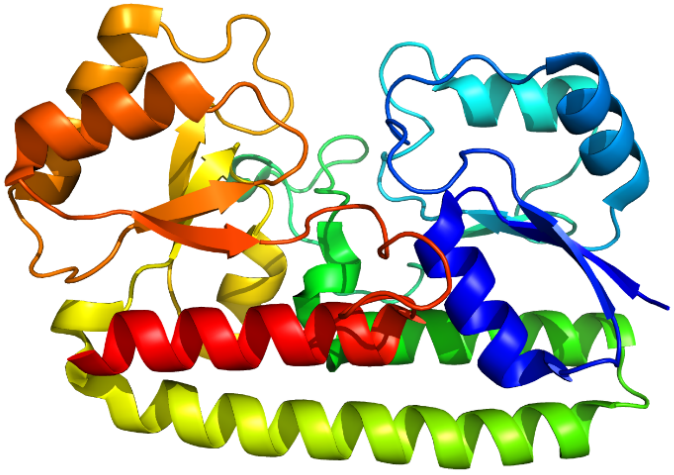


Fig. 6.1.1: the overall structure of the apoprotein form of PsaA (PDB 3ZK7) with the lobe domains and connecting helix highlighted (progression from N- to C-terminal shown by the blue to red colouring).

The crystal structures of the apoprotein, manganese-bounds and zinc-bound forms of PsaA have been resolved (PDB accession codes 3ZK7[37], 3ZTT[38] and 1PSZ[39] respectively). The overall structure of the protein consists of two $(\beta/\alpha)_4$ sandwich domains, designated as the N- and C-terminal lobes[37]. These domains are connected by a novel inter-domain helix backbone (previously unseen in ABC transporters)[39]; it is here that the protein is found to bend (residues 190-194)[37]. The metal-binding site is located in a cleft between the lobes, which is rich in negatively charged residue side chains. The specific coordinating residues are His67, His139, Glu205 and Asp280[37, 39].

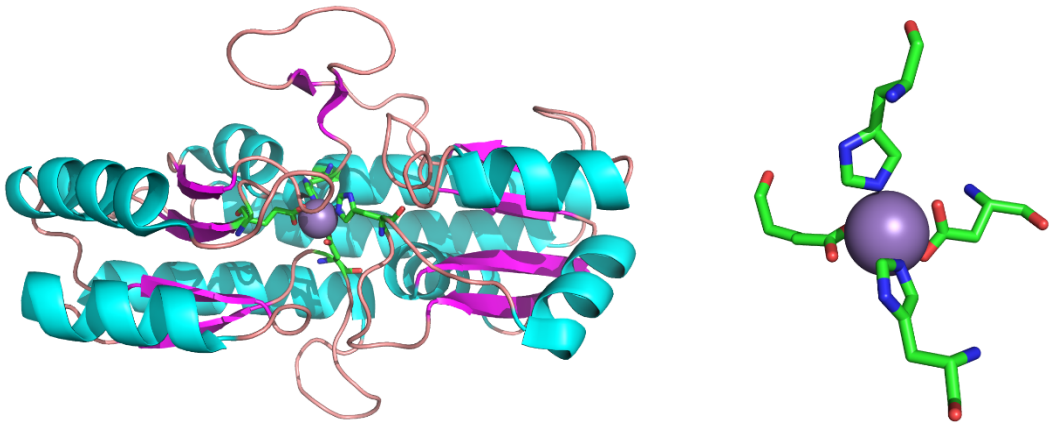


Fig. 6.1.2: the overall structure of the Mn(II)-bound form of PsaA (PDB 3ZTT) with metal-coordinating ligands highlighted, left, and close-up view of Mn(II) coordinated by the ligands, right.

Whilst divalent manganese is the natural substrate of PsaA, Zn^{2+} is known to also form a complex with the protein. It is believed that zinc is a competitive inhibitor of manganese uptake in this system; manganese is known to be important for the survival, proliferation and virulence of bacteria. In particular, Mn(II) plays a central role in the regulation of oxidative stress as a cofactor in the superoxide dismutase (SOD) enzyme, where it catalyses the reaction of superoxide to yield oxygen and hydrogen peroxide[40-41]. Zinc-mediated Mn^{2+}-starvation would leave the bacteria susceptible to oxidative damage[40]. Interestingly, there is also evidence that the host immune response involves an increase in Zn^{2+} concentration at the site of pathogen infection[37, 39-40].

The crystal structure of the Mn(II)-bound form of PsaA shows four ligands coordinating the metal ion in a tetrahedral arrangement. This is unusual in manganese-binding but would be expected in zinc-coordination[37, 39]. Reported comparisons of the Mn(II)- and Zn(II)-bound crystal structures have shown only very small differences[37, 40]; it is therefore not readily apparent as to how PsaA is manganese-specific. The first explanation for Mn(II)-specificity lies in the preferred ligands for each of these metal ions; Zn^{2+}-binding is favoured by three N ligands (i.e. three histidines), whilst Mn^{2+}-binding is favoured by the combination of two N and two O ligands (as found here in the form of two imidazole rings and two carboxylates)[39]. In addition, the derived affinity constant (or equilibrium dissociation constant, K_d) of PsaA for Mn(II) has been reported as 3.3±1.0 nM, whilst for Zn(II) it is two orders of magnitude higher at 231±1.9 nM (for 1:1 metal ion to protein ratio)[41]. Moreover, the entropic contributions of zinc-coordination are highly unfavourable in comparison to those of manganese-binding[37, 39]. Conflictingly, according to the Irving-Williams series, which predicts the stability of metal ion complexes, zinc would be the preferred substrate.

The unusual and unfavourable tetrahedral binding geometry of Mn(II) (which prefers a six-coordinate octahedral geometry) has been theorised to allow the metal ion to bind reversibly; the transient binding facilitates the transfer of Mn(II) from PsaA to the rest of complex for relocation across the membrane to the cytoplasm. Furthermore, since tetrahedral coordination is the preferred arrangement for Zn(II), this may bind more stably to PsaA thus disrupting the manganese-uptake system[37]. However, this is at odds with the K_d values reported above; a lower K_d indicates a ‘tighter’ binding. As a result of this contradictory information several related questions can be raised:

- How is Mn^{2+} binding transient, despite the high affinity binding?
- How is Zn^{2+} binding irreversible, despite the lower affinity binding?
- Why is Mn^{2+} binding competitively inhibited by a substrate with lower affinity?

6.2 Proposed crystallographic model

A ‘Venus flytrap’ mechanism has been suggested for various other substrate binding domains (SBDs), wherein the substrate molecule is caught between the SBD’s lobes which rotate about a flexible hinge region[42]. The degree of rotation between the lobes and hinge has been shown to vary greatly for different SBDs, with extreme examples being the amino acid-binding LivJ, which models show rotating up to 60° [43], and Zn(II)-binding TroA, which exhibits more restricted rotation of only 4° [44]. Due to the lack of high-resolution structural data available in relation to SBD conformations and the disparity in lobe-hinge rotational freedom, it is difficult to determine how ubiquitous the ‘Venus Flytrap’ mechanism is.

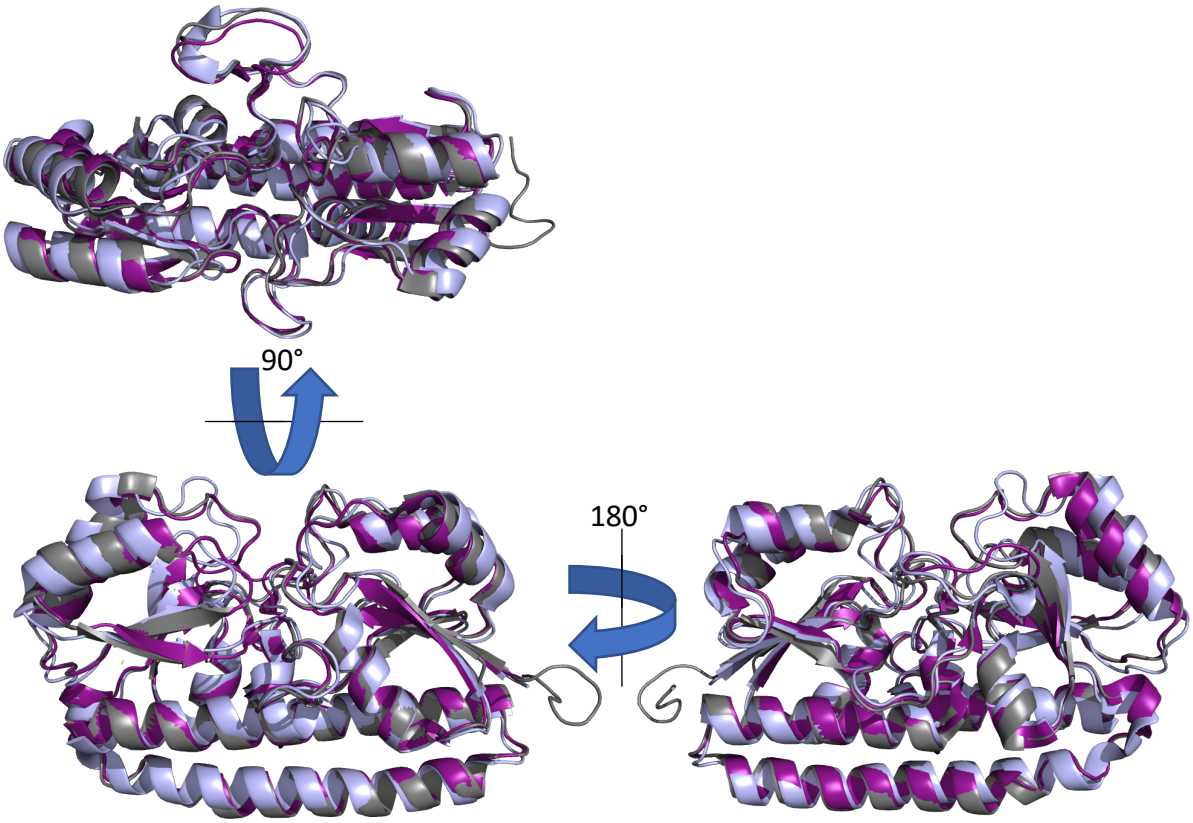


Fig. 6.1.3: overlaid crystal structures of PsaA in its different states from different aspects: pale blue is the apoprotein, purple is manganese-bound, and grey is zinc-bound (PDB 3ZK7, 3ZTT and 1PSZ respectively). Metal ions have been left out for clarity.

Based on the various crystal structures which have been resolved for PsaA (in the apo, Mn(II)-bound and Zn(II)-bound forms; PDB 3ZK7, 3ZTT and 1PSZ respectively), an alternative or adapted version of this model for the binding mechanisms has been built. This proposes that, as the substrate interacts with the coordinating side chains in the binding pocket the two lobes move towards one another (largest differences between apo and metal-bound form crystal structures,

see above, fig. 6.1.3); it is posited that the N-terminal lobe remains relatively fixed and the C-terminal lobe rotates, clasping the metal in place. It is then suggested that a section of the helical backbone, and the hydrogen bonds that cause it to adopt this fold, acts as a spring. In the case of Mn^{2+}-binding, which is transient in nature, the 'spring' bends and partially uncoils; some of the hydrogen bonds are broken in the process. These are thought to reform to aid metal release. However, upon Zn^{2+}-binding, the 'spring' is thought to bend further and unwind at the end of the helix, beyond the point where the hydrogen bonds can reform. This stretching of the spring is thought to be at least one of the factors that prevents the substrate release and causes zinc to be a competitive inhibitor of PsaA.

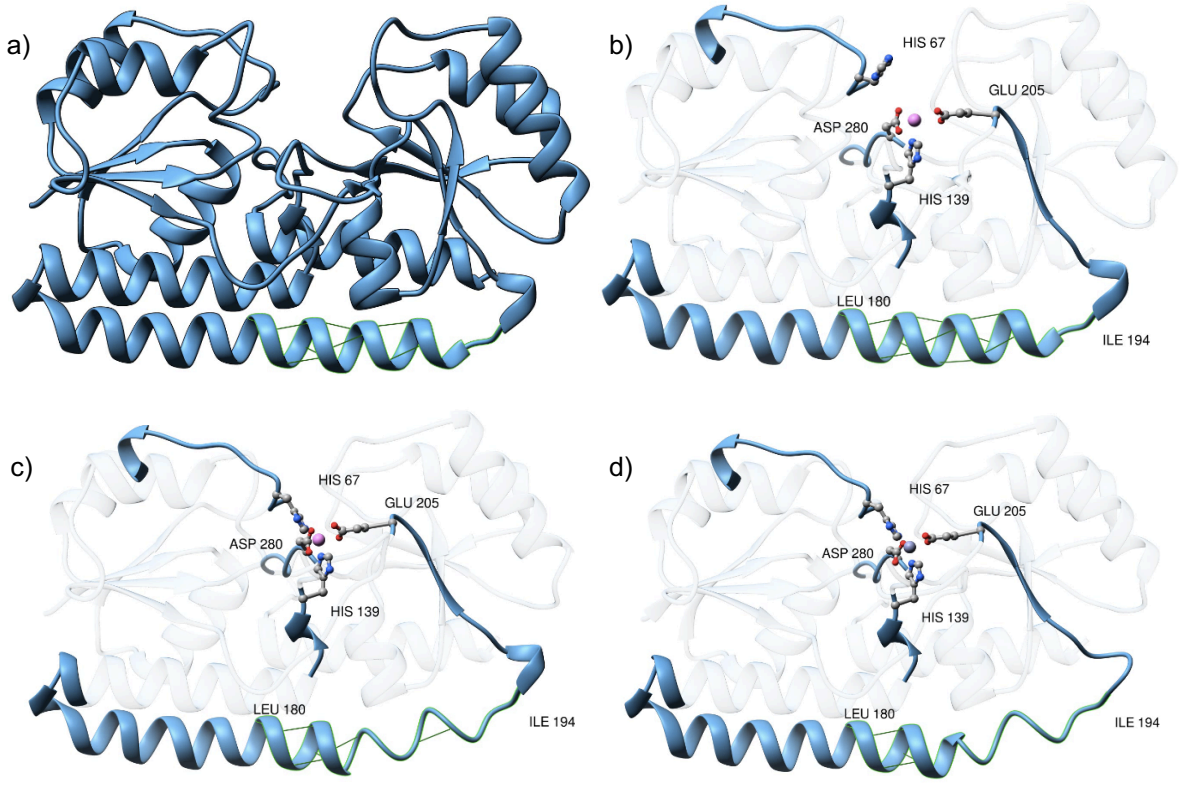


Fig. 6.2.1: scheme of the proposed mechanism of PsaA metal binding; a) apo PsaA; b) open Mn(II)-bound (metal coordinated, lobes moving together); c) closed, Mn(II)-bound (spring stretched, lobes closed); d) closed, Zn(II)-bound (spring unwound, lobes closed).

6.3 Results and Discussion

in silico spin labelling

Five single-cysteine (L56C, S58C, I125C, I236C, S266C) and five double-cysteine (L56C/I236C, L56C/S266C, S58C/S266C, S58C/I125C, I125C/S266C) containing variants of PsaA were designed to investigate the proposed mechanism using EPR spectroscopy. These cysteine mutations can be created *in silico* and a spin label computationally attached to it using the crystal structure. The degree of motional freedom the attached spin label has is then calculated and a library of rotamers produced. In the case of computational attachment of two or more spin labels, the distances between the labels are calculated, and the distance distributions (i.e. weighted probabilities of distances populated between the spin label rotamers) can be predicted.

The rotamer libraries were calculated for each of the labelling positions in PsaA under each condition (apo, Mn^{2+}-bound and Zn^{2+}-bound) from the crystal structures; these were calculated at ambient temperature (298 K) using Multiscale Modelling of Macromolecular Systems (MMM) software package that runs via MATLAB[12].

Table 6.3.1: rotamer libraries of the five labelling positions of PsaA in the apo, Mn^{2+}-bound and Zn^{2+}-bound forms, calculated from the crystal structures (PDB 3ZK7, 3ZTT and 1PSZ, respectively) at 298 K.

Variant	Calculated Rotamer Library (298 K)		
	Apo	Mn^{2+}-bound	Zn^{2+}-bound
L56C	2	11	12
S58C	1	12	23
I125C	21	23	24
I236C	15	34	31
S266C	16	33	27

Though this method is generally a good indicator of the motional freedom a spin label will have at a particular site, results must be assessed remembering that the program takes the crystal structure input as a rigid body. This means the software does not take into account the movements of biological molecules in solution, such as motion of the protein backbone; it is possible, therefore, that a spin label attached to a protein in solution may actually have more space to sample (i.e. more accessible rotamer conformations) than predicted. The predicted distances and distance distributions for the double label pairs were then calculated.

Table 6.3.2: the predicted distances between the pairs of spin labels in the double-labelled variants of PsaA as extracted from the predicted distance distributions, calculated at 298 K. In cases where there are multiple modes in the predicted distance distributions, the predominant value is shown in bold.

Variant	**Predicted inter-label distance (Å) (MMM @ 298 K)**		
	Apo	**Mn^{2+}-bound**	**Zn^{2+}-bound**
L56C/I236C	50.0	45.5	39.5, **44.5**
L56C/S266C	42.0	37.0, **40.0**, 43.5	37.0, **40.0**
S58C/I125C	35.5	**34.0**, 37.0	**33.5**, 36.0
S58C/S266C	34.0	31.5	31.0
I125C/I236C	**31.0**, 36.0	27.5, **32.0**, 36.0	31.5 (v broad), **38.0**

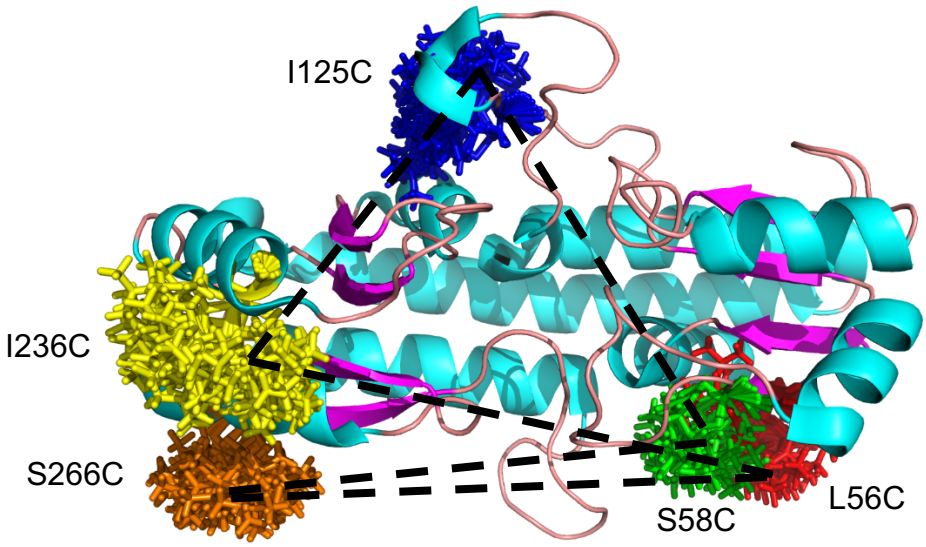


Fig. 6.3.1: the crystal structure of the apo form of PsaA (PDB 3ZK7) with MMM-generated rotamer libraries highlighted, at the residue positions where cysteines will be introduced. The double-cysteine containing variants are indicated by dashed lines connecting the respective component rotamer libraries.

Distances between the spin labels and bound substrate could then also be predicted by taking measurements from the modified crystal structure (now containing the rotamer libraries for the labelling sites).

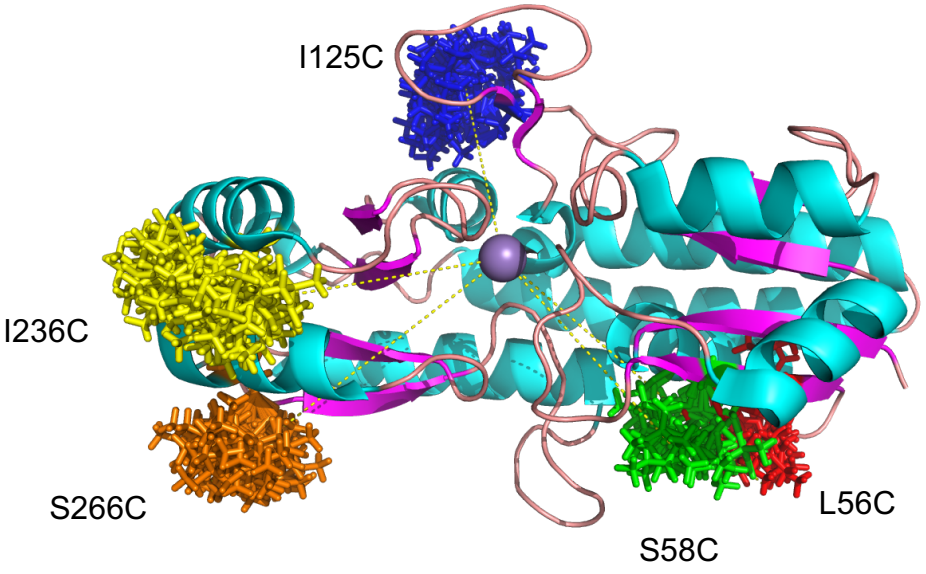


Fig. 6.3.2: the crystal structure of the Mn2+-bound form of PsaA (PDB 3ZTT) with MMM-generated rotamer libraries highlighted, at the residue positions where cysteines will be introduced, with the measurements between those rotamer libraries and the bound Mn^{2+} ion represented by dashed lines.

Table 6.3.3: the predicted distances between the spin labels and the Mn^{2+} ion in the variants of PsaA as extracted from the crystal structure (PDB 37TT) containing the MMM-generated rotamer libraries, as calculated at 298 K.

Variant	**Predicted spin label to Mn^{2+} ion distance (Å)**
L56C	26.5
S58C	20.0
I125C	20.2
I236C	24.9
S266C	26.3

As discussed above, it is important to remember that the crystal structure is considered a rigid body by the software, and so real data may vary from the predictions.

cw-EPR at X-band

The single- and double-cysteine containing variants of PsaA were spin labelled with MTSL (details in section 13.1 'Site directed spin labelling', in accordance with protocols outlined in previous literature[45]). The labelling efficiency was calculated using a calibration curve of spin label at known concentrations (double integration of the first derivative signal gives area under the curve, which is related to the concentration of 'spin', or radical species, in a sample), and was consistently above 80% in both the single- and double-labelled variants. The room temperature cw-EPR spectra of the single- (fig. 6.3.3) and double-labelled (fig. 6.3.4) variants of PsaA were recorded at X-band.

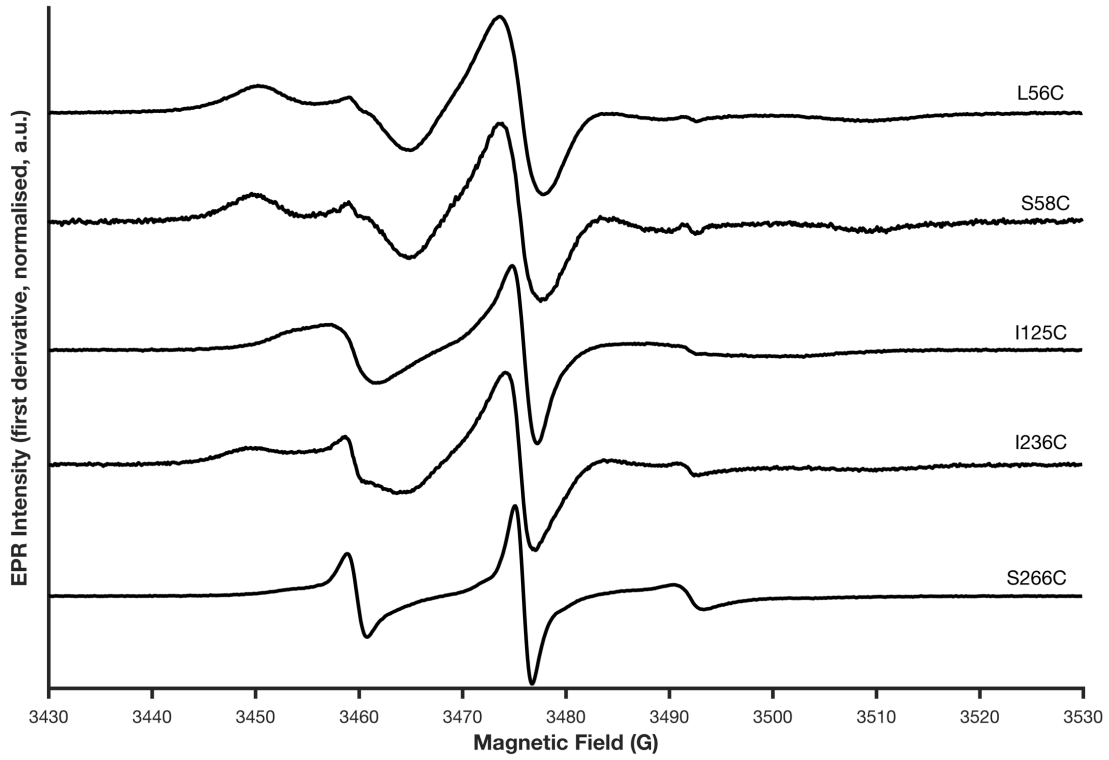


Figure 6.3.3: the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants of PsaA at X-band; from top to bottom: L56C, S58C, I125C, I236C and S266C.

The spectra of the labelled single variants of PsaA generally show low degrees of motional freedom for the attached label; the lineshapes are generally broad and asymmetric. As a rough approximation based on the appearances of the spectra, i.e. generally broad central peaks and asymmetric outer spectral features, the τ_c values that would be expected in the 1-50 ns range (see section 2.2 'Correlation times'). The S266C variant appears to be the most mobile labelling site, as the spectrum exhibits relatively sharp and symmetric peaks, though the intensity ratio of these (around 3:5:1) indicates that the label is indeed attached to a large

molecule and is not free label in solution. The L56C and S58C variants, on the other hand, have spectra that demonstrate much more restricted labelling sites, as the characteristics of central peak width and spectral breadth begin to resemble those of a frozen solution spectrum. The outer peaks also appear to demonstrate a significant degree of anisotropy. These spectra will require simulation and fitting of the experimental and simulated spectra in order to extract numerical parameters, such as rotational correlation time, τ_c, in order to fully analyse the results in a quantitative manner (see below, section 6.3 'Simulations').

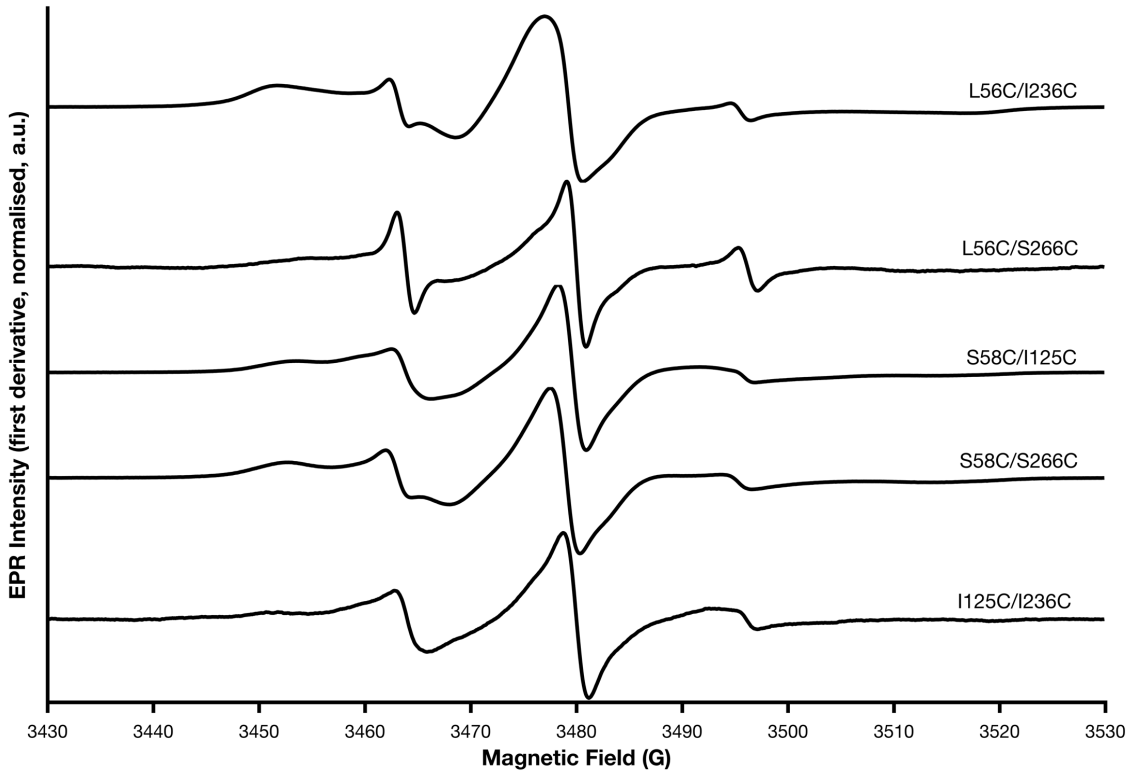


Figure 6.3.4: the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled double cysteine variants of PsaA at X-band; from top to bottom: L56C/I236C, L56C/S266C, S58C/I125C, S58C/S266C and I125C/S266C.

The spectra of the labelled double-variants exhibit lineshape characteristics visible in each of their component single-labelled variants' spectra. For example, the S58C/S266C variant spectrum displays a very broad central peak, and outer peaks that resemble those of the S58C single-labelled variant; these features appear to be overlain with some sharper signals, characteristic of the S266C single-labelled variant's spectrum.

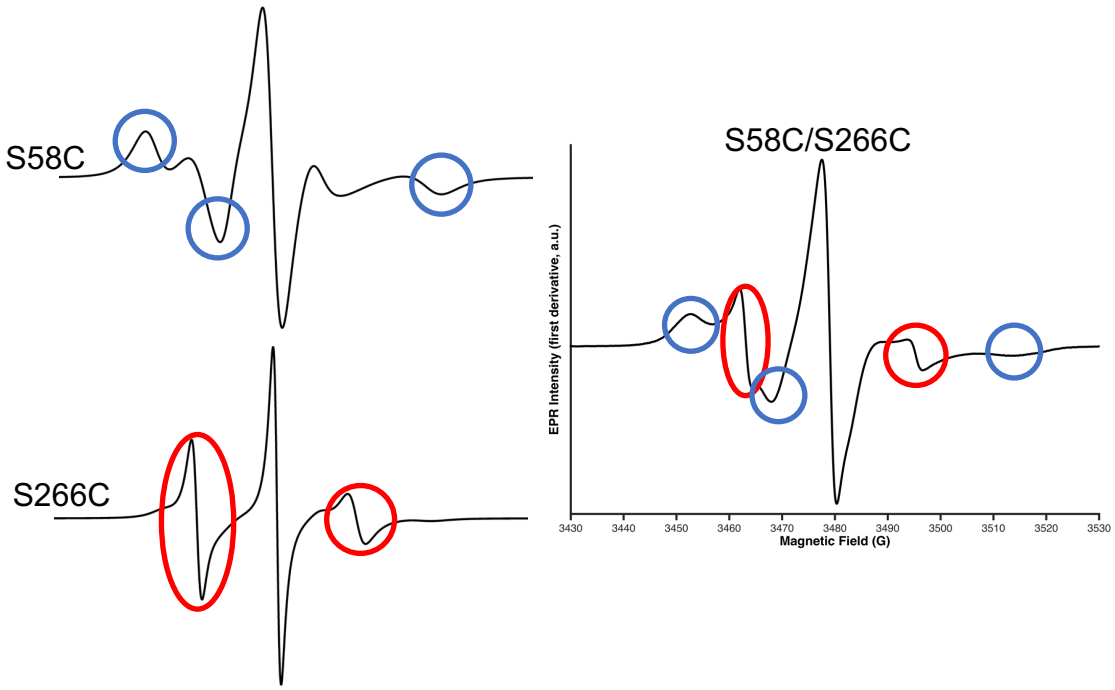


Fig. 6.3.5: the simulations of room temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants S58C and S266C of PsaA at X-band, with characteristic features circled (blue and red for S58C and S266C respectively). The same features are visible in the room temperature cw-EPR spectrum of the apoprotein, MTSL-labelled double cysteine variant S58C/S266C of PsaA at X-band and are circled in the respective colours.

Theoretically, the simulations of the single-labelled variants may be combined and fitted with the double-labelled variant experimental data, with the relative weightings varied; this will be explored in more detail in a later section (see below, section 6.3 ‘Simulations’).

The X-band cw-EPR spectra of the single-labelled variants of PsaA were also recorded at low temperature (20 K).

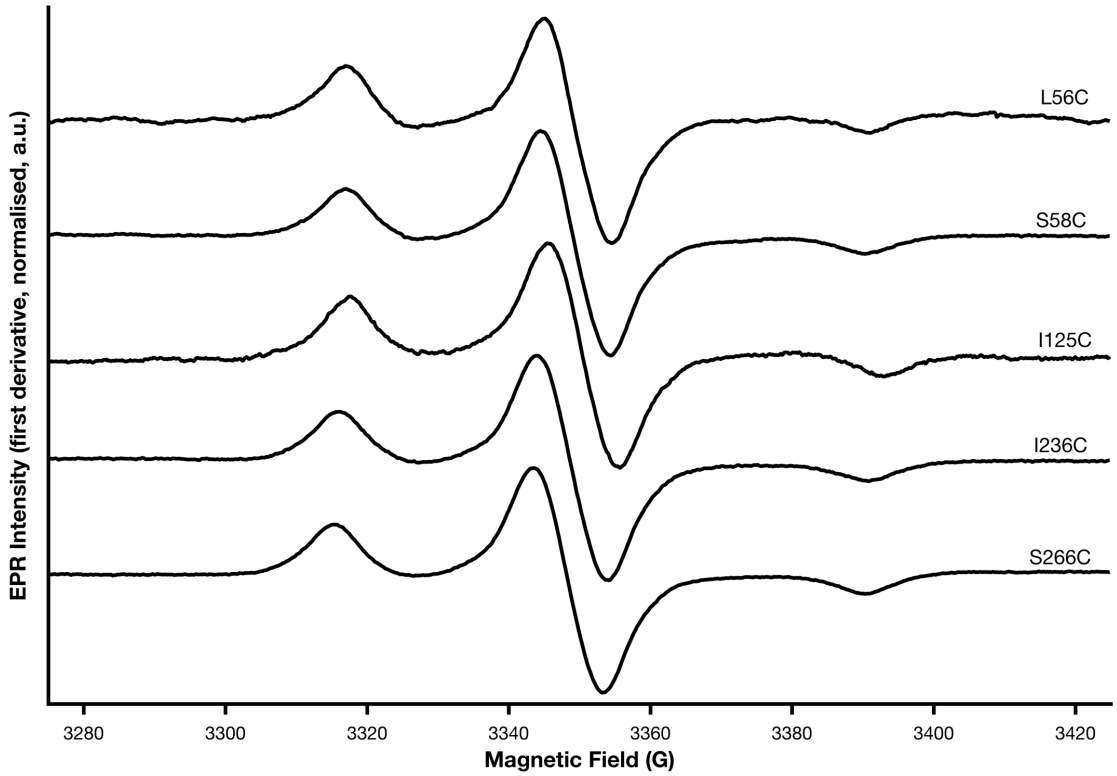


Fig. 6.3.6: the low temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants of PsaA at X-band; from top to bottom: L56C, S58C, I125C, I236C and S266C.

Obtaining the low temperature cw-EPR spectra of the double-labelled variants at X-band would identify any distances between spin labels below about 20 Å. Any evidence of short-range dipolar interaction between the two labels would be detected through broadening of the spectra. However, the predicted inter-label distances calculated from the *in silico* labelling experiments are all outside of that short distance range, in all three binding states (i.e. apoprotein, Mn^{2+}-bound and Zn^{2+}-bound; see below for distance data, section 6.3 'Pulsed EPR').

cw-EPR at Q-band

Low temperature (20 K) cw-EPR spectra of the single-labelled variants of PsaA were also recorded at Q-band.

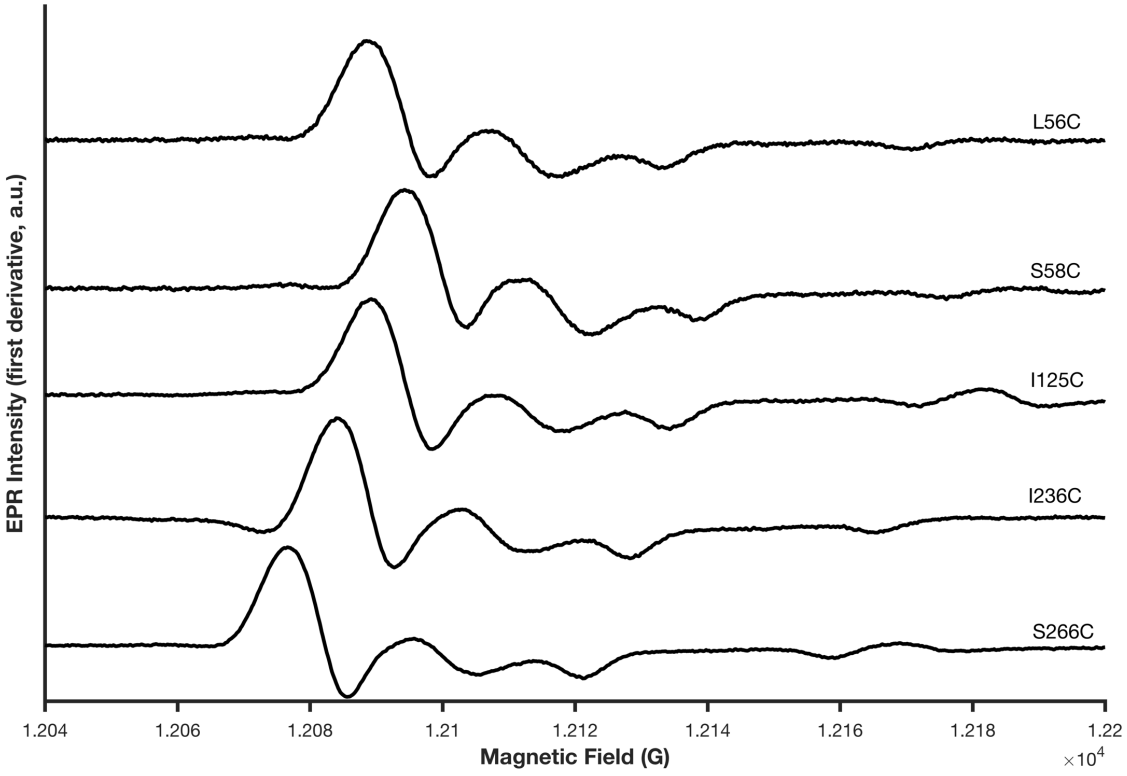


Fig. 6.3.7: the low temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants of PsaA at Q-band; from top to bottom: L56C, S58C, I125C, I236C and S266C.

The increased spectral resolution (most notably in the g-tensor) granted by measurement at higher microwave frequencies has previously been explained (see section 2.2 ‘Multifrequency EPR’). These spectra were recorded to further characterise the PsaA variants and to aid in the fitting of simulated data to experimental data, as further discussed below.

A second set of low temperature cw-EPR spectra for some of the single- and double-labelled variants of PsaA were recorded (as well as that of the wild-type PsaA, containing no spin label), following incubation with Mn(II), with parameters selected to pick out the six resonance lines characteristic of this metal (formally forbidden transitions are suppressed by the higher field – all spectra shown in fig. 6.3.9). Relatively high microwave power was used, saturating the spin label contribution to the spectra (though this has not been fully suppressed and still dominates the spectrum). Certain characteristics of the manganese signals were assessed: the width of the first peak (i.e. the difference in magnetic field between

the peak and trough of the left-most first derivative signal), and the overall spectral width (i.e. the difference in magnetic field between the left-most and right-most first derivative signals), as shown below (fig. 6.3.8).

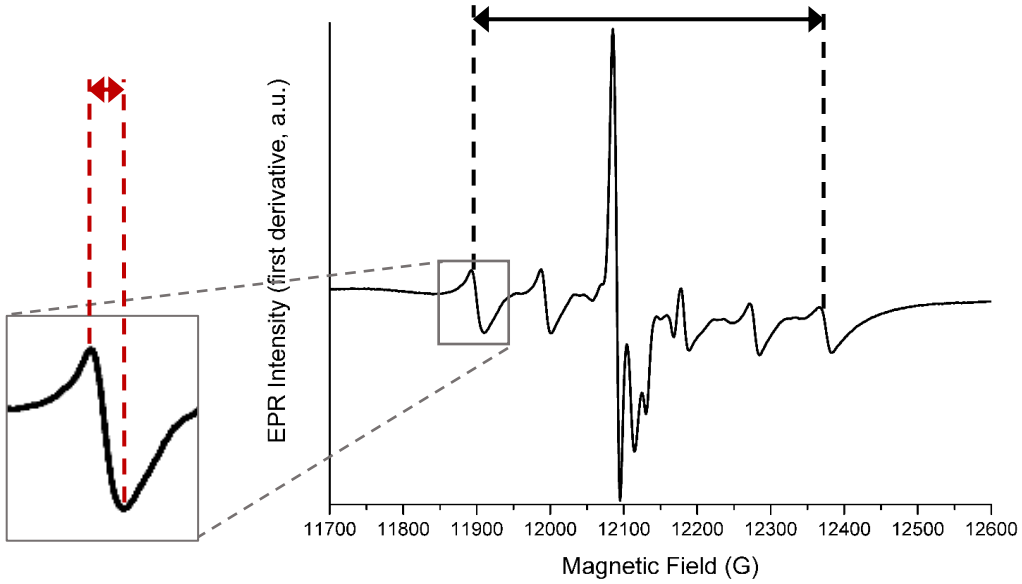


Fig. 6.3.8: an example of the low temperature cw-EPR spectrum of an Mn^{2+}-incubated, MTSL-labelled single cysteine containing variant of PsaA at Q-band, highlighting measurements of the spectral width in black and the first peak width in red.

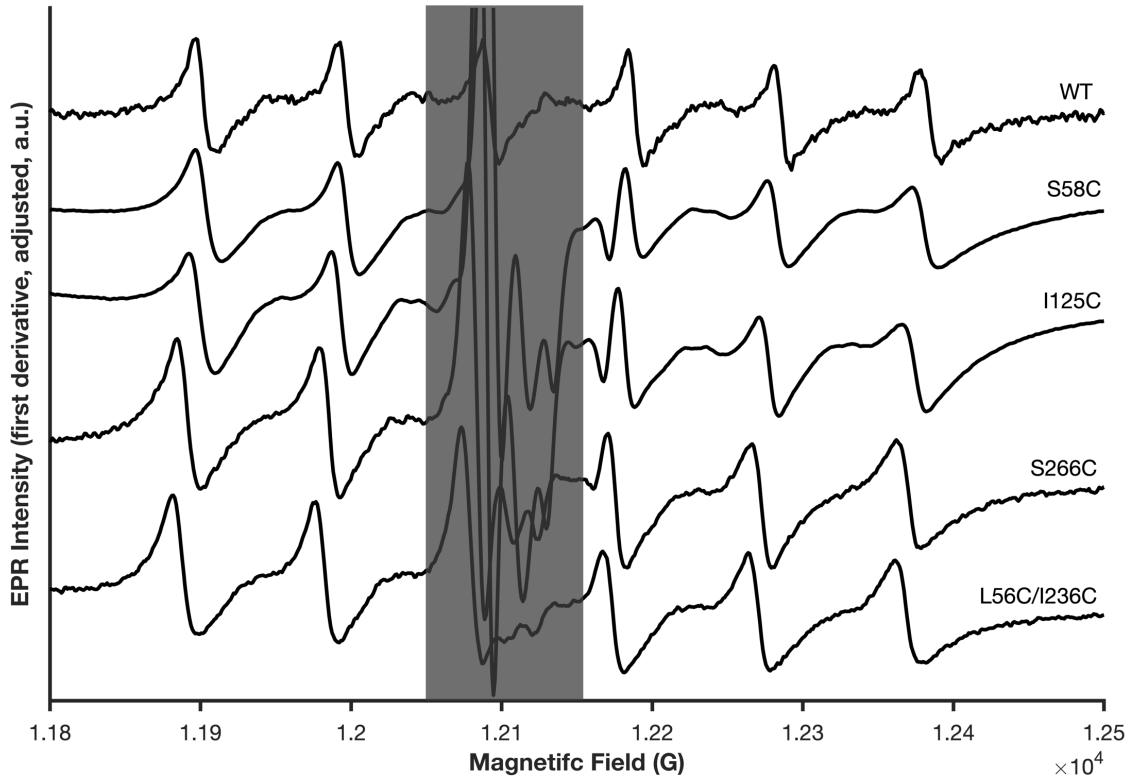


Fig. 6.3.9: the low temperature (20 K) cw-EPR spectra of Mn^{2+}-incubated variants of PsaA (from top to bottom): WT PsaA, MTSL-labelled single cysteine containing variants S58C, I125C and S266C, and MTSL-labelled double cysteine containing variant L56C/I236C at Q-band. The intensity of each spectrum has been adjusted to make the manganese-characteristic peaks roughly equal and the field position of the L56C/I236C variant has been adjusted for better overlap. The MTSL-characteristic features have been covered with a partially transparent box for clarity.

Table 6.3.4: the first peak width and overall spectral breadth measured from the low temperature (20 K) cw-EPR spectra of a selection of Mn^{2+}-incubated, MTSL-labelled single and double cysteine containing variants of PsaA at Q-band, compared to the predicted distances between the spin labels and the Mn^{2+} ion as extracted from the crystal structure (PDB 37TT) containing the MMM-generated rotamer libraries, as calculated at 298 K.

Variant	**1st line width (G)**	**Spectral width (G)**	**Distance between Mn(II) ion and modelled spin label (Å)**
WT	9.78	482.90	-
S58C	18.57	477.03	20.0
I125C	16.62	474.09	20.2
S266C	15.64	478.01	26.3
L56C/I236C	17.60	479.96	26.5/24.9

There are interesting differences in these manganese spectral characteristics for the wild-type and labelled variants; the measured width of the first resonance line is significantly broader in the spin-labelled PsaA spectra while the overall spectral width is narrower. The width of the individual spectral line of the WT here matches well with that of another protein with coordinated Mn(II) (around 10 G)[46].

Though the data is limited by the number of variants used, it could be proposed that the extent of broadening of the individual first derivative peaks is correlated with the distance between the paramagnetic species, indicating it is dipolar coupling that causes this phenomenon. The distances given above are predicted measurements using the *in silico* calculated spin label rotamer libraries modelled into the crystal structure. In general, these predictions work well with the theory that the extent of dipolar interaction is related to the broadening of the individual resonance lines. The one exception to this is the double cysteine containing variant, L56C/I236C. Both of these labels are predicted to be further from the Mn^{2+} ion than that of the single cysteine containing variant I125C, however, the first spectral peak is more significantly broadened. This may be a result of it being a

more complex case, as there are three paramagnetic species present and therefore three dipolar couplings.

The reason for the overall narrowing of the characteristic manganese sextet in the spin labelled variants, however, is less intuitive. All of the values here do agree with the well documented hyperfine coupling constants of hexacoordinate Mn^{2+} (around 95 G)[46-48]. The coordination geometry at the binding site is generally accepted to be tetrahedral, as presumed from the crystal structure data. However, a recent publication highlighted that the interatomic distances could not rule out the possibility of a hexacoordinate binding sphere[49] if the glutamate and aspartate ligands were to be bidentate.

Simulations

All simulations and fittings were performed using the EasySpin package[13] for MATLAB. The aim of fitting simulated data with experimental spectra is to extract numerical descriptions of various parameters intrinsic to the species being measured, such as the g- and A-tensors. The process here is especially iterative, given that there are multiple spectra (across temperatures and frequencies) for a single PsaA variant. One of the reasons for collecting this variety of data is to help make these simulations more accurate; the goal is to use multiple inputs of experimental spectra to converge on a consistent set of 'best fit' simulated data (table 6.3.5). Here, a set of g- and A-tensors has been fitted simultaneously to the low temperature X- (fig. 6.3.10) and Q-band (fig. 6.3.11) data for the single-labelled variants.

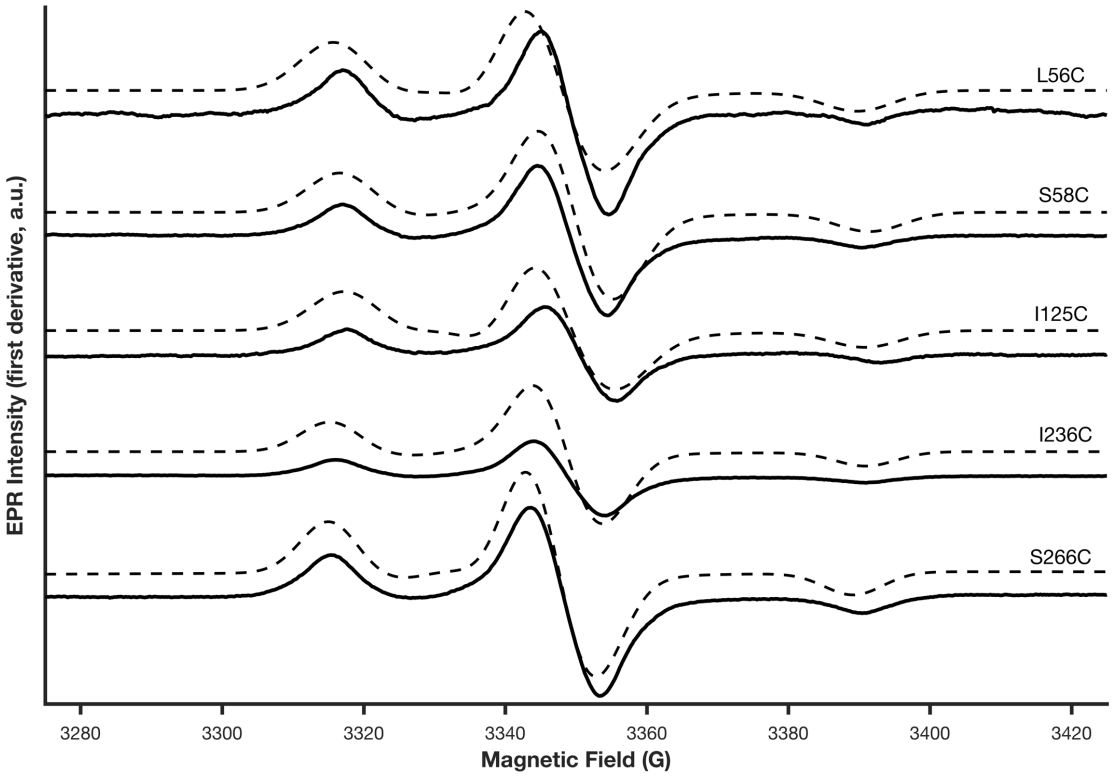


Fig. 6.3.10: the fitted simulations of the low temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants of PsaA at X-band; from top to bottom: L56C, S58C, I125C, I236C and S266C. Experimental data are shown as solid lines, fitted simulations are shown as dotted lines.

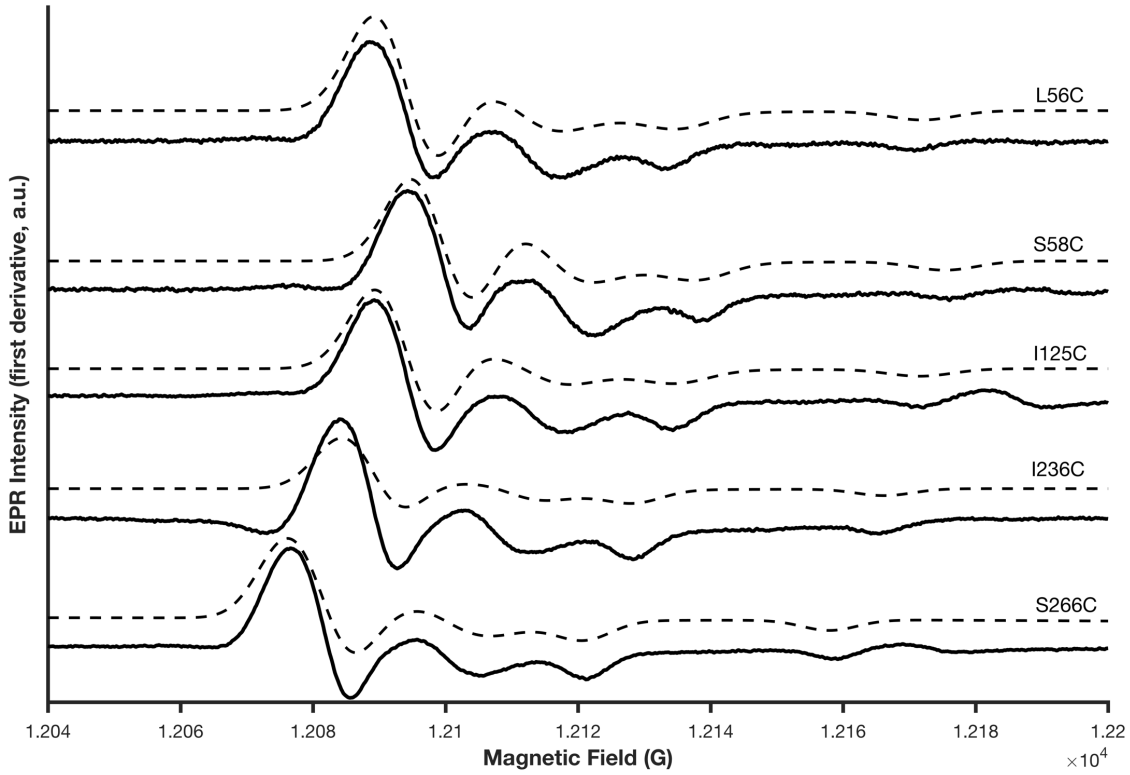


Fig. 6.3.11: the fitted simulations of the low temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants of PsaA at Q-band; from top to bottom: L56C, S58C, I125C, I236C and S266C. Experimental data are shown as solid lines, fitted simulations are shown as dotted lines.

Table 6.3.5: best fit simulation values for the x, y and z components of each of the g- and A-tensors obtained from simultaneous simulation/fitting of the concatenated low temperature X- and Q-band experimental data. The RMSD of each fitting is given.

Variant	g-tensor			A-tensor (MHz)			RMSD
	x	y	z	x	y	z	
L56C	2.0079	2.0072	2.0024	20.54	16.96	103.24	0.069368
S58C	2.0063	2.0059	2.0016	17.18	10.05	104.42	0.066775
I125C	2.0075	2.0074	2.0023	20.75	21.28	102.61	0.074424
I236C	2.0077	2.0055	2.0021	11.69	11.36	105.71	0.118646
S266C	2.0083	2.0061	2.0024	16.25	14.42	103.53	0.053947

The individual components of the g- and A-tensors are more easily observed (and therefore more easily calculated) at higher frequencies/fields; the g-values scale with applied field, whilst their associated A-values do not. Ideally, this experiment would be performed at W-band to resolve all components of the g- and A-tensors, but this is considerably more difficult more the equivalent experiment performed at Q-band, which in turn is more complex than it is at X-band. Spectroscopy is often a matter of balance; whilst a more advanced method may provide enhanced spectral resolution, this usually comes at the cost of requiring higher quality samples (higher concentrations of labelled protein) and more experimental or operational constraints. Often, the method of choice is simply that which is 'sufficient'. Low temperature cw-EPR at X-band is generally sufficient for observing g_{zz} and A_{zz}, whilst the same experiment at Q-band begins to resolve g_{yy}.

The RMSD values for most of the fits indicate fairly good compliance between the simulations and the experimental spectra. The worst of these, I236C, visibly fits less well than the others. It is unclear why this might be as the data appears to be of similar quality to the rest of the set, but the 'best fit' values are still within the range expected for an aqueous-soluble protein labelled with MTSL[14-15].

These values can then be taken as a fair starting point for simulating and fitting against the room temperature X-band data in order to extract the rotational correlation times for the single-labelled variants. (fig. 6.3.12 and table 6.3.6)

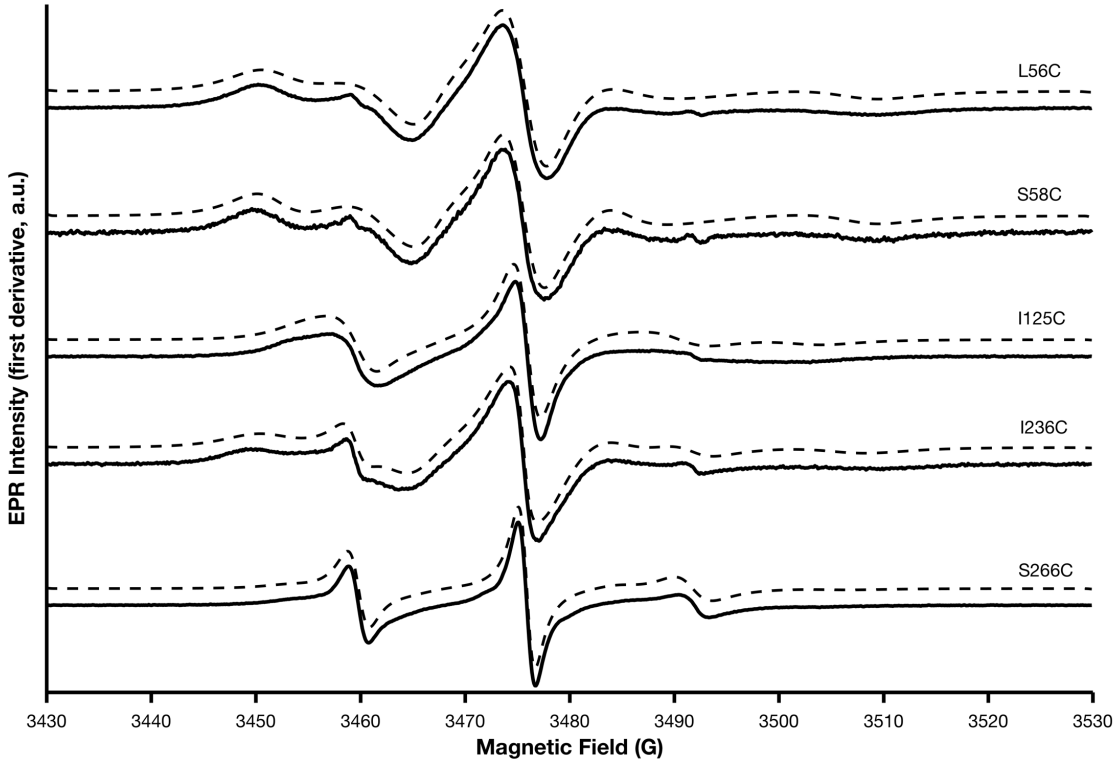


Fig. 6.3.12: the fitted simulations of the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled single cysteine variants of PsaA at X-band; from top to bottom: L56C, S58C, I125C, I236C and S266C. Experimental data are shown as solid lines, fitted simulations are shown as dotted lines.

Table 6.3.6: best fit simulation values for the x, y and z components of each of the g- and A-tensors, as well as linewidth and averaged correlation times obtained from simulation/fitting of the room temperature X-band experimental data of the MTSL-labelled single variants of PsaA. Some of these spectra contained two separate ‘ species’ of spin label; the values of the parameters extracted for each of these is provided, with their respective, relative weightings. The RMSD of each fitting is given.

Variant	Component	g-tensor			A-tensor (MHz)			Linewidth (peak-to-peak, mT)	τ_c (ns)	Component weighting	Simulation RMSD
		x	y	z	x	y	z				
L56C		2.0071	2.006	2.0021	12	27	96	0.01	36.4	-	0.030387
S58C		2.0069	2.0061	2.0020	12	27	96	0.01	34.5	-	0.030369
I125C	1	2.0073	2.0043	2.0023	14	16	102	0.01	25.9	0.64	0.033915
	2	2.0073	2.0043	2.0023	14	15	108	0.01	2.3	0.36	
I236C	1	2.0073	2.005	2.0025	19	22	94	0.01	35.6	0.88	0.014252
	2	2.0073	2.005	2.0025	22	28	84	0.01	3.3	0.12	
S266C	1	2.0052	2.005	2.0024	19	26	91	0.01	2.8	0.65	0.010366
	2	2.006	2.005	2.0013	22	26	87	0.01	9	0.35	

The most notable difference between the results of these simulations/fittings and those reported above for the multifrequency low temperature cw-EPR data is the A_{zz} values. The g- and A-values for room temperature solution EPR are usually reported as g_{iso} and A_{iso}, i.e. an average of the three cartesian components of the respective tensors. However, this fitting method, which allows for the τ_c to be defined, requires the input of anisotropic g- and A-tensors. However, the regime of motion being measured here is not sufficiently slow to cause anisotropy to the extent needed to resolve any individual components of these tensors, and therefore the accuracy of these reported individual 'best fit' values should be considered cautiously; indeed, the results for these values from the simultaneous multifrequency simulation/fitting above are far more accurate.

The main goal of the simulation and fitting of the room temperature cw-EPR spectra was to extract the values for τ_c. These vary quite dramatically from the previously reported values[45].

Table 6.3.7: comparison of the reported rotational correlation time, τ_c, for the apoprotein, MTSL-labelled single cysteine variants of PsaA as extracted from their room temperature cw-EPR spectra at X-band. Some of these spectra contained two separate 'species' of spin label; the τ_c extracted for each of these is provided, with their respective, relative weightings.

[a] – data reported in publication by Deplazes et al., 2015[45]

		Literature[a]		**Simulation/fitting**	
Variant	**Component**	**τ_c (ns)**	**Relative weighting**	**τ_c (ns)**	**Relative weighting**
L56C		7.6	-	36.4	-
S58C		9.3	-	34.5	-
I125C	1	2.7	-	25.9	0.64
	2	-	-	2.3	0.36
I236C	1	7.2	0.87	35.6	0.88
	2	1.2	0.13	3.3	0.12
S266C	1	0.97	-	2.8	0.65
	2	-	-	9.0	0.35

Only approximate input values, and no output values other than τ_c are provided for the simulations/fittings reported in the publication (which were also performed using EasySpin); it is therefore not possible to fully replicate these simulations and attempt to fit them to the experimental data collected here. Additionally, as there is no RMSD assessment of the fittings it is difficult to assess the accuracies of these results in comparison with those detailed here.

Having obtained the best fitted simulations of the single variants, it was then attempted to combine them into their respective pairs to allow the simulating and fitting the double-labelled variant spectra. However, this approach encountered multiple problems. Firstly, no way was found in the EasySpin simulation/fitting routine to take two separate systems, that are themselves composed of various different label species, and then independently fit them whilst keeping their component parts consistent. Regardless of this problem, it is then computationally difficult to simulate a system of 3 or more component parts and fit all of their various parameters to the experimental data; even to calculate the g-tensors (3 input values each) would involve fitting 12-16 variables per spectrum. Considering this, and the wealth of information extracted from the single cysteine variants above, as well as the demonstrable contributions of these individually labelled sites to the experimental spectra of their respective double cysteine variants, full computational simulation and fitting in the same manner is not essential for the aims of this investigation. Therefore, in order to generate adequate simulations for the double cysteine variants, no fitting was performed. The fitted simulations of the single variants were added together, assuming equal labelling of each cysteine position, and normalised to the scale of the experimental data.

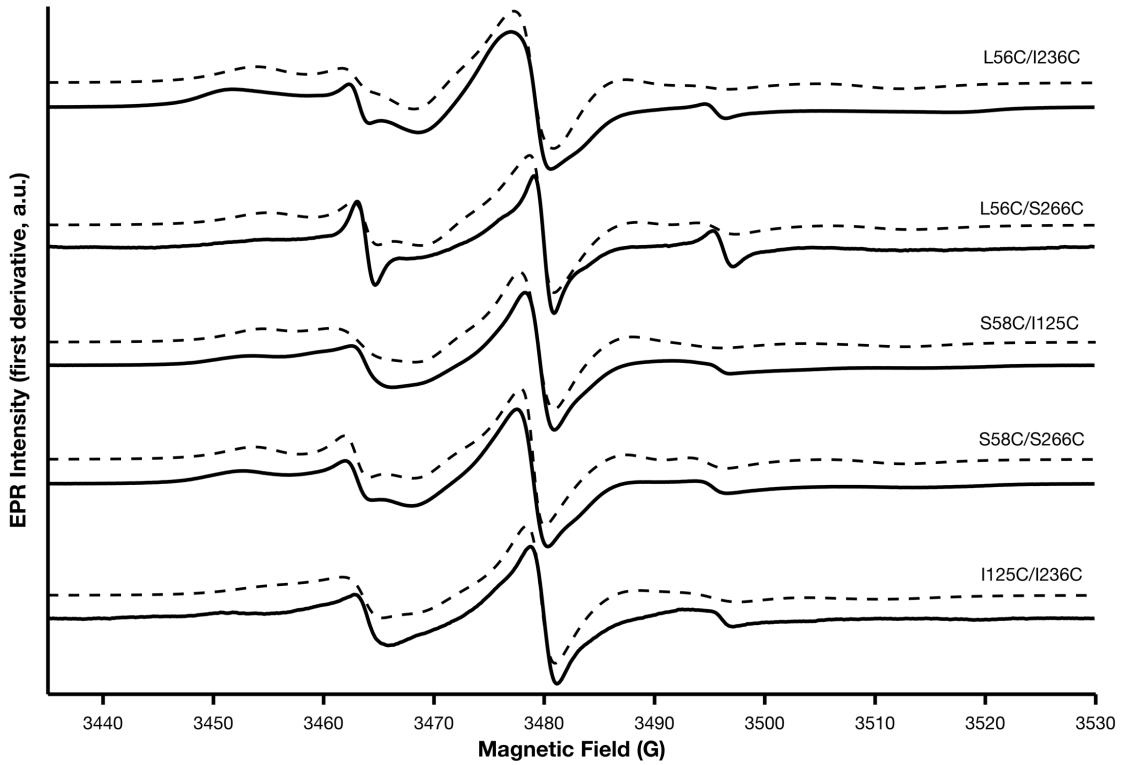


Fig. 6.3.13: the fitted simulations of the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled double cysteine variants of PsaA at X-band; from top to bottom: L56C/I236C, L56C/S266C, S58C/I125C, S58C/S266C and I125C/I236C. Experimental data are shown as solid lines, simulations are shown as dotted lines.

There is no RMSD assessment of the quality of the simulations; instead, the 'fit' can only be visually appraised. For the most part, these simulations appear to be good representations of the experimental data they are attempting to reproduce. In the cases where the alignment of the experimental and simulated data is not as good, it is likely that the experimental spectrum represents a sample in which the two labelling sites are not equally labelled. Alternatively, where two very different environments are being labelled (i.e. L56C/S266C; the labelled L56C variant has a very broad/asymmetric spectrum and a τ_c of 36 ns, whereas the labelled S266C variant's spectrum is relatively sharp/symmetric and a dominant τ_c of 3 ns) there may be an imbalance in the spectrum's appearance; i.e. the component represented by sharper, narrower peaks seems to dominate the spectrum visually, though this is not necessarily an accurate representation of the real label populations in the sample.

Pulsed EPR

The PELDOR data for four of the six double-labelled variants under all conditions (apo, Mn^{2+}-bound and Zn^{2+}-bound) has been collected. These data have been processed using the DeerAnalysis package[18] that runs on MATLAB. The data was largely collected prior to my involvement in the project by Dr Jessica van Wonderen, UEA; any outstanding or incomplete datasets were completed, and all datasets processed and analysed by me.

In order to simplify the range of distance distribution data presented herein, the doubly-labelled variants can be divided into two categories based on the domain movements they are being used to assess: lobe-lobe (i.e. the cysteine mutation sites each sit on a separate lobe domain: L56C/I236C, L56C/S266C, and S58C/S266C) and loop-lobe (i.e. one of the cysteine mutation sites sits on a lobe domain and the other on the loop: S58C/I125C, and I125C/I236C).

The movement of the two lobe domains toward and away from one another is the typical conformational change observed in the Venus flytrap mechanism for SBPs. In the context of the PELDOR spectroscopic measurements, this would be expected to be relatively simple to demonstrate, with an open form of lower dipolar frequency (where spin labels are separated by a greater distance), and a closed form of higher dipolar frequency (where the spin labels are closer together). As presented by the PsaA variants being studied here, it is not always so simple.

The first of the lobe-lobe variants, L56C/I236C, shows a multimodal distance distribution in the apoprotein form (fig. 6.3.15). These modes, at 36.9, 44.7 and 52.7 Å, have very different populations, with the peak at 36.9 Å dominating the profile; the bands of confidence would indicate that this distance is the most accurately calculated from the time trace data and the longer distance peaks may be inaccurate (fig. 6.3.14). However, in the Mn^{2+}-bound form, there is a single distance population at 32.9 Å, with very small features at 42.4 and 53.0 Å, which are most likely artefacts as result of the noise (both are classified as low-confidence calculations); the measured dipolar evolution time is not sufficient to observe oscillations at frequencies that would result from such long inter-label distances. The dominant distance of the Mn^{2+}-bound state is replicated in the Zn^{2+}-bound form of L56C/I236C (at 33.1 Å), but similar to the apoprotein form, there are smaller modes at longer distances (40.4 and 48.2 Å). The raw time traces of the apoprotein and Zn^{2+}-bound forms are relatively complex; there is a very long

decay profile with shallow oscillations. This indicates that there is more than one frequency of dipolar coupling present, i.e. the less populated modes observed in their respective distance distributions are unlikely to be artefacts. These separate populations may arise either from distinct conformational states, perhaps that the protein is sampling in an infrequent, transient manner, or from different orientations of the labels relative to one another whilst the protein conformation remains ‘static’.

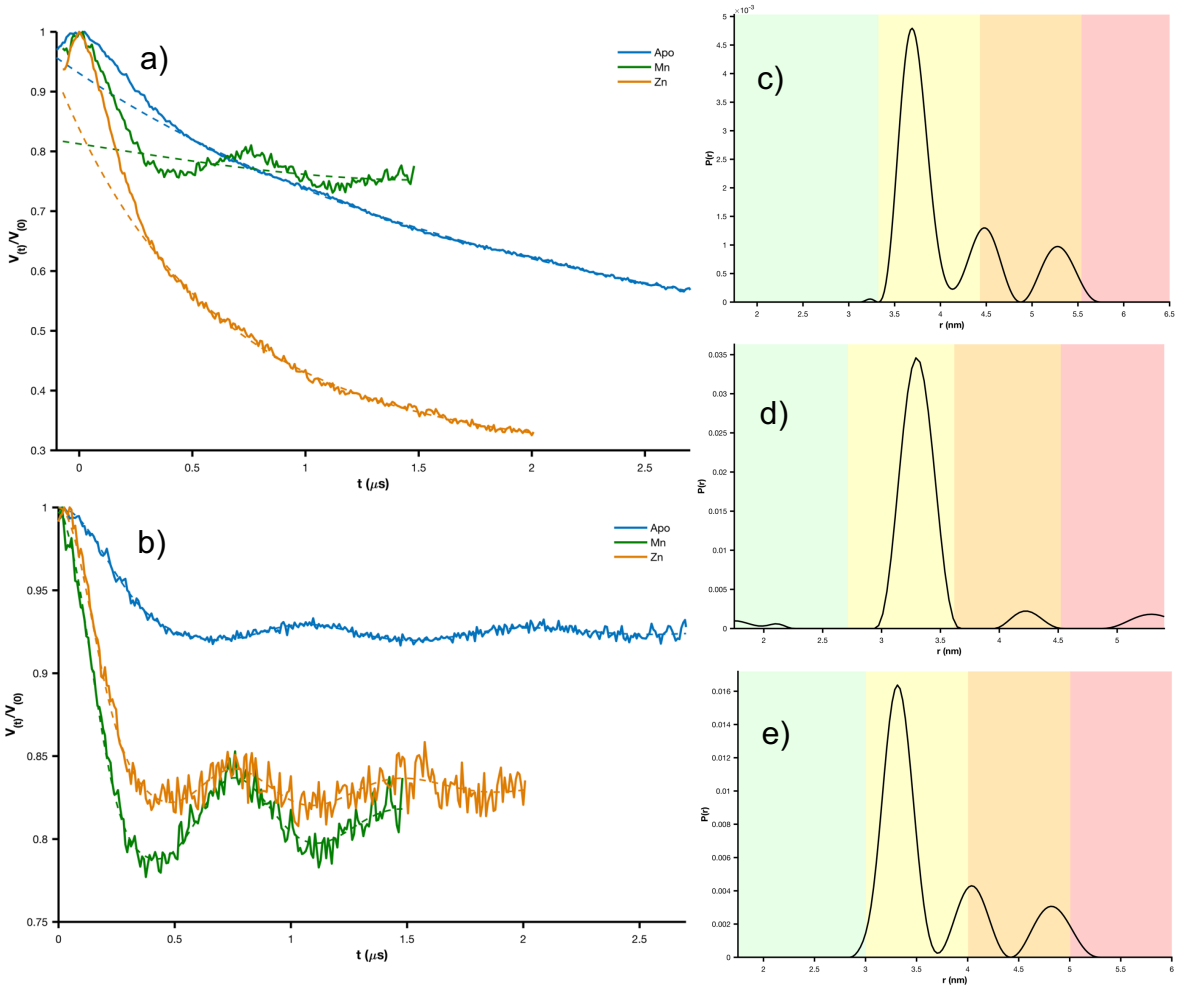


Fig. 6.3.14: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled L56C/I236C variant of PsaA, a) ‘raw’ data (trace of the echo decay and oscillation in the time domain) of apoprotein form, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange, b) background factor subtracted time trace (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange), c) apoprotein distance distribution output, d) Mn^{2+}-bound distance distribution output, e) Zn^{2+}-bound distance distribution output.

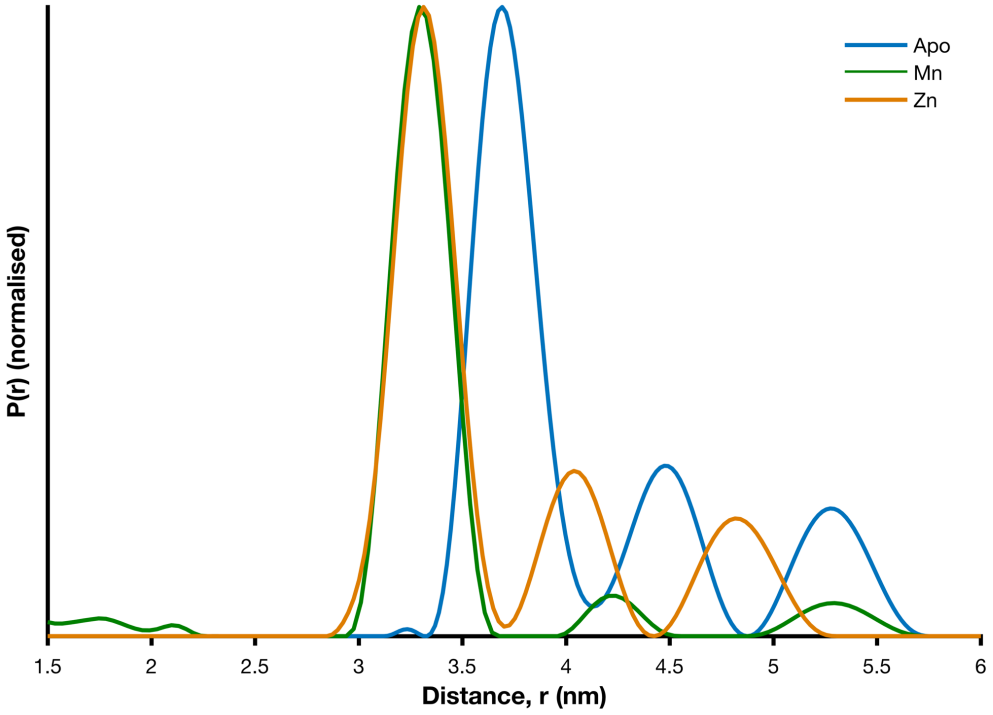


Fig. 6.3.15: scale-normalised overlay of the distance distributions resulting from 4P PELDOR of the MTSL-labelled L56C/I236C variant of PsaA (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange).

The next lobe-lobe variant, L56C/S266C, shows very little change in the inter-label distance measurements between the different binding states (fig. 6.3.16). In the apoprotein form, the distance distribution has a single mode, peaking at 35.5 Å. This feature, however, is broad, indicating that there may be a range of protein conformations being represented here. The Mn^{2+}-bound form has a much narrower peak at 34.7 Å. Though this is a small change when looking solely at the distances, the difference in distributions suggests a more notable difference in the conformations of these two states. The Zn^{2+}-bound form yields a distance distribution that resembles a hybrid of those of the apoprotein and Mn^{2+}-bound states; the peak is at 34.1 Å and is broader than the feature in the Mn^{2+}-bound distribution, but narrower than that in the apoprotein distribution (fig. 6.3.17).

Looking at the raw time trace data, and then the background corrected time traces, all three forms have relatively similar profiles (fig. 6.3.16). The main difference in shape is their decays; the apoprotein time trace decays more slowly than the other forms, and background correction results in a less obvious oscillation, which is represented by the broadness of the modal feature in the distance distribution. This is often observed (but not always) in cases of more mobile spin label environments. It has already been shown by room temperature cw-EPR that the

S266C labelling position is fairly unrestricted (the majority of the label population has a τ_c of around 3 ns), and the room temperature cw-EPR spectrum of the labelled, apoprotein L56C/S266C variant reflects this. Since there is no equivalent data as yet for the metal-bound states of these PsaA variants, it is not possible to say for certain that this is the reason for the broader distance distributions. Other than this, the modulation depths are very different; this may be caused by differences in sample concentration (as modulation depth is related to 'spin' concentration).

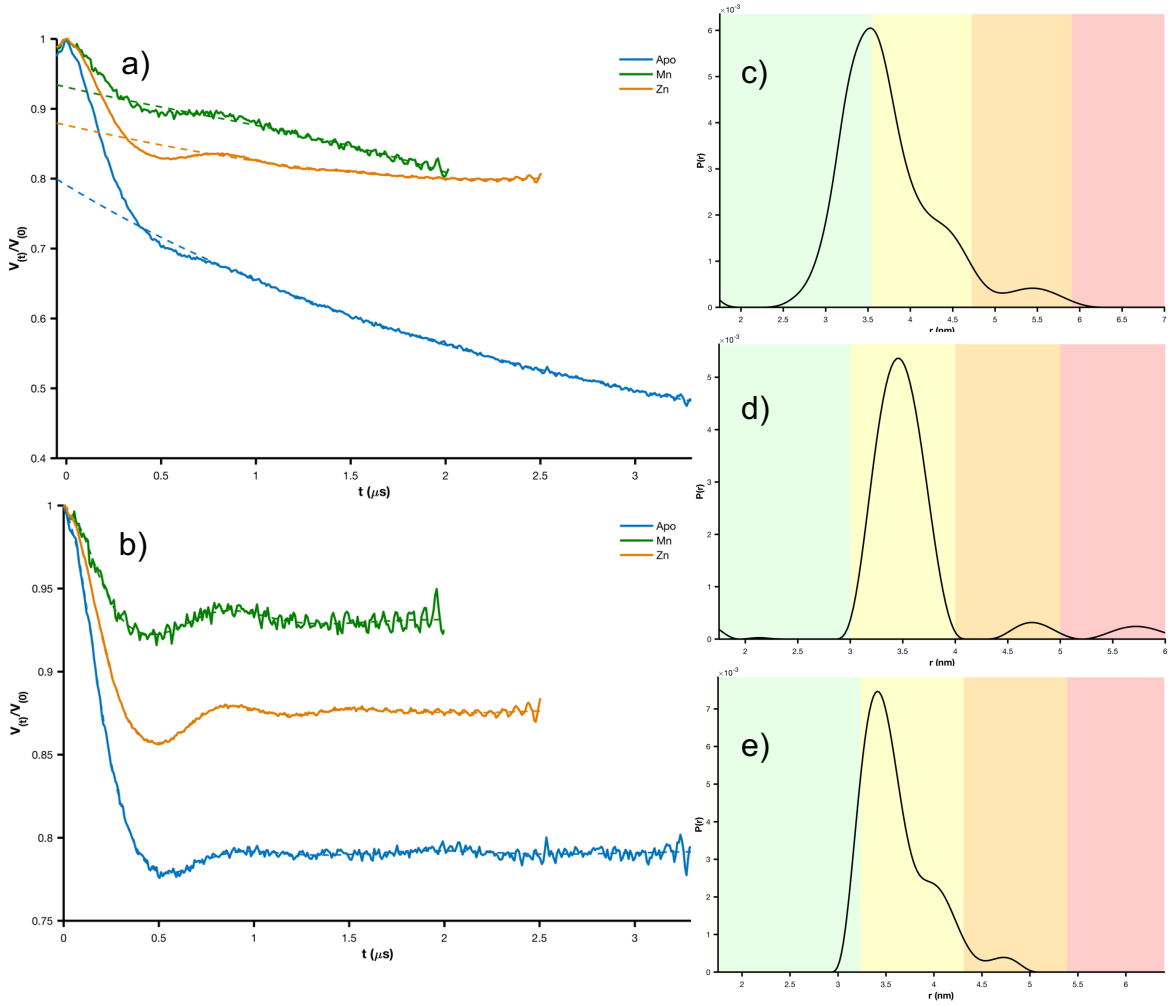


Fig. 6.3.16: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled L56C/S266C variant of PsaA, a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange, b) background factor subtracted time trace (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange), c) apoprotein distance distribution output, d) Mn^{2+}-bound distance distribution output, e) Zn^{2+}-bound distance distribution output.

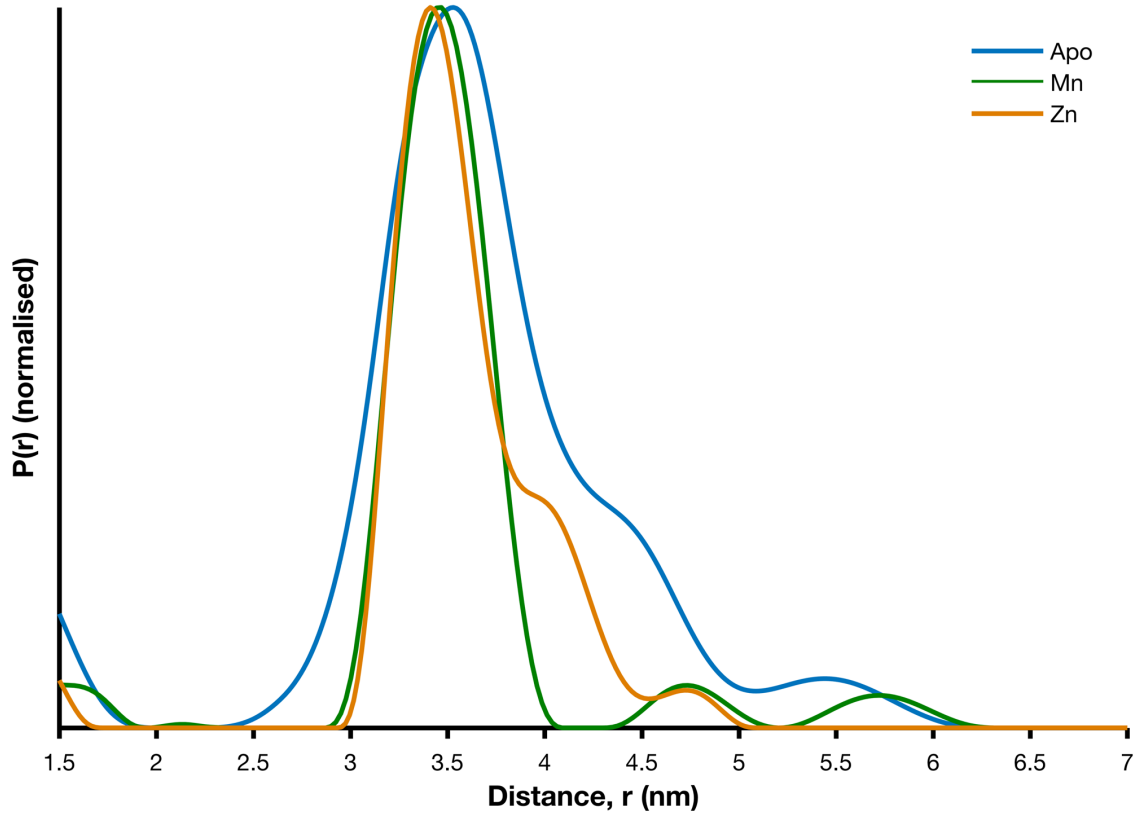


Fig. 6.3.17: scale-normalised overlay of the distance distributions resulting from 4P PELDOR of the MTSL-labelled L56C/S266C variant of PsaA (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange).

The final lobe-lobe variant is S58C/S266C; again, there is little change in the predominant distances measured (fig. 6.3.18), but more information can be gathered from the relative populations and shapes of the distributions (fig. 6.3.19). The apoprotein distribution shows a broad overlap of two populations: one large feature peaking at 36.6 Å and a smaller one at 46.5 Å. Similar to the L56C/S266C variant above, the Mn^{2+}-bound state exhibits a single, narrower peak at 35.7 Å. The Zn^{2+}-bound form also presents a dominant feature at 35.7 Å, which is broader in shape than that of the Mn^{2+}-bound form, and also a small population at 23.2 Å.

Referring to the raw data (fig. 6.3.18), as with the previous variant, L56C/S266C, the apoprotein time trace has a long decay profile which, when background corrected, gives shallow oscillations. Towards the end of the measured signal, the noise increases, potentially masking another oscillation frequency, and causing uncertainty in the processing routine. This may be the source of the smaller feature observed in the apoprotein's distance distribution, which falls within a low confidence band. The similarity between the data collected for the apoprotein form of the two variants L56C/S266C and S58C/S266C may lend some credence to the posited relationship between label mobility and dipolar oscillation profiles. The

background correction of the time trace data for the Zn^{2+}-bound form enhances the noise, and this is likely the source of the short distance seen in the distribution.

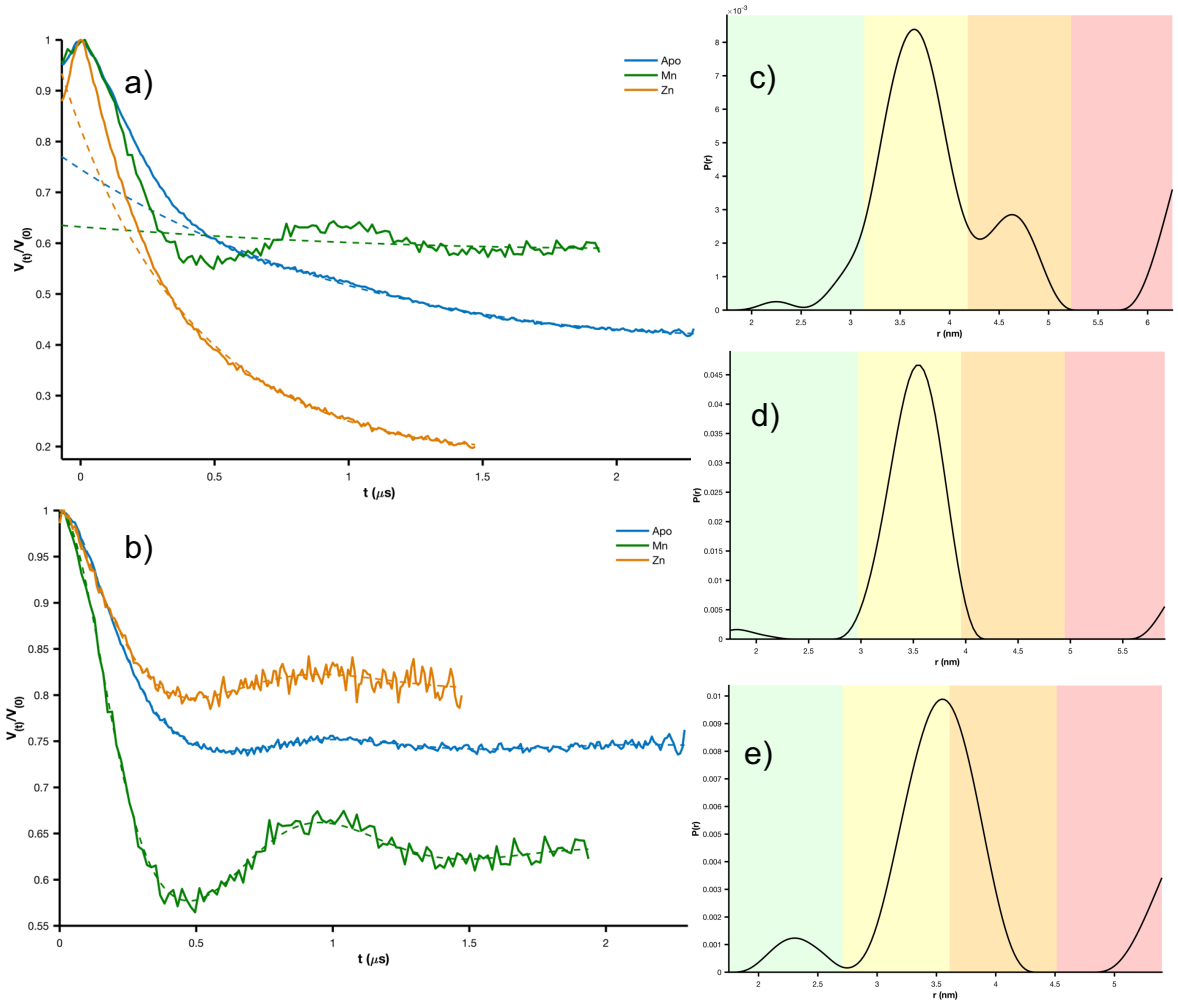


Fig. 6.3.18: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled S58C/S266C variant of PsaA, a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange, b) background factor subtracted time trace (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange), c) apoprotein distance distribution output, d) Mn^{2+}-bound distance distribution output, e) Zn^{2+}-bound distance distribution output.

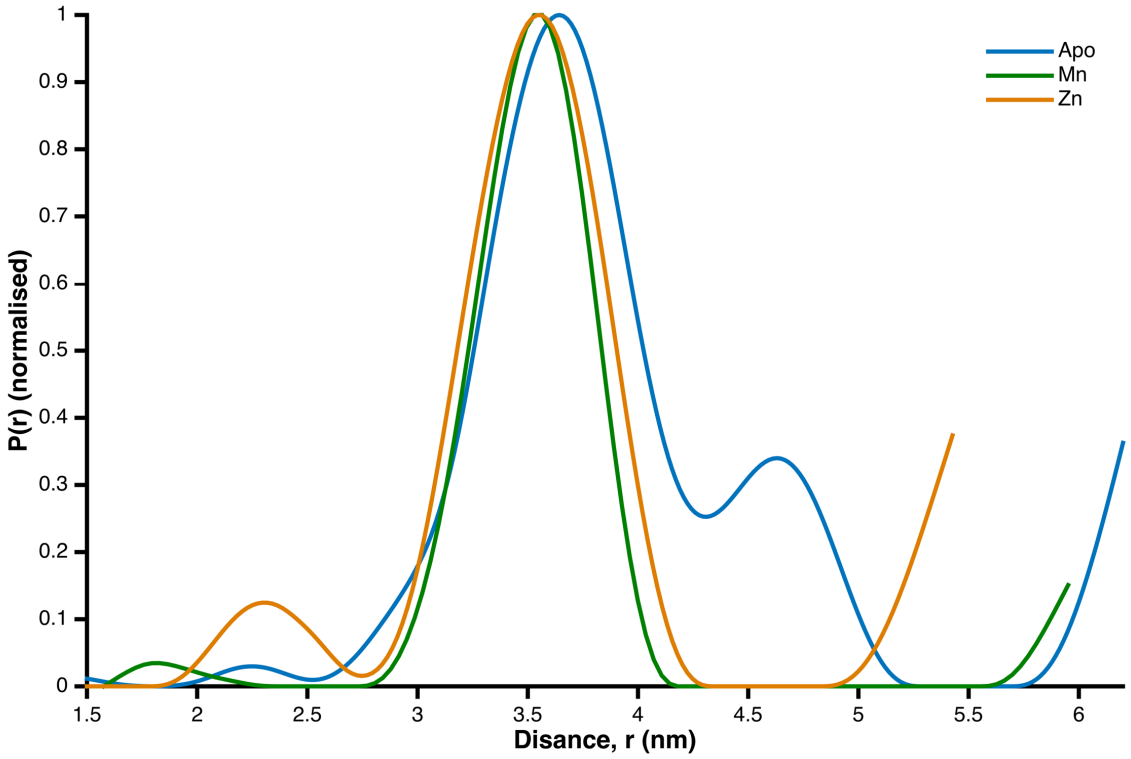


Fig. 6.3.19: scale-normalised overlay of the distance distributions resulting from 4P PELDOR of the MTSL-labelled S58C/S266C variant of PsaA (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange).

The first loop-lobe variant is S58C/I125C. All of the different binding states of this variant show a predominant feature at the same distance of around 33.3-34.0 Å (only a 0.7 Å deviation; fig. 6.3.21). In the apoprotein form, this feature, at 34.0 Å, is the only population (there is a small, low confidence peak at 49.4 Å which should be disregarded). This is interesting considering how flexible loop regions are expected to be, as it might be anticipated to observe more modes, or broader features, in the distance distribution. In the Mn^{2+}-bound state, there are more small features at shorter distances: 21.3 and 28.4 Å. There is also a small, shorter distance feature in the distance distribution of the Zn^{2+}-bound form, at 23.5 Å. This is smaller and broader than the features in the Mn^{2+}-bound state, which may indicate that while there are conformations here similar to those in the Mn^{2+}-bound state, they are less distinct and less frequently sampled.

The short distances in the Mn^{2+}-bound form may be artefacts due to the noise of the data (fig. 6.3.20), but they do appear within the high confidence bands. However, the profile of the this data would suggest that the shortest distance is not to be believed; such close proximity of two spin labels (21.3 Å) would be evident in a very sharp feature at the start of the trace, and would also probably yield very noticeable oscillations across the measured dipolar evolution time. The influence

of the noise in this time trace is heightened by the fact that the modulation depth is so small. This may be caused by a difference in sample concentration or may be the effect of the paramagnetic Mn^{2+} ion enhancing the relaxation of the nitroxide spins. In order for this to occur, the spin labels would have to be in relatively close proximity to the Mn^{2+} ion; in the case of S58C/I125C here, both are predicted to be about 20 Å away (see fig. 6.3.2 and table 6.3.3), which agrees with this hypothesis.

Similarly, background corrected time trace of the Zn^{2+}-bound form is fairly noisy (fig. 6.3.20), which may have led to the presence of a short distance population in the distribution. If these populations are to be believed, they would suggest a greater degree of flexibility or range of conformational positions of the loop in the bound states, but they may be artefacts introduced by noise and processing routine. This is probably the case, considering the similarity of the background corrected time domain data of the Zn^{2+}-bound and apoprotein forms.

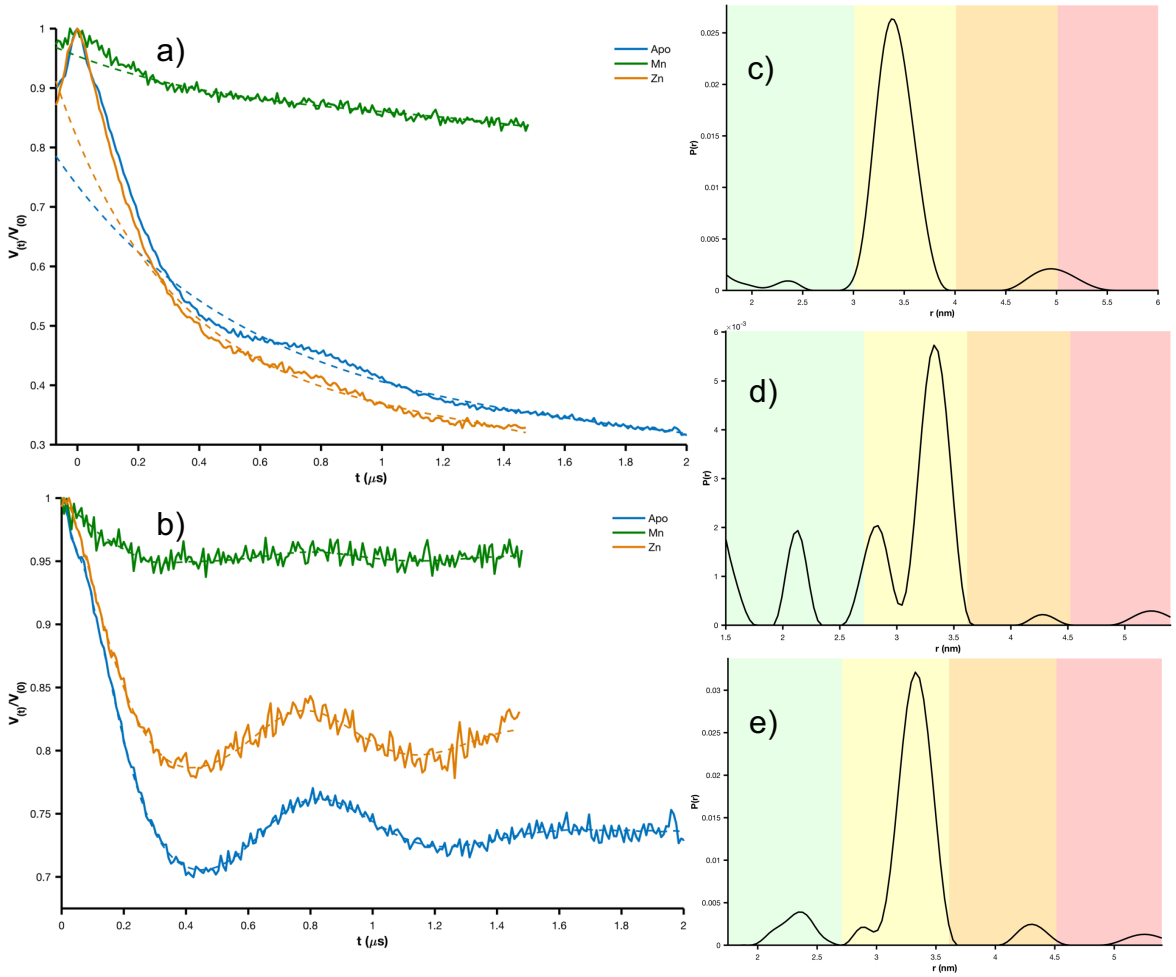


Fig. 6.3.20: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled S58C/I125C variant of PsaA, a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange, b) background factor subtracted time trace (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange), c) apoprotein distance distribution output, d) Mn^{2+}-bound distance distribution output, e) Zn^{2+}-bound distance distribution output.

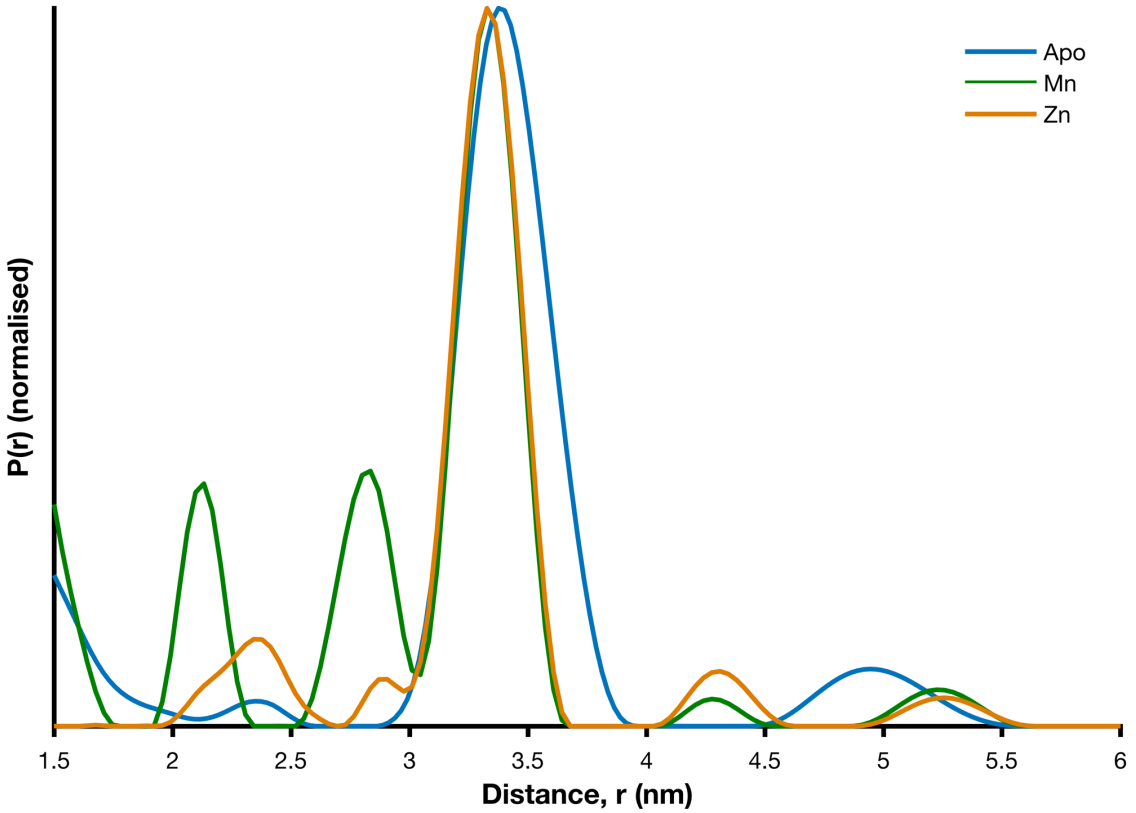


Fig. 6.3.21: scale-normalised overlay of the distance distributions resulting from 4P PELDOR of the MTSL-labelled S58C/I125C variant of PsaA (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange).

The other loop-lobe variant, I125C/I236C, shows a range of distances in each of the binding states, but a predominant mode peaking at around 29.2-29.8 Å is observed in all of them. In the apoprotein form, this feature (at 29.2 Å) is fairly even in population with a peak at 35.2 Å, with smaller populations at 42.7 and 51.8 Å, which can be disregarded; the former has a negligible population and the latter is very low confidence. In the Mn^{2+}-bound form, there is also a feature at 34.3 Å, but it is significantly smaller than the predominant peak at 29.4 Å. This state also has small populations at 23.2 and 43.5 Å, and one below 20 Å which would need to be verified by low temperature cw-EPR. The Zn^{2+}-bound state gives a multimodal distance distribution, with the dominant feature at 29.8 Å, and smaller, fairly even peaks at 35.7 and 43.5 Å; however, the longer distance peak falls into the low confidence band, and measured dipolar evolution time would not be sufficient to extract such a long distance.

The modulation depth observed for the Mn^{2+}-bound form is very low, similar to the S58C/I125C variant discussed above. Again, this may have an effect on the data processing by enhancing the influence of noise, giving artefacts in the short distance areas of the distribution.

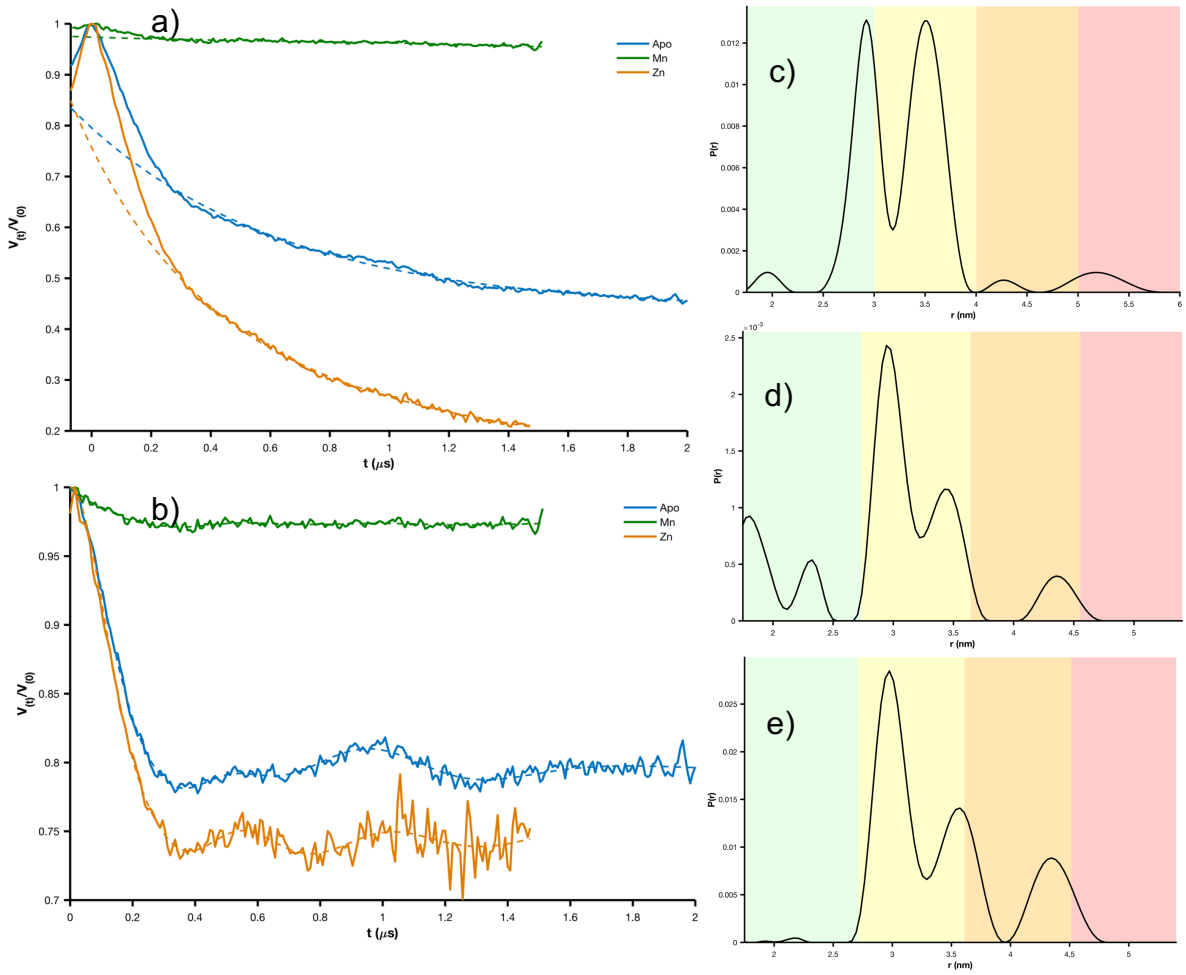


Fig. 6.3.22: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled I125C/I236C variant of PsaA, a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange, b) background factor subtracted time trace (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange), c) apoprotein distance distribution output, d) Mn^{2+}-bound distance distribution output, e) Zn^{2+}-bound distance distribution output.

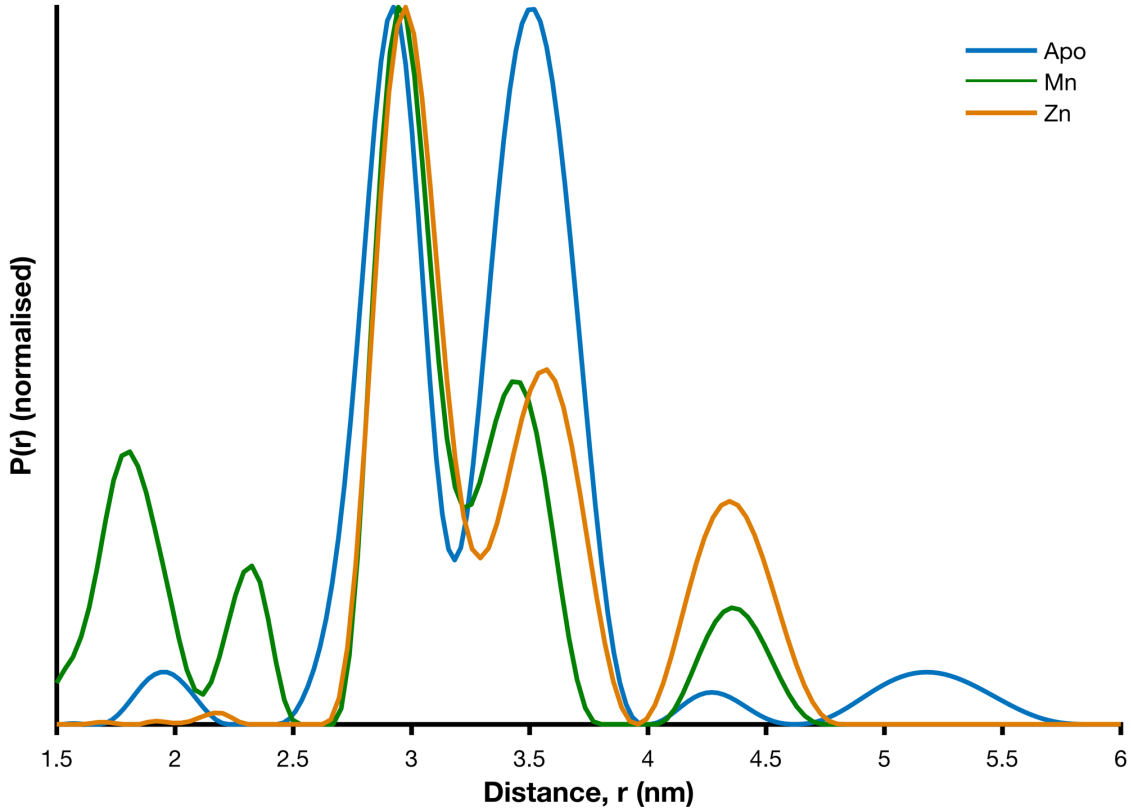


Fig. 6.3.23: scale-normalised overlay of the distance distributions resulting from 4P PELDOR of the MTSL-labelled I125C/I236C variant of PsaA (apoprotein, blue, Mn^{2+}-bound, green, and Zn^{2+}-bound, orange).

Table 6.3.8: the measured distances between the pairs of spin labels in the double-labelled variants of PsaA as extracted from the distance distributions obtained from the PELDOR experiments. In cases where there are multiple modes in the distance distributions, the predominant value is shown in bold.

Variant	**Measured inter-label distance (Å)**		
	Apo	**Mn^{2+}-bound**	**Zn^{2+}-bound**
L56C/I236C	**36.9**, 44.7	32.9	**33.1**, 40.4
L56C/S266C	35.5 (broad)	34.7	34.1 (broad)
S58C/S266C	**36.6**, 46.5	35.7	35.7
S58C/I125C	34.0	28.4, **33.3**	33.3
I125C/I236C	29.2, 35.2 (even)	**29.4**, 34.3	**29.8**, 35.7

From this data (table 6.3.8), a model can be built of the conformational dynamics of PsaA in its different binding states. In the apoprotein form, the distances labelling positions on the lobes indicate that they rest open, with some degree of flexibility that allows multiple or broad populations to occur in the distance distributions; this flexibility is also suggested in the work that used room temperature cw-EPR at X-band and MD simulations[45]. This flexibility is suggested to be asymmetric, i.e. one lobe moves more than the other. This is supported by the loop-lobe distance data; the S58C/I125C apoprotein distribution is unimodal, whilst that of the I125C/I236C is bimodal, with fairly even distribution between the two populations. Therefore, the loop and N-terminal lobe (containing positions L56 and S58) likely occupy a single conformation in this binding state, whilst the C-terminal lobe (I236 and S266) has some range of conformations.

When Mn^{2+} binds, the lobe-lobe distances all display unimodal distance distributions, with shorter distances populated than in the apoprotein state. However, the difference in distances between the apoprotein and Mn^{2+}-bound are not consistent: L56C/I236C changes by 4.0 Å, L56C/S266C by 0.8 Å and S58C/S266C by 0.9 Å. Both of the S266C-containing variants show smaller distance change between apoprotein and Mn^{2+}-bound forms than that containing the I236C mutation site. This can be interpreted as twisting or rotating of the lobes during closure, leading to a greater distance change between N- and C-terminal lobes on one side than the other.

In both loop-lobe cases the distance distributions have more population modes at generally shorter distances in the Mn^{2+}-bound form, though the dominant distances exhibit only small changes. Since the lobe-lobe variants show single distance populations, these differences seen in the loop-lobe variants are likely caused by conformational change in the loop itself. If the short distances in the distributions are to be believed, there may be multiple conformations sampled. There is very little change in the dominant distance of the S58C/I125C (0.7 Å), which supports the idea that the N-terminal lobe shows little movement. If the mode at 28.4 Å is not an artefact, this would suggest the loop samples an ‘open’ and ‘closed’ position, with the latter significantly less populated, possibly due to the transient nature of Mn^{2+} binding. The two modes seen in the apoprotein I125C/I236C distance distribution (29.2 and 35.2 Å) are present in the Mn^{2+}-bound form (29.4 and 34.3 Å) with the majority population now at the shorter distance and with a change in the longer distance of 0.9 Å. If the C-terminal lobe is not

sampling conformations in a similar manner to the apoprotein, as suggested by the unimodal distribution of the L56C/I236C variant, then the loop is. Interestingly, in both variants the difference between the dominant and next-largest population is 4.9 Å. The consistency of this value suggests that the loop is sampling two discrete conformations, and that the distances between these conformations and that of the C-terminal lobe in the apoprotein form are just coincidental.

The data for the Zn^{2+}-bound form appears to indicate that is shares attributes of both the apoprotein (flexibility) and Mn^{2+}-bound (shorter distances) states. The Zn^{2+}-bound L56C/I236C shows a distribution very similar in shape/features to the apoprotein, but the distances are much shorter; the dominant distance of 33.1 Å is much closer to the distance of the Mn^{2+}-bound form (32.9 Å) than the dominant population of the apoprotein (36.9 Å). Much like in the apoprotein form, the multimodal distribution of L56C/I236C and broader populations of the L56C/S266C and S58C/S266C variants suggest that the C-terminal lobe is asymmetrically flexible. However, the shorter distances seen here compared to the apoprotein (35.5 vs 34.1 Å for L56C/S266C, and 36.6 vs 35.7 Å for S58C/S266C for the apoprotein vs Zn^{2+}-bound forms, respectively) indicate the lobes are ‘closed’ around the substrate, much like the Mn^{2+}-bound form.

The data presented for the loop-lobe variants echo this; S58C/I125C exhibits a single conformation, like the apoprotein, but with the same distance of the dominant feature in Mn^{2+}-bound form (33.3 Å; 0.7 Å difference compared to the apoprotein). The Zn^{2+}-bound I125C/I236C variant has a bimodal distance distribution, with a difference of 5.9 Å between the two populations. This is more similar to the apoprotein, where there is 6.0 Å difference between the equal populations, than to the Mn^{2+}-bound form, which has a difference of 4.9 Å between the two dominant features. As the S58C/I125C variant indicates there is no movement of the loop in the Zn^{2+}-bound state, the bimodal distribution probably arises from the movement of the C-terminal lobe, which aligns with the interpretation of the lobe-lobe variant data. The 0.7 Å change in loop position (as extracted from the S58C/I125C data) from the apoprotein to the Zn^{2+}-bound state likely accounts for difference in absolute distances measured between the spin labels exhibited in the I125C/I236C variant (29.2 & 35.2 Å vs 29.8 & 35.7 Å for apoprotein and Zn^{2+}-bound form respectively; this is a difference of 0.6 and 0.5 Å).

A speculative interpretation of this model could be that the binding of Zn^{2+} does not result in the correct or complete closure of the protein, which may then prevent it from reopening in the way that it does when Mn^{2+} binds. This would explain how this substrate binds irreversibly, despite the lower affinity of PsaA for Zn^{2+} compared to Mn^{2+} and the less favourable coordination sphere.

Comparison with model

The adapted 'Venus flytrap' model taken from the three different crystal structures posited an asymmetric movement of the lobes toward each other in order to occlude the central pocket upon substrate binding; the N-terminal lobe is proposed to move very little, if at all, whilst it is suggest that the C-terminal lobe does the majority of the 'closing' movement.

The PELDOR results for the lobe-lobe measurements (variants L56C/I236C, L56C/S266C and S58C/S266C) generally agree with the 'Venus flytrap' model of SBD movement. When comparing the distance distributions given from measurement of the apoprotein forms against those of the Mn^{2+}-bound forms of these variants, there is a general trend in which the distances between the pairs of spin labels decrease.

The comparable results for the loop-lobe variants (S58C/I125C and I125C/I236C) reveal asymmetric inter-label distance changes, wherein the N-terminal lobe-loop variant S58C/I125C demonstrates little change between the apoprotein and Mn^{2+}-bound states (Δ<1 Å when considering the majority populations in the distributions). In contrast, the C-terminal lobe-loop variant I125C/I236C exhibits more variation; an even, bimodal distribution of inter-label distance populations in the apoprotein form becomes dominated by a single distance in the Mn^{2+}-bound form.

The extent of the distance changes between the two lobes and the shapes of the distributions are very different across the three variants. Two of the variants (L56C/S266C and S58C/S266C) exhibit relatively small inter-label distance changes between the apoprotein and Mn^{2+}-bound states of ~1 Å. Meanwhile, the other variant (L56C/I236C) demonstrates a much more prominent movement of ~4 Å. Considering the proximity of the L56 and S58 sites on the N-terminal lobe, the source of the uneven distance change is unlikely to be any large-scale motion of this lobe. As such the C-terminal lobe, containing positions I236 and S266, is responsible for the majority of the distance changes, in agreement with the crystal structure-based model.

The other important feature of this model is the role of the helical backbone, a segment of which is posited to act as a spring. Upon Mn^{2+}-binding the C-terminal-lobe end of the helix is thought to bend and partially uncoil in a reversible manner.

However, in the case of Zn^{2+}-binding, the ‘spring’ is purported to be overstretched, preventing it from recoiling and releasing the bound metal. The results presented here are not able to report on this proposal, either in support or opposition. In order to further investigate the validity of the proposed model, studies involving the helix region would be required.

The model presented by the crystal structures does not specifically reference the loop region, this is potentially because the intrinsically flexible nature of loops makes it difficult to speculate. However, the data obtained from the *in silico* labelling predicted changes in the conformation. The results of the PELDOR distance measurements reinforced this and showed that the loop had varying degrees of flexibility based on the binding state. It may be that the loop has some involvement in transiently ‘locking in’ the cognate ligand prior to transport.

FRET data

Recent studies involving FRET on PsaA[11] did not study the range of the variants discussed here, though the E74C/K237C and V76C/K237C are not too dissimilar to the L56C/I236C variant used for PELDOR studies above; the E74 and V76 positions sit on the N-terminal lobe. However, the results reported in the FRET study contribute to the understanding of the function and dynamics of PsaA with its cognate ligand and its dysfunction in the presence of Zn^{2+}.

Table 6.3.9: comparison of the measured and reported distances for PsaA variants E74C/K237C, V76C/K237C and L56C/I236C in a range of conditions. This includes reported FRET efficiencies (of the fluorescent-labelled PsaA variants E74C/K237C and V76C/K237C and the distance measurements consisting of: the distance between the C_α atoms of the mutation positions as calculated from the crystal structures, the distances between the spin labels in the double-labelled as predicted from the simulated distance distributions calculated at 298 K using spin label MTSL, and the measured distances between the spin labels in the double-labelled L56C/I236C as extracted from the distance distributions obtained from the PELDOR experiments. In cases where there are multiple modes in the distance distributions, the predominant value is shown in bold. Dashes indicate where there are missing data points.

[a] – data reported in publication by de Boer et al., 2019[11]

Variant		**Distance (Å)**			**FRET Efficiency, E* (ALEX sm-FRET)[a]**	**FRET-determined Distance Change (Å, vs Apo)[a]**
		C_α-C_α	**MMM**	**PELDOR**		
L56C/I236C	Apo	42	50	**36.9**, 44.7	-	-
	+ Mn^{2+}	39.1	45.5	32.9	-	-
	+Zn^{2+}	39.3	44.5	**33.1**, 40.4	-	-
E74C/K237C	Apo	44.6	49.4	-	0.518 ± 0.003	-
	+ Mn^{2+}	41.1	45.5	-	0.567 ± 0.003	3.2 ± 0.3
	+Zn^{2+}	41.1	48.2	-	0.570 ± 0.003	3.2 ± 0.3
V76C/K237C	Apo	44	50.4	-	0.615 ± 0.003	-
	+ Mn^{2+}	40.2	41.6	-	0.681 ± 0.004	4.0 ± 0.4
	+Zn^{2+}	40.7	45.7	-	0.688 ± 0.004	4.1 ± 0.4

FRET measurements for E74C/K237C exhibited distance changes of 3.2 Å for both Mn^{2+}- and Zn^{2+}-binding, and V76C/K237C was reported to change by 4.0 and 4.1 Å upon binding of Mn^{2+} and Zn^{2+} respectively[11]. Similarly, the distance change measured for the L56C/I236C variant using PELDOR was 4.0 Å and 3.8 Å with Mn^{2+}- and Zn^{2+}-binding (for the majority population of the distance distributions).

However, the FRET histograms were reported to be unimodal, whereas the PELDOR results reported more than one population of inter-label distance in both the apoprotein and Zn^{2+}-bound forms. Considering that I236 and K237 are neighbouring positions on the C-terminal lobe, which is proposed to undergo the majority of the movement in the SBD closure process and is expected to be somewhat flexible based on the PELDOR model described above, a multimodal FRET histogram would be anticipated. This discrepancy, however, may be due to the constraints of FRET; by using larger probes (vs spin labels in EPR), further away from the local environment which they are being used to study (i.e. longer linker chains) there may be some dampening of the effect on the sensitivity of the measurements because of the motional freedom of the fluorescent label. Also, if the conformational dynamics of these movements are faster than the millisecond timescale the histograms may not show discrete populations, whereas PELDOR uses a flash-frozen snapshot of the protein in ensemble.

Additionally, EDTA was found to scavenge metal from Mn^{2+}-PsaA but not Zn^{2+}-PsaA through measurement of the lifetime of the closed conformations[11]. This suggests that the irreversible binding of divalent zinc is explained by slow opening kinetics of PsaA. A binding variant (mutation of the D280 position to asparagine), which causes the SBD to release Zn^{2+} like WT-PsaA releases Mn^{2+}, was found to allow zinc transport. Background efflux was prevented by knockout of the Zn^{2+}-efflux pump gene. In-cell Zn^{2+} accumulation occurred in the bacterial strain containing the D280N PsaA gene, but not in the strain with WT-PsaA. Therefore, it is the PsaA subunit that confers metal transport specificity to the PsaBCA complex.

7 Summary & Conclusions

A range of different ABC-type transporter SBDs were studied here with a general theme of using various EPR spectroscopy techniques to investigate how they bind their substrates, both cognate and non-cognate, and the conformational dynamics they undergo.

The SBD subunit of the maltose transporter MalEFGK$_2$, MalE (also referred to as MBP), was first taken as a well-documented example of this group of proteins. Site directed spin labelling (SDSL) was performed on a variant in which two cysteine sites were introduced, one on each lobe of the SBD (T36C/S352C). Room temperature cw-EPR was used to verify it had been successfully labelled to about 50 % efficiency, and the data simulated and fitted to extract certain characterising parameters. This revealed two different label species, which was taken to indicate the two labelling sites were significantly different in terms of rotational freedom for the attached label. Low temperature cw-EPR was also used on the apoprotein form of MalE to further characterise the spin labelled variant.

PELDOR studies of the spin labelled variant, in both the absence and presence of the substrate maltose, revealed distances (52.9 Å and 40.5 Å for apoprotein and maltose-bound respectively) similar to those that were expected from the crystal structures (51.0 Å and 40.8 Å) and *in silico* labelling experiments (52.5 Å and 40.0 Å). These results also agreed well with reported FRET results for the same (large increase in FRET efficiency upon maltose binding[11]) and a very similar variant (K34C/R354C; 58.3 Å and 49.0 Å[10]) of MalE. This demonstrated the complementarity of the PELDOR and FRET distance measurement techniques, which was explored further using the next system.

Following on from the studies of the MalE protein, another system that had also been investigated recently using FRET was the amino acid transporter GlnPQ. This has two SBDs, known as SBD1 and SBD2. As with the previous studies, SDSL was performed on a two-cysteine containing variant of each protein (Q87C/T159C for SBD1, and T369C/S451C for SBD2). These variants were spin labelled to about 94 % best efficiency in both cases.

SBD1 binds glutamine and asparagine with very different K_d values; glutamine is bound fairly ‘loosely’ with a K_d of 92 μM and asparagine is bound much more ‘tightly’ with a K_d of 0.9 μM. No significant differences between the room

temperature or low temperature cw-EPR spectra of the three binding states (apoprotein, Gln-bound and Asn-bound). The lack of discernible difference in the room temperature spectra means that binding of either substrate does not significantly alter the environments of the attached spin labels. In the case of the low temperature cw-EPR, no dipolar broadening was exhibited, meaning none of the binding states had inter-label distances of less than about 25 Å.

The room temperature cw-EPR spectra of the three binding states were also measured as a function of incident microwave power (i.e. a power saturation experiment). The results of these were largely inconclusive, as no real difference was found between the extracted $P_{1/2}$ values of the binding states through the two-trendline method. However, the more complex method fitting the data gave vastly different results for both $P_{1/2}$ and the measure of saturation homogeneity in the apoprotein state compared with those of the substrate-bound states. It is unclear as to why this is, but the introduction of errors into the original datasets were discussed, which would therefore affect the fitting outputs.

The equivalent room and low temperature cw-EPR studies of SBD2 were only performed on the apoprotein form of the spin labelled variant, though the protein binds glutamine and glutamate. However, comparison of the room temperature cw-EPR spectrum of this and that of the apoprotein SBD1 variant showed that the SBD2 spin labels were slightly more rotationally restricted. This agreed well with the projected rotamer library populations of the individual mutation sites in each protein as calculated by the *in silico* labelling experiments. The SBD1 Q87C/T159C variant was predicted to have 85 and 115 rotamers at each respective labelling site, whilst SBD2 T369C/S451C was predicted to have 69 and 47 respectively. This matched well with the τ_c values extracted by best-fit simulations of the experimental data, in which the SBD1 spectrum was dominated by the species with a relatively short τ_c of 2.7 ns. In contrast the SBD2 simulation had fairly even weighting of two species, one with a τ_c of 5.0 ns and the other at 9.8 ns.

PELDOR measurements of SBD1 in the absence of substrate (52.3 Å), with glutamine (43.5 Å) and with asparagine (39.9 Å) agreed well with the trend set out by the FRET efficiencies[11] and the aforementioned K_d values. In the case of SBD2, the apoprotein form was found to have a bimodal distance distribution (44.3 and 56.3 Å, with the longer distance more highly populated), which was not

observed in the equivalent FRET data. Similar distances were measured in the Glu-bound form (44.2 and 54.6 Å), but the populations of the two distances were inverted (dominated by the shorter distance). This was taken to suggest that the apoprotein form was in some way contaminated by the presence of a binding amino acid, and the Glu-bound form perhaps did not have sufficient substrate for full closure. The Gln-bound form of SBD2 was found to have a single population in its distance distribution (43.0 Å). These data, too, agreed well with the FRET efficiencies trend[11].

The SBD of the manganese importer system of *S. pneumoniae* (the PsaBCA complex), PsaA, was the main focus and most in-depth of the studies presented in this chapter. Five single-cysteine containing variants and five double-cysteine containing variants were investigated here in a range of different EPR methods in order to characterise the system and probe Mn^{2+}- and Zn^{2+}-binding and any potential differences between the resulting conformational states.

First, the room temperature cw-EPR spectra of the spin labelled single cysteine containing variants were recorded. Next, a full set of low temperature cw-EPR spectra were recorded for the singly labelled variants were recorded both at X-band and at Q-band. Attempts were made to simultaneously simulate and fit these data for each variant in order to achieve a set of convergent g- and A-tensors to characterise these systems. It was found that this was best done across the low temperature cw-EPR datasets and fairly good convergence with the experimental data was achieved (low RMSDs). The results of these best-fit simulations were used as starting points for the equivalent spectra recorded at room temperature at X-band.

These studies revealed a conflict with previously reported results[45] for the rotational correlation time, τ_c, for all of the variants. In the published literature, all but one of the simulations were performed using one spin label ‘species’, and the resulting τ_c values were consistently smaller than those obtained by the investigations shown here. Since the published simulations do not have any RMSD associated with them, and only report the τ_c values, it was difficult to draw any conclusions from this. However, it could be argued that even through simple qualitative assessment of the spectral lineshapes (in both cases) that the τ_c values would be significantly larger than those previously reported. Following on from this, the best-fit simulations of the room temperature cw-EPR spectra of the single-

labelled variants were combined in a 1:1 ratio for the respective double cysteine containing variants. This was shown to qualitatively match the experimental room temperature cw-EPR spectra of the double-labelled variants well.

The low temperature cw-EPR studies at Q-band also looked at the spectral characteristics of the manganese substrate. Certain single- and double-labelled variants were studied here, where the distances between the attached spin labels and the metal ion are predicted to be < 30 Å (from the *in silico* labelling experiments), as well as unlabelled wild-type (WT) PsaA. Both the width of the first resonance line and the overall spectral breadth were measured and were found to be significantly different in the labelled variants compared to the WT. In general, the extent of line broadening seemed to be related to the projected distance between the spin label and the metal binding site; i.e. the closer the two paramagnetic species, the greater the broadening. This makes sense given what is known about the effects of dipolar coupling and resonance line broadening between two spin labels. The overall width of the manganese $I = 5/2$ sextet was narrower in the spin labelled variants than the WT protein. However, across the dataset the corresponding hyperfine coupling values agreed well literature values for 6-coordinate Mn(II), which would oppose the general consensus from crystal data that PsaA has a tetrahedral binding geometry. However, if the carboxylic acid groups of the coordinating Glu and Asp residues were bidentate, the binding sphere would indeed be hexacoordinate.

The extensive PELDOR data of the double-labelled variants were divided to interpret the lobe-lobe (L56C/I236C, L56C/S266C and S58C/S266C) and loop-lobe (S58C/I125C and I125C/I236C) movements. A model for the conformational dynamics of PsaA upon Mn^{2+}- vs Zn^{2+}-binding was built on these measurements.

Firstly, the apoprotein form showed some one-sided lobe flexibility, in agreement with the conclusions of the literature based on room temperature cw-EPR at X-band[45]. The C-terminal lobe (which carries the I236 and S266 sites) was posited to move more upon substrate binding that the N-terminal lobe (containing L56 and S58). It was found that Mn^{2+}-binding resulted in single populations in the distance distributions, suggesting that a single conformation exists in terms of lobe-lobe measurements. Additionally, two of the lobe-lobe variants showed greater distance change than the remaining variant, potentially signifying a twisting of the lobes relative to one another as they close. In contrast, Zn^{2+}-binding had similar flexibility

characteristics to the apoprotein form, but with the shorter distances (comparable to those in the Mn^{2+}-bound state) dominating the distance distributions.

The degree of loop flexibility also appeared to be dependent on the binding state. When accounting for the flexibility of the C-terminal lobe, the loop appears to occupy a single conformation in the apoprotein form. When Mn^{2+} binds, the loop populates several positions relative to the lobes, suggesting that it is flexible between ‘open’ and ‘closed’ conformations. This is likely related to the transient nature of the binding of the cognate ligand. However, Zn^{2+}-binding causes the loop to only occupy the ‘closed’ position (the same distance as seen in the Mn^{2+}-bound form); again, this was interpreted taking into account the C-terminal lobe’s movements.

The data obtained during these studies were not directly comparable with the results of FRET studies on PsaA[11] due to the difference in the sites where cysteines were introduced (E74C/K237C and V76C/K237C). However, the positions of the cysteine sites in L56C/I236C are the closest for the sake of comparison. The PELDOR data for this variant showed bimodal distributions in the apoprotein and Zn^{2+}-bound forms, which were not replicated in the corresponding FRET histograms. Nonetheless, the distance changes that were reported match well with the distance changes measured using PELDOR, when disregarding the lesser populated distances. The discrepancies between these results may be due to the differences in experimental/sample conditions for the two methods. In addition, the E74 and V76 positions of the FRET variants are located in a more surface-exposed part of the N-terminal lobe than L56, which may cause these positions to be less sensitive to the lobe flexibility as any attached probes will experience large degrees of rotational freedom.

8 Outlook

There are many lines of investigation still open on all of the systems presented here. Outlined below are many avenues yet to be explored, which may help to complete the studies presented here.

In order to verify or refute the apparent short distance seen in the maltose-bound MalE PELDOR, low temperature cw-EPR experiments should be performed. Sub-saturating amounts of ligand could also be titrated in and the changes in conformation monitored in order to investigate if there are two discrete populations in the distance distributions, which would demonstrate whether the SBD samples two conformations (i.e. open and closed) or not (a range of conformations between fully-open and fully-closed).

In the FRET study discussed extensively here, many different conformations of MalE were found using many different substrates[11], with the main finding being that the extent of closure of MalE seems to depend on the size of the substrate. Non-cognate ligands were also found to bind to MalE, but the protein did not close around them to the same extent as the cognate ligands. For a more in-depth comparison of EPR and FRET as tools for investigating conformational dynamics, it would be interesting to perform these same binding experiments and measure the closure of MalE using PELDOR.

In the case of the GlnPQ SBDs, ligand titration experiments were hindered by the difficulties in production and purification of fully active protein. In the case of SBD1, where one substrate is known to bind more tightly, a displacement titration could also be performed and the resultant changes in conformation monitored through PELDOR. If the problem of contaminated ‘apoprotein’ samples were to persist with new purifications, unlabelled SBD could be added to the samples in order to scavenge the substrate. However, this would need to be done with a high concentration of labelled SBD to start with as the dilution caused by adding unlabelled protein would affect the signal intensity. In addition, completion of the crystal structure and K_d datasets would help to confirm or disregard many of the proposed details presented here, particularly in the case of SBD2 and glutamate binding.

Further investigations into PsaA and metal binding could be done by recording the low temperature cw-EPR spectra of the single-labelled variants incubated with

Mn(II) and Zn(II). Since Mn(II) is paramagnetic, any short-range distances between the metal and a spin label would likely result in dipolar broadening. Investigation of the relationship between the extent of spectral broadening of the spin label characteristics and the distance between the metal and spin probe could be done using these spectra. This could help elucidate the relative positions of each of the labelling sites (i.e. the two lobes and loop) in relation to the central binding cleft.

As with the GlnPQ SBDs, sub-K_d amounts of Mn(II) and Zn(II) could be sequentially titrated into labelled double-cysteine variants to track the changes substrate displacement has on the conformation of PsaA.

In order to investigate the idea of the ‘spring-stretch’ from the proposed binding model (taken from the crystal data), a variant involving the end of the helix linker region should be designed. Simple room temperature cw-EPR spectroscopy at X-band may be able to verify this, as manganese-binding should cause a slight increase in label mobility, whilst zinc-binding should cause a more extensive increase in rotational freedom. These same experiments on the existing single- and double-cysteine variants would also be valuable. This may help to account for the broadness of certain features seen in some of the PELDOR distance distributions (i.e. more mobile label environments often result in broader distributions).

References

1. Bordignon, E.; Grote, M.; Schneider, E., The Maltose ATP-Binding Cassette Transporter in the 21st Century - Towards a Structural Dynamic Perspective on Its Mode of Action. *Mol. Microbiol.* **2010,** *77* (6), 1354-1366.

2. Scheepers, G. H.; Lycklama, J. A.; Poolman, B., An Updated Structural Classification of Substrate-Binding Proteins. *Febs Letters* **2016,** *590* (23), 4393-4401.

3. Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A., Crystallographic Evidence of a Large Ligand-Induced Hinge-Twist Motion between the 2 Domains of the Maltodextrin Binding-Protein Involved in Active-Transport and Chemotaxis. *Biochemistry* **1992,** *31* (44), 10657-10663.

4. Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E., Extensive Features of Tight Oligosaccharide Binding Revealed in High-Resolution Structures of the Maltodextrin Transport Chemosensory Receptor. *Structure* **1997,** *5* (8), 997-1015.

5. Bao, H.; Duong, F., ATP Alone Triggers the Outward Facing Conformation of the Maltose ATP-Binding Cassette Transporter. *Journal of Biological Chemistry* **2013,** *288* (5), 3439-3448.

6. Oldham, M. L.; Khare, D.; Quiocho, F. A.; Davidson, A. L.; Chen, J., Crystal Structure of a Catalytic Intermediate of the Maltose Transporter. *Nature* **2007,** *450* (7169), 515-U7.

7. Khare, D.; Oldham, M. L.; Orelle, C.; Davidson, A. L.; Chen, J., Alternating Access in Maltose Transporter Mediated by Rigid-Body Rotations. *Mol. Cell* **2009,** *33* (4), 528-536.

8. Bao, H.; Duong, F., Discovery of an Auto-Regulation Mechanism for the Maltose ABC Transporter MalFGK(2). *Plos One* **2012,** *7* (4), 13.

9. Mullen, A.; Hall, J.; Diegel, J.; Hassan, I.; Fey, A.; Macmillan, F., Membrane Transporters Studied by EPR Spectroscopy: Structure Determination and Elucidation of Functional Dynamics. *Biochem. Soc. Trans.* **2016,** *44*, 905-915.

10. Kim, E.; Lee, S.; Jeon, A.; Choi, J. M.; Lee, H. S.; Hohng, S.; Kim, H. S., A Single-Molecule Dissection of Ligand Binding to a Protein with Intrinsic Dynamics. *Nature Chemical Biology* **2013,** *9* (5), 313-+.

11. De Boer, M.; Gouridis, G.; Vietrov, R.; Begg, S. L.; Schuurman-Wolters, G. K.; Husada, F.; Eleftheriadis, N.; Poolman, B.; Mcdevit, C. A.; Cordes, T., Conformational and Dynamic Plasticity in Substrate-Binding Proteins Underlies Selective Transport in ABC Importers. *eLife* **2019,** *8*, 28.

12. Polyhach, Y.; Bordignon, E.; Jeschke, G., Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. *Physical Chemistry Chemical Physics* **2011,** *13* (6), 2356-2366.

13. Stoll, S.; Schweiger, A., Easyspin, a Comprehensive Software Package for Spectral Simulation and Analysis in EPR. *Journal of Magnetic Resonance* **2006,** *178* (1), 42-55.

14. Zhang, Z. W.; Fleissner, M. R.; Tipikin, D. S.; Liang, Z. C.; Moscicki, J. K.; Earle, K. A.; Hubbell, W. L.; Freed, J. H., Multifrequency Electron Spin Resonance Study of the

Dynamics of Spin Labeled T4 Lysozyme. *Journal of Physical Chemistry B* **2010,** *114* (16), 5503-5521.

15. Haugland, M. M.; Anderson, E. A.; Lovett, J. E., Tuning the Properties of Nitroxide Spin Labels for Use in Electron Paramagnetic Resonance Spectroscopy through Chemical Modification of the Nitroxide Framework. In *Electron Paramagnetic Resonance: Volume 25*, The Royal Society of Chemistry: 2017; Vol. 25, pp 1-34.

16. Klare, J.; Steinhoff, H., Spin Labeling EPR. *Photosynthesis Research* **2009,** *102* (2-3), 377-390.

17. Owenius, R.; Engstrom, M.; Lindgren, M.; Huber, M., Influence of Solvent Polarity and Hydrogen Bonding on the EPR Parameters of a Nitroxide Spin Label Studied by 9-GHz and 95-GHz EPR Spectroscopy and Dft Calculations. *Journal of Physical Chemistry a* **2001,** *105* (49), 10967-10977.

18. Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H., DeerAnalysis2006 - a Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006,** *30* (3-4), 473-498.

19. Lerner, E.; Cordes, T.; Ingargiola, A.; Alhadid, Y.; Chung, S.; Michalet, X.; Weiss, S., Toward Dynamic Structural Biology: Two Decades of Single-Molecule Forster Resonance Energy Transfer. *Science* **2018,** *359* (6373), 288-+.

20. Mao, B.; Pear, M. R.; Mccammon, J. A.; Quiocho, F. A., Hinge-Bending in L-Arabinose-Binding Protein - the Venus-Flytrap Model. *Journal of Biological Chemistry* **1982,** *257* (3), 1131-1133.

21. Quiocho, F. A.; Ledvina, P. S., Atomic Structure and Specificity of Bacterial Periplasmic Receptors for Active Transport and Chemotaxis: Variation of Common Themes. *Mol. Microbiol.* **1996,** *20* (1), 17-25.

22. Felder, C. B.; Graul, R. C.; Lee, A. Y.; Merkle, H. P.; Sadee, W., The Venus Flytrap of Periplasmic Binding Proteins: An Ancient Protein Module Present in Multiple Drug Receptors. *AAPS Pharmsci* **1999,** *1* (2), 28.

23. Schuurman-Wolters, G. K.; Poolman, B., Substrate Specificity and Ionic Regulation of GlnPQ from Lactococcus lactis - an ATP-Binding Cassette Transporter with Four Extracytoplasmic Substrate-Binding Domains. *Journal of Biological Chemistry* **2005,** *280* (25), 23785-23790.

24. Fulyani, F.; Schuurman-Wolters, G. K.; Zagar, A. V.; Guskov, A.; Slotboom, D. J.; Poolman, B., Functional Diversity of Tandem Substrate-Binding Domains in ABC Transporters from Pathogenic Bacteria. *Structure* **2013,** *21* (10), 1879-1888.

25. Schuurman-Wolters, G. K., Guskov, A., Poolman, B., Crystal Structure of Substrate Binding Domain 1 (Sbd1) of ABC Transporter Glnpq in Complex with Asparagine. *RCSB PDB* **TO BE PUBLISHED**.

26. Gouridis, G.; Schuurman-Wolters, G. K.; Ploetz, E.; Husada, F.; Vietrov, R.; De Boer, M.; Cordes, T.; Poolman, B., Conformational Dynamics in Substrate-Binding Domains Influences Transport in the ABC Importer GlnPQ. *Nat. Struct. Mol. Biol.* **2015,** *22* (1), 57-64.

27. Altenbach, C.; Greenhalgh, D. A.; Khorana, H. G.; Hubbell, W. L., A Collision Gradient-Method to Determine the Immersion Depth of Nitroxides in Lipid Bilayers - Application to Spin-Labeled Mutants of Bacteriorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **1994,** *91* (5), 1667-1671.

28. Vogel, S. S.; Van Der Meer, B. W.; Blank, P. S., Estimating the Distance Separating Fluorescent Protein FRET Pairs. *Methods* **2014,** *66* (2), 131-138.

29. Bogaert, D.; De Groot, R.; Hermans, P. W., Streptococcus pneumoniae Colonisation: The Key to Pneumococcal Disease. *Lancet Infect Dis* **2004,** *4* (3), 144-54.

30. Jedrzejas, M. J., Pneumococcal Virulence Factors: Structure and Function. *Microbiol Mol Biol Rev* **2001,** *65* (2), 187-207 ; first page, table of contents.

31. Ogunniyi, A. D.; Mahdi, L. K.; Jennings, M. P.; Mcewan, A. G.; Mcdevitt, C. A.; Van Der Hoek, M. B.; Bagley, C. J.; Hoffmann, P.; Gould, K. A.; Paton, J. C., Central Role of Manganese in Regulation of Stress Responses, Physiology, and Metabolism in Streptococcus pneumoniae. *J Bacteriol* **2010,** *192* (17), 4489-97.

32. Rice, A. J.; Park, A.; Pinkett, H. W., Diversity in ABC Transporters: Type I, II and III Importers. *Crit. Rev. Biochem. Mol. Biol.* **2014,** *49* (5), 426-437.

33. Davidson, A. L.; Chen, J., ATP-Binding Cassette Transporters in Bacteria. *Annu Rev Biochem* **2004,** *73*, 241-68.

34. Berntsson, R. P.; Smits, S. H.; Schmitt, L.; Slotboom, D. J.; Poolman, B., A Structural Classification of Substrate-Binding Proteins. *FEBS Lett* **2010,** *584* (12), 2606-17.

35. Eijkelkamp, B. A.; Mcdevitt, C. A.; Kitten, T., Manganese Uptake and Streptococcal Virulence. *Biometals* **2015,** *28* (3), 491-508.

36. Davidson, A. L.; Dassa, E.; Orelle, C.; Chen, J., Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol Mol Biol Rev* **2008,** *72* (2), 317-64, table of contents.

37. Couñago, R. M.; Ween, M. P.; Begg, S. L.; Bajaj, M.; Zuegg, J.; O'mara, M. L.; Cooper, M. A.; Mcewan, A. G.; Paton, J. C.; Kobe, B.; Mcdevitt, C. A., Imperfect Coordination Chemistry Facilitates Metal Ion Release in the Psa Permease. *Nat Chem Biol* **2014,** *10* (1), 35-41.

38. Mcdevitt, C. A.; Ogunniyi, A. D.; Valkov, E.; Lawrence, M. C.; Kobe, B.; Mcewan, A. G.; Paton, J. C., A Molecular Mechanism for Bacterial Susceptibility to Zinc. *Plos Pathogens* **2011,** *7* (11).

39. Lawrence, M. C.; Pilling, P. A.; Epa, V. C.; Berry, A. M.; Ogunniyi, A. D.; Paton, J. C., The Crystal Structure of Pneumococcal Surface Antigen PsaA Reveals a Metal-Binding Site and a Novel Structure for a Putative ABC-Type Binding Protein. *Structure* **1998,** *6* (12), 1553-61.

40. Eijkelkamp, B. A.; Morey, J. R.; Ween, M. P.; Ong, C. L.; Mcewan, A. G.; Paton, J. C.; Mcdevitt, C. A., Extracellular Zinc Competitively Inhibits Manganese Uptake and Compromises Oxidative Stress Management in Streptococcus pneumoniae. *PLoS One* **2014,** *9* (2), e89427.

41. Reed, G. H.; Cohn, M., Electron Paramagnetic Resonance Spectra of Manganese (II)-Protein Complexes. Manganese (II)-Concanavalin A. *J Biol Chem* **1970,** *245* (3), 662-4.

42. Felder, C. B.; Graul, R. C.; Lee, A. Y.; Merkle, H. P.; Sadee, W., The Venus Flytrap of Periplasmic Binding Proteins: An Ancient Protein Module Present in Multiple Drug Receptors. *Aaps Pharmsci* **1999,** *1* (2), art. no.-2.

43. Trakhanov, S.; Vyas, N. K.; Luecke, H.; Kristensen, D. M.; Ma, J. P.; Quiocho, F. A., Ligand-Free and -Bound Structures of the Binding Protein (LivJ) of the Escherichia coli ABC Leucine/Isoleucine/Valine Transport System: Trajectory and Dynamics of the Interdomain Rotation and Ligand Specificity. *Biochemistry* **2005,** *44* (17), 6597-6608.

44. Lee, Y. H.; Dorwart, M. R.; Hazlett, K. R. O.; Deka, R. K.; Norgard, M. V.; Radolf, J. D.; Hasemann, C. A., The Crystal Structure of Zn(II)-Free Treponema Pallidum TroA, a Periplasmic Metal-Binding Protein, Reveals a Closed Conformation. *Journal of Bacteriology* **2002,** *184* (8), 2300-2304.

45. Deplazes, E.; Begg, S. L.; Van Wonderen, J. H.; Campbell, R.; Kobe, B.; Paton, J. C.; Macmillan, F.; Mcdevitt, C. A.; O'mara, M. L., Characterizing the Conformational Dynamics of Metal-Free PsaA Using Molecular Dynamics Simulations and Electron Paramagnetic Resonance Spectroscopy. *Biophysical Chemistry* **2015,** *207*, 51-60.

46. Kass, H.; Macmillan, F.; Ludwig, B.; Prisner, T. F., Investigation of the Mn Binding Site in Cytochrome C Oxidase from Paracoccus denitrificans by High-Frequency EPR. *Journal of Physical Chemistry B* **2000,** *104* (22), 5362-5371.

47. Espe, M. P.; Hosler, J. P.; Fergusonmiller, S.; Babcock, G. T.; Mccracken, J., A Continuous-Wave and Pulsed EPR Characterization of the Mn2+ Binding-Site in Rhodobacter-Sphaeroides Cytochrome-C-Oxidase. *Biochemistry* **1995,** *34* (23), 7593-7602.

48. Haltia, T., Reduction of Cua Induces a Conformational Change in Cytochrome-C-Oxidase from Paracoccus-Denitrificans. *Biochim. Biophys. Acta* **1992,** *1098* (3), 343-350.

49. Gagnon, D. M.; Hadley, R. C.; Ozarowski, A.; Nolan, E. M.; Britt, R. D., High-Field EPR Spectroscopic Characterization of Mn(II) Bound to the Bacterial Solute-Binding Proteins MntC and PsaA. *Journal of Physical Chemistry B* **2019,** *123* (23), 4929-4934.

Secondary Active Transporters

9 *Vibrio cholerae* 'I'm Not Dead Yet'

9.1 Background

The 'I'm Not Dead Yet' (INDY) genes encode secondary active transporters of the SLC13 family. The unique name is a result of the gene's proposed influence on lifespan and a reference to a scene in *Monty Python and the Holy Grail*, in which a 'corpse' being carried out for collection during the plague says "I'm not dead yet!".

Reduced expression of the gene was found to increase the longevity (i.e. not-dead-ness) of model organisms (mice and *Drosophila melanogaster*) fed a high calorie diet, in a manner similar to restricting the caloric intake of control groups[1-2], though these results are considered controversial[3]. The gene has been connected to delaying the onset of diet- and age-related symptoms, such as hepatic insulin resistance. As such, this gene and the transporter it encodes is of high medical relevance in fighting ageing and diet-related diseases such as type 2 diabetes.

VcINDY (*Vibrio cholerae* 'I'm Not Dead Yet') symports sodium cations and succinate, a metabolic intermediate of the Krebs Cycle, which is involved in a wide range of metabolic processes, including the metabolism of carbohydrates, amino acids, fatty acids, cholesterol, and haem[4]. The solutes are co-transported in a stoichiometry of 3:1 respectively[5]. Whilst the cognate substrate of VcINDY is known to be succinate, it has been found that the transporter will bind other 4- and 5-carbon dicarboxylates as competitive inhibitors; these include fumarate, malate and citrate[5-6].

This is a bacterial homologue of the SLC13 family with between 26-33% sequence identity with 3 human homologues and was the first of the SLC13/DASS family to have its crystal structure resolved to high resolution[6]. It functions as a homodimer, and each monomer is comprised of 11 TM helices (fig. 9.1.1).

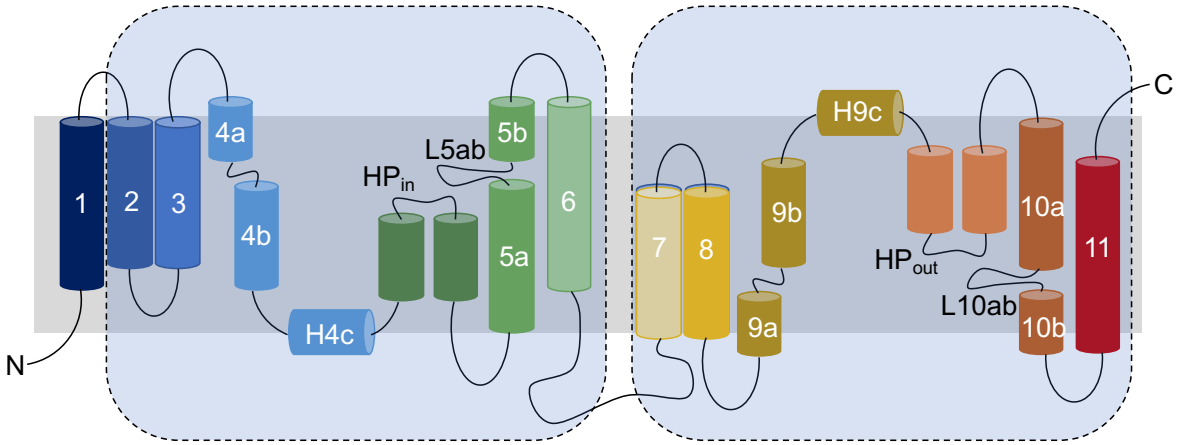


Fig. 9.1.1: the topology of the 11 TM helices and connecting loops of VcINDY in the membrane (grey), displaying the inverted two-fold symmetry of the repeats, which are shown in separate boxes with dashed outlines; repeat 1 consists of TMs 2-6 and repeat 2 consists of TMs 7-11. The colours show the progression of the chain from the N-terminal (blue) to C-terminal (red) ends.

These helices are arranged into two repeats with inverted twofold symmetry. The first repeat, at the N-terminal end, is comprised of TMs 2-6, of which 4 and 5 are segmented. At the C-terminal end, the second repeat is made up of TMs 7-11, and 9 and 10 are segmented. Helical hairpins (HPs) sit between these segmented helices; HP_{in} is connected to TM4a & b via H4c at one side and TM5a & b on the other, whilst HP_{out} is connected to TM9a & b via H9c on one side and TM10a & b on the other[7]. These HP regions are believed to play a role in the functional mechanism of the transporter and are associated with the cation binding sites[6].

The crystal structure of VcINDY (PDB 4F35[6]) was assigned as 'inward facing' and had a bound molecule of citrate and one Na^+ ion (fig. 9.1.2). The sodium binding site (Na1) consists of residues S146 (side chain and peptide backbone), S150 (backbone), and N151 (side chain) in the HP_{in} region, and G199 (backbone) in loop L5ab (see below, fig. 9.1.3). Mutation experiments that targeted these residues revealed their importance in the rate of transport. It would be expected, considering the symmetry of the transporter's topology, that the second Na^+ binding site would sit in the equivalent HP_{out} region. Both HP regions are highly conserved across INDY proteins, further supporting this theory.

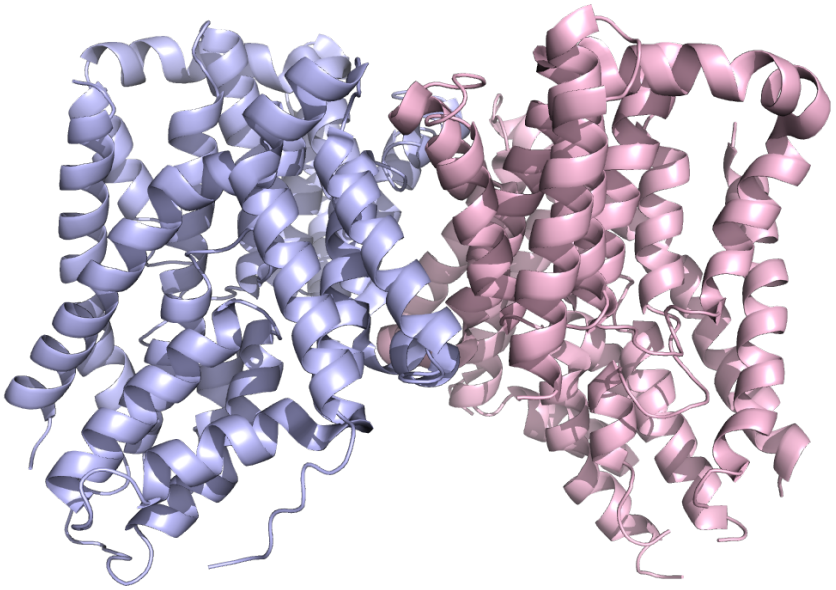


Fig. 9.1.2: crystal structure of homodimeric VcINDY in the inward-facing state; different colours are used to distinguish the separate monomers (PDB 4F35).

The remaining Na^+ binding site is not revealed in this crystal structure, which was resolved using both Na^+ and Li^+ ions; the electron density of a lithium ion bound at a Na^+ binding site would likely be obscured at this resolution (3.2 Å). The citrate binding site was found to be in close proximity to the Na1 site, in a pocket between HP_{in}, HP_{out} and TMs 5 and 10. The carboxyl group at one end of the citrate molecule (5-carboxyl) is oriented towards the S150 and N151 residues of the Na1 site. The other carboxyl moiety (1-carboxyl group) interacts with S377, N378, T379 in the HP_{out} region, and T421 in the loop L10ab[6] (see below, fig. 9.1.3).

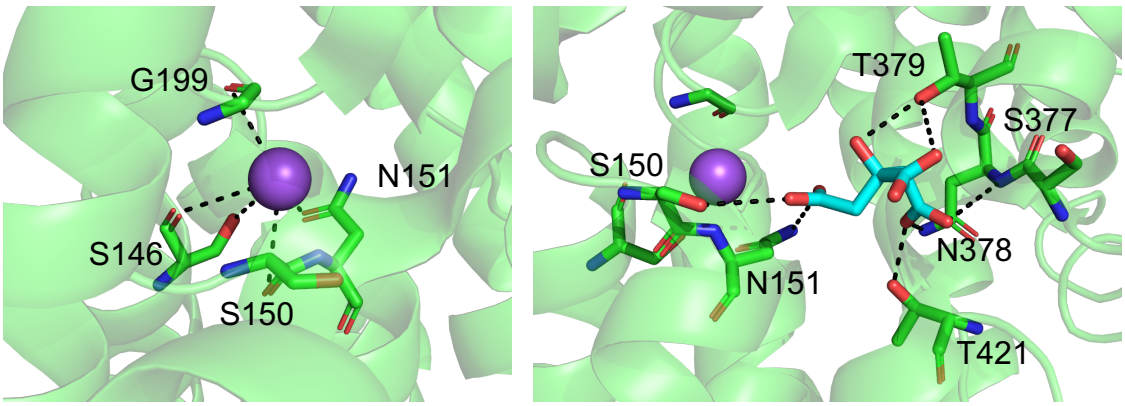


Fig. 9.1.3: the coordination of the sodium ion at site Na1 (left) and of citrate (right) in the crystal structure of inward-facing VcINDY (PDB 4F35). Hydrogen bonds are shown as dashed lines, Na^+ is shown as a purple sphere and citrate is shown in cyan.

A more recent crystal structure was determined at slightly higher resolution (2.8 Å), where two Na^+ binding sites were resolved and with the cognate substrate succinate in the place of citrate (PDB 5UL7[8]). The Na1 site found in the previous structure was confirmed, and the second site sodium binding site (Na2) was found to consist of residues T373 (backbone and side chain), A376 (backbone), and N378 (backbone) in the HP_{out} region, and A420 (backbone) in loop L10ab. With succinate bound, there were interactions of the side chains of N151 and T152 to one carboxylate (HP_{in}/Na1 region), and S377 and N378 with the other(HP_{out}/Na2 region)[9] (see below, fig. 9.1.4).

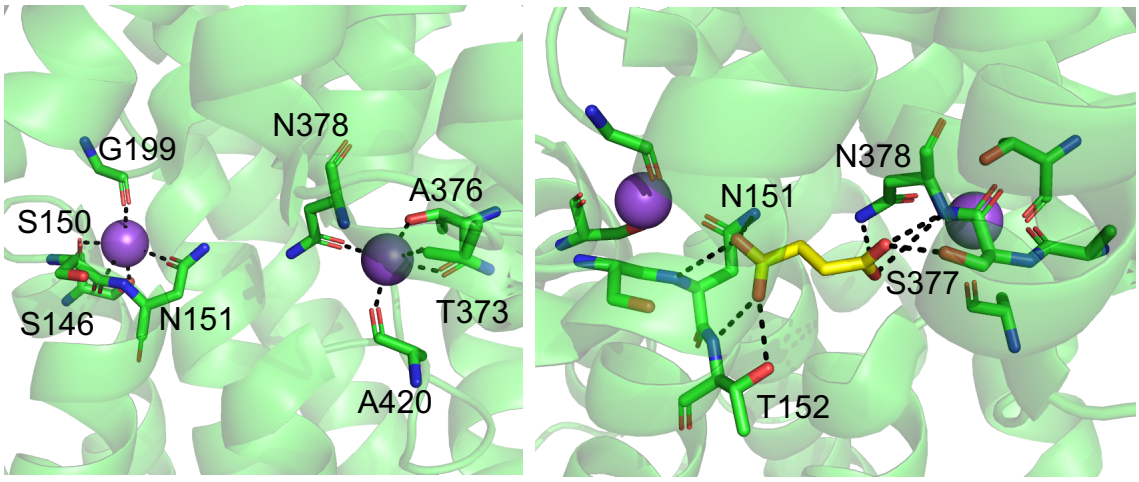


Fig. 9.1.4: the coordination of the sodium ions at sites Na1 and Na2 (left) and of succinate (right) in the crystal structure of inward-facing VcINDY (PDB 5UL7). Hydrogen bonds are shown as dashed lines, Na^+ is shown as a purple sphere and succinate is shown in yellow.

Though the mechanism by which this transporter functions is as yet unconfirmed, it is generally believed to operate though an elevator-type model of alternating access (previously explained in 1.2 'Secondary Active Transport'). The topology of the VcINDY monomers can be divided into two domains: the transport domain containing binding sites and providing access across the membrane, and the scaffold domain containing oligomerisation contacts and remaining relatively motionless. In order to undergo the conformational change from inward- to outward-facing, consistent with an elevator-type mechanism, the transport domain is proposed to undergo 15 Å translation and 43° rotation relative to the scaffold.

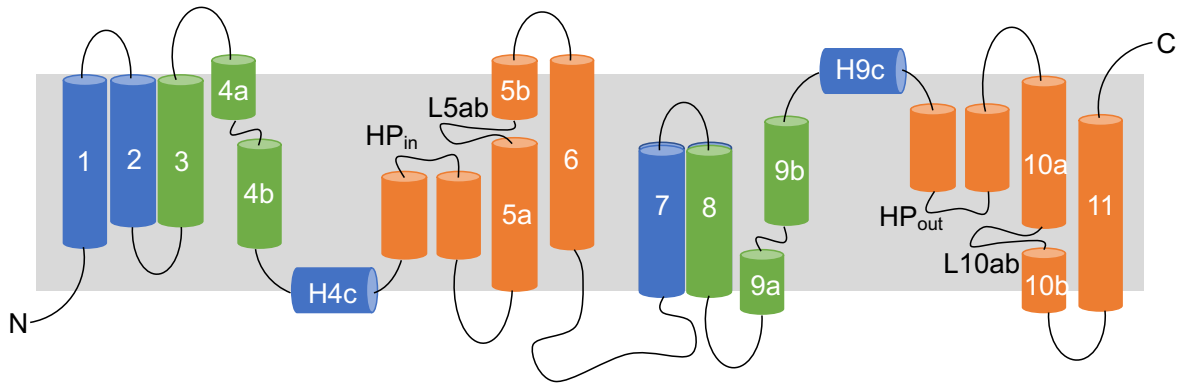


Fig. 9.1.5: the topology of the 11 TM helices and connecting loops of VcINDY in the membrane (grey), with each domain colour-coded: scaffold domain is shown in blue, oligomerisation domain shown in green, and transport domain shown in orange.

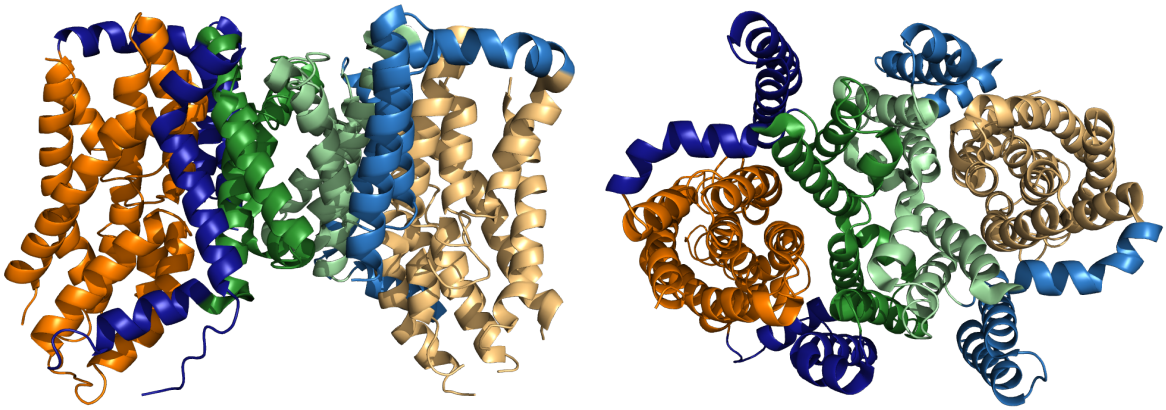


Fig. 9.1.6: crystal structure of homodimeric VcINDY in the inward-facing state (PDB 4F35) from side-on (left) and top-down (right) views, with each domain colour-coded: scaffold domain is shown in blue, oligomerisation domain shown in green, and transport domain shown in orange; different shades are used to distinguish the separate monomers.

The Na^+ ions binding at sites Na1 and Na2 are thought to aid binding of the divalent anion substrate. Though there are no direct interactions between the ions and substrate themselves, they share certain coordinating residues, e.g. N151, N378. From the crystal structure (PDB 5UL7) it was posited that, due to the respective accessibilities of the binding sites, the substrate is released before dissociation of the sodium ions from sites Na1 and Na2[9].

Work presented here is preliminary and ongoing, with huge scope to develop and many ideas for future investigation (see section 11 ‘Outlook’).

9.2 Results and Discussion

in silico spin labelling

A range of cysteine-containing variants of VcINDY were designed to study its functional dynamics using EPR spectroscopy; since it functions as a homodimer, the variants were designed such that a single cysteine would be inserted per monomer. These cysteine mutations were generated *in silico* and a spin label computationally attached to it using the crystal structure. The degree of rotational freedom of the attached spin label can then be calculated and a library of rotamers produced. The distances between the labels (one per monomer) are calculated, and the distance distributions (i.e. weighted probabilities of distances populated between the spin label rotamers) can be predicted.

The rotamer libraries were calculated for each of the labelling positions in VcINDY from the crystal structure of the inward-facing form (4F35[6]) and an MD simulation of the protein with substrate (i.e. Na^+ and succinate, also referred to as 'outward facing'). These were calculated at ambient temperature (298 K) using Multiscale Modelling of Macromolecular Systems (MMM) software package that runs via MATLAB[10].

Table 9.2.1: rotamer libraries of the labelling positions on each monomer chain of VcINDY in the inward- and outward-facing forms, calculated from the crystal structure (PDB 4F35) and MD simulation respectively at 298 K.

	Calculated Rotamer Library (298 K)			
Variant	**Inward-facing**		**Outward-facing**	
	Chain A	**Chain B**	**Chain A**	**Chain B**
E42C	142	136	68	68
F79C	71	55	67	70
W148C	5	8	99	99
T177C	5	102	9	9
A208C	2	1	108	108
G211C	136	137	1	1
E394C	48	42	90	90
S436C	61	142	50	50

This method can be a good general indicator of the motional freedom a spin label may have at a particular site and is often used to assess the potential accessibility of the site (i.e. whether a particular residue, if mutated to a cysteine, will label well). However, it is important to bear in mind that the program considers the input structure as a rigid body and therefore it does not allow for the movements of biological molecules in solution. A spin label attached to a protein in solution may therefore have a greater degree of flexibility than predicted. The predicted distances and distance distributions for the spin label pairs were then calculated.

Table 9.2.2: the predicted distances between the pairs of spin labels in the variants of VcINDY as extracted from the predicted distance distributions, calculated at 298 K.

Variant	Predicted inter-label distance (Å) (MMM @ 298 K)	
	Inward-facing	**Outward-facing**
E42C	55.5	54.0
F79C	44.1	46.7
W148C	49.8	67.1
T177C	64.9	52.1
A208C	23.8	50.4
G211C	26.9	78.5
E394C	35.4	64.2
S436C	73.0	49.4

Some of these variants were selected due to the large-scale changes they are expected to exhibit, which will give rise to dramatic differences in their respective PELDOR data. For example, the G211C variant in the apoprotein state (26.9 Å predicted distance) would be anticipated to have a PELDOR trace with relatively high frequency oscillations that would be fairly quick to collect as it would only need to be measured over a short dipolar evolution time. In contrast, the substrate bound state (78.5 Å predicted distance) would be expected to have either very low frequency or unmeasurable oscillations. This is because such a long distance is towards the upper limit of the sensitivity of PELDOR, particularly with membrane proteins which are typically at low concentration already, unlikely to be labelled to 100% efficiency and would require further dilution with deuterated cryoprotectant and buffer solution. Other variants were designed to monitor the ‘scaffold’ region of the homodimer, where little conformational change is expected to occur. These variants, E42C and F79C, are predicted to have only 1-3 Å inter-label distance change.

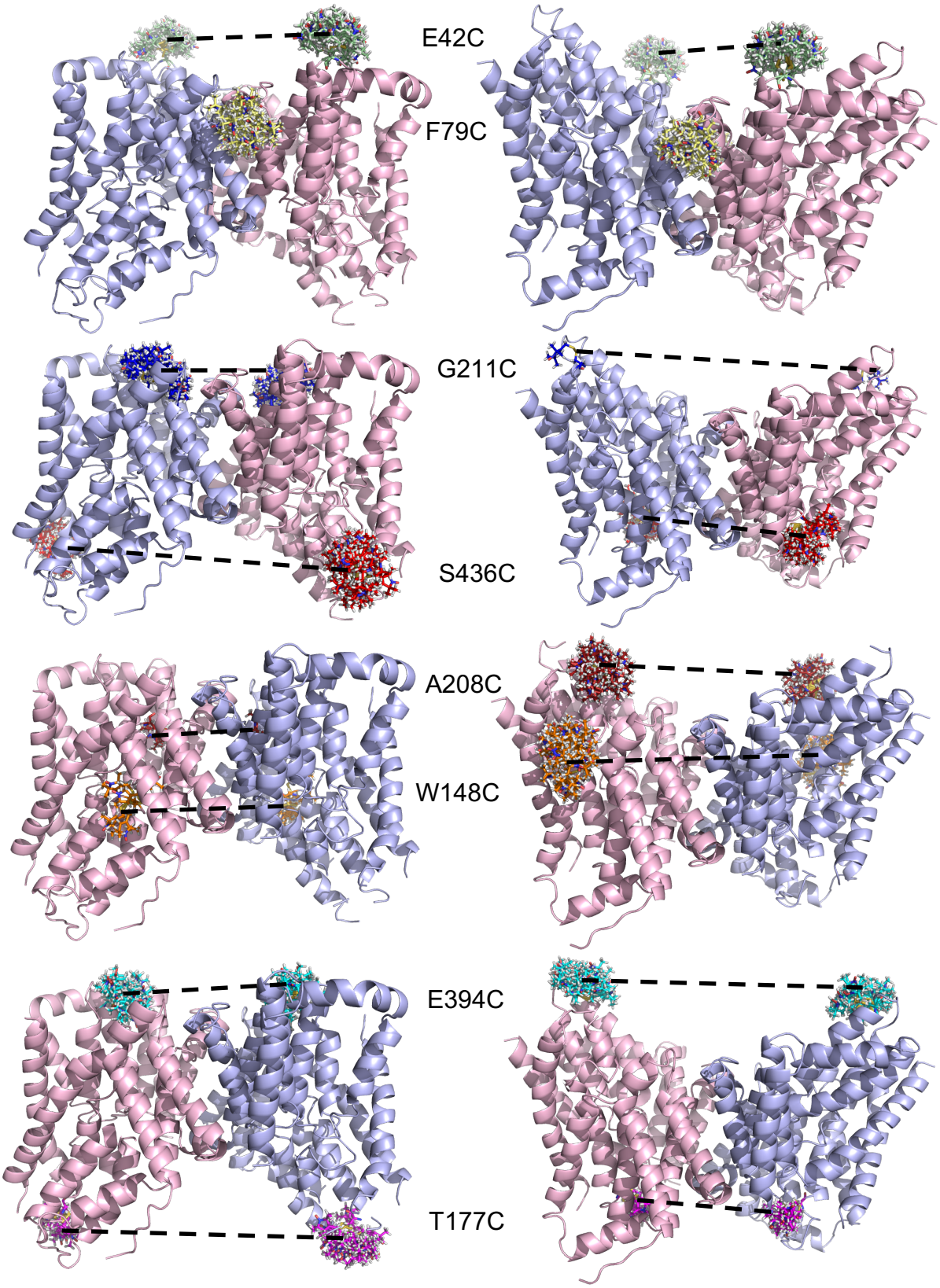


Fig. 9.2.1: the crystal structure of the inward-facing form of VcINDY (PDB 4F35, left) and the MD simulated structure of the outward-facing form (not published, right), with MMM-generated rotamer libraries highlighted, at the residue positions where cysteines will be introduced, with the measurements between the equivalent rotamer libraries on each monomer represented by dashed lines.

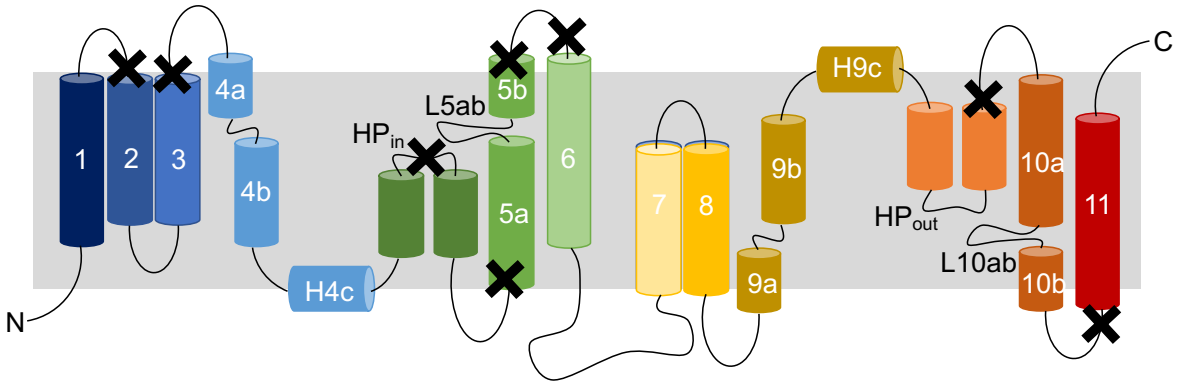


Fig. 9.2.2: the topology of the 11 TM helices and connecting loops of VcINDY in the membrane (grey), displaying the approximate residue positions where cysteines will be introduced (marked with crosses). The colours show the progression of the chain from the N-terminal (blue) to C-terminal (red) ends; thus, in order from left to right, the residue sites are E42 (TM2), F79 (TM3), W148 (HP_{in}), T177 (TM5a), A208 (TM5b), G211 (TM6), E394 (HP_{out}) and S436 (TM11).

cw-EPR & simulations

Room temperature X-band

Here, the results will be reported for four of the eight designed variants. Variants A208C, G211C, E394C and S436C were spin labelled with MTSL (details in section 13.1 ‘Site directed spin labelling’). The labelling efficiencies and protein concentrations of these samples varied across the variants:

- A208C: 28 μM, 168%
- G211C: 50 μM, 241 %
- E394C: 6 μM, 68 %
- S436C: 40 μM, 219 %

Labelling efficiency was calculated using a calibration curve of spin label at known concentrations (double integration of the first derivative signal gives the area under the curve, which is related to the concentration of the radical species in the sample). The room temperature cw-EPR spectra of these spin-labelled VcINDY variants were recorded at X-band in the apoprotein state (fig. 9.2.3).

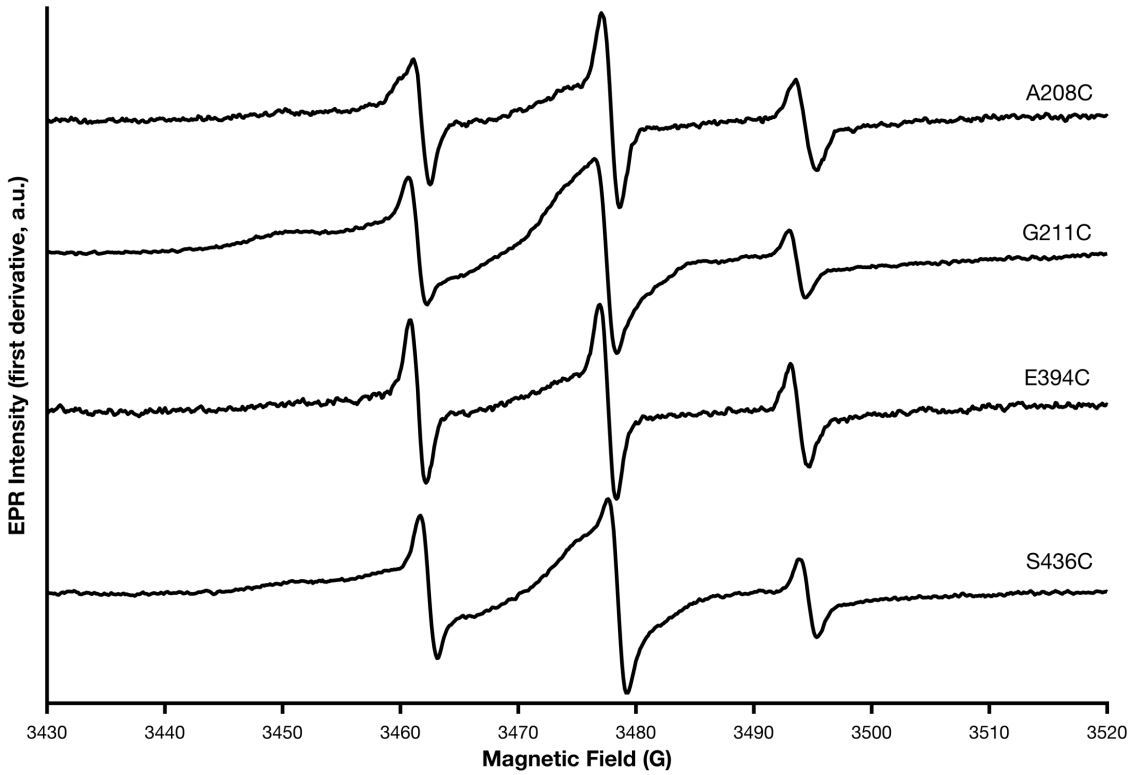


Fig. 9.2.3: the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled variants of VcINDY at X-band; from top to bottom: A208C, G211C, E394C and S436C. All spectra have been normalised to same scale.

Across these labelled variants, there is a wide range of spectral ‘shapes’, with features characteristic of free label exhibited to varying extents. The variants in which this is most obvious are A208C and E394C.

It is unclear exactly as to why this is but may be an effect of the accessibility of the labelling sites in the apoprotein/inward-facing state, as this is the conformation the protein will be in during the labelling process. The *in silico* labelling experiments predict variants G211C and S436C to have large rotamer libraries in the inward-facing form (136 for G211C and 61 & 142 for S436C). In contrast, the predictions for A208C and E394C in the same form are much lower (1 for A208C and 42 & 48 for E394C). As a result, much of the label may not be able to access the cysteine sites, and excess free label is then held in the hydrophobic detergent micelles. The presence of excess free label is further confirmed by several of the labelling efficiencies of >100 %; i.e. where there is more signal from spin label than there are cysteine sites (as determined by protein concentration multiplied by the number of cysteine sites per protein). Extra washing steps may reduce the amount of free label, but certain methods may lead to loss of protein.

These spectra were then simulated and the simulations were fitted to the experimental data using the EasySpin package[10] for MATLAB (figure 9.2.4). At least two species will need to be simulated here for each variant considering the significant influence of potentially free label on the spectral characteristics as discussed previously.

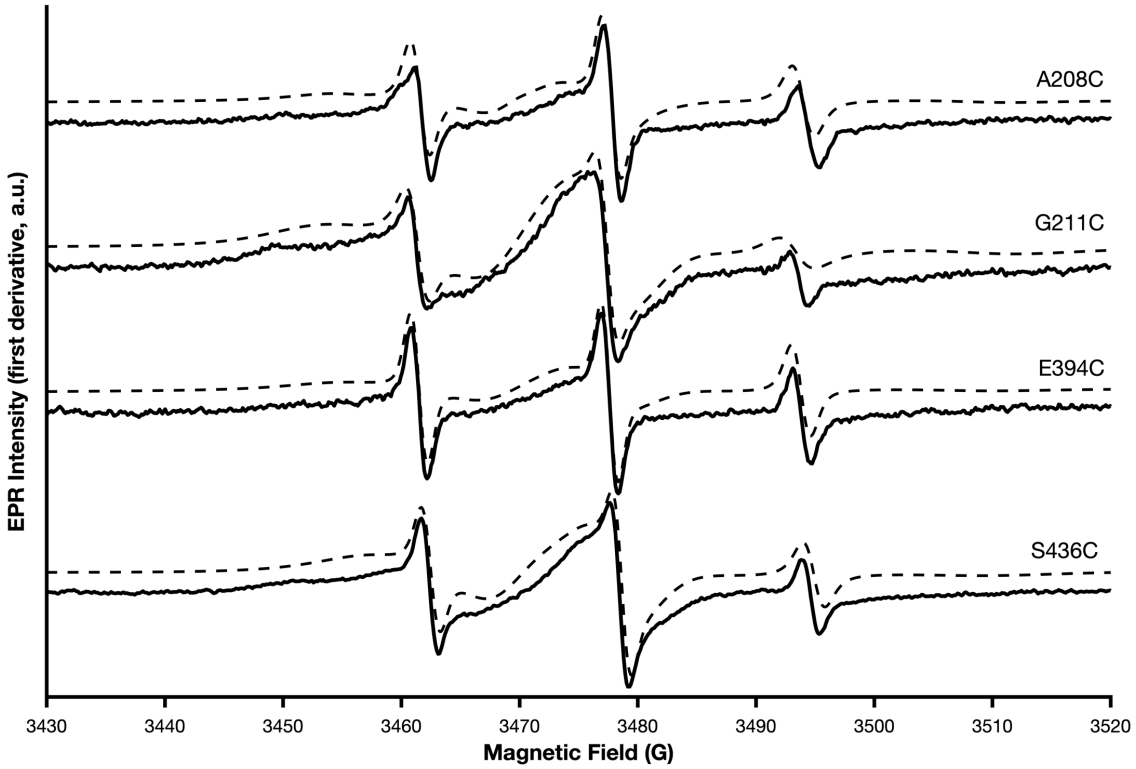


Fig. 9.2.4: the fitted simulations of the room temperature cw-EPR spectra of the apoprotein, MTSL-labelled variants of VcINDY at X-band; from top to bottom: A208C, G211C, E394C and S436C. Experimental data are shown as solid lines, fitted simulations are shown as dotted lines.

Table 9.2.3: best fit simulation values for the x, y and z components of each of the g- and A-tensors, as well as linewidth and averaged correlation times for each of the room temperature X-band experimental data of the MTSL-labelled VcINDY. The respective, relative weightings are provided for each spin label ' species' component. The RMSD of each fitting is given.

Component	g-tensor			A-tensor (MHz)			Linewidth (peak-to-peak, mT)	τ_c (ns)	Relative weighting	Simulation RMSD
	x	y	z	x	y	z				
1	2.0083	2.0057	2.0018	14	14.5	95.5	0.01	41.3	0.77	0.053694
2	2.0076	2.0059	2.0019	15.1	14.2	106.8	0.11	0.7	0.23	
1	2.0079	2.0063	2.0019	14.1	15.5	101.5	0.3	43.5	0.9	0.052429
2	2.008	2.0059	2.0021	14.3	14.1	107.4	0.08	0.9	0.1	
1	2.0079	2.0063	2.0021	14.2	15.6	96.8	0.3	9	0.85	0.043708
2	2.0081	2.0056	2.002	15.3	15.1	105.7	0.11	0.3	0.15	
1	2.0079	2.0063	2.0021	13.5	14.4	98.9	0.05	68.2	0.91	0.046069
2	2.0078	2.0056	2.002	13.6	14.7	108.2	0.11	0.5	0.09	

The best fit results (table 9.2.3) across the different variants are all fairly different, as would be expected from the range of spectral features observed. As previously discussed, room temperature cw-EPR spectra do not allow for the resolving of individual components of the g- and A-tensors, though the isotropic equivalents of these (g_{iso} and A_{iso}) can be calculated from the values extracted in the fitted simulations. However, the main focus of room temperature cw-EPR studies on spin labelled protein is the extracted rotational correlation time, τ_c.

All of the fittings were done with two spin label species simulated: one with a very short τ_c (in the range of < 1 ns, i.e. 'free label') and one with a much longer τ_c (in the range of tens of ns, i.e. 'bound label'). The two variants that showed the sharper features characteristic of free label more obviously, A208C and E394C, had higher relative contributions of the short τ_c compared to the two variants that demonstrated broader and more asymmetric 'bound' label features (G211C and S436C). Considering the amount of free label that appears to exist in most of these samples, it is hard to compare the results and these spectra beyond this semi-quantitative assessment.

VcINDY variant G211C was selected as the most viable to pursue for more in-depth investigations, as it showed the most success in the labelling experiments. In addition, this variant had a relatively short predicted inter-label distance to measure in the inward-facing state, with a dramatic difference expected upon conformational change.

To begin with, the room temperature cw-EPR spectra were recorded at X-band for spin labelled G211C under different binding conditions: in the apoprotein state, then with excess sodium, and finally with excess sodium and succinate (fig. 9.2.5).

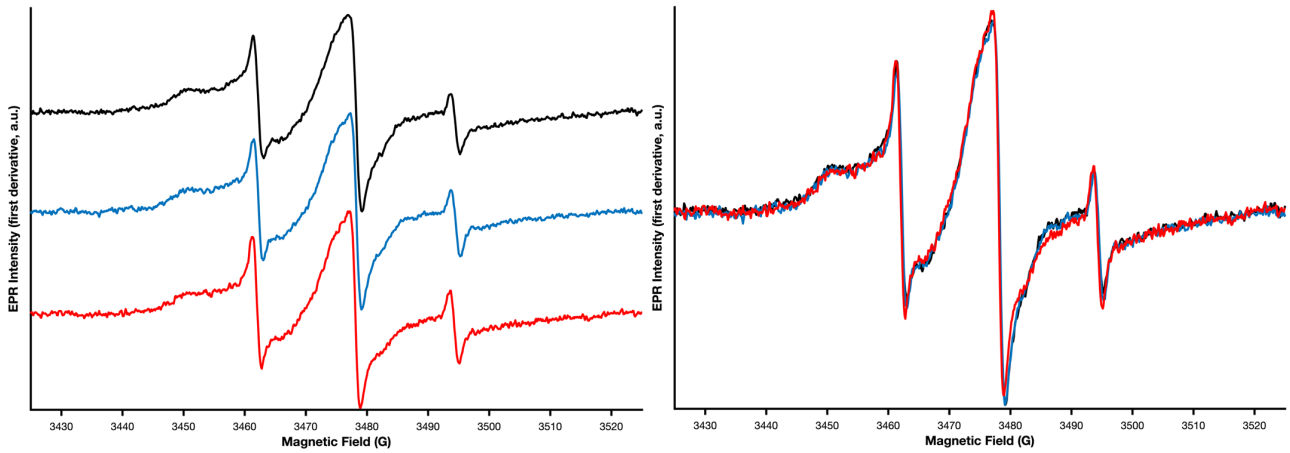


Fig. 9.2.5: the room temperature cw-EPR spectra of the MTSL-labelled G211C variant of VcINDY at X-band under various conditions; apoprotein in black, with 50 mM sodium in blue, and with 50 mM sodium and 0.8 mM succinate in red. On the left the spectra are stacked to show them individually, and on the right are overlain to highlight the similarities across the binding states. All spectra have been normalised to same scale.

Addition of excess sodium and substrate would be expected to drive conformational change from the inward- to outward-facing arrangement. However, if this is the case it did not cause any visible changes in the spectral features seen here. There was no significant change to the overall shape, indicating there was no change in the T_c values of the attached spin labels for each of the three different binding states.

Considering the *in silico* labelling experiments, the conformational change is expected to significantly alter the number of rotamers at the 211 position. In the inward-facing state there are 136 rotamers predicted, whereas the outward-facing conformation is expected to have very restricted rotational freedom with only 1 rotamer predicted. This may not be exhibited in the spectra for a number of reasons, either relating to the predictions and therefore structural data, or relating to the experimental conditions (or both).

Issues with the predictions stem from the assumptions made in assessing structural data. For example, the inward-facing conformation is based a crystal structure and the outward-facing conformation is based on a molecular dynamics (MD) simulation (which itself is based on the crystal structure of the inward-facing form). These may or may not be accurate snapshots of the physiological conformations VcINDY actually samples. The *in silico* experiments also take these inputs to be rigid bodies and therefore may not generate libraries of rotamers that

are reflective of the true movements of the attached spin labels in detergent-solubilised protein.

The experimental conditions may also contribute to the results not reflecting the predictions; detergent solubilised protein may not act in the same way as protein in its native state. Additionally, the attachment of spin labels may in some way affect the normal function of the protein and may hinder the dynamic processes it undergoes. Though the concentration of the substrate used in this experiment was expected to be enough to cause the protein to change conformation (as it was at least 2 orders of magnitude in excess), it may not have been sufficient for the outward-facing form to exist long enough, or in high enough proportion, to observe with this method.

Low temperature X-band

The spin labelled G211C variant was then studied using low temperature (50 K) cw-EPR at X-band (fig. 9.2.6). Again, the influence of substrate was investigated by comparing the spectrum of the apoprotein against that of the protein incubated with excess sodium and succinate. Due to the lack of observable changes in the room temperature cw-EPR spectra across the binding states, a greater excess (5 mM) of substrate was used here. This was an attempt to 'force' the conformational change such that all VcINDY dimer units in the sample would be saturated with substrate, and there would be observable differences in the spectra.

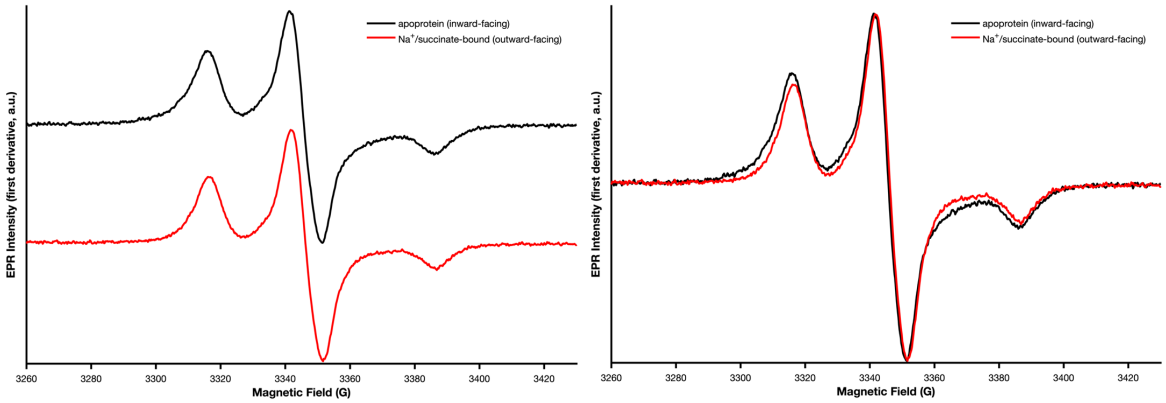


Fig. 9.2.6: the low temperature (50 K) cw-EPR spectra of the MTSL-labelled G211C variant of VcINDY at X-band under various conditions; apoprotein in black, and with 5 mM sodium succinate in red. On the left the spectra are stacked to show them individually, and on the right are overlain to highlight the relative broadness of the spectra across the binding states. All spectra have been normalised to same scale.

The spectrum of the apoprotein (inward-facing) form appears to be significantly broadened in comparison to that of the Na^+/succinate-bound (outward-facing) state. This is most likely due to the dipolar interaction of the two spin labels at a distance of less than approximately 25 Å.

The simulation and fitting of the data was therefore performed on the spectrum of the Na^+/succinate-bound (outward-facing) form of spin-labelled G211C, using the EasySpin package[10] for MATLAB as before (fig. 9.2.7 and table 9.2.4). The results for the g- and A-tensors of the fitted simulations of the room temperature cw-EPR spectra at X-band were taken as starting values here.

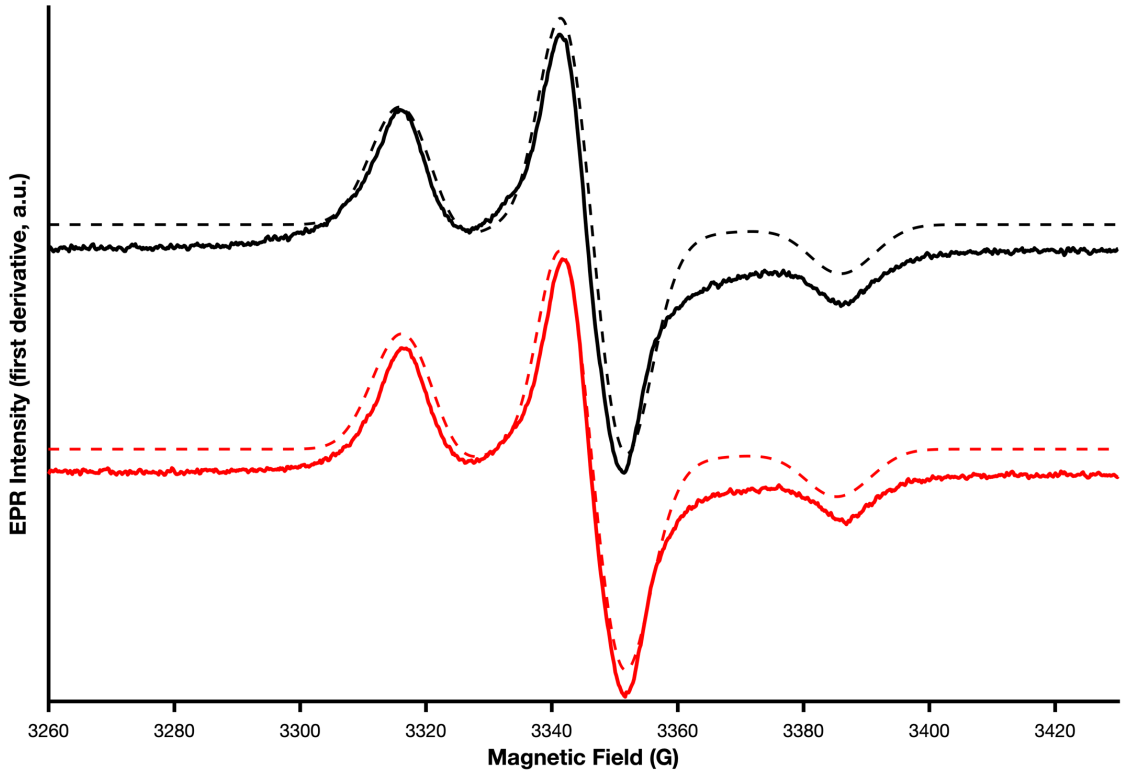


Fig. 9.2.7: the fitted simulations of the low temperature (50 K) cw-EPR spectra of the MTSL-labelled G211C variant of VcINDY at X-band in various binding states: apoprotein in black, and with 5 mM sodium succinate in red; experimental data are shown as solid lines, fitted simulations are shown as dotted lines. All spectra have been normalised to same scale.

Table 9.2.4: best fit simulation values for the x, y and z components of each of the g- and A-tensors and linewidth for the low temperature X-band experimental data of the MTSL-labelled G211C variant of VcINDY in various binding states: apoprotein and with 5 mM sodium succinate. Two spin label ‘ species’ are simulated for the fitting of a single spectrum; the respective, relative weightings are provided for each component. The RMSD of the fitting is given.

Conformation	**Component**	**g-tensor**			**A-tensor (MHz)**			**Linewidth (peak-to-peak, mT)**	**Relative weighting**	**Simulation RMSD**
		x	y	z	x	y	z			
Inward-facing (apoprotein)	1	2.0079	2.0055	2.0017	14.4	15.7	95.1	0.9	0.78	0.066205
	2	2.0082	2.0057	2.0015	14.6	15.1	109.2	0.8	0.22	
Outward-facing (Na^+/succinate -bound)	1	2.0079	2.0055	2.0017	14.4	15.7	93.5	0.9	0.76	0.037663
	2	2.0082	2.0057	2.0015	14.6	15.1	107.5	0.8	0.24	

A reasonably good fit (low RMSD) was achieved for the simulation and experimental data of the Na^+/succinate-bound form. The linewidths and the g- and A-tensors were mostly unvaried when used to simulate and fit the experimental data of the apoprotein state. However, the A_{zz} components of the A-tensors and the relative weightings of both spin label 'species' were allowed to vary to account for the effect of the dipolar broadening.

The best fit values for the relative weightings (±0.02) and the general scale of the A_{zz} values remain relatively consistent between the two binding states. However, the apoprotein A_{zz} values for both components are shifted higher by about 1.6 to 1.7 MHz. This is likely an effect of the fitting routine attempting to match the simulation to the broadened features of the experimental spectrum. It is also worth noting the much higher RMSD value for the apoprotein fitting compared to that of the Na^+/succinate-bound form, as a good fit cannot be achieved without accounting for the actual dipolar interactions of spin labels in close proximity.

The spectra were then fit using a different software package, pyDipfit[11], which allows for extraction of an estimated distance based on the dipolar coupling. This method requires the fitting of the spectrum of a singly labelled species first. For this, the spectrum of the outward-facing (Na^+/succinate-bound) form was used, as there is no dipolar broadening effect observed (fig. 9.2.8, table 9.2.5).

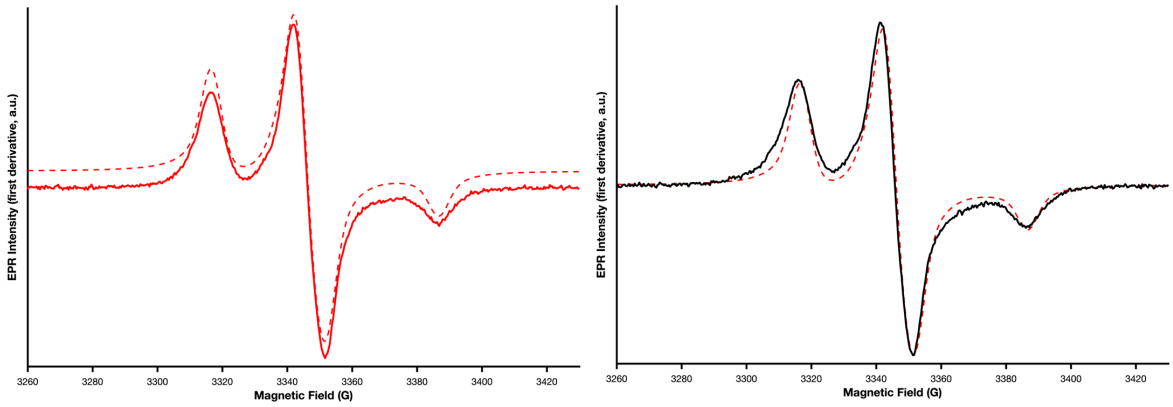


Fig. 9.2.8: left, the fitted simulation of the low temperature (50 K) cw-EPR spectrum of the MTSL-labelled G211C variant of VcINDY with 5 mM sodium succinate (at X-band). Right, the same simulation (red) of the 5 mM sodium succinate-bound sample overlain with the low temperature (50 K) cw-EPR spectrum of the MTSL-labelled G211C variant of apoprotein VcINDY (at X-band, black). Experimental data are shown as solid lines, fitted simulations are shown as dotted lines. All spectra have been normalised to same scale.

The results of this fitting are then mostly preserved for fitting the spectrum of the doubly labelled species (fig. 9.2.9, table 9.2.5). In this case, this is the inward-facing (apoprotein) form.

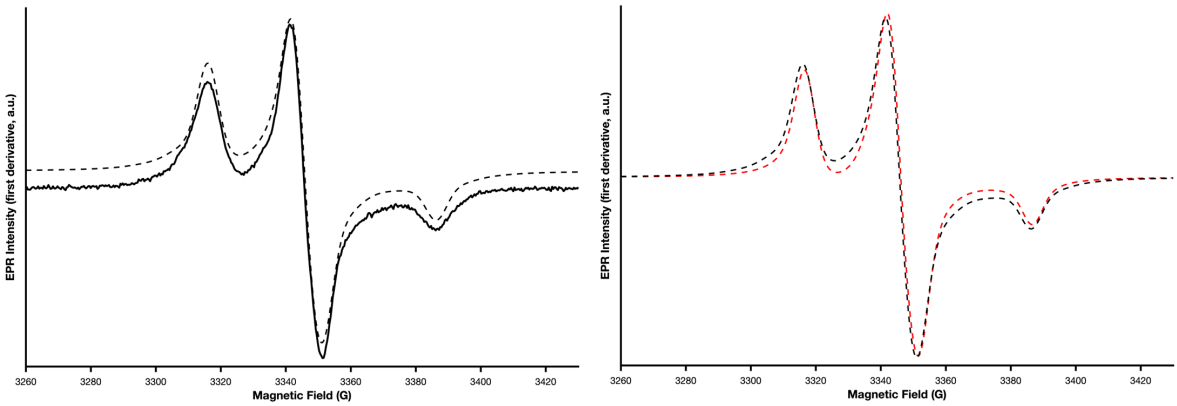


Fig. 9.2.9: left, the fitted simulation of the low temperature (50 K) cw-EPR spectrum of the MTSL-labelled G211C variant of apoprotein VcINDY (at X-band). Right, the simulation (red) of the 5 mM sodium succinate-bound sample overlain with the simulation of the apoprotein sample (black). Experimental data are shown as solid lines, fitted simulations are shown as dotted lines. All spectra have been normalised to same scale.

Table 9.2.5: best fit simulation values for the x, y and z components of each of the g- and A-tensors and linewidths for the low temperature X-band experimental data of the MTSL-labelled G211C variant of VcINDY in various binding states: apoprotein and with 5 mM sodium succinate. The RMSD of the fitting is given.

Conformation	**g-tensor**			**A-tensor (G)**			**Linewidths (G)**			**Simulation RMSD**
	x	y	z	x	y	z	Lorentzian	Gaussian	Ratio	
Inward-facing (apoprotein)	2.0087	2.0061	2.0024	0.77	5.15	35.03	7.9	3	0.41	0.24
Outward-facing (Na^+/succinate -bound)	2.0087	2.0061	2.0024	0.77	5.15	34.96	5.6	2.7	0.59	0.23

Only certain parameters should be varied here, such as A_{zz} in case the polarity of the environment of the spin labels changes between the two binding states. Different estimates of the distance between the spin labels are calculated based on different approximations. The distance calculated with the strong-coupling approximation was 10.5 Å, and with the weak-coupling approximation was 9.2 Å. This is significantly different to the *in silico* labelling experiment prediction, which gave an inter-label distance of 26.9 Å in the inward-facing conformation.

It is important to point out that the estimates from pyDipfit were calculated based on a few assumptions:

- the spectrum of the Na^+/succinate-bound form was indeed unaffected by any dipolar interactions of the two spin labels (i.e. there are no spin label pairs interacting at distances less than around 25 Å).
- all cysteine sites were labelled (i.e. 100% labelling efficiency), and all of these labels were interacting in pairs at distances less than around 25 Å in the apoprotein form.

The latter is a large assumption to make. From the room temperature cw-EPR experiments, it has already been calculated that the labelling efficiency is above 100% (241 %). However, considering that this efficiency inaccurate due to the presence of excess free label in the sample, it is likely that there will be some units of VcINDY dimer with two attached spin labels (which can interact), some with one label, and some with none. This value was not varied for the fitting, as attempts to fit gave a ‘fraction’ of > 1 (i.e. the fraction of interacting spins at < 25 Å with respect to all spins in the sample), which is not by definition viable.

Additionally, the room temperature cw-EPR spectrum of G211C showed some evidence of free label. This is significant with respect to the first assumption; if there is free label, potentially in the detergent micelles, that are close enough to the units of VcINDY protein that are solubilised, they may interact with the labels attached at the G211C site.

Pulsed EPR

Three of the VcINDY variants (A208C, G211C and S436C) were investigated with PELDOR under different binding conditions, with the majority of the work done so far using variant G211C. All PELDOR data were processed using the DeerAnalysis[12] software package for MATLAB.

First, the experiment was performed using the samples investigated with low temperature cw-EPR (i.e. with sodium buffer and 5 mM succinate) with added d_8-glycerol (fig. 9.2.10).

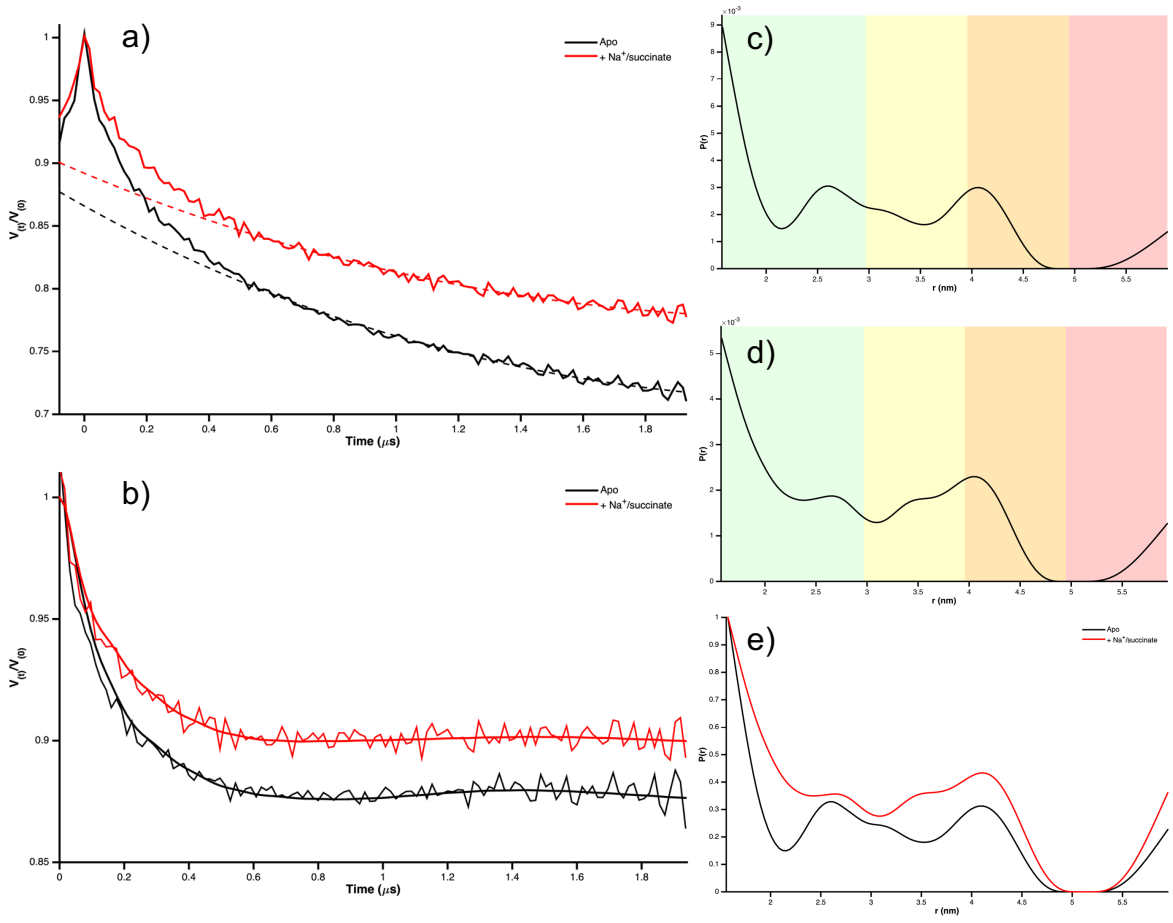


Fig. 9.2.10: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled G211C VcINDY: a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and Na^+/succinate-bound form (5 mM), red, b) background factor subtracted time trace (apoprotein in black, Na^+/succinate-bound in red), c) apoprotein distance distribution output, d) Na^+/succinate-bound distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black Na^+/succinate-bound in red).

Assessment of just the raw time trace data shows an initially steep drop-off in echo intensity with no immediately observable oscillations over the recorded time. This sharp feature at the zero-time often indicates a distance of less than about 20 Å between the two spin labels; this is demonstrated in the distance distributions.

Background factor correction shows there may be a very shallow, long oscillation towards the end of the measured time. However, with the level of signal-to-noise achieved here and the short dipolar evolution time measured (2 μs), the level of confidence in this analysis should be low. This is reiterated in the colour-coded distance distributions, which place any populations above about 35 Å in the low confidence band.

The short distance indicated by the initial steep decay in echo intensity and sharp feature at zero-time agrees with the low temperature cw-EPR analysis, where there is a distance of < 20 Å (9.2 to 10.5 Å as calculated using pyDipfit[11], using the weak- and strong-coupling approximations, respectively) in the inward-facing (apoprotein) form. However, the persistence of this PELDOR feature in the outward-facing (Na^+/succinate-bound) form is unexpected, as it had been assumed that there was no dipolar broadening in the equivalent low temperature cw-EPR spectrum.

Additionally, the *in silico* labelling experiments had predicted G211C to go from 26.9 Å to 78.5 Å in the inward- and outward-facing forms, respectively. A distance of close to 80 Å should be close to the limits of detection for routine PELDOR experiments at X-band. However, in both the raw and background-corrected time traces both samples appear to decay in much the same way, with little difference in the extracted distance distributions, except that the features are broader in the outward-facing form (which cannot be trusted as accurate considering the signal-to-noise and measurement time, as previously explained).

Next, this experiment was repeated but with higher substrate concentration in the outward-facing sample (20 mM succinate). Also, the protonated buffer was exchanged deuterated buffer (see section 13.1 ‘Buffer deuteration’) to attempt to make the signal last longer and increase the measurement time (fig. 9.2.11).

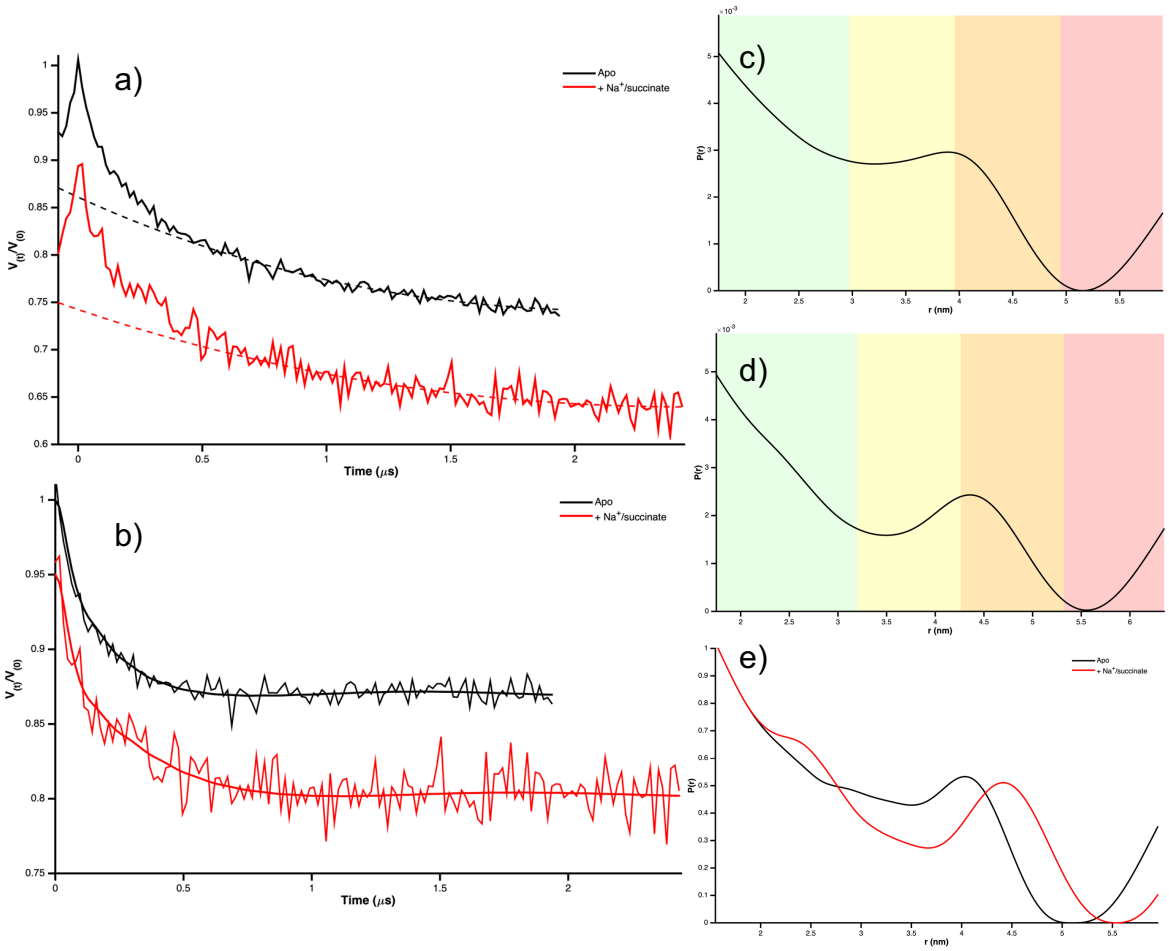


Fig. 9.2.11: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled G211C VcINDY in deuterated buffer: a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and Na^+/succinate-bound form (20 mM), red, b) background factor subtracted time trace (apoprotein in black, Na^+/succinate-bound in red), c) apoprotein distance distribution output, d) Na^+/succinate-bound distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black Na^+/succinate-bound in red).

This appeared to make little difference; the time traces both show the steep feature at zero-time with decays that demonstrate little-to-no observable oscillation. The experiment for the outward-facing form was measured with a slightly longer dipolar evolution time (2.5 μs) to check for a low frequency oscillation, but this affected the achievable signal-to-noise over the same sort of time frame as the inward-facing form (i.e. several days). Again, the two distance distributions match one another and no discernible difference between the conformations is seen.

It was noted that with the poor signal-to-noise and lack of obvious oscillations it was very difficult to apply background correction without artificially introducing shape to the time traces. In addition, the L-curves yielded by Tikhinov regularisation did not exhibit their characteristic shape, making it hard to select a

reasonable regularisation parameter. This results in either over-resolving or over-smoothing of the distance distributions. This should be considered when assessing the distance distributions and results presented here, where in general they err on the side of over-smoothing (i.e. broad features across a range of distances, rather than multiple sharp populations at distinct distances).

Equivalent experiments were performed on the S436C and A208C variants (fig. 9.2.12 and fig 9.2.13, respectively). In both of these cases, deuterated buffer was not used, as in the former G211C experiment. This was decided based on the balance of sample manipulation time and cost against the lack of significant changes seen in the G211C PELDOR traces. The first of these variants to be addressed is S436C; this is at the opposite face of the dimer unit to G211C, so the opposite trend of distance changes upon conformational change is expected.

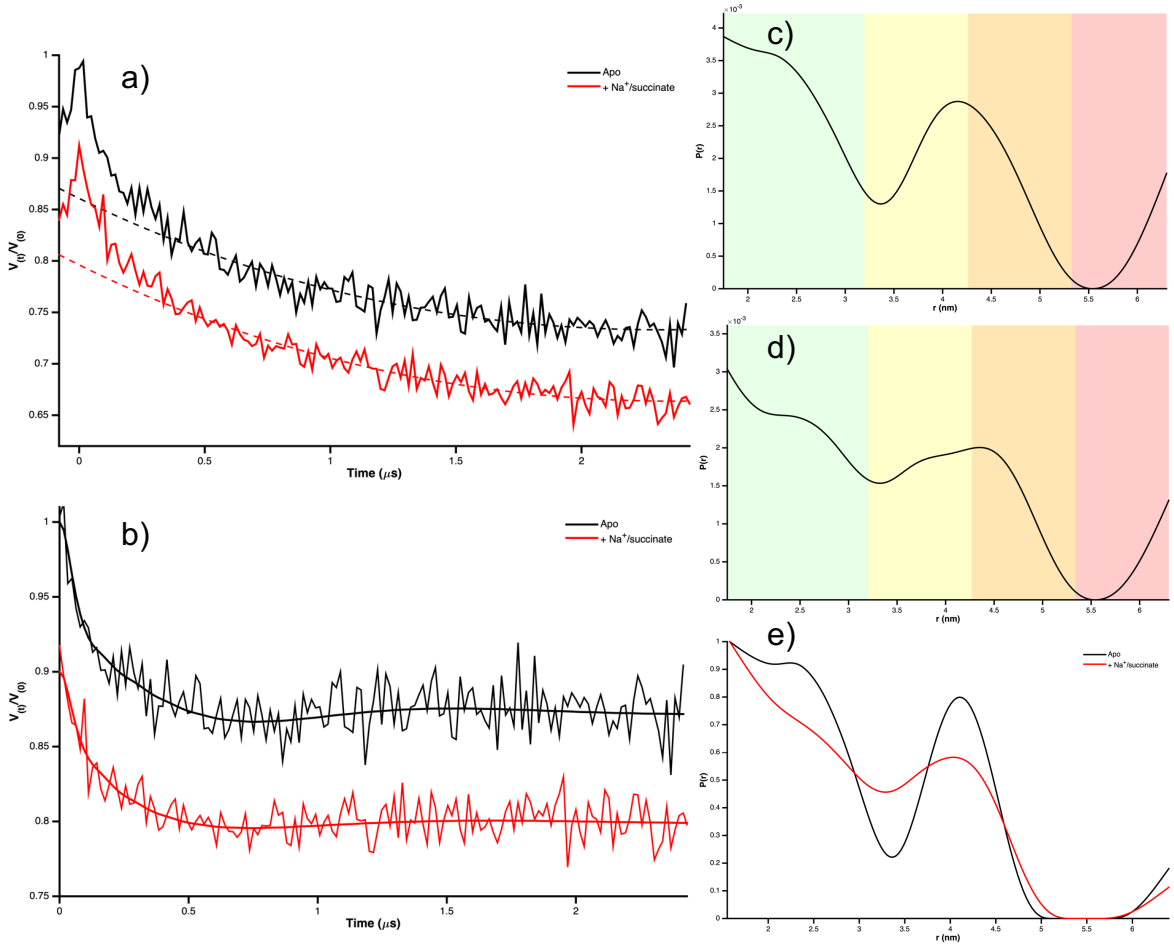


Fig. 9.2.12: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled S436C VcINDY: a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and Na^+/succinate-bound form (20 mM), red, b) background factor subtracted time trace (apoprotein in black, Na^+/succinate-bound in red), c) apoprotein distance distribution output, d) Na^+/succinate-bound distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black Na^+/succinate-bound in red).

Here, the raw time trace data show even greater noise levels than the equivalent data recorded for the G211C variant; both samples were recorded for similar time periods, but differences in protein and effective spin label concentration will affect this. A similarly sharp feature at the start of the trace (which is further emphasised upon background correction) is observed here, which results in a short distance (< 20 Å) in the extracted distance distributions. The poor signal-to-noise here may overexaggerate this feature and hide any oscillations across the measured time (2.5 µs). It is therefore not possible to discern any real differences in the data between the two binding states.

The *in silico* labelling experiments predicted a change from 73.0 Å to 49.5 Å for the inward- to outward-facing forms, respectively. Here, the dipolar evolution time measured (2.5 µs) is not likely to be sufficient to observe oscillations at the

corresponding frequencies. In order to achieve this, far higher concentration of efficiently labelled protein will be necessary, along with deuterated buffer and cryoprotectant.

Finally, initial PELDOR data for the A208C variant was measured (fig. 9.2.13) for both binding states. The concentration of protein (approximately 28 μM prior to addition of d_8-glycerol) severely limited the echo intensity and signal lifetime such that a dipolar evolution time of only 750 ns was achieved. As a result, the distance distributions have a low level of confidence for populations beyond about 25 Å.

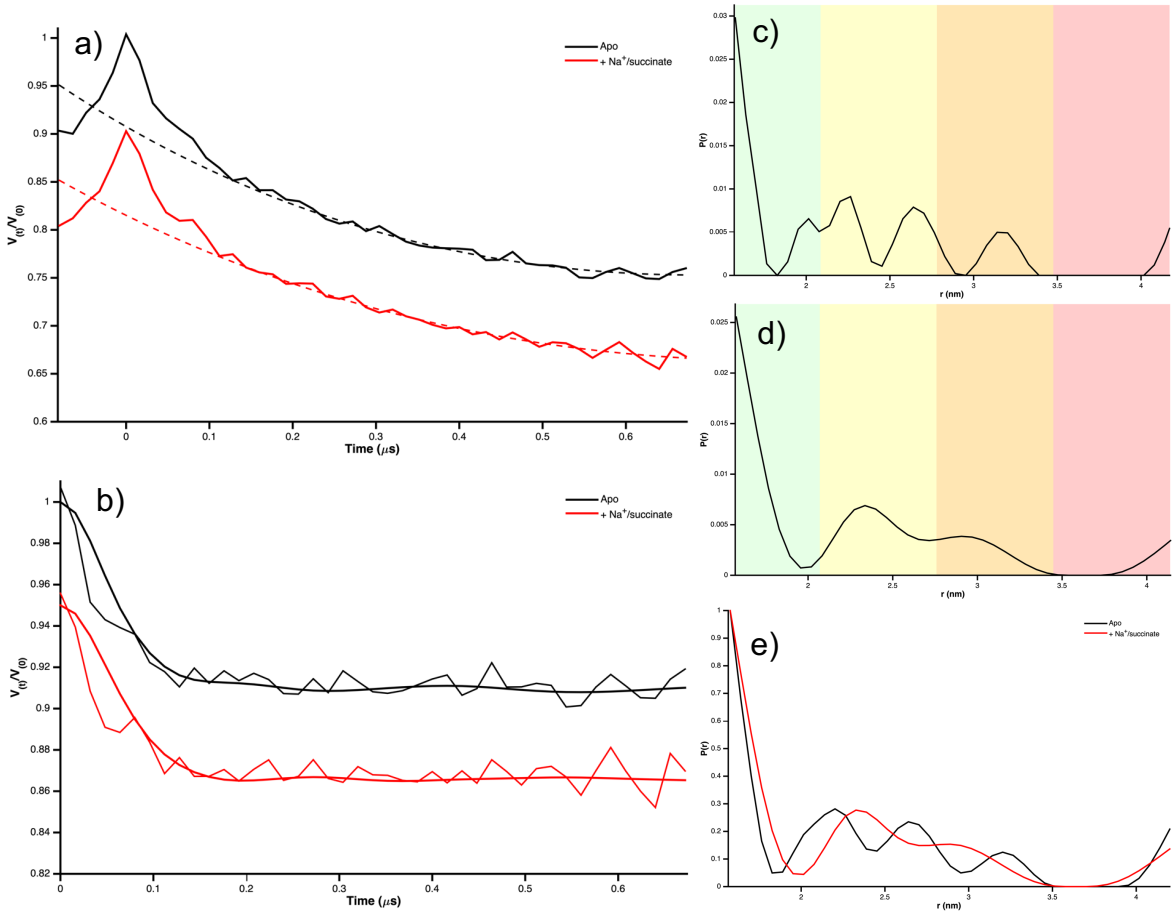


Fig. 9.2.13: the 4P-PELDOR data and resulting distance distributions of MTSL-labelled A208C VcINDY: a) 'raw' data (trace of the echo decay and oscillation in the time domain) of apoprotein form, black, and Na^+/succinate-bound form (20 mM), red, b) background factor subtracted time trace (apoprotein in black, Na^+/succinate-bound in red), c) apoprotein distance distribution output, d) Na^+/succinate-bound distance distribution output, e) scale-normalised overlay of the distance distributions (apoprotein in black Na^+/succinate-bound in red).

Much like both previous variants presented here, the A208C data show sharp features at the start of the time traces, further exaggerated by the short measurement time (750 ns). Here, however, it is expected for the inward-facing (apoprotein) form considering the *in silico* labelling experiment prediction of 23.8

Å. This is predicted to change to 50.4 Å upon conformational change to the outward-facing (Na^+/succinate-bound) form. This certainly would not be observable at the over the dipolar evolution time measured here. The lack of data points measured here, along with the difficulties in processing the data as explained previously, also likely rationalises the sharp shapes observed in the distance distribution for the inward-facing form here.

As described above for each variant, there is very little observable change in the PELDOR results regardless of the concentration of incubated substrate. Also, the persistence of the sharp and steep initial decay feature indicates a problem across the variants. It may be the case that the 'free' label observed in the room temperature cw-EPR experiments are causing issues with the dipolar coupling measured here using PELDOR.

It is generally well known that membrane transporters can be difficult to work with, in that they typically are produced to significantly lower yields (i.e. lower units of protein per litre of growth medium) than their soluble protein counterparts. Additionally, there are often other complications in the labelling process, such as being solubilised in detergent. This can make the labelling process difficult as spin labels can get 'stuck' in these artificial hydrophobic environments.

The steps that are often taken to increase PELDOR signal lifetime in other systems are not necessarily viable here. Since liposomes and detergents are not normally deuterated (the financial cost of this would probably be very high), there is a lot of opportunity for surrounding protons in the local environment to influence the relaxation of the system and echo modulation, despite deuterating the aqueous buffer (which was shown to have limited impact for the G211C variant). Achieving sufficient labelled protein concentration for PELDOR is already complicated by the risk of the solubilised protein (either in detergent or liposomes) precipitating out. Further dilution by addition of more deuterated glycerol (i.e. to a higher v/v %) has to be considered in balance with the potential increase in signal lifetime.

10 Summary & Conclusions

The studies performed here for VcINDY are largely preliminary, with the most in depth investigation being for the G211C variant. The room temperature cw-EPR studies showed varying degrees of successful labelling of variants A208C, G211C, E394C and S436C, with certain variants showing greater degrees of label attachment (G211C and S436C) than the others (A208C and E394C). The G211C variant was studied under various conditions: in non-binding K^+ salt buffer without substrate, in binding Na^+ salt buffer without substrate, and in binding Na^+ salt buffer with substrate. No spectral changes were observed across these samples, despite the *in silico* labelling experiment predictions that VcINDY changing conformations to the outward-facing form (i.e. upon Na^+/succinate binding) would significantly alter the rotational freedom of the attached spin label.

Low temperature cw-EPR measurements, however, did exhibit observable differences for G211C under different binding conditions. There was a significant degree of dipolar broadening in the inward-facing (apoprotein) form of VcINDY. This was vastly reduced upon conformational change (Na^+/succinate binding) to the outward-facing form, thereby supporting the *in silico* labelling experiment predictions.

Further to this, PELDOR measurements were done using the same samples; the time trace data showed no clear oscillations within the measurable dipolar evolution time. Considering the previous results of the low temperature cw-EPR studies, the presence of a short distance was expected from the inward-facing form, and was observable as a sharp, steep initial decay in echo intensity. However, this feature persisted in the presence of excess substrate. In an effort to gain greater signal lifetime and force the conformational change to the outward-facing form, deuterated buffer and further excess of substrate was used. This had little effect on either the measured dipolar evolution time or yielded no difference in the time trace data.

Beyond this, the other variants measured thus far using PELDOR spectroscopy, S436C and A208C, also displayed this sharp initial feature. It is therefore unclear what is causing this, and all of the PELOR results must be reassessed for further experiments.

11 Outlook

As previously described, the room temperature cw-EPR studies showed varying degrees of success in terms of labelling efficiency, and sample quality (i.e. how much of the free label removed during the washing process). This protocol should be refined such that sufficient efficiency (roughly >70 %) is consistently achieved, along with reliably removing excess free label and retaining as much protein as possible. This may be achieved by labelling the protein during the purification process (i.e. when it is bound to a purification column) to minimise protein losses and incorporate the reduction, label incubation and washing steps.

Following this, the room temperature cw-EPR studies on G211C should be repeated with the equivalent substrate balance as in the low temperature cw-EPR studies, as conformational change appeared to be evident in the latter experiments. The expectation, based on the *in silico* labelling experiment predictions, would be that the rotational mobility would be severely reduced (and therefore the τ_c value, extracted by simulation and fitting, would increase significantly). Further to this, the same experiments should be performed on the other designed variants that are expected to significantly change in terms of label environment upon conformational change, most prominently:

- A208C (~1 or 2, and 108 rotamers in the inward- and outward-facing forms respectively)
- W148C (~ 5 to 8, and 99 rotamers in the inward- and outward-facing forms respectively)

Similarly, upon improvement of the SDSL process for VcINDY, the low temperature cw-EPR experiments should also be performed with the A208C variant, for which the *in silico* labelling experiments yielded an inter-label distance prediction of 23.8 Å in the inward-facing conformation. Development of the simulation and fitting routine may be possible to account for dipolar interactions for distances below about 25 Å with newer releases of the MATLAB package EasySpin[10], in which case the G211C data should be reanalysed and the results compared with the distances extracted using pyDipfit[11].

Interpretation of the PELDOR results is difficult considering the poor signal-to-noise and measurement of only short dipolar evolution times. It appears that issues exist around these measurement, likely due to VcINDY being a solubilised

membrane protein. It is not clear whether the presence of free label and/or detergent is interfering with the measurement, or if these measurements are made difficult simply by virtue of the nature of membrane proteins, the limitations of which have been explained previously (concentration, signal lifetime etc). Either way, it is likely that the refinement of the SDSL process will also yield better results here.

Further to this, reconstitution of the labelled protein into proteoliposomes would be desirable as a more physiologically relevant medium for VcINDY to be studied in. However, this may present new problems such as liposome homogeneity and controlling the orientation of the functional dimer units. In addition, the process of reconstitution often decreases the effective concentration of the protein, making the PELDOR measurement even more challenging.

Continuation of the studies on this system based on the designed variants and these preliminary results will be undertaken by another PhD candidate, with a view to branch into NMR spectroscopy for STD studies. This may help to elucidate the third Na^+ binding site, binding order, monomer cooperativity/independence and provide a point of comparison for the reported K_d value.

In order to check the effect of spin labelling on the protein, transport assays of the labelled variants could be performed. This may be done with proteoliposomes, wherein the external media contains the cognate cation Na^+ whilst the internal media contains non-cognate K^+ to cause unidirectional transport (and thereby account for the orientation of the VcINDY units in the liposome membrane). Radiolabelled (^{14}C-containing, for example) 'hot' ligand could be used and the accumulated concentration monitored and compared with unlabelled variant and wild-type VcINDY. The same method of transport where a Na^+/K^+ gradient is created across the proteoliposomes membranes could be utilised for the PELDOR measurements.

References

1. Rogina, B.; Reenan, R. A.; Nilsen, S. P.; Helfand, S. L., Extended life-span conferred by cotransporter gene mutations in Drosophila. *Science* **2000,** *290* (5499), 2137-2140.

2. Birkenfeld, A. L.; Lee, H. Y.; Guebre-Egziabher, F.; Alves, T. C.; Jurczak, M. J.; Jornayvaz, F. R.; Zhang, D. Y.; Hsiao, J. J.; Martin-Montalvo, A.; Fischer-Rosinsky, A.; Spranger, J.; Pfeiffer, A. F.; Jordan, J.; Fromm, M. F.; Konig, J.; Lieske, S.; Carmean, C. M.; Frederick, D. W.; Weismann, D.; Knauf, F.; Irusta, P. M.; De Cabo, R.; Helfand, S. L.; Samuel, V. T.; Shulman, G. I., Deletion of the Mammalian INDY Homo log Mimics Aspects of Dietary Restriction and Protects against Adiposity and Insulin Resistance in Mice. *Cell Metabolism* **2011,** *14* (2), 184-195.

3. Toivonen, J. M.; Walker, G. A.; Martinez-Diaz, P.; Bjedov, I.; Driege, Y.; Jacobs, H. T.; Gems, D.; Partridge, L., No influence of Indy on lifespan in Drosophila after correction for genetic and cytoplasmic background effects. *Plos Genetics* **2007,** *3* (6), 973-983.

4. Tretter, L.; Patocs, A.; Chinopoulos, C., Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis. *Biochimica Et Biophysica Acta-Bioenergetics* **2016,** *1857* (8), 1086-1101.

5. Mulligan, C.; Fitzgerald, G. A.; Wang, D. N.; Mindell, J. A., Functional characterization of a Na+-dependent dicarboxylate transporter from Vibrio cholerae. *Journal of General Physiology* **2014,** *143* (6), 745-759.

6. Mancusso, R.; Gregorio, G. G.; Liu, Q.; Wang, D. N., Structure and mechanism of a bacterial sodium-dependent dicarboxylate transporter. *Nature* **2012,** *491* (7425), 622-+.

7. Bai, X. Y.; Moraes, T. F.; Reithmeier, R. A. F., Structural biology of solute carrier (SLC) membrane transport proteins. *Molecular Membrane Biology* **2017,** *34* (1-2), 1-32.

8. Nie, R. X.; Stark, S.; Symersky, J.; Kaplan, R. S.; Lu, M., Structure and function of the divalent anion/Na+ symporter from Vibrio cholerae and a humanized variant. *Nature Communications* **2017,** *8*, 10.

9. Lu, M.; Symersky, J.; Radchenko, M.; Koide, A.; Guo, Y.; Nie, R.; Koide, S., Structures of a Na+-coupled, substrate-bound MATE multidrug transporter. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (6), 2099-2104.

10. Stoll, S.; Schweiger, A., EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of Magnetic Resonance* **2006,** *178* (1), 42-55.

11. Steinhoff, H.; Radzwill, N.; Thevis, W.; Lenz, V.; Brandenburg, D.; Antson, A.; Dodson, G.; Wollmer, A., Determination of interspin distances between spin labels attached to insulin: Comparison of electron paramagnetic resonance data with the x-ray structure. *Biophysical Journal* **1997,** *73* (6), 3287-3298.

12. Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H., DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance* **2006,** *30* (3-4), 473-498.

Materials & Methods

12 Biological Material

The majority of the biological material used for the studies reported in this thesis were supplied by various collaborators.

12.1 Received materials

The MalE and GlnPQ variants were supplied by the research group of Prof Thorben Cordes (originally at the Zernike Institute for Advanced Materials, University of Groningen, The Netherlands and later at the Ludwig-Maximilians-Universität München, Germany). These were produced using methods reported in previous literature[1-2].

For later studies, several attempts were made to produce and purify SBDs 1 and 2 of the GlnPQ system in-house, using plasmid supplied by the group of Prof Bert Poolman[1-7] at the University of Groningen, The Netherlands. This involved extracting plasmid both from dry samples and from a storage strain (DH5α) of *Escherichia coli.*

The PsaA variants were supplied by the group of Dr Chris McDevitt (at the Research Centre for Infectious Diseases at the University of Adelaide, Australia, now at Department of Microbiology and Immunology at the University of Melbourne, Australia) using protocols outlined in published literature[8].

Finally, the VcINDY variants were supplied as part of an ongoing collaboration with Dr Chris Mulligan at the University of Kent, United Kingdom[9-11].

12.2 LB growth media and LB-agar plates

Lysogeny broth (LB) prepared using 10:10:5 g/L of peptone, NaCl and yeast extract in ddH_2O was used throughout the following protocols as media for bacterial culture. All LB was sterilised by autoclaving prior to use.

LB-agar plates were prepared using LB and 10 g/L agar. The mixture was autoclaved and stored sealed for later use. Prior to pouring plates, the mixture was heated in a microwave until the agar had melted. This was mixed by swirling to distribute the contents evenly as it cooled to ‘hand-hot’ (50-60°C). Under sterile conditions, 100 μg/mL ampicillin was then added, and the mixture poured into

plastic plates (15-20 mL per plate). These were immediately covered and left to cool. Once set, the plates were inverted to avoid contamination and stored in a sealed bag.

12.3 Plasmid extraction

Dry plasmid samples were extracted from filter paper by soaking and mixing thoroughly in 100 µL Milli-Q Ultrapure water. 10 µL of supernatant were added to 100 µL TOP10 (*E. coli*) competent cells and mixed on ice for 30 mins. The cells were transformed by thermal shock, heating at 42°C for 45 seconds and immediately cooling on ice for 1 min. 500 µL of sterile LB (no antibiotic, pre-warmed to 37°C) was added, and this was then incubated at 37°C for 1 hour with shaking at 200 rpm. The culture was spread onto an LB-agar plate (with 100 µg/mL ampicillin), sealed and incubated at 37°C. After 15 mins the plate was inverted for incubation overnight. A single colony was selected for further culture, as described below (section 12.4 ‘Bacterial culture and protein overexpression’).

An alternative of storage strain ‘stabs’ were provided, consisting of DH5α *E. coli* in small aliquots of LB-agar with 100 µg/mL ampicillin. Sterile LB-agar plates (100 µg/mL ampicillin) were streaked under sterile conditions (three individual streaks with sterilisation of the streaking implement between each one; each streak passed through the previous streak once) with bacteria from these cultures, inverted immediately and incubated at 37°C overnight.

Plates were removed from the incubator and stored in at 4°C for 8 hours. Under sterile conditions, a single colony was selected for each further culture in 5 mL sterile LB with 100 µg/mL ampicillin. This was incubated overnight at 37°C with shaking at 200 rpm. Cells were pelleted by centrifugation at 4 000 xg, 15 mins, 4°C. The supernatant was discarded, and the pellets stored at -20°C for 5 hours. The plasmid in the cell pellets were then extracted and amplified using commercially available kits (QIAprep Spin Miniprep Kit, QIAGEN) according to manufacturer protocols, and the resulting purified plasmid stored at -20°C.

The plasmid was thawed and transformed into MC1061 *E. coli* (expression strain, 2% v/v ratio plasmid to cell) by thermal shock, heating at 42°C for 45 seconds and immediately cooling on ice for 1 min. 500 µL of sterile LB (no antibiotic, pre-warmed to 37°C) was added, and this was then incubated at 37°C for 1 hour with shaking at 200 rpm. The culture was spread onto a LB-agar plate (with 100 µg/mL

ampicillin), sealed and incubated at 37°C. After 15 mins the plate was inverted for incubation overnight. A single colony was selected for a starter culture in 15 mL sterile LB with 100 µg/mL ampicillin. This was incubated at 37°C for 7 hours with shaking at 200 rpm. 400 µL aliquots of starter culture were taken and mixed with glycerol (20% v/v). These were then flash frozen in liquid nitrogen and stored at -80°C.

12.4 Bacterial culture and protein overexpression

Grow-ups from glycerol stocks required a starter culture of 100 mL sterile LB with 100 µg/mL ampicillin, incubated overnight at 37°C with shaking at 200 rpm. Inoculation was performed under a sterile environment.

Growth cultures of 1 L of sterile LB with 100 µg/mL ampicillin were inoculated with 10 mL overnight culture. This was incubated at 37°C with shaking at 200 rpm shaking and the optical density at 600 nm (OD_{600}) monitored for culture growth. At OD_{600} = ~0.6 the cultures were induced with L-arabinose (0.0002% or 0.002 % w/v) and incubated for a further 2 hours at 30°C with shaking at 200 rpm. Cells were harvested by centrifugation at 6 000 xg, 20 mins, 4°C. The pellets were collected, flash frozen in liquid nitrogen and stored at -80°C.

12.5 Protein purification

Cell pellets that were previously stored at -80°C were thawed on ice. Lysis buffer (50 mM KP_i, pH 7.5) was prepared with 1 mM phenylmethylsulfonyl fluoride (PMSF; a serine protease inhibitor) and 0.1 mg/mL DNase. This was added to each cell pellet, thoroughly mixed and left on ice for 30 mins. The mixtures were prepared for sonication with the addition of 20% v/v glycerol. Two cycles of sonication, of 8 minutes each at 50% power, were undertaken keeping the sample cooled in ice (the suspensions were rotated to prevent overheating). The lysates were centrifuged at 45 000 xg for 45 mins at 4°C to remove the cell debris.

His-tag purification resin (TALON cobalt ion immobilised metal affinity chromatography (IMAC) resin; Takara Bio/Clontech) was prepared and incubated with the lysis buffer (50 mM KP_i, pH 7.5) with 10 mM imidazole, pH 8.0. This was mixed thoroughly and allowed to settle, and the supernatant removed. The lysates were applied to the resin, with 10 mM imidazole (pH 8.0) added, and mixed for 1 hour, 4°C. The resin mixtures were run into columns, and the flow collected. 40 bed volumes of wash 1 buffer (50 mM KP_i, pH 8.0, 200 mM KCl, 10 mM imidazole)

was added, followed by 5 bed volumes of wash 2 buffer (50 mM KP_i, pH 8.0, 200 mM KCl, 20 mM imidazole). The flows of these were collected separately. 5 bed volumes of elution buffer (50 mM KP_i, pH 8.0, 200 mM KCl, 500 mM imidazole) was used to remove the protein, and the flow collected together for further use. The column was washed with 10 bed volumes elution buffer, and the flow collected.

5 L of dialysis buffer (50 mM Tris-HCl, pH 8.0, 0.5 mM EDTA) was prepared and 1 mM DTT added. Dialysis membranes were soaked (SnakeSkin Dialysis Tubing, 7K MWCO, 22 mm dry diameter, Thermo Scientific) in Milli-Q Ultrapure water and then suspended in the dialysis medium with one end sealed. The protein fractions of column eluant were added to the membrane bags with TEV protease (100:1 v/v ratio of protein to cleavage protease). The dialysis bags were sealed and suspended in the dialysis media overnight with gentle stirring at 4°C.

Spin concentrators (10 K MWCO, 2 mL, Vivaspin from Sartorius or Amicon from Merck) were prepared with buffer (50 mM KP_i, pH 8.0, 200 mM KCl) as per the manufacturer's instructions (4 000 xg, 5 mins, 4°C) and the flow discarded. The tag-cleaved protein samples were added in aliquots and concentrated (4 000 xg, 10 mins at a time, 4°C; repeated until all ~10 mL of dialysed protein were transferred). The concentrated protein was washed with 1 mL buffer (4 000 xg, 10 mins 4°C) twice. All flows of the protein concentration and washing processes were collected separately. The final protein samples were collected by inverting the concentrators (as per manufacturer protocols) and centrifuged at 1 000 xg, 2 mins, 4°C.

TALON resin was prepared with buffer (50 mM KP_i, pH 8.0, 200 mM KCl) and 10 mM imidazole (pH 8.0) as before, and the supernatant removed. The tag-cleaved protein was mixed with the resin and buffer (50 mM KP_i, pH 8.0, 200 mM KCl) with 10 mM imidazole (pH 8.0) for 30 mins, 4°C. This was run through a column and the flow collected. The protein concentration was checked using UV/VIS spectrophotometry. The protein was then mixed with 20 % v/v glycerol, flash frozen with liquid nitrogen and stored at -80°C until use.

12.6 SDS-PAGE

Sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE) uses two different gels. The resolving gel (for separating contents of a sample by

molecular weight) was composed of 2.5 mL 1.5 M Tris-HCl buffer, pH 8.0, 4 mL of 30%-0.8% acrylamide-bisacrylamide, 0.1 mL of 10% SDS, 0.1 mL of 10% ammonium persulphate (APS) and 3.29 mL Milli-Q Ultrapure water. 10 µL of N,N,N',N'-tetramethylene diamine (TEMED) was added to begin polymerising the gel, and the solution run between two BioRad glass plates to set. The stacking gel (for loading samples in separate lanes) was made from 1.25 mL 0.5 M Tris-HCl, pH 6.8, 0.84 mL 30%-0.8% acrylamide-bisacrylamide, 50 µL of 10 % SDS, 50 µL of 10 % APS and 2.81 mL Milli-Q Ultrapure water. 5 µL of TEMED was added to begin polymerisation. The solution was then poured on top of the set resolving gel and a BioRad Teflon comb used to create running lanes. Once set, the gels were loaded into a BioRad running tank, and the tank filled with running buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1 mM EDTA, 0.5% SDS).

For the post-purification gels, samples collected throughout the purification protocol (15 µL) were mixed with loading dye (5 µL of 4x concentration dye; 1M Tris-HCl, pH 6.8, 8% w/v SDS, 40% v/v glycerol, 4% v/v β-mercaptoethanol, 0.5 M EDTA, 0.08% w/v bromophenol blue) The comb was then removed, and 15 µL of each sample with stain loaded into their respective lanes.

For the expression gel (see Appendix A)1 mL aliquots of cell culture were taken throughout the culture and overexpression and stored at -80°C overnight. Cells were pelleted by centrifugation (14 000rpm, 10 mins, room temperature) in a benchtop microcentrifuge. The supernatant was removed, and the cells resuspended in 60 µL Milli-Q Ultrapure water and 20 µL loading dye (4x concentration, as above). The samples were placed on a heating plate at 100-120°C until loading (at least 10 mins; larger cell pellet samples were loaded last).

For the denaturing/reducing experiment gel the boiled samples were left on a heating plate set to 100°C for 10 mins before loading (15 µL protein, 5 µL 4x concentration loading dye). The following denaturant/reductant stocks were made in MilliQ Ultrapure water: 8M urea, 6M Gdn.HCl, 1M L-arginine, 250 mM DTT and 250 mM TCEP. For the samples containing urea, Gdn.HCl, L-arginine and TCEP, the following amounts were used: 10 µL protein sample, 5 µL stock and 5 µL 4x concentration loading dye. For the samples containing DTT, the following amounts were used: 15 µL protein sample, 0.4 µL stock, 5 µL 4x concentration loading dye.

See Appendix A for gel details.

12.7 Buffer formulations

The following buffers were used for any biochemical manipulations of the protein systems where the compositions are not otherwise described:

- MalE: 50 mM KP_i, pH 7.0-8.0, 150-200 mM KCl, or 50 mM Tris-HCl, pH 7.4, 50 mM KCl
- SBD1 and SBD2: 50 mM KP_i, pH 7.0-8.0, 150-200 mM KCl, or 50 mM Tris-HCl, pH 7.4, 50 mM KCl
- PsaA: 20 mM MOPS, pH 7.2, 100 mM NaCl
- VcINDY: 20 mM TRIS-HCl pH 8.0, 100 mM NaCl or KCl, 0.05 % DDM

Buffers were prepared using Milli-Q Ultrapure water, with concentrated (1 M or 5 M) HCl, and NaOH or NH_4OH to adjust for pH. Later they were autoclaved for sterility.

13 Spectroscopic Analysis

13.1 Sample Preparation

Site directed spin labelling

Aliquots of protein (usually 350 - 700 µL, 5-500 µM) were reduced with 1 mM final concentration dithiothreitol (DTT, from 250 mM stock in deionised water) for four hours at 4°C with gentle mixing. The excess DTT was removed using a Zeba Spin Desalting Column (7K MWCO, 2 mL from Thermo Scientific), prepared following manufacturer instructions; the protein fraction elutes at 1000 x g, 4°C, 2 mins. The protein was then incubated with a 10:1 (spin label:cysteine residues) molar excess of S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro- 1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL, from Toronto Research Chemicals, Ontario, Canada; from 50-250 mM stocks in dimethyl sulfoxide or isopropanol) overnight, in the dark, at 4°C with gentle shaking. The excess MTSL was removed using Zeba Spin column; elutes at 1 000 xg, 4°C, 2 mins. The protein concentration was then compared to the spin label concentration to determine labelling efficiency. This was determined using its double integration value compared to those of standards of MTSL.

Spin concentration

Aliquots of the proteins (~ 1.75 µM, MW 108415 Da) were concentrated using Vivaspin or Amicon spin concentrators (10 K, 30 K or 100 K MWCO as appropriate for their respective molecular weights; Vivaspin from Sartorius or Amicon from Merck) according to manufacturer instructions.

This method was also used for buffer exchange (i.e. protonated to deuterated buffer). The samples were concentrated, the volume made back up with the new buffer and repeated for a total of 3 washes to ensure sufficient exchange.

Buffer deuteration

Pre-prepared buffer was deuterated by lyophilising (freeze-drying) it until completely dry, and re-dissolving in d_2O. This was repeated for a total of three lyophilisation cycles.

Substrate additions

Large excesses of maltose for MalE and asparagine, glutamine or glutamate for the GlnPQ SBDs were used. For the metal incubations of PsaA, a 1:1 stoichiometry of metal to protein was used. For VcINDY, different buffers were used to introduce Na^+ into the system (i.e. using KCl or NaCl salt), and a large excess of succinate was used. In all cases, the substrates were mixed in thoroughly and incubated with the protein for sufficiently long to allow binding (at least 1 hour at 4°C, or on ice).

13.2 Room temperature cw-EPR X-band

Instrumentation/set-up

The room temperature set-up for X-band cw-EPR was comprised of a Bruker E500 Elexsys spectrometer fitted with the ER 4123D (dielectric RT cw-EPR, Bruker) resonator. 5-10 µL aliquots of the aqueous samples were loaded into quartz capillaries (0.8 mm o.d., CMScientific, sealed at one end). These were then sheathed in larger quartz capillaries (1.5 mm o.d., CMScientific, sealed at one end). Long capillary tubes were purchased and then cut and sealed in-house.

General Parameters

For spin labelled protein samples, the following parameters are usually sufficient across all experiments: modulation frequency 100 kHz, modulation amplitude 1 G, time constant and conversion time both 20.48 or 40.96 ms. The centre of the spectrum depends on the g-factor of the system and the microwave frequency used; here 3480 G is typical of a system around g = 2 using microwaves of around 9.7 GHz, with a sweep width of 100-150 G. The power of the incident microwave radiation is usually selected to be such that it achieves maximal signal but is non-saturating (i.e. just below $P_{1/2}$; around 0.2-1.0 mW but optimised on a sample by sample basis). The number of scans is generally dependent on the signal to noise; lower signal quality necessitates a greater number of averaged scans.

For power saturation experiments, all parameters were kept the same except the incident microwave power. The attenuation was varied between 6-54 dB (50 mW to 0.8 µW) in 6 dB increments. Typically, fewer averaged scans per slice of this experiment are used (around 4 scans).

13.3 Low temperature cw-EPR X-band

Instrumentation/set-up

Low temperatures are achieved using a liquid helium transfer set-up with a turbomolecular vacuum pump (PT-80 Dry Divac, Oerlikon Leybold), a proportional-integral-derivative (PID) type temperature controller (Intelligent Temperature Controller 503, Oxford Instruments) and a gas flow controller (VC41, Oxford Instruments). The transfer equipment consists of a 4122SHQE (super-high Q) resonator (Bruker) with the ESR900 low cryostat (Oxford Instruments), in a Bruker E500 Elexsys spectrometer. This set up involves a quartz Dewar and sample holder in which a sample is held in a continuous flow of temperature-controlled helium gas. The helium transfer line into the cryostat is a gas-shielded transfer tube; cold helium from the large Dewar is shielded by the gas that passes back from the cryostat. Helium flow is in part controlled manually by a needle valve on the transfer line and by the gas flow controller.

Aliquots of around 70 µL of the aqueous samples were loaded into quartz tubes (4 mm o.d., 250 mm length, Wilmad, Sigma-Aldrich) and flash frozen with liquid nitrogen. Some samples (those to be measured with pulsed EPR methods, or those that had already been measured with pulsed EPR methods) contained 30-50% v/v d_8-glycerol prior to freezing.

General Parameters

For spin labelled protein samples, the following parameters are usually sufficient across all experiments: modulation frequency 100 kHz, modulation amplitude 1-2 G, time constant and conversion time both 20.48 or 40.96 ms. To measure spin label species (g = ~2) 3350 G is typical using microwaves of around 9.3 GHz, with a sweep width of 150-200 G. The incident microwave power depends on the system and the temperature, but <0.05 mW is typical at 20 K.

To measure metal species often the spectrum is far broader (up to 1500-2000 G width) and longer time constants and conversion times as a result (40.96 or 81.92 ms), with higher incident microwave powers up to around 4 mW (at 20 K). As the linewidths of metal species are typically large, modulation amplitudes of up to 10 G are reasonable.

The number of averaged scans is optimised for each sample dependent on signal quality.

13.4 Low temperature cw-EPR Q-band

Instrumentation/set-up

Low temperature Q-band cw-EPR spectra were obtained using a home-built closed cycle refrigeration ‘displex’ and a Bruker E560 Elexsys spectrometer with a Bruker ER 5106 QTW resonator. The displex set-up was comprised of an ARS-2HW compressor (Advanced Research Systems), vacuum-shrouded cryostat, TurboLab 80 pump (Oerlikon Leybold) and a gaseous helium flow system, with a LakeShore 332 temperature controller (LakeShore Cryonics, Inc.).

Aliquots of around 30 µL of the aqueous samples were loaded into quartz tubes (3 mm o.d., 159 mm length, Wilmad, Sigma-Aldrich) and flash frozen with liquid nitrogen.

General parameters

For spin labelled protein samples, the following parameters are usually sufficient across all experiments: modulation frequency 10 kHz, modulation amplitude 1-5 G, time constant and conversion time both 40.96 ms. For measuring spin label species (g = ~2) 12100 G is typical using microwaves of around 34.0 GHz, with a sweep width of around 200 G. The incident microwave power depends on the system and the temperature, but <0.01 mW is typical at 20 K.

To measure metal species often the spectrum is far broader (up to 1500-2000 G width) and longer time constants and conversion times as a result (40.96 or 81.92 ms), with higher incident microwave powers up to around 0.05 mW (at 20 K). As the linewidths of metal species are typically large, modulation amplitudes of up to 10 G are reasonable.

The number of averaged scans is optimised for each sample dependent on signal quality.

13.5 Pulsed EPR X-band

Instrumentation/set-up

The low temperatures for pulsed X-band EPR (using ER 4118 flexline resonator, Bruker) are achieved using a closed cycle refrigeration 'displex'. This set-up is comprised of a compressor (Advanced Research Systems ARS-4HW), vacuum-shrouded cryostat, turbomolecular pump (TurboLab 80, Oerlikon Leybold), a gaseous helium flow and LakeShore 331 temperature controller (LakeShore Cryonics, Inc.).

Aliquots of around 50-70 µL of the aqueous samples, with 30-50% v/v d_8-glycerol were loaded into quartz tubes 4 mm o.d., 250 mm length, Wilmad, Sigma-Aldrich) and flash frozen with liquid nitrogen.

General Parameters

FSEs of the spin labelled protein samples were recorded at 50 K with microwaves at around 9.7 GHz. 150 G sweep width was centred about g = 2 (field position of around 3440 G), 1-5 dB attenuation, 42-66 dB video gain, 50 shots per point, 400 points, 1-3 ms shot repetition time, 4 scans. These settings were optimised for each sample/experiment (particularly for the attenuation and video gain) and are the basis for the T_2 experiment below.

The T_2 experiment takes the field point at which there is greatest echo intensity (typically around 3440 G). As above, this experiment was performed at 50 K with microwaves around 9.7 GHz. 1-5 dB attenuation, 42-66 dB video gain, 50 shots per point, 600 points, time increment 8 ns, 1-3 ms shot repetition time, 4 scans.

4P-PELDOR experiments were also conducted at 50 K with microwaves around 9.7 GHz, with the observer field position optimised for maximum signal (as found through the FSE experiment; typically around 3440 G). The attenuation and video gain were typically 1-5 dB and 42-66 dB respectively, 50 shots per point, number of points was optimised for each experiment, time increment 8 ns, 1-3 ms shot repetition time, number of scans was optimised for each experiment. The ELDOR pump position was set around 25 G down-field (~70 MHz frequency difference vs the observe frequency). The power of the pulses in the ELDOR channel were optimised to give maximum echo inversion, typically 0-5 dB.

The parameters were set using a custom 4 pulse ELDOR program in PulseSPEL, which also employs phase cycling of the pulse sequence and nuclear modulation suppression.

References

1. Gouridis, G.; Schuurman-Wolters, G. K.; Ploetz, E.; Husada, F.; Vietrov, R.; de Boer, M.; Cordes, T.; Poolman, B., Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nat. Struct. Mol. Biol.* **2015,** *22* (1), 57-64.

2. Husada, F.; Gouridis, G.; Vietrov, R.; Schuurman-Wolters, G. K.; Ploetz, E.; de Boer, M.; Poolman, B.; Cordes, T., Watching conformational dynamics of ABC transporters with single-molecule tools. *Biochem. Soc. Trans.* **2015,** *43*, 1041-1047.

3. de Boer, M.; Gouridis, G.; Vietrov, R.; Begg, S. L.; Schuurman-Wolters, G. K.; Husada, F.; Eleftheriadis, N.; Poolman, B.; McDevit, C. A.; Cordes, T., Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **2019,** *8*, 28.

4. Fulyani, F.; Schuurman-Wolters, G. K.; Zagar, A. V.; Guskov, A.; Slotboom, D. J.; Poolman, B., Functional Diversity of Tandem Substrate-Binding Domains in ABC Transporters from Pathogenic Bacteria. *Structure* **2013,** *21* (10), 1879-1888.

5. Poolman, B.; Slotboom, D.-J., Substrate Capture by ABC Transporters. In *Encyclopedia of Biophysics*, Roberts, G. C. K., Ed. Springer: Berlin, Heidelberg, 2013; pp 2503-2509.

6. Scheepers, G. H.; Lycklama, J. A.; Poolman, B., An updated structural classification of substrate-binding proteins. *Febs Letters* **2016,** *590* (23), 4393-4401.

7. Schuurman-Wolters, G. K.; Poolman, B., Substrate specificity and ionic regulation of GlnPQ from Lactococcus lactis - An ATP-binding cassette transporter with four extracytoplasmic substrate-binding domains. *Journal of Biological Chemistry* **2005,** *280* (25), 23785-23790.

8. Deplazes, E.; Begg, S. L.; van Wonderen, J. H.; Campbell, R.; Kobe, B.; Paton, J. C.; MacMillan, F.; McDevitt, C. A.; O'Mara, M. L., Characterizing the conformational dynamics of metal-free PsaA using molecular dynamics simulations and electron paramagnetic resonance spectroscopy. *Biophysical Chemistry* **2015,** *207*, 51-60.

9. Mulligan, C.; Fenollar-Ferrer, C.; Fitzgerald, G. A.; Vergara-Jaque, A.; Kaufmann, D.; Li, Y.; Forrest, L. R.; Mindell, J. A., The bacterial dicarboxylate transporter VcINDY uses a two-domain elevator-type mechanism. *Nat. Struct. Mol. Biol.* **2016,** *23* (3), 256-263.

10. Mulligan, C.; Fitzgerald, G. A.; Mindell, J. A., Assessing Protomer Independence of the Dimeric C4-Dicarboxylate Transporter, Vcindy. *Biophysical Journal* **2015,** *108* (2), 461A-461A.

11. Mulligan, C.; Fitzgerald, G. A.; Wang, D. N.; Mindell, J. A., Functional characterization of a Na+-dependent dicarboxylate transporter from Vibrio cholerae. *J. Gen. Physiol.* **2014,** *143* (6), 745-759.

Appendix A – Supplementary Data

A.1 GlnPQ SBDs SDS-PAGE

To test the amount of L-arabinose needed for inducing protein overexpression, two concentrations were selected: 0.002 and 0.0002% w/v. Expression gels were performed using samples taken at various stages throughout the bacterial culture and protein overexpression of the XXX variant of SBD1, grown in *E. coli*, using 0.002% (flask 1) or 0.0002% (flask 2) w/v L-arabinose to induce overexpression.

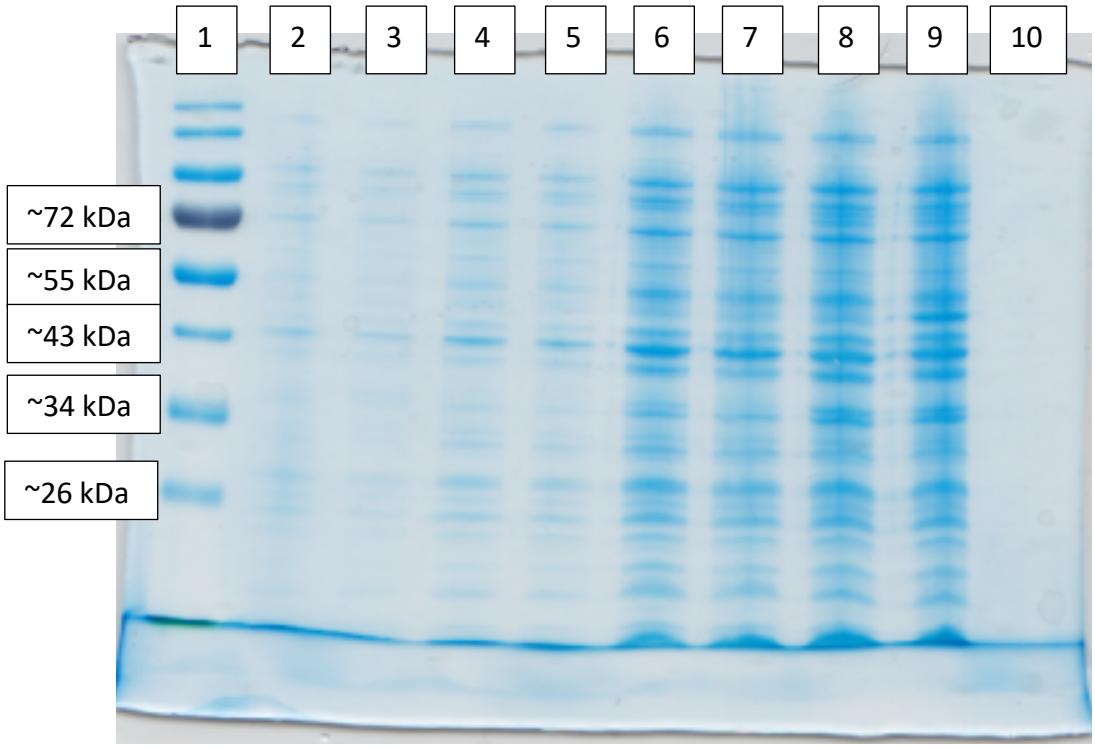


Figure A1.1: SDS-PAGE of the various stages throughout the bacterial culture and protein overexpression of the Q87C/T159C variant of SBD1 grown in E. coli, induced with 0.002% (flask 1) or 0.0002% (flask 2) w/v L-arabinose. Lane contents: 1) molecular weight marker ladder; 2) flask 1 (pre-inducing, OD_{600} = 0.176); 3) flask 2 (pre-inducing, OD_{600} = 0.152); 4) flask 1 (pre-inducing, OD_{600} = 0.566); 5) flask 2 (pre-inducing, OD_{600} = 0.502); 6) flask 1 (1hr after inducing); 7) flask 2 (1hr after inducing); 8) flask 1 (2hr after inducing, pre-harvest); 9) flask 2 (2hr after inducing, pre-harvest); 10) empty.

To investigate further, the SDS-PAGE gels containing samples taken throughout the purification protocol were run.

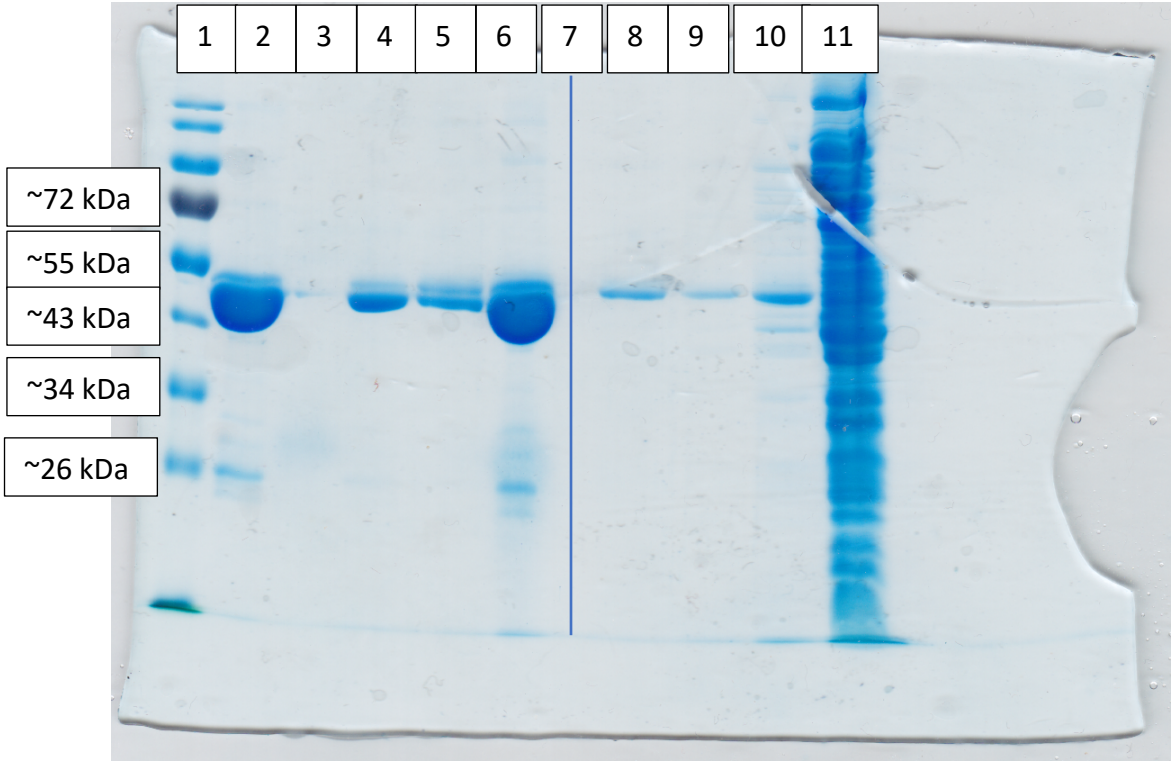


Figure A1.2: SDS-PAGE of the purification steps of the Q87C/T159C variant of SBD1 grown in E. coli, induced with 0.002% w/v L-arabinose. Lane contents: 1) molecular weight marker ladder; 2) post-purification SBD1; 3) post-dialysis spin concentrator flow-through; 4) post-dialysis spin concentrator membrane wash; 5) post-dialysis column elution wash; 6) dialysis bag contents; 7) empty; 8) post-elution column wash; 9) column wash 2 flow-through; 10) column was 1 flow-through; 11) cell lysate. A line marks the empty lane for clarity.

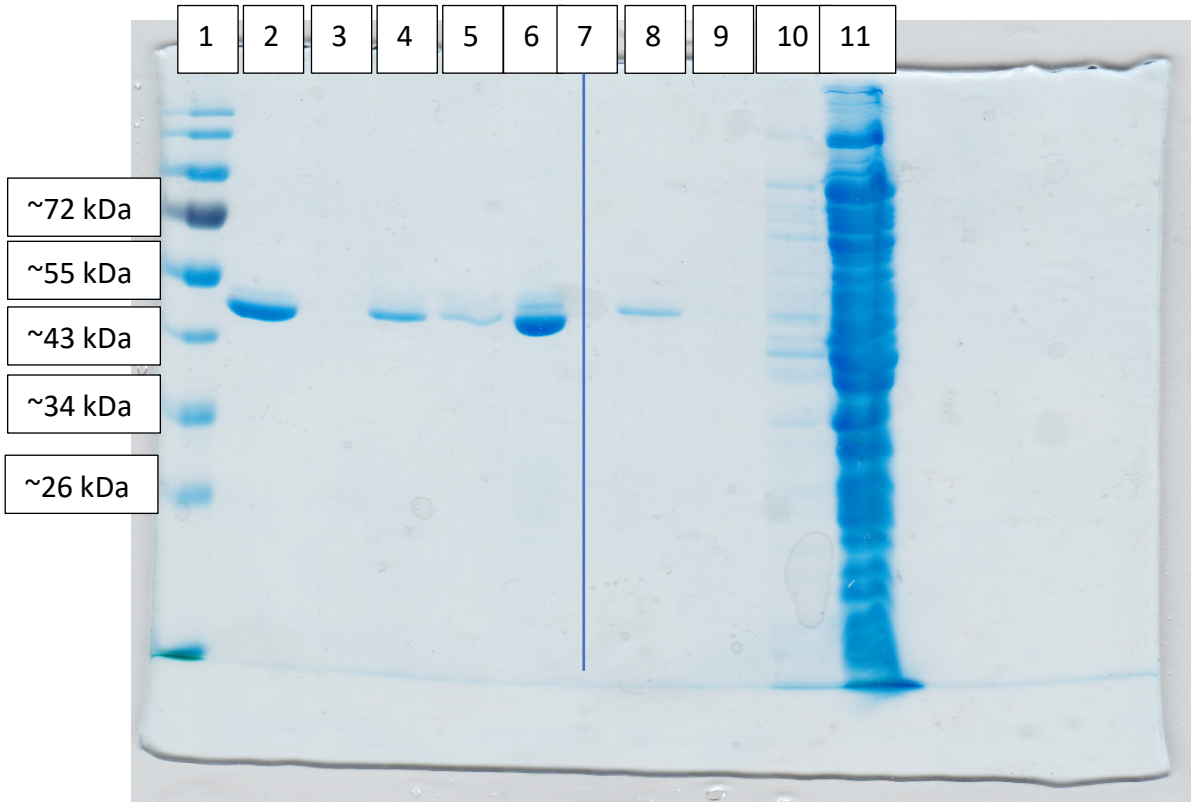


Figure A1.3: SDS-PAGE of the purification steps of the Q87C/T159C variant of SBD1 grown in E. coli, induced with 0.0002% w/v L-arabinose. Lane contents: 1) molecular weight marker ladder; 2) post-purification SBD1; 3) post-dialysis spin concentrator flow-through; 4) post-dialysis spin concentrator membrane wash; 5) post-dialysis column elution wash; 6) dialysis bag contents; 7) empty; 8) post-elution column wash; 9) column wash 2 flow-through; 10) column was 1 flow-through; 11) cell lysate. A line marks the empty lane for clarity.

The 0.002% w/v L-arabinose induced sample clearly yields more product, but in both cases, the purified SBD1 protein is showing at between around 43 and 55 kDa. SBD1 is known to have a molecular weight of 27.6 kDa (as calculated from its primary sequence). It was thought that perhaps aggregation was being seen, in which two SBD1 proteins were in some way stuck together (even with SDS denaturant). Therefore, samples of each purified SBD1 sample (from the two purifications above) were exposed to various further denaturing or reducing conditions.

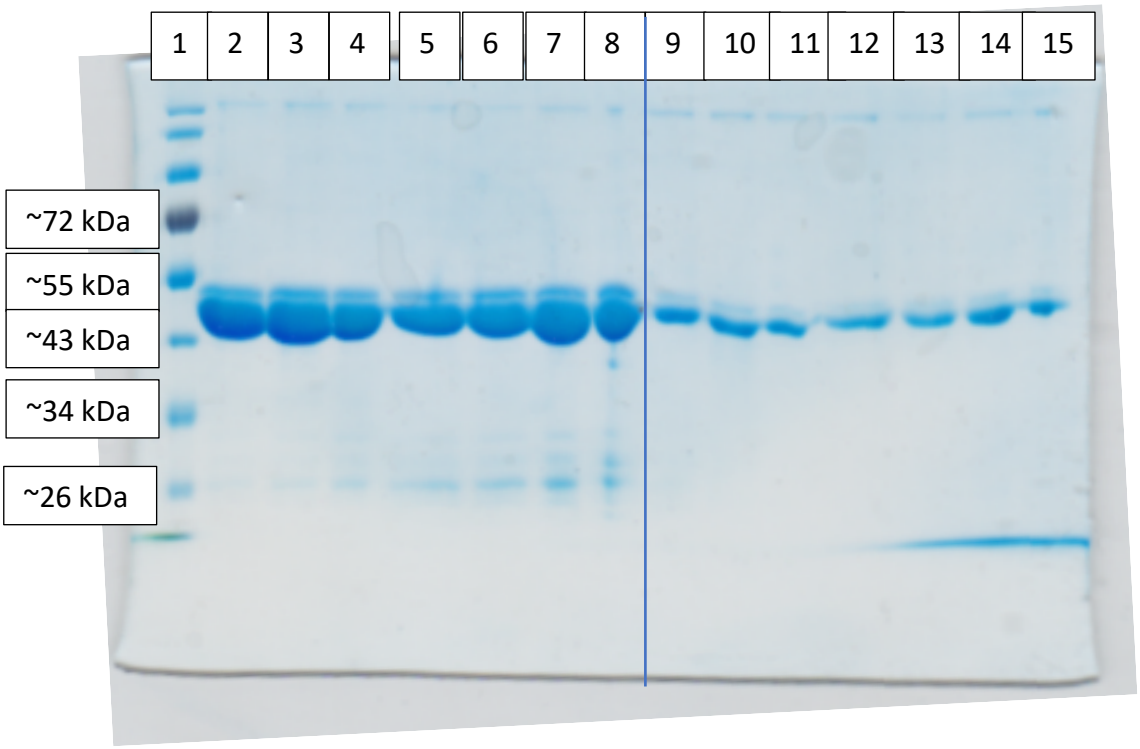


Figure A1.4: SDS-PAGE of the denaturing experiments of the Q87C/T159C variant of SBD1 grown in E. coli, induced with 0.002% (lanes 2-8) or 0.0002% (lanes 9-15) w/v L-arabinose. Lane contents: 1) molecular weight marker ladder; 2) & 9) purified SBD1; 3) & 10) boiled (100°C, 10 mins); 4) & 11) + 2 M urea; 5) & 12) + 1.5 M Gdn.HCl; 6) & 13) + 250 mM L-arginine; 7) & 14) + 5 mM DTT; 8) & 15) + 50 mM TCEP. A line bisects the gel between the 0.002% and 0.0002% w/v L-arabinose induced samples for clarity.

No difference between the 'purified' SBD1 samples and those exposed to the various denaturing or reducing conditions is observed. It is therefore still unclear as to why the molecular weight of the purified protein does not match that which is calculated from the primary sequence.

Further to this, the SDS-PAGE gel containing samples taken throughout the purification protocol of overexpressed XXX variant of SBD2 was run. The protocol for the overexpression used 0.0002% w/v L-arabinose to induce.

Figure A1.5: SDS-PAGE of the purification steps of the T369C/S451 variant of SBD2 grown in E. coli, induced with 0.0002% w/v L-arabinose. Lanes are largely indistinguishable except for that containing the molecular weight marker ladder and those.

It appears that there is little-to-no protein in the loaded gel sample, even at the 'cell lysate' stage. It is unclear as to why, as the bacterial cultures were monitored and induced at the correct 'log phase' of growth (indicated by OD_{600} = ~0.6).

A.2 4P-PELDOR set-up experiments

The experiments usually recorded to accurately set up a 4P-PELDOR experiment (i.e. FSE and T_2 measurements) are shown below where acquired.

MalE

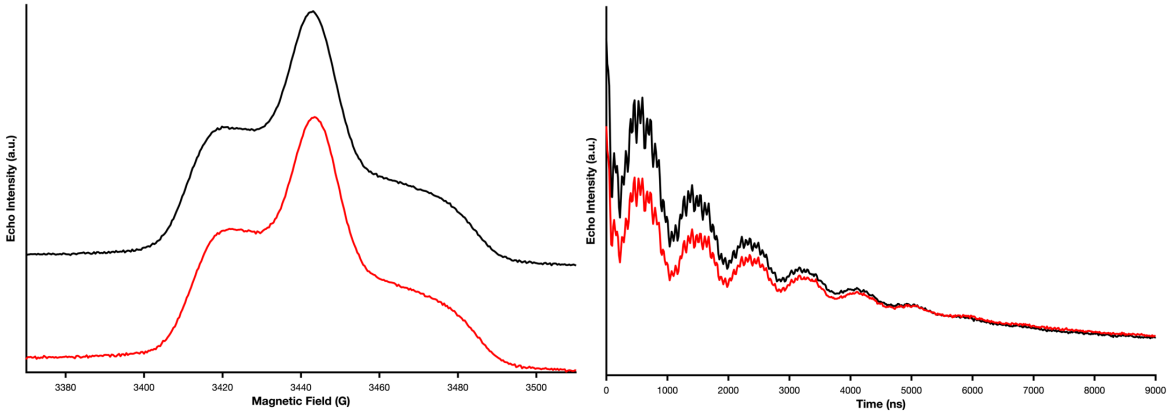


Figure A2.1: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled T36C/S352C *variant of MalE in the apoprotein (black) and maltose-bound (red) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled T36C/S352C variant of MalE in the apoprotein (black) and maltose-bound (red) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 60 K.*

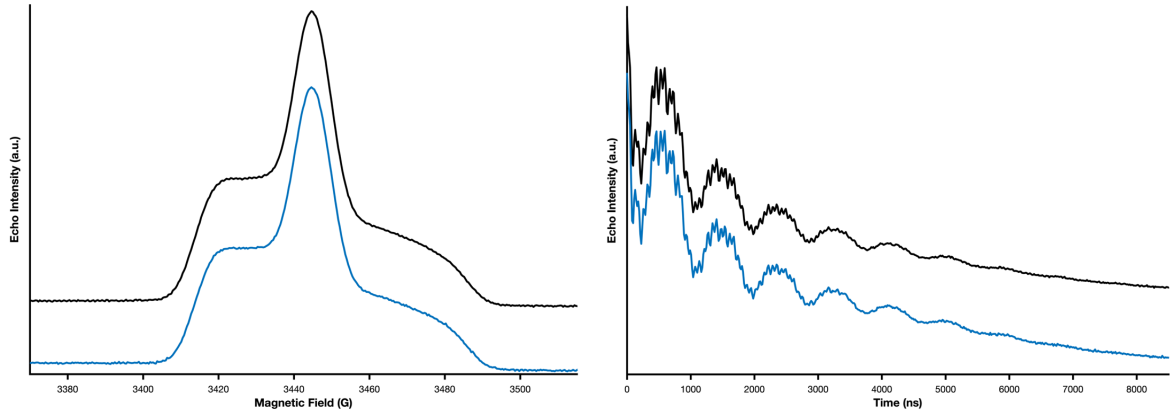


Figure A2.2: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled Q87C/T159C variant of SBD1 (later 'non-binding' prep, provided by collaborators at Ludwig-Maximilians-Universität München) in the apoprotein form (black) and with excess Gln (blue), measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled Q87C/T159C variant of SBD1 (later 'non-binding' prep, provided by collaborators at Ludwig-Maximilians-Universität München) in the apoprotein form (black) and with excess Gln (blue), measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

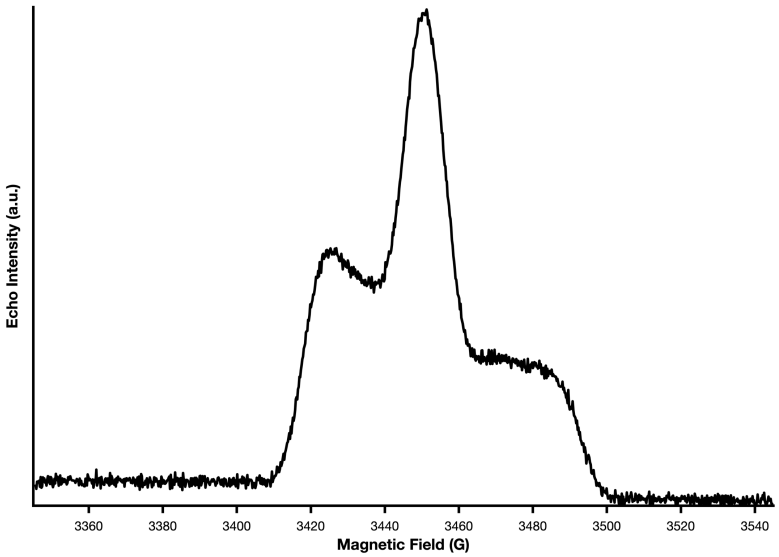


Figure A2.3: the Field Swept Echo (FSE) spectrum of the MTSL-labelled T369C/S451 variant of SBD2 (earlier 'contaminated prep', provided by collaborators at Zernike Institute for Advanced Materials, University of Groningen) in the apoprotein form, measured using a standard Hanh echo recorded across a range of magnetic fields. Measured at X-band, at 50 K.

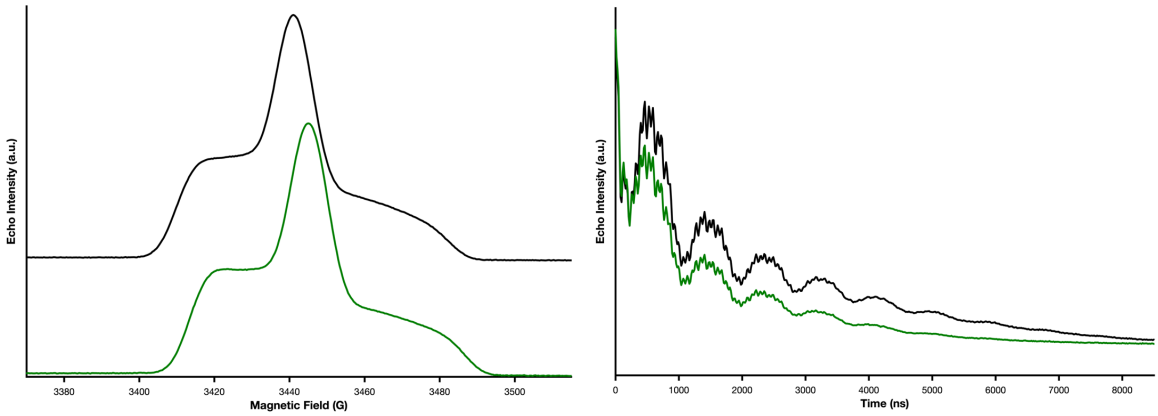


Figure A2.4: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled T369C/S451 variant of SBD2 (later 'non-binding' prep, provided by collaborators at Ludwig-Maximilians-Universität München) in the apoprotein form (black) and with excess Gln (green), measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled T369C/S451 variant of SBD2 (later 'non-binding' prep, provided by collaborators at Ludwig-Maximilians-Universität München) in the apoprotein form (black) and with excess Gln (green), measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

PsaA

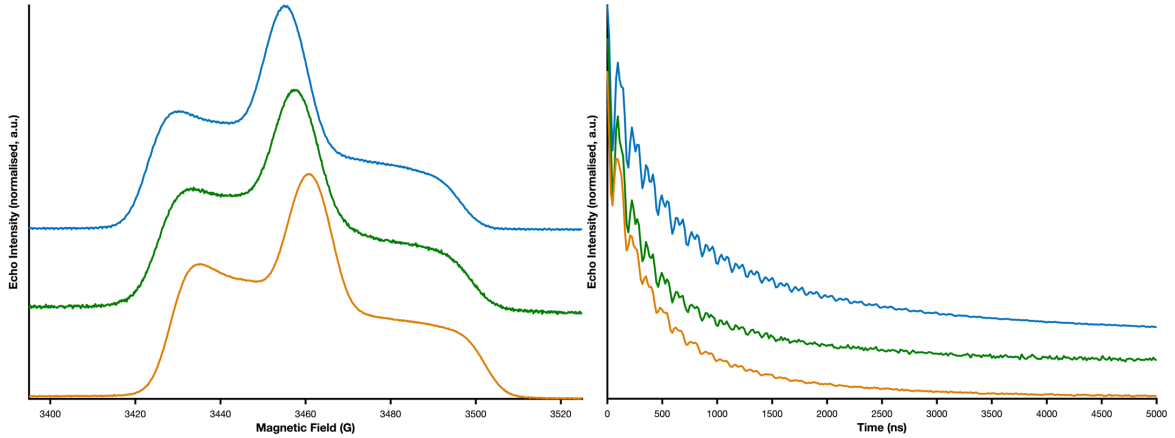


Figure A2.5: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled L56C/I236C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled L56C/I236C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

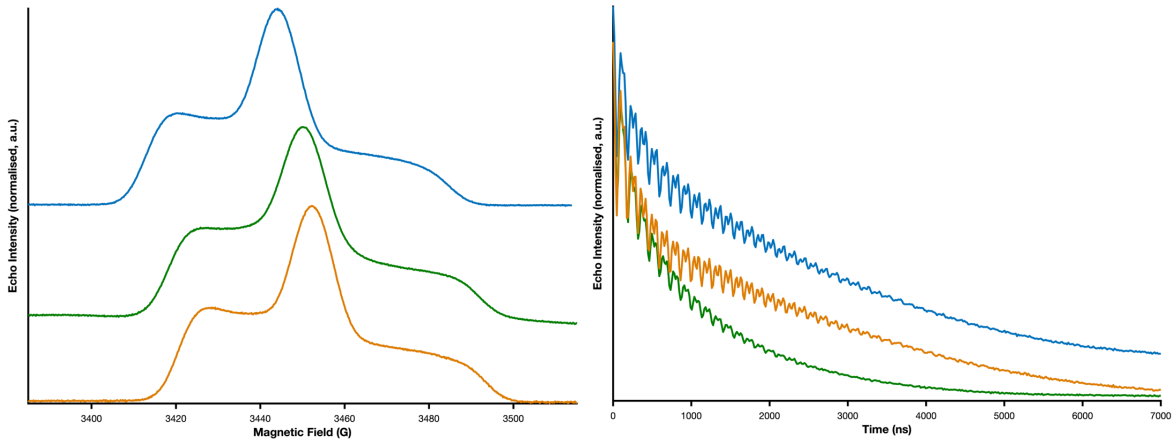


Figure A2.6: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled L56C/S266C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled L56C/S266C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

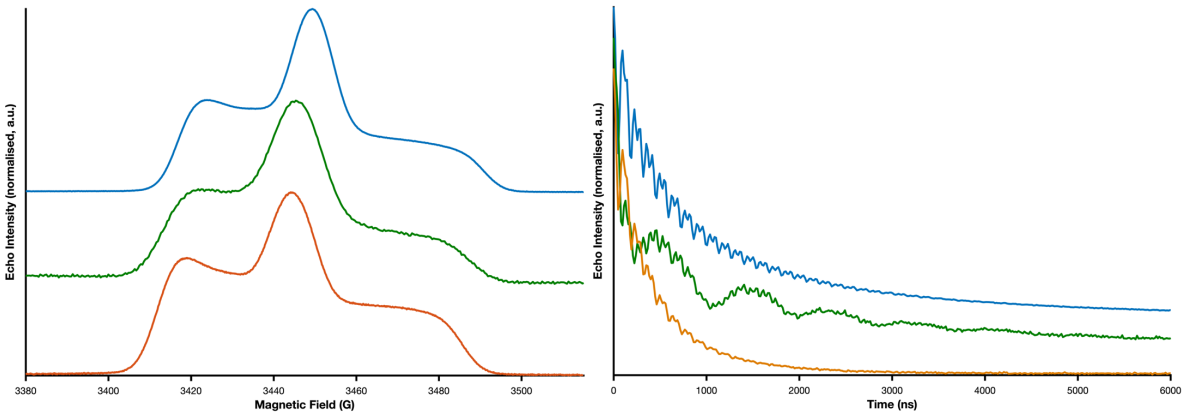


Figure A2.7: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled S58C/S266C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled S58C/S266C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

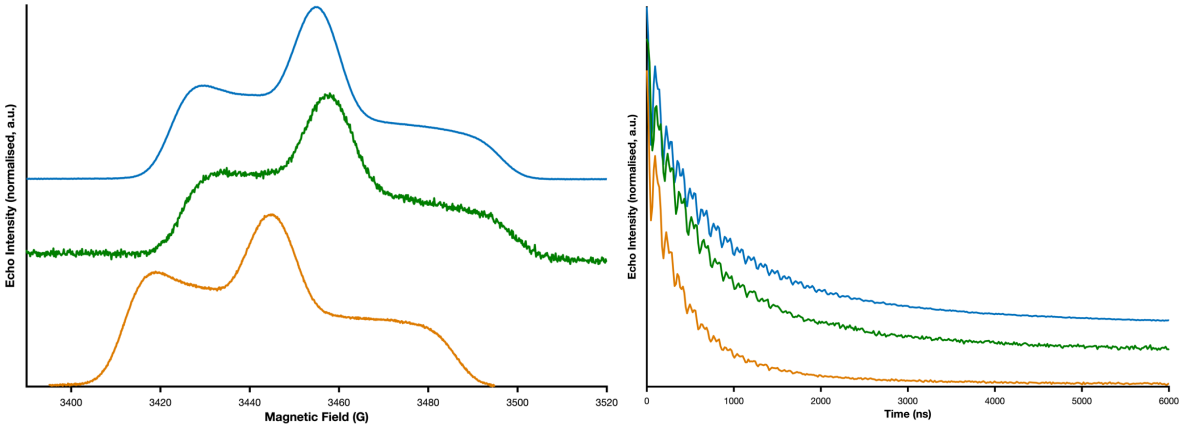


Figure A2.8: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled S58C/I125C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled S58C/I125C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

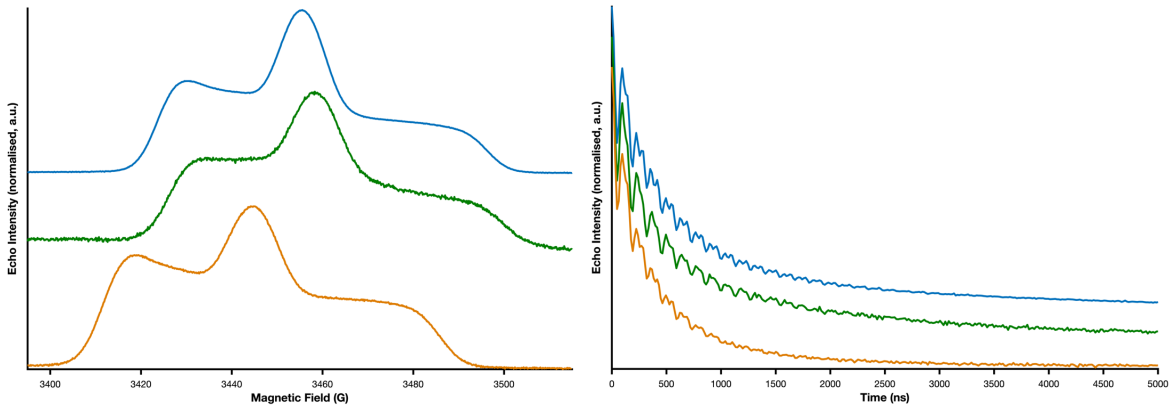


Figure A2.9: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled I125C/I236C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled I125C/I236C variant of PsaA in the apoprotein (blue), Mn^{2+}-bound (green) and Zn^{2+}-bound (orange) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

VcINDY

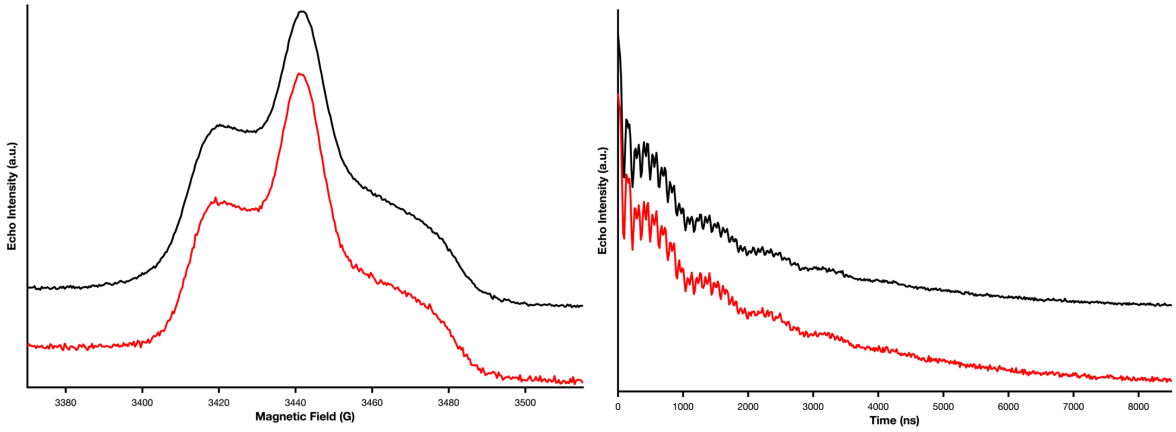


Figure A2.10: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled G211C variant of VcINDY in the apoprotein (black) and Na^+/succinate-bound (5 mM, red) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled G211C variant of VcINDY in the apoprotein (black) and Na^+/succinate-bound (5 mM, red) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

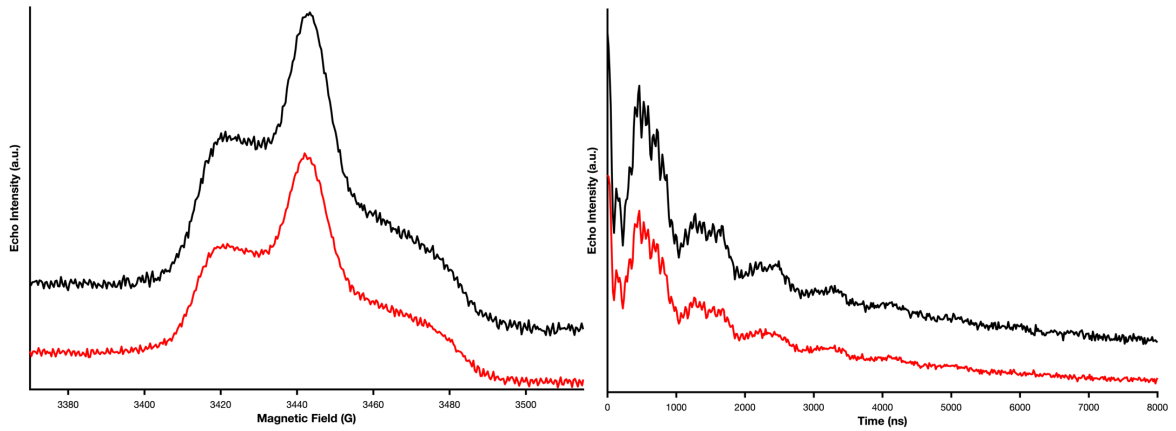


Figure A2.11: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled G211C variant of VcINDY in deuterated buffer, in the apoprotein (black) and Na^+/succinate-bound (20 mM, red) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled G211C variant of VcINDY in deuterated buffer, in the apoprotein (black) and Na^+/succinate-bound (20 mM, red) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

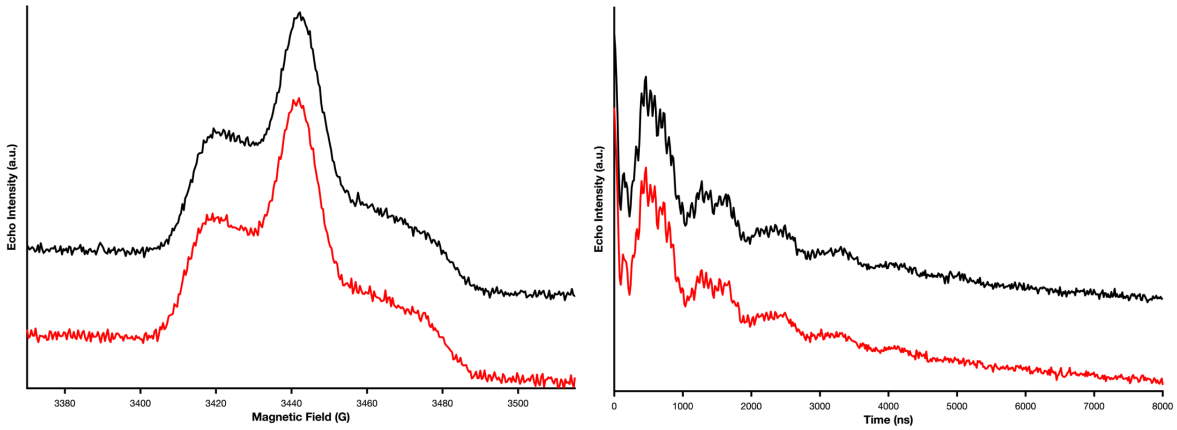


Figure A2.12: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled S436C variant of VcINDY in deuterated buffer, in the apoprotein (black) and Na^+/succinate-bound (20 mM, red) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled S436C variant of VcINDY in deuterated buffer, in the apoprotein (black) and Na^+/succinate-bound (20 mM, red) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

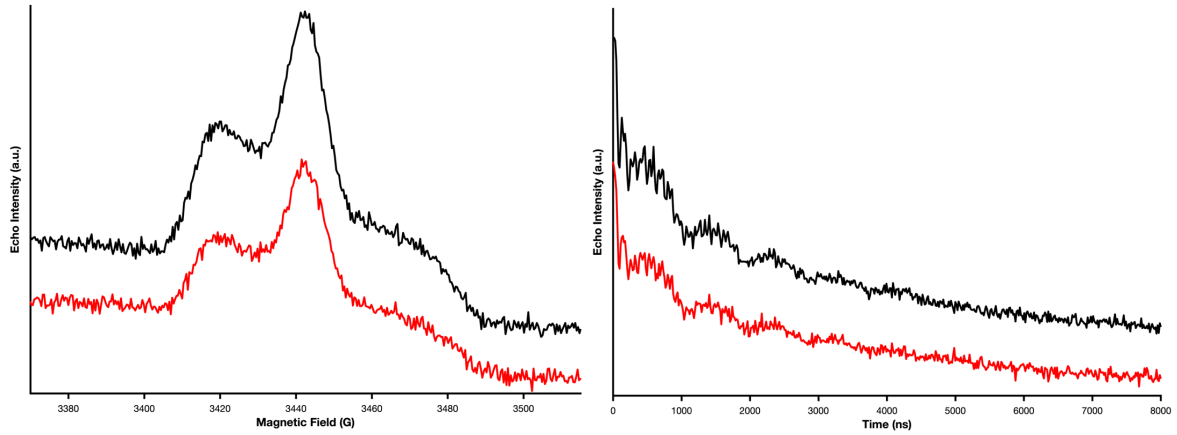


Figure A2.13: left, the Field Swept Echo (FSE) spectra of the MTSL-labelled A208C variant of VcINDY in deuterated buffer, in the apoprotein (black) and Na^+/succinate-bound (20 mM, red) forms, measured using a standard Hanh echo recorded across a range of magnetic fields. Right, the Hanh echo decay (T_2) spectra of the MTSL-labelled S208C variant of VcINDY in deuterated buffer, in the apoprotein (black) and Na^+/succinate-bound (20 mM, red) forms, measured using a standard Hanh echo with the separation, τ, between the π/2- and π-pulses increased between each point. Both measured at X-band, at 50 K.

Paper contributions

Over the course of my PhD studies, I had the opportunity to utilise my skills in EPR spectroscopy in order to contribute to various collaborative projects outside the scope of the research primarily presented within this thesis. In the remaining appended sections, the relevant portions of the peer-reviewed outputs of these collaborations are presented.

Within the published literature presented here are a range of studies involving multifrequency cw-EPR at both room and low temperature. These have been performed on a range of systems: a naturally derived protein, an *ab initio* designed 'synthetic' protein, and an abiotic catalyst. Specialist knowledge was required not only to measure but to interpret the data, and to disseminate to the non-spectroscopist collaborative partners the key findings.

Appendix B – *De novo* Enzyme

The following is a summary of the research published in the paper by Watkins *et al.*, 2017[1] (see original publication for Supplementary Information). Contributions included acquisition and analysis of the EPR spectra.

Being able to produce custom protein catalysts that mimic the natural functions of enzymes may be crucial in the development of novel, cost-effective, more environmentally-friendly options for industrial biocatalysis[2-4]. Though there have been some prominent accomplishments in producing *de novo* enzymes from natural protein scaffolds[5-7], it has been shown that arguably simpler manmade proteins can also replicate certain functions of natural metalloenzymes[8-11]]. Maquettes (4-α-helix bundles[4]) use relatively simple yet hardy blueprints that are particularly amenable to the incorporation of catalytic elements, as previously demonstrated by the successes involving complex functions such as oxygen binding[12] and intermolecular electron transfer[4]. Despite these achievements, no enzymatic activity has yet been integrated into a maquette. Here, the maquette approach is used to construct a highly efficient, promiscuous, thermostable and chemically resistant artificial enzyme. This catalyses an array of oxidations coupled to the reduction of H_2O_2, with kinetics similar to, and even exceeding, those of certain natural peroxidases. The maquette retains its activity in organic solvents and allows the probing of catalytic intermediates common to natural haem-containing enzymes.

It was recently demonstrated that the post-translational cytochrome *c* maturation machinery, found in the periplasmic space of *E. coli*, can be harnessed to process maquettes in order to implant haem into the helix bundle[13-14]. The resulting *c*-type cytochrome maquettes (CTMs) retain the aforementioned functionalities of the maquettes (e.g. oxygen binding), but also allow for further engineering. As these CTMs are completely produced *in vivo* there is an opportunity to study *de novo* protein function within living cells. Additionally, the use of directed evolution may be employed to improve incipient function. The strategy used here is based on a simple, two-step conversion of a non-catalytic CTM to produce a catalytically active holoprotein from *E. coli*.

The coupled substrate oxidation/H_2O_2 reduction (peroxidase) functionality was selected for inclusion into a CTM. Natural peroxidase enzymes typically contain

haem B or C axially coordinated by a single histidine side chain[15]; it was therefore hypothesised that replicating this with a monohistidine-ligated CTM would achieve catalytic turnover of hydrogen peroxide within a maquette scaffold. In order to construct this, a CTM with haem covalently appended to helix 4 (C4) was selected for its high thermal stability[14]. Next, to increase the conformational rigidity and improve hydrophobic core packing both histidines of the second, non-covalent tetrapyrrole-binding site were replaced with phenylalanine (C46) (Fig. B.1a). This removed the ability to bind a second tetrapyrrole and increased the melting transition temperature (T_m) by 31 to 84°C. The distal histidine on helix 2 was subsequently replaced with phenylalanine, yielding the monohistidine-ligated CTM C45. The resulting maquette is a predominantly helical protein that retains the hyperthermophilic characteristics of its precursor (T_m(C45) = 86 °C) and reversibly refolds following thermal denaturation (Fig. B.1b, c).

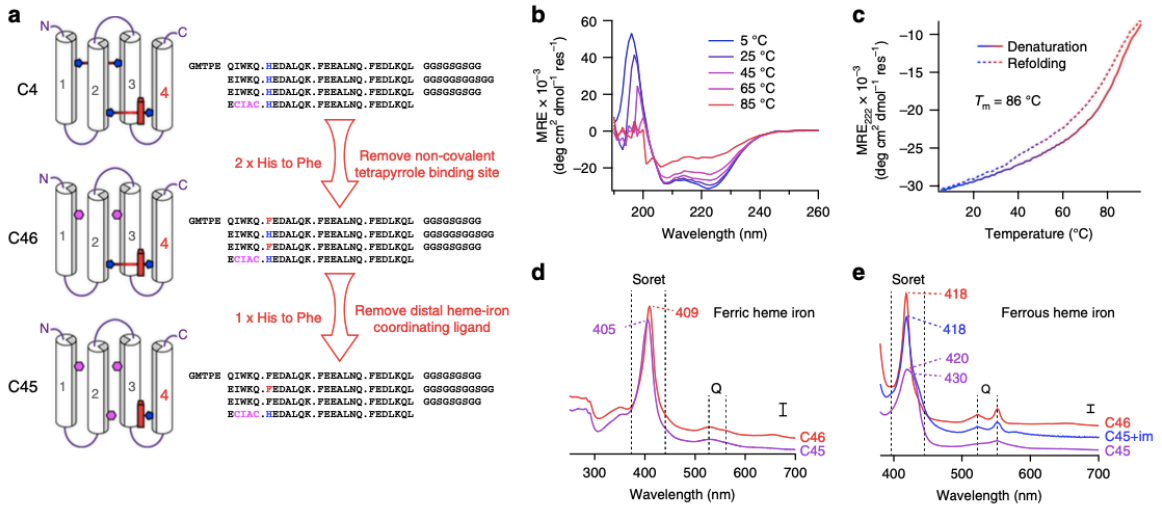


Fig. B.1: Design and characterisation of an artificial peroxidase. a) The design process begins with C4, a CTM containing a bis-histidine haem C-binding site with the consensus motif for covalent haem incorporation on helix 4 with the distal histidine haem ligand on helix 2, and a second non-covalent tetrapyrrole-binding site (occluded by helix 2) consisting of a bis-histidine pair on helices 1 and 3. Mutation of both histidines of the non-covalent-binding site to phenylalanine produces the CTM, C46. Subsequent mutation of the distal haem C-ligating histidine ligand results in the mono-histidine ligated C45. b) Far-UV circular dichroism spectra of C45 with varying temperature collected in 100 mM KCl, 20 mM CHES, pH 8.58. c) Temperature dependence of the CD signal monitored at 222 nm during denaturation (solid line) and refolding (dashed line). d) & e) UV/visible spectra of ferric (d) and ferrous (e) C45 (purple) and C46 (red). The spectrum of ferrous C45 with added exogenous imidazole is displayed in blue. Scale bars represent optical densities of 0.1 (taken from Watkins et al., 2017[1]).

The UV/visible spectra of the C45 construct indicates the presence of both ferric and ferrous iron species (Fig. B.1d, e), similar to those exhibited by the monohistidine-ligated cytochrome *c′* [16] and horse heart cytochrome *c* distal methionine mutants[17]. This suggests either some mixture of both low- and high-spin states exists at the haem iron[18] or a titratable water molecule occupies the 'vacant' axial coordination site[19-20]. The ferric C45 spectrum is pH dependent; a pK_a of 8.11 is observed for the transition between spectra at neutral and alkaline pH (pH 6.5–11), similar to the acid-alkaline transition pK_a of horse heart myoglobin (pK_a = 8.9)[19]. This suggests the presence of a haem C-bound water molecule occupying the available axial site. C45 displays a distinct pH-dependence in its haem redox potential, increasing from −208 mV at pH 10 to −174 mV at pH 8.6. At pH 7.5 two redox potentials are observed, likely reflecting the aforementioned mixture of spin states and the water/hydroxide equilibrium at the distal coordination site. C45 also retains the diatomic ligand-binding functionality of its CTM precursors[13-14] (both carbon monoxide and oxygen are readily bound by C45

in the ferrous form, and cyanide in its ferric form). The oxygen-binding capability suggests that C45 has retained the relatively 'dry' and conformationally stable oxygen-binding site on the distal haem face from previous designs[12-13].

When rapidly mixed with hydrogen peroxide, the monohistidine-ligated ferric C45 reacts with the classical peroxidase substrate ABTS (2,2′-azino-bis(3-ethylbenzothiazo- line-6-sulfonic acid))[15] to form the green ABTS radical cation (Fig. B.2a). At the optimum pH of 8.58 the reaction follows the ping-pong steady-state kinetics typical of natural peroxidases[21] (Fig. B.2b), with exceptional catalytic efficiency for the electron transfer between peroxide-activated C45 and ABTS. This catalytic efficiency is among the highest observed in a *de novo* enzyme and matches well that of horseradish peroxidase (HRP) operating at its optimal pH[22]. Despite the high efficiency of the electron transfer step, the activation of peroxide by C45 is less catalytically efficient than in HRP[23], and is likely due to the lack of peroxide-activating amino acid side chains in the vicinity of the haem. Compared to other *de novo* haem-containing proteins that exhibit peroxidase activity, C45 is markedly more catalytically efficient and does not require helix-stabilising additives such as trifluoroethanol for catalytic activity[24-26]. C45 is remarkably resilient to both elevated temperature (Fig. B.2c) and the presence of organic solvents and retains impressive catalytic efficiency close to its denaturation temperature.

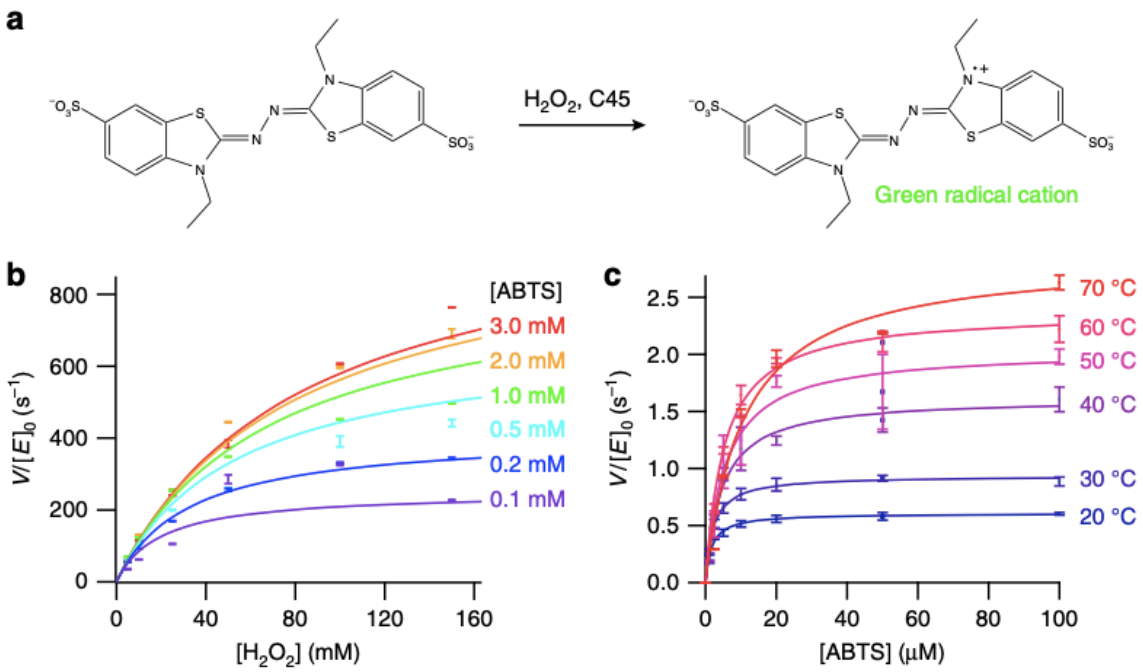


Fig. B.2: Steady-state kinetics of ABTS and hydrogen peroxide turnover by C45. a) Transformation of ABTS to a green radical cation as catalysed by natural peroxidases and C45. b) Steady-state kinetics plot for ABTS turnover by C45 at varying [ABTS] and [H_2O_2]. Data are fit to a ping-pong steady-state kinetics model. c) The thermal stability of C45 enables classical Michaelis–Menten kinetics to be determined up to 70 °C. All data were recorded in triplicate and error bars represent the standard deviation. Kinetic data were collected in 20 mM CHES, 100 mM KCl, pH 8.58 with 0.1 μM C45 (taken from Watkins et al., 2017[1]).

Previous work has illustrated that haem maquettes can undergo rapid interprotein electron transfer with natural proteins such as cytochrome *c*, facilitated by complementary surface electrostatic interactions[27]. Since C45 presents similar electrostatic surfaces to these maquettes, it could be reasoned that ferrous cytochrome *c* would act as an electron donor in the peroxidase reaction with C45, reproducing the activity of the natural cytochrome *c* peroxidases[15]. C45 was indeed found to be capable of functioning as an impressively efficient artificial cytochrome *c* peroxidase at limiting H_2O_2 concentration, compared to the natural yeast cytochrome *c* peroxidase[28]. This demonstrates the ability of C45 to couple interprotein electron transfer to a chemical transformation at a protein-bound haem, an intrinsic feature of respiratory enzymes such as cytochrome oxidase[29].

Much like the natural haem-containing peroxidases[15], C45 exhibits significant substrate promiscuity: it catalyses the oxidation of peroxidase substrates including guaiacol, p-anisidine, o-phenylenediamine, 5-aminosalicylic acid, luminol, reactive blue 4, reactive black 5, and the anti-tuberculosis prodrug, isoniazid. This catalytic promiscuity also extends to the oxidative dehalogenation of halogenated phenols

such as 2,4,6-trichlorophenol (TCP), its bromo- and fluoro-analogues, and 4-bromophenol. As before, at a limiting hydrogen peroxide concentration, C45 was found to have far greater catalytic efficiency compared to the natural dehaloperoxidase (from *Amphitrite ornate*[30]). Furthermore, kinetics at non-limiting H_2O_2 concentrations are also predicted to be significantly higher than for the natural enzyme.

To further probe the artificial peroxidase mechanism, ferric C45 was rapidly mixed with hydrogen peroxide or organic peracids in a stopped flow spectrophotometer. Approximately 5s later, a spectroscopic intermediate (Fig. B.3a, b) was observed resembling cytochrome *c* peroxidase compound I[28]. This was subsequently examined using electron paramagnetic resonance (EPR) spectroscopy and a new narrow EPR signal centred around g=2.0032 was observed with a line width of 1.1 mT (peak to peak). Microwave power saturation experiments at X-band and Q-band (34 GHz) suggest that this is an isolated species indicative of an amino acid-based radical. The lack of any discernible g-anisotropy at 34 GHz indicates a tryptophan side chain as the origin of this radical species (most likely W43 which is closest to the haem), analogous to the characteristics of cytochrome *c* peroxidase compound I[31-32].

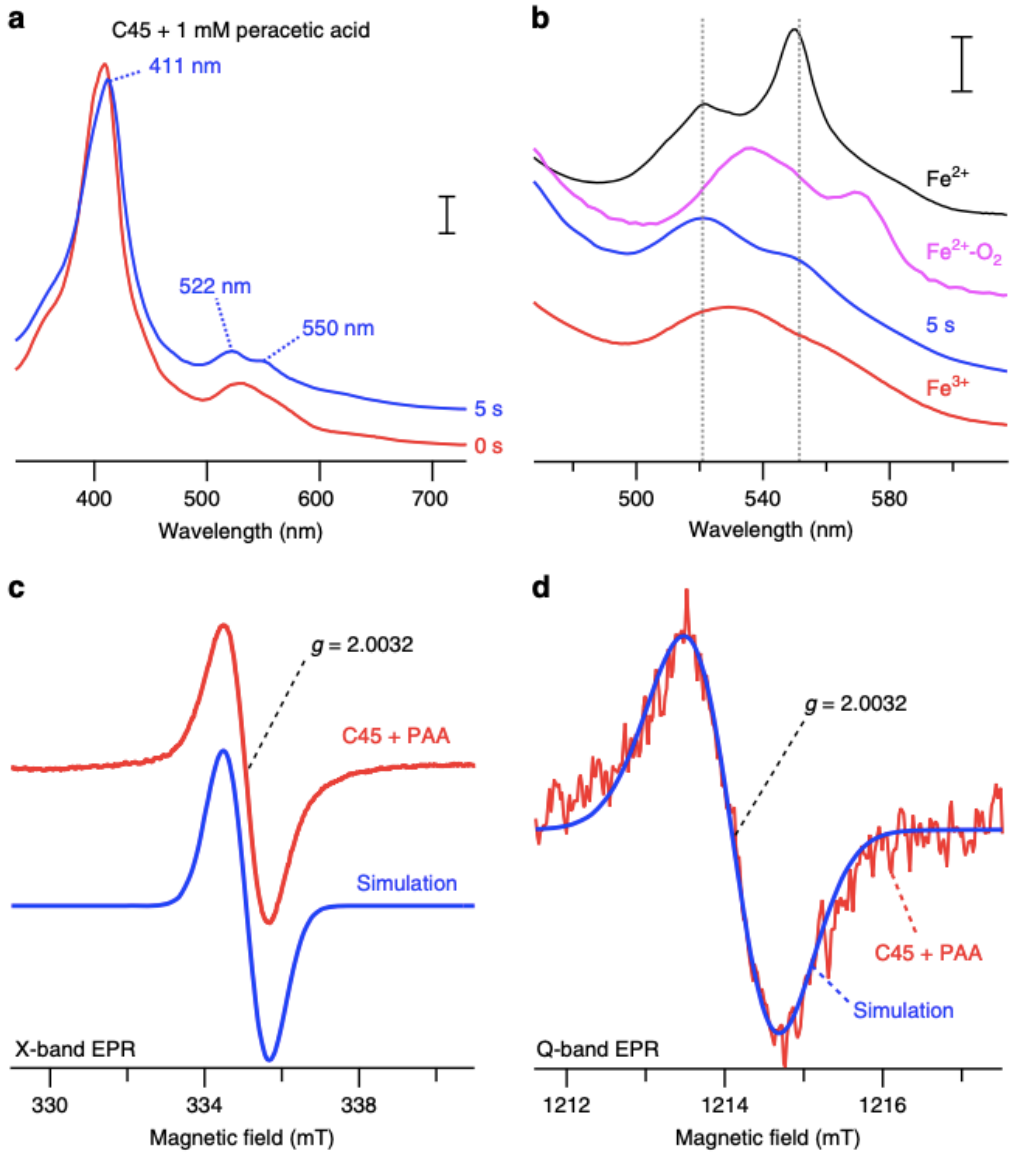


Fig. B.3: Isolation of high-valent oxo-iron species in a CTM. a) & b) UV/visible spectra of ferric C45 (red) and peracetic acid-treated C45 (blue) obtained by rapid mixing experiments in a stopped-flow spectrophotometer. Ferrous (black), ferric (red), and oxyferrous C45 (magenta) spectra are displayed for comparison. The putative C45 compound I species were generated by mixing 2 mM peracetic acid with 20 µM ferric C45 in 100 mM KCl, 20 mM CHES, pH 8.58. Scale bars represent optical densities of 0.05 (a) and 0.02 (b). c) X-band cw-EPR spectrum of C45 mixed with peracetic acid (red) indicates the formation of a radical species with g = 2.0032. Simulated data of a tryptophan radical species within C45 are presented in blue. Spectra were obtained by mixing 1 mM peracetic acid with C45 (700 µM) in 100 mM KCl, 20 mM CHES, pH 8.58. Experimental conditions: EPR microwave frequency = 9.3933 GHz, microwave power = 1 mW, modulation amplitude = 0.3 mT, temperature = 12 K. d) Q-band cw-EPR spectrum of C45 mixed with peracetic acid (red, conditions as for the X-band EPR data) and simulated data of a tryptophan radical species in C45 (blue). The lack of observable g-anisotropy of the radical signal indicates the presence of an amino acid side chain-based radical species. Experimental conditions: EPR microwave frequency = 34.027 GHz, microwave power = 3 µW, modulation amplitude = 0.3 mT, temperature = 50 K (taken from Watkins et al., 2017[1]).

The study goes on to further probe the system using nuclear magnetic resonance (NMR) spectroscopy. The NMR data indicate that ferric C45 exists in a conformationally heterogeneous and dynamic state. However, during O_2-binding, C45 forms a discrete, monomeric, and thermally stable structure with scant evidence of conformational heterogeneity. These structural characteristics and rather unusual ^{1}H-^{15}N spectra are observed in many haem-containing maquettes[4, 13] and *de novo* proteins[33], and evidently do not inhibit efficient peroxidase activity in C45. Indeed, there are notable examples of natural and engineered enzymes that exhibit analogous flexibility either globally or locally at the active site[34-36]. In these cases, the substrate is thought to confer structural rigidity through the induced-fit mechanism.

To further investigate the structure and substrate interactions of C45, a computational model of C45 was created using Chimera[37] and molecular dynamics (MD) simulations run using GROMACS[38] with the CHARMM27 forcefield[39]. ABTS and TCP binding was probed using BUDE (Bristol University Docking Engine)[40].

The models indicate that ABTS may bind in up to 10 sites on the surface of the protein with similar binding energies. However, there are relatively few hotspots for TCP binding despite the variation in amino acid side chain conformation in the starting C45 poses, and the highest calculated binding energy for TCP is for a cleft close to the haem (<10 Å), which is ideal for the rapid electron tunnelling necessary to support the observed catalytic rate[41]. This binding mode is stable in MD simulations. The two computational methods concur in identifying the binding site for TCP, suggestive of binding with good affinity and the indications are that this substrate-binding site is relevant for catalytic activity, offering potential for enhancing specify through designed modification. Despite the differences between the ABTS and TCP-binding interactions with C45, the substrate binding hotspots are all found at the protein surface. It can therefore be surmised that C45 behaves in a similar manner to lignin peroxidase[42] and other natural peroxidases, presenting an interaction surface for substrate to dock prior to rapid electron transfer to the reactive haem intermediates.

The simple, systematic construction of a functional *de novo* enzyme demonstrated here highlights the power and utility of the maquette approach to functional oxidoreductase design. Similar to the construction of the original oxygen-binding

maquette[12], the engineering steps required to attain the desired catalytic function in this simple protein scaffold are few in number, reflecting the relative ease with which evolutionarily naive proteins can be rationally engineered in a tractable, iterative process. In this case, the concomitant restriction of conformational flexibility and removal of a haem ligand were sufficient to attain efficient, multistep catalysis within a heat-resistant and solvent-resistant maquette chassis, representing just three amino acid substitutions from the oxygen-binding progenitor.

Despite the relatively low-catalytic efficiency of hydrogen peroxide turnover (104 M^{-1} s^{-1}), the enzymatic rate constant k_{cat}(ABTS) is among the highest yet observed in a *de novo* enzyme, and k_{cat}/K_m(ABTS) is greater than the catalytic efficiencies of the best reported *de novo* enzymes[6, 9]. Given the lack of a highly specific substrate-binding site, it is perhaps unsurprising that C45 exhibits broad substrate promiscuity. However, this catalytic promiscuity may be a general feature of primitive, nascent enzymes[43]. Since C45 is fully assembled *in vivo*, there is now a tantalising opportunity to employ both rational protein design and directed evolution methodologies to optimise the catalytic chassis toward a selected substrate or chemical mechanism. As such the descendants of this *de novo* designed enzyme may act as cheap, green, and catalytically proficient catalysts for industrial biotechnology and versatile bioblocks for synthetic biology, while providing insight into the rules underpinning the engineering of natural and designed enzymes.

The EPR samples were prepared using hydrogen peroxide or peracetic acid (1 mM), added to C45 (700 μM in 100 mM KCl, 20 mM CHES, pH 8.58) in suprasil quartz sample tubes and flash frozen in liquid nitrogen. X-band cw-EPR spectra were recorded on a Bruker eleXsys E500 spectrometer using a standard rectangular Bruker EPR cavity (ER4102T) equipped with an Oxford helium cryostat (ESR900). Experimental parameters: microwave power, 1 mW; field modulation amplitude, 3 G; field modulation frequency, 100 kHz; measuring time 160 s; temperature 12 K. Q-band cw-EPR spectra were performed on a Bruker eleXsys E-560 spectrometer using a ER 5106QT-W1 resonator equipped with a home-built ARS cryogen-free cryostat. The measured spectra were corrected for an offset against a known g standard [1,1-diphenyl-2-picrylhydrazyl, DPPH, g = 2.00351 ± 0.00002]. Spectral simulations were performed using the Matlab-based Easyspin package[44].

References

1. Watkins, D. W.; Jenkins, J. M. X.; Grayson, K. J.; Wood, N.; Steventon, J. W.; Le Vay, K. K.; Goodwin, M. I.; Mullen, A. S.; Bailey, H. J.; Crump, M. P.; Macmillan, F.; Mulholland, A. J.; Cameron, G.; Sessions, R. B.; Mann, S.; Anderson, J. L. R., Construction and in Vivo Assembly of a Catalytically Proficient and Hyperthermostable De Novo Enzyme. *Nat. Commun.* **2017,** *8*, 9.

2. Nanda, V.; Koder, R. L., Designing Artificial Enzymes by Intuition and Computation. *Nature Chemistry* **2010,** *2* (1), 15-24.

3. Watkins, D. W.; Armstrong, C. T.; Anderson, J. L. R., De Novo Protein Components for Oxidoreductase Assembly and Biological Integration. *Current Opinion in Chemical Biology* **2014,** *19*, 90-98.

4. Farid, T. A.; Kodali, G.; Solomon, L. A.; Lichtenstein, B. R.; Sheehan, M. M.; Fry, B. A.; Bialas, C.; Ennist, N. M.; Siedlecki, J. A.; Zhao, Z.; Stetz, M. A.; Valentine, K. G.; Anderson, J. L. R.; Wand, A. J.; Discher, B. M.; Moser, C. C.; Dutton, P. L., Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nature Chemical Biology* **2013,** *9* (12), 826-+.

5. Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; Dechancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D., Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008,** *453* (7192), 190-U4.

6. Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Gruetter, M. G.; Privett, H. K.; Mayo, S. L.; Hilvert, D., Precision Is Essential for Efficient Catalysis in an Evolved Kemp Eliminase. *Nature* **2013,** *503* (7476), 418-+.

7. Kiss, G.; Celebi-Oelcuem, N.; Moretti, R.; Baker, D.; Houk, K. N., Computational Enzyme Design. *Angewandte Chemie-International Edition* **2013,** *52* (22), 5700-5725.

8. Reig, A. J.; Pires, M. M.; Snyder, R. A.; Wu, Y.; Jo, H.; Kulp, D. W.; Butch, S. E.; Calhoun, J. R.; Szyperski, T. A.; Solomon, E. I.; Degrado, W. F., Alteration of the Oxygen-Dependent Reactivity of De Novo Due Ferri Proteins. *Nature Chemistry* **2012,** *4* (11), 900-906.

9. Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L., Hydrolytic Catalysis and Structural Stabilization in a Designed Metalloprotein. *Nature Chemistry* **2012,** *4* (2), 118-123.

10. Tebo, A. G.; Pecoraro, V. L., Artificial Metalloenzymes Derived from Three-Helix Bundles. *Current Opinion in Chemical Biology* **2015,** *25*, 65-70.

11. Nastri, F.; Chino, M.; Maglio, O.; Bhagi-Damodaran, A.; Lu, Y.; Lombardi, A., Design and Engineering of Artificial Oxygen-Activating Metalloenzymes. *Chemical Society Reviews* **2016,** *45* (18), 5020-5054.

12. Koder, R. L.; Anderson, J. L. R.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L., Design and Engineering of an O-2 Transport Protein. *Nature* **2009,** *458* (7236), 305-U64.

13. Anderson, J. L. R.; Armstrong, C. T.; Kodali, G.; Lichtenstein, B. R.; Watkins, D. W.; Mancini, J. A.; Boyle, A. L.; Farid, T. A.; Crump, M. P.; Moser, C.

C.; Dutton, P. L., Constructing a Man-Made C-Type Cytochrome Maquette in Vivo: Electron Transfer, Oxygen Transport and Conversion to a Photoactive Light Harvesting Maquette. *Chemical Science* **2014,** *5* (2), 507-514.

14. Watkins, D. W.; Armstrong, C. T.; Beesley, J. L.; Marsh, J. E.; Jenkins, J. M. X.; Sessions, R. B.; Mann, S.; Anderson, J. L. R., A Suite of De Novo C-Type Cytochromes for Functional Oxidoreductase Engineering. *Biochimica Et Biophysica Acta-Bioenergetics* **2016,** *1857* (5), 493-502.

15. Poulos, T. L., Heme Enzyme Structure and Function. *Chemical Reviews* **2014,** *114* (7), 3919-3962.

16. Zahn, J. A.; Arciero, D. M.; Hooper, A. B.; Dispirito, A. A., Cytochrome C' of Methylococcus Capsulatus Bath. *European Journal of Biochemistry* **1996,** *240* (3), 684-691.

17. Wallace, C. J. A.; Clarklewis, I., Functional-Role of Heme Ligation in Cytochrome-C - Effects of Replacement of Methionine-80 with Natural and Nonnatural Residues by Semisynthesis. *Journal of Biological Chemistry* **1992,** *267* (6), 3852-3861.

18. Weiss, R.; Gold, A.; Terner, J., Cytochromes C ': Biological Models for the S = (3)/(2),(5)/(2) Spin-State Admixture? *Chemical Reviews* **2006,** *106* (6), 2550-2579.

19. Antonini, E., *Hemoglobin and Myoglobin in Their Reactions with Ligands. [by] Eraldo Antonini and Maurizio Brunori*. North-Holland Pub. Co: Amsterdam, 1971.

20. Lu, Y.; Casimiro, D. R.; Bren, K. L.; Richards, J. H.; Gray, H. B., Structurally Engineered Cytochromes with Unusual Ligand-Binding Properties - Expression of Saccharomyces-Cerevisiae Met-80- Ala Iso-1-Cytochrome-C. *Proceedings of the National Academy of Sciences of the United States of America* **1993,** *90* (24), 11456-11459.

21. Dunford, H. B., Horseradish Peroxidase: Structure and Kinetic Properties. In *Peroxidase in Chemistry and Biology, Vol. 2*, Everse, J.; Everse, K. E.; Grisham, M. B., Eds. CRC Press: 1991.

22. Savenkova, M. I.; Kuo, J. M.; De Montellano, P. R. O., Improvement of Peroxygenase Activity by Relocation of a Catalytic Histidine within the Active Site of Horseradish Peroxidase. *Biochemistry* **1998,** *37* (30), 10828-10836.

23. Rodriguezlopez, J. N.; Smith, A. T.; Thorneley, R. N. F., Role of Arginine 38 in Horseradish Peroxidase a Critical, Residue for Substrate Binding and Catalysis. *Journal of Biological Chemistry* **1996,** *271* (8), 4023-4030.

24. Moffet, D. A.; Certain, L. K.; Smith, A. J.; Kessel, A. J.; Beckwith, K. A.; Hecht, M. H., Peroxidase Activity in Heme Proteins Derived from a Designed Combinatorial Library. *Journal of the American Chemical Society* **2000,** *122* (31), 7612-7613.

25. Faiella, M.; Maglio, O.; Nastri, F.; Lombardi, A.; Lista, L.; Hagen, W. R.; Pavone, V., De Novo Design, Synthesis and Characterisation of Mp3, a New Catalytic Four-Helix Bundle Hemeprotein. *Chemistry-a European Journal* **2012,** *18* (50), 15960-15971.

26. Vitale, R.; Lista, L.; Cerrone, C.; Caserta, G.; Chino, M.; Maglio, O.; Nastri, F.; Pavone, V.; Lombardi, A., An Artificial Heme-Enzyme with Enhanced Catalytic Activity: Evolution, Functional Screening and Structural Characterization. *Organic & Biomolecular Chemistry* **2015,** *13* (17), 4859-4868.

27. Fry, B. A.; Solomon, L. A.; Dutton, P. L.; Moser, C. C., Design and Engineering of a Man-Made Diffusive Electron-Transport Protein. *Biochimica Et Biophysica Acta-Bioenergetics* **2016,** *1857* (5), 513-521.

28. Murphy, E. J.; Metcalfe, C. L.; Basran, J.; Moody, P. C. E.; Raven, E. L., Engineering the Substrate Specificity and Reactivity of a Heme Protein: Creation of an Ascorbate Binding Site in Cytochrome C Peroxidase. *Biochemistry* **2008,** *47* (52), 13933-13941.

29. Yoshikawa, S.; Shimada, A., Reaction Mechanism of Cytochrome C Oxidase. *Chemical Reviews* **2015,** *115* (4), 1936-1989.

30. Le, P.; Zhao, J.; Franzen, S., Correlation of Heme Binding Affinity and Enzyme Kinetics of Dehaloperoxidase. *Biochemistry* **2014,** *53* (44), 6863-6877.

31. Casadei, C. M.; Gumiero, A.; Metcalfe, C. L.; Murphy, E. J.; Basran, J.; Concilio, M. G.; Teixeira, S. C. M.; Schrader, T. E.; Fielding, A. J.; Ostermann, A.; Blakeley, M. P.; Raven, E. L.; Moody, P. C. E., Neutron Cryo-Crystallography Captures the Protonation State of Ferryl Heme in a Peroxidase. *Science* **2014,** *345* (6193), 193-197.

32. Harvey, J. N.; Bathelt, C. M.; Mulholland, A. J., Qm/Mm Modeling of Compound I Active Species in Cytochrome P450, Cytochrome C Peroxidase, and Ascorbate Peroxidase. *J Comput Chem* **2006,** *27* (12), 1352-1362.

33. Murphy, G. S.; Greisman, J. B.; Hecht, M. H., De Novo Proteins with Life-Sustaining Functions Are Structurally Dynamic. *Journal of Molecular Biology* **2016,** *428* (2), 399-411.

34. Vamvaca, K.; Vogeli, B.; Kast, P.; Pervushin, K.; Hilvert, D., An Enzymatic Molten Globule: Efficient Coupling of Folding and Catalysis. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (35), 12860-12864.

35. Schulenburg, C.; Stark, Y.; Kuenzle, M.; Hilvert, D., Comparative Laboratory Evolution of Ordered and Disordered Enzymes. *Journal of Biological Chemistry* **2015,** *290* (15), 9310-9320.

36. Forouhar, F.; Anderson, J. L. R.; Mowat, C. G.; Vorobiev, S. M.; Hussain, A.; Abashidze, M.; Bruckmann, C.; Thackray, S. J.; Seetharaman, J.; Tucker, T.; Xiao, R.; Ma, L.-C.; Zhao, L.; Acton, T. B.; Montelione, G. T.; Chapman, S. K.; Tong, L., Molecular Insights into Substrate Recognition and Catalysis by Tryptophan 2,3-Dioxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (2), 473-478.

37. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., Ucsf Chimera - a Visualization System for Exploratory Research and Analysis. *J Comput Chem* **2004,** *25* (13), 1605-1612.

38. Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R., Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Computer Physics Communications* **1995,** *91* (1-3), 43-56.

39. Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E., Implementation of the Charmm Force Field in Gromacs: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *Journal of Chemical Theory and Computation* **2010,** *6* (2), 459-466.

40. Mcintosh-Smith, S.; Price, J.; Sessions, R. B.; Ibarra, A. A., High Performance in Silico Virtual Drug Screening on Many-Core Processors. *International Journal of High Performance Computing Applications* **2015,** *29* (2), 119-134.

41. Moser, C. C.; Anderson, J. L. R.; Dutton, P. L., Guidelines for Tunneling in Enzymes. *Biochimica Et Biophysica Acta-Bioenergetics* **2010,** *1797* (8), 1573-1586.

42. Miki, Y.; Calvino, F. R.; Pogni, R.; Giansanti, S.; Ruiz-Duenas, F. J.; Jesus Martinez, M.; Basosi, R.; Romero, A.; Martinez, A. T., Crystallographic, Kinetic, and Spectroscopic Study of the First Ligninolytic Peroxidase Presenting a Catalytic Tyrosine. *Journal of Biological Chemistry* **2011,** *286* (17), 15525-15534.

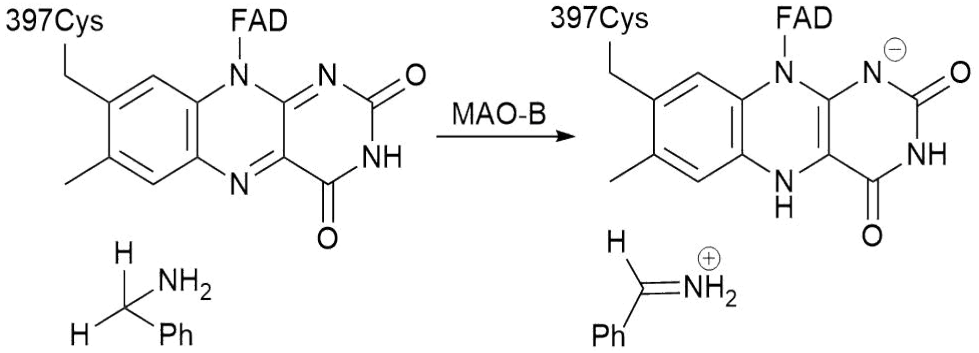
43. Jensen, R. A., Enzyme Recruitment in Evolution of New Function. *Annual Review of Microbiology* **1976,** *30*, 409-425.

44. Stoll, S.; Schweiger, A., Easyspin, a Comprehensive Software Package for Spectral Simulation and Analysis in Epr. *Journal of Magnetic Resonance* **2006,** *178* (1), 42-55.

Appendix C – Monoamine Oxidase B

The following is an excerpt of a peer-reviewed publication reproduced in part with permission from Jones *et al.*, 2019[1] (see the original publication for the full study and Supporting Information). Contributions here include acquisition and analysis of EPR spectra, fig. C.2 and a portion of the text describing and explaining the EPR data.

Human monoamine oxidase B (MAO-B) catalyses the oxidative deamination of amines, by electron transfer, via its flavin adenine dinucleotide (FAD) cofactor (Scheme C.1)[2] . Inhibition of this enzyme is used for the treatment of both Parkinson's disease and depression[3-4]. It is located in the outer mitochondrial membrane[5], as a dimer, with bipartite substrate binding and active site cavities[2].



Scheme C.1: General reaction catalysed by MAO-B (figure taken from Jones et al, 2019[1])

Though this enzyme is of undeniable medical importance, its catalytic mechanism remains unclear, and novel means of targeting this process could prove invaluable. Throughout literature there is evidence of a radical in either the transition state or the resting state of MAO-B, which is suggested to be either a flavin semiquinone or a tyrosyl radical, or both. Presented here is evidence of a resting state flavin semiquinone, which was determined via absorption redox studies and electron paramagnetic resonance (EPR) spectroscopy.

The reduction of the FAD has been shown to proceed by a tunnelling mechanism via primary (1°) and secondary (2°) kinetic isotope effect (KIE) studies[6]; these studies have suggested that MAO-B catalysis is not linked to fast (pico/nanosecond) dynamics[6-7]. MAO-B catalysis has been investigated via experimental and computational studies, with at least seven different proposed mechanisms, including polar nucleophilic[8], radical[9], direct hydride transfer[10-12] and

two-step hydride transfer[13]. The original direct single electron transfer (SET) radical mechanism was previously discounted, both experimentally[14] and by quantum mechanics/molecular mechanics (QM/MM)[15]. A separate radical mechanism was then proposed by Murray *et a*[16], which used a model small molecule reaction that mimics MAO-B to provide evidence for a neutral semiquinone flavin that can be formed aerobically. It was suggested that a neutral semiquinone flavin is the reactive species for the oxidation of benzylamine (BZA)[16]. This led to the hypothesis of a radical mechanism whereby MAO-B forms a neutral semiquinone flavin via a proximal tyrosine radical (Y398). The presence of a stable anionic semiquinone flavin and tyrosyl radical intermediate is also reported in MAO-A and MAO-N[17-18].

Previous observation of a stable anionic semiquinone FAD in MAO-B[19-20] led to a proposition of half-site reactivity, which posits that one monomer of the MAO-B dimer is inaccessible to oxygen and the substrate, resulting in the formation of the stable semiquinone species, while the other contains oxidised FAD. Electrons are then shuttled to the semiquinone species, allowing for the oxidation of the reduced FAD upon turnover[20]. It is suggested that this might arise from electron shuttling between specific amino acids. Potentially, such a mechanism might require significant conformational change associated with turnover in each monomer. Indeed, conformational changes have been found to be associated with MAO-B turnover, with a MD study demonstrating the potential for the membrane to regulate access to the active site entrance via two gating loops (residues 85−119 and 155−165)[21]. Other MD studies have been carried out on MAO-B, both with and without a membrane environment, with the focus on identifying or improving inhibitors for MAO-B[22-26] or ascertaining how MAO-B binds to the membrane[27].

Here, MAO-B was expressed and purified in *Pichia pastoris*[28]. Previous studies in which MAO-B was purified from bovine liver had attributed observations of oxidised/semiquinone MAO-B FAD to reactive oxygen species (ROS) that form upon the disruption of the mitochondrial membrane, which were not observed upon purification from *P. pastoris*[28]. The presence of purified MAO-B was established by SDS-PAGE and via electrospray ionization quadrupole time-of-flight (ESI-Q TOF) mass spectrometry in combination with the MASCOT server[29].

The absorption spectrum of the purified MAO-B is shown in Figure 1A. The spectra share the characteristics feature at ~415 nm[30] of an anionic semiquinone

FAD (FAD•−). Multiple preparations of MAO-B were completed using production and purification protocols of MAO-B were replicated from those used by Newton-Vinson *et al*[28], with this ~415 nm spectral feature consistently present. The effect of anaerobic incubation with BZA on the MAO-B feature at ~415 nm was monitored temporally (Figure C.1A, inset). From Figure 1A, inset, there is a clear decrease in absorption at 415 nm with respect to time upon incubation with BZA, suggesting the ~415 nm spectral feature is redox-sensitive with a natural MAO-B substrate.

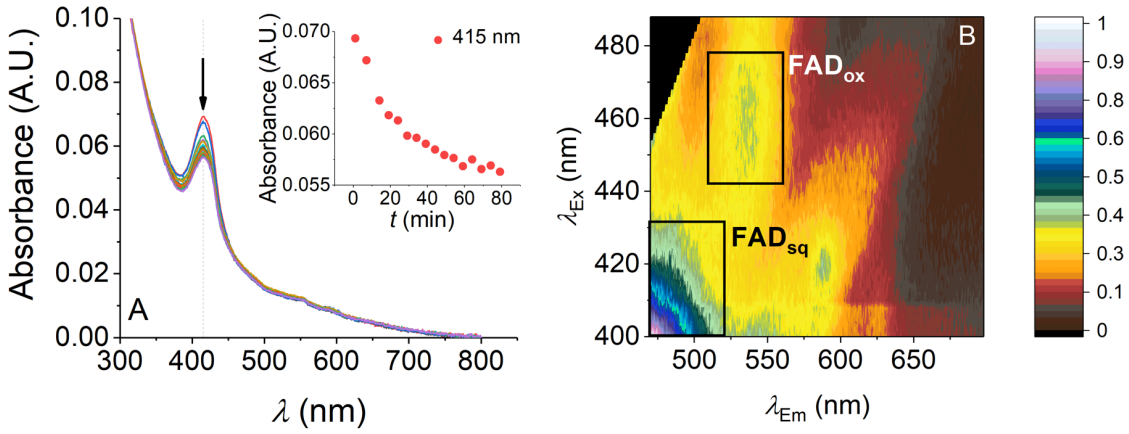


Figure C.1. (A) The absorbance spectra of MAO-B recorded at various time points after treatment with BZA. (Inset) The effect on the 415 nm peak over time. (B) Fluorescence excitation/emission matrix resolves oxidised and semiquinone flavin states (highlighted in solid black boxes). Scale bar is the relative intensity (figure taken from Jones et al, 2019[1]). Conditions: 50 mM HEPES 0.5% w/v Triton X-100, 20 °C; (for absorption experiments) 40 mM BZA and anaerobic conditions as shown in Materials and Methods.

The absorption spectrum shown in Figure C.1A lacks a defined feature around ~450 nm where one would expect oxidised flavin, as well as spectral features at ~475 nm that would also characterise a putative anionic semiquinone. There is no clear explanation for the lack of these features and the complexity of the absorption spectrum given that the protein is well purified and identified as MAO-B by mass spectrometry and is shown to be active with the natural substrates (Figure C.1A, inset).

The electronic environment around both flavins is highly complex, with three proximal Tyr residues (Y60, 398, and 435 positioned 3.1, 3.3, and 3.5 Å from the alloxan moiety, respectively). This may result in a composite absorption spectrum arising from different electronic environments for a subset of active site conformational states; this relies on MAO-B to be able to explore different conformational states, which is considered in more detail below. Given the

complexity of the absorption spectrum, fluorescence and electron paramagnetic resonance (EPR) spectroscopy were used to provide more specific evidence for the oxidation state(s) of the MAO-B flavin.

The ~415 nm feature may be attributable to a tyrosyl radical, which has a characteristic absorbance peak at 410 nm[31], or a mixture of both an FAD semiquinone and tyrosyl radical, as seen as intermediates in MAO-A[17-18]. It was postulated that the reactive semiquinone FAD was formed by a proximal tyrosyl radical (Y398), but the UV/vis absorption spectrum of Y398F MAO-B also exhibits the feature at ~415 nm, indicating that this suggestion is likely incorrect. The Y398F variant also shows a slight increase in K_m, similar in magnitude to those previously reported[32].

Previous steady-state fluorescence spectroscopy of MAO-B revealed two distinct chromophores in the resting-state[33]. It was concluded that these two species were consistent with separate oxidised and semiquinone flavins. To establish if this was the case here, the fluorescence excitation−emission matrix was monitored (Figure 1B). Similar to the previous study, at least two different emission peaks were resolved, suggesting the presence of two oxidation states of the flavin with λ_{Ex}~400−420 and ~460 nm. Notably, the excitation/emission profile at λ_{Ex}~460 and λ_{Em}~540 nm (Figure 1B) suggests the presence of oxidised flavin.

The low temperature X-band cw-EPR spectrum of the wild-type MAO-B (wtMAO-B) clearly indicates that the signal arises from a semiquinone radical (Figure C.2), in agreement with the UV/vis and fluorescence spectroscopy (Figure C.1). This is identifiable in the measured spectra, as they lack the defining features that would characterise the signal as arising from a tyrosyl radical: the typical “wings” or “shoulders” around the central signal at around g = 2[17, 34] at X-band (Figure C.2) and an increased g-value anisotropy at a higher frequency (and therefore resonant field), i.e., Q-band (shown in Figure C.2B and C.2C). Furthermore, the signal persists in the Y398F variant, confirming that it is not caused by this proximal tyrosine. Computational simulation and fitting of the experimental X-band data of the WT and Y398F MAO-B (Table 1) suggest that the semiquinone radical species is anionic; the hyperfine environment of a neutral semiquinone radical would contain an additional contribution from a hydrogen atom[20, 35], which is not the case for the signals seen in the X-band spectra.

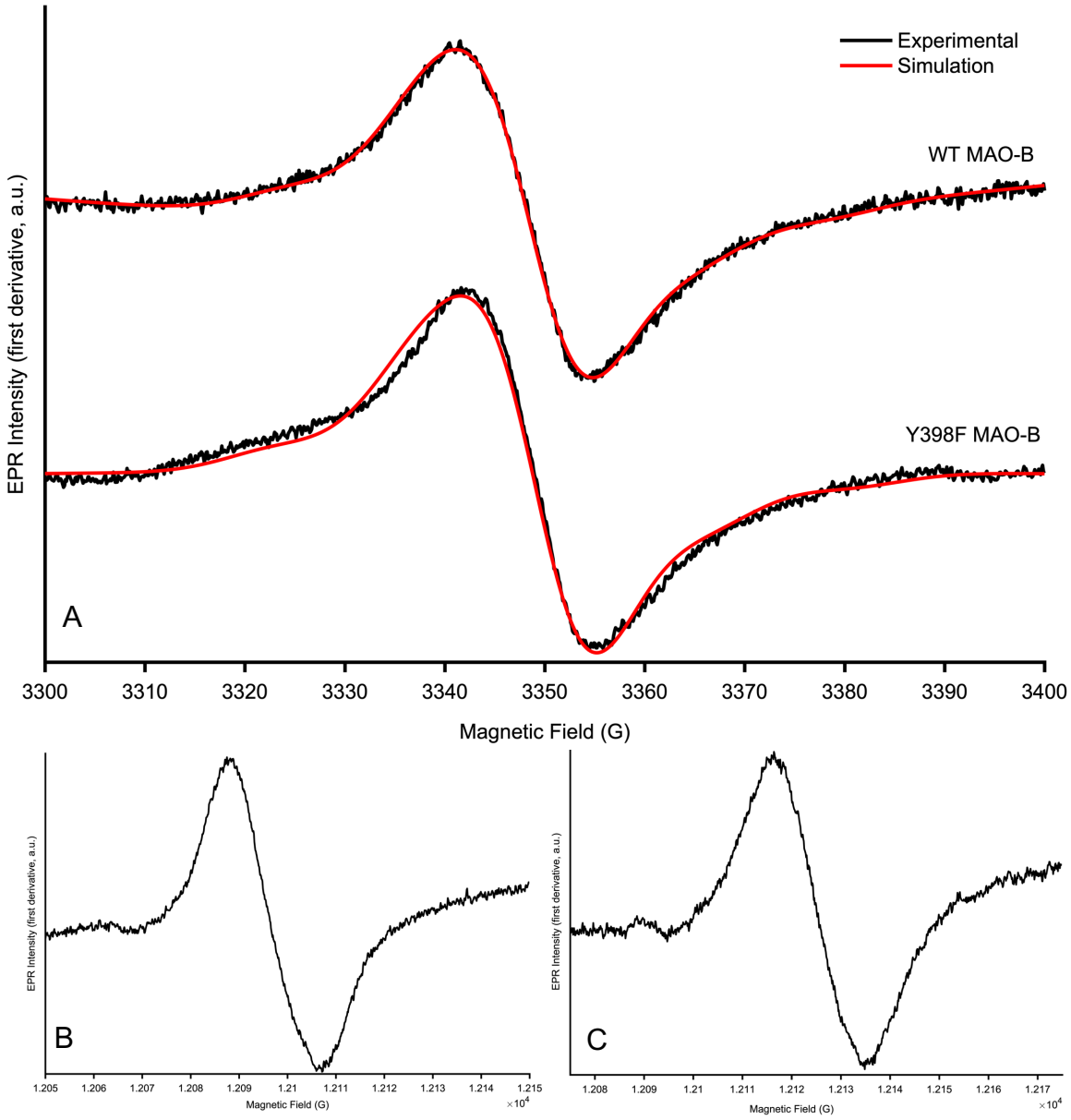


Figure C.2: (A) low temperature X-band cw-EPR spectra of WT (top, black) and Y398F (bottom, black) MAO-B, with their respective fitted simulations (red). EPR microwave frequency = 9.3916 GHz (WT) and 9.3926 GHz (Y398F), microwave power = 0.2 mW, modulation amplitude = 0.5 mT, temperature = 16 K (taken from Jones et al, 2019[1]). (B) low temperature Q-band cw-EPR spectrum of WT MAO-B (C)) low temperature Q-band cw-EPR spectrum of Y398F MAO-B.). EPR microwave frequency = 33.94365 GHz (WT) and 34.021907 GHz (Y398F), microwave power = 0.05 mW and modulation amplitude = 0.2 mT (WT) and 0.7 mT (Y398F), temperature = 20 K.

Table C.1: Parameters of the best fit simulations of the low temperature X-band cw-EPR spectra of WT and Y398F MAO-B (taken from Jones et al, 2019[1]).

Variant	**g-tensor**				**A(^{14}N) (MHz)**		**A(^{14}N) (MHz)**		**Linewidth (mT)**	**Simulation RMSD**
	x	y	z	iso	$A_{//}$	$A_{\perp}$	$A_{//}$	$A_{\perp}$		
WT MAO-B	2.00444	2.00429	2.00191	2.00355	39.3	0	34.1	0	1.1	0.0255
Y398F MAO-B	2.00444	2.00429	2.00191	2.00355	38.4	0	40.1	0	1.1	0.044

Previous studies showed that the membrane environment is important for facilitating the normal enzymatic activity of MAO-B[21]. Spectroscopic studies were conducted to determine whether this putative semiquinoid species was also stably present in an artificial membrane environment (using l-α-phosphatidylcholine styrene maleic acid copolymer, or SMA, nanodiscs[36]). The absorption feature at ~415 nm was present in both reduced Triton X-100 and nanodisc environments, implying that the alleged anionic semiquinone is not an artefact of the buffer system used and that its presence is not affected by the specific membrane environment used. It is clear from these data (Figures 1 and 2) that the FAD of resting-state MAO-B is able to stably occupy oxidation states other than fully oxidised FAD. Evidence for this is recurring in the literature, demonstrated by Raman[19], fluorescence[33] and EPR[20, 28] spectroscopy. These different observations may be rationalised by an equilibrium of energetically similar conformational states, which allow differently stable oxidation states.

The study goes on to explore the influence of the membrane environment on MAO-B turnover using enzyme kinetic investigations, enzyme variant experiments, and MD simulations. The MD simulations reveal a previously undocumented entrance for substrate binding, membrane modulated substrate access, and indications for half-site reactivity, wherein only one active site is accessible to binding at a time. This study combines both experimental and computational evidence to illustrate the subtle interplay between enzyme activity and protein dynamics and the immediate membrane environment. Understanding key biomedical enzymes to this level of detail may prove crucial to inform strategies (and binding sites) for rational drug design for these targets.

It was found that MAO-B in SMA nanodiscs has a significantly different heat capacity of catalysis ($\Delta C_P^{\ddagger}$) compared to that in reduced Triton X-100. This reflects differences in the distribution and frequency of vibrational modes between the ground (reactant) and transition states and implies that the membrane environment is affecting the global protein dynamics of MAO-B and that these dynamics influence the thermodynamics of enzyme turnover. Indeed, k_{cat} is ~5 times faster in the nanodisc environment versus reduced Triton X-100, further hinting at the role of the specific lipid composition and fine structure of the membrane to tune MAO-B turnover efficiency. These findings are corroborated by studying enzyme variants that are predicted to alter the network of flexible motion in the enzyme but do not affect the overall structure. These variants, which are

distal to the active site, all lead to an inactive enzyme, indicating the critical role of "optimised" global protein dynamics of MAO-B.

Through atomistic MD simulations with and without a substrate bound, two substrate entrance/product exit channels were found that are mediated by membrane interaction, one of which was not previously described. The simulations further indicate an asymmetry in substrate access in the MAO-B dimer. Specifically, either one or the other monomer may allow substrate access at any one time, with active site occupation in one monomer preferentially allowing substrate access to the other. Moreover, the gating loop dynamics appear to be highly sensitive to the global enzyme dynamics, potentially reflecting long-range networks of enzyme motion.

Taken together, the study suggests that the global protein dynamics of MAO-B are "tuned" by the specific immediate membrane environment. These protein dynamics have a major effect on MAO-B function, through tuning fluctuations linked to enzyme turnover, including controlling the opening and closing of substrate/product channels. The finding that two different channels mediated by the membrane environment are present in MAO-B illustrates the potential to exploit novel small molecule binding sites with a rational drug design. Therefore, when searching for novel small molecule binding sites, not only the static structure of the system in isolation should be considered but also time-dependent changes in the population of conformational substrates and in the "native" environment[37-38]. The finding that there may be long-range networks of motions that can, in particular, affect the gating loop also indicates the potential for allosteric inhibitors. Further, as MAO-B resides in the outer mitochondrial membrane of cells, the finding of a solvent- accessible entrance is important for rational drug design efforts. That is, inhibitors that target MAO-B may not need to enter the mitochondrial membrane in order to access the active site of MAO-B.

EPR measurements were performed using WT and Y398F variant MAO-B (provided by the group of Dr Pudney, University of Bath, UK) using the protocols described in the original publication[1]) in 50 mM HEPES 0.5% Triton, pH 7.5, flash-frozen in liquid nitrogen in Suprasil quartz sample tubes. X-band cw-EPR spectra were recorded on a Bruker eleXsys E500 spectrometer using a standard rectangular Bruker EPR cavity (ER4102T) equipped with an Oxford helium cryostat (ESR900). The following experimental parameters were used: microwave

power, 0.2 mW; field modulation amplitude, 5 G; field modulation frequency, 100 kHz; measuring time, 10s; temperature, 16 K. Q-band cw-EPR spectra were performed on a Bruker eleXsys E-560 spectrometer using a ER 5106QT-W1 resonator equipped with a home-built ARS cryogen-free cryostat. The following experimental parameters were used: microwave power, 0.05 mW; field modulation amplitude, 2 G (WT) and 7 G (Y398F); field modulation frequency, 100 kHz; measuring time, 41 s; temperature, 20 K. Spectral simulations were performed using the Matlab-based Easyspin package[39].

References

1. Jones, H. B. L.; Crean, R. M.; Mullen, A.; Kendrick, E. G.; Bull, S. D.; Wells, S. A.; Carbery, D. R.; Macmillan, F.; Van Der Kamp, M. W.; Pudney, C. R., Exposing the Interplay between Enzyme Turnover, Protein Dynamics, and the Membrane Environment in Monoamine Oxidase B. *Biochemistry* **2019,** *58* (18), 2362-2372.

2. Binda, C.; Newton-Vinson, P.; Hubalek, F.; Edmondson, D. E.; Mattevi, A., Structure of Human Monoamine Oxidase B, a Drug Target for the Treatment of Neurological Disorders. *Nature Structural Biology* **2002,** *9* (1), 22-26.

3. Knoll, J.; Ecseri, Z.; Kelemen, K.; Nievel, J.; Knoll, B., Phenylisopropylmethylpropinylamine (E-250) a New Spectrum Psychic Energizer. *Archives Internationales De Pharmacodynamie Et De Therapie* **1965,** *155* (1), 154-&.

4. Riederer, P.; Laux, G., Mao-Inhibitors in Parkinson's Disease. *Experimental neurobiology* **2011,** *20* (1), 1-17.

5. Schnaitman, C.; Erwin, V. G.; Greenawalt, J. W., Submitochondrial Localization of Monoamine Oxidase - an Enzymatic Marker for Outer Membrane of Rat Liver Mitochondria. *Journal of Cell Biology* **1967,** *32* (3), 719-+.

6. Jonsson, T.; Edmondson, D. E.; Klinman, J. P., Hydrogen Tunneling in the Flavoenzyme Monoamine-Oxidase-B. *Biochemistry* **1994,** *33* (49), 14871-14878.

7. Nagel, Z. D.; Klinman, J. P., Tunneling and Dynamics in Enzymatic Hydride Transfer. *Chemical Reviews* **2006,** *106* (8), 3095-3118.

8. Macmillar, S.; Edmondson, D. E.; Matsson, O., Nitrogen Kinetic Isotope Effects for the Monoamine Oxidase B-Catalyzed Oxidation of Benzylamine and (1,1-H-2(2))Benzylamine: Nitrogen Rehybridization and Ch Bond Cleavage Are Not Concerted. *Journal of the American Chemical Society* **2011,** *133* (32), 12319-12321.

9. Silverman, R. B., Radical Ideas About Monoamine-Oxidase. *Accounts of Chemical Research* **1995,** *28* (8), 335-342.

10. Akyuz, M. A.; Erdem, S. S., Computational Modeling of the Direct Hydride Transfer Mechanism for the Mao Catalyzed Oxidation of Phenethylamine and Benzylamine: Oniom (Qm/Qm) Calculations. *Journal of Neural Transmission* **2013,** *120* (6), 937-945.

11. Kurtz, K. A.; Rishavy, M. A.; Cleland, W. W.; Fitzpatrick, P. F., Nitrogen Isotope Effects as Probes of the Mechanism of D-Amino Acid Oxidase. *Journal of the American Chemical Society* **2000,** *122* (51), 12896-12897.

12. Fitzpatrick, P. F., Oxidation of Amines by Flavoproteins. *Archives of Biochemistry and Biophysics* **2010,** *493* (1), 13-25.

13. Vianello, R.; Repic, M.; Mavri, J., How Are Biogenic Amines Metabolized by Monoamine Oxidases? *European Journal of Organic Chemistry* **2012,** (36), 7057-7065.

14. Walker, M. C.; Edmondson, D. E., Structure-Activity-Relationships in the Oxidation of Benzylamine Analogs by Bovine Liver Mitochondrial Monoamine-Oxidase-B. *Biochemistry* **1994,** *33* (23), 7088-7098.

15. Abad, E.; Zenn, R. K.; Kaestner, J., Reaction Mechanism of Monoamine Oxidase from Qm/Mm Calculations. *Journal of Physical Chemistry B* **2013,** *117* (46), 14238-14246.

16. Murray, A. T.; Dowley, M. J. H.; Pradaux-Caggiano, F.; Baldansuren, A.; Fielding, A. J.; Tuna, F.; Hendon, C. H.; Walsh, A.; Lloyd-Jones, G. C.; John, M. P.; Carbery, D. R., Catalytic Amine Oxidation under Ambient Aerobic Conditions: Mimicry of Monoamine Oxidase B. *Angewandte Chemie-International Edition* **2015,** *54* (31), 8997-9000.

17. Rigby, S. E. J.; Hynson, R. M. G.; Ramsay, R. R.; Munro, A. W.; Scrutton, N. S., A Stable Tyrosyl Radical in Monoamine Oxidase A. *Journal of Biological Chemistry* **2005,** *280* (6), 4627-4631.

18. Dunn, R. V.; Munro, A. W.; Turner, N. J.; Rigby, S. E. J.; Scrutton, N. S., Tyrosyl Radical Formation and Propagation in Flavin Dependent Monoamine Oxidases. *Chembiochem* **2010,** *11* (9), 1228-1231.

19. Yue, K. T.; Bhattacharyya, A. K.; Zhelyaskov, V. R.; Edmondson, D. E., Resonance Raman-Spectroscopic Evidence for an Anionic Flavin Semiquinone in Bovine Liver Monoamine-Oxidase. *Archives of Biochemistry and Biophysics* **1993,** *300* (1), 178-185.

20. Derose, V. J.; Woo, J. C. G.; Hawe, W. P.; Hoffman, B. M.; Silverman, R. B.; Yelekci, K., Observation of a Flavin Semiquinone in the Resting State of Monoamine Oxidase B by Electron Paramagnetic Resonance and Electron Nuclear Double Resonance Spectroscopy. *Biochemistry* **1996,** *35* (34), 11085-11091.

21. Allen, W. J.; Bevan, D. R., Steered Molecular Dynamics Simulations Reveal Important Mechanisms in Reversible Monoamine Oxidase B Inhibition. *Biochemistry* **2011,** *50* (29), 6441-6454.

22. Is, Y. S.; Durdagi, S.; Aksoydan, B.; Yurtsever, M., Proposing Novel Mao-B Hit Inhibitors Using Multidimensional Molecular Modeling Approaches and Application of Binary Qsar Models for Prediction of Their Therapeutic Activity, Pharmacokinetic and Toxicity Properties. *Acs Chemical Neuroscience* **2018,** *9* (7), 1768-1782.

23. Braun, G. H.; Jorge, D. M. M.; Ramos, H. P.; Alves, R. M.; Da Silva, V. B.; Giuliatti, S.; Sampaio, S. V.; Taft, C. A.; Silva, C. H. T. P., Molecular Dynamics, Flexible Docking, Virtual Screening, Admet Predictions, and Molecular Interaction Field Studies to Design Novel Potential Mao-B Inhibitors. *Journal of Biomolecular Structure & Dynamics* **2008,** *25* (4), 347-355.

24. La Regina, G.; Silvestri, R.; Artico, M.; Lavecchia, A.; Novellino, E.; Befani, O.; Turini, P.; Agostinelli, E., New Pyrrole Inhibitors of Monoamine Oxidase: Synthesis, Biological Evaluation, and Structural Determinants of Mao-a and Mao-B Selectivity. *Journal of Medicinal Chemistry* **2007,** *50* (5), 922-931.

25. Maccioni, E.; Alcaro, S.; Cirilli, R.; Vigo, S.; Cardia, M. C.; Sanna, M. L.; Meleddu, R.; Yanez, M.; Costa, G.; Casu, L.; Matyus, P.; Distinto, S., 3-Acetyl-2,5-Diaryl-2,3-Dihydro-1,3,4-Oxadiazoles: A New Scaffold for the Selective Inhibition

of Monoamine Oxidase B. *Journal of Medicinal Chemistry* **2011,** *54* (18), 6394-6398.

26. Chimenti, F.; Maccioni, E.; Secci, D.; Bolasco, A.; Chimenti, P.; Granese, A.; Befani, O.; Turini, P.; Alcaro, S.; Ortuso, F.; Cirilli, R.; La Torre, F.; Cardia, M. C.; Distinto, S., Synthesis, Molecular Modeling Studies, and Selective Inhibitory Activity against Monoamine Oxidase of 1-Thiocarbamoyl-3,5-Diaryl-4,5-Dihydro-(1h)-Pyrazole Derivatives. *Journal of Medicinal Chemistry* **2005,** *48* (23), 7113-7122.

27. Fowler, P. W.; Balali-Mood, K.; Deol, S.; Coveney, P. V.; Sansom, M. S. P., Monotopic Enzymes and Lipid Bilayers: A Comparative Study. *Biochemistry* **2007,** *46* (11), 3108-3115.

28. Newton-Vinson, P.; Hubalek, F.; Edmondson, D. E., High-Level Expression of Human Liver Monoamine Oxidase B in Pichia Pastoris. *Protein Expression and Purification* **2000,** *20* (2), 334-345.

29. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S., Probability-Based Protein Identification by Searching Sequence Databases Using Mass Spectrometry Data. *Electrophoresis* **1999,** *20* (18), 3551-3567.

30. Kao, Y.-T.; Saxena, C.; He, T.-F.; Guo, L.; Wang, L.; Sancar, A.; Zhong, D., Ultrafast Dynamics of Flavins in Five Redox States. *Journal of the American Chemical Society* **2008,** *130* (39), 13132-13139.

31. Candeias, L. P.; Turconi, S.; Nugent, J. H. A., Tyrosine Y-Z and Y-D of Photosystem Ii - Comparison of Optical Spectra to Those of Tyrosine Oxidised by Pulsed Radiolysis. *Biochimica Et Biophysica Acta-Bioenergetics* **1998,** *1363* (1), 1-5.

32. Geha, R. M.; Chen, K.; Wouters, J.; Ooms, F.; Shih, J. C., Analysis of Conserved Active Site Residues in Monoamine Oxidase a and B and Their Three-Dimensional Molecular Modeling. *Journal of Biological Chemistry* **2002,** *277* (19), 17209-17216.

33. Woo, J. C. G.; Silverman, R. B., Observation of 2 Different Chromophores in the Resting State of Monoamine-Oxidase-B by Fluorescence Spectroscopy. *Biochemical and Biophysical Research Communications* **1994,** *202* (3), 1574-1578.

34. Aubert, C.; Brettel, K.; Mathis, P.; Eker, A. P. M.; Boussac, A., Epr Detection of the Transient Tyrosyl Radical in DNA Photolyase from Anacystis Nidulans. *Journal of the American Chemical Society* **1999,** *121* (37), 8659-8660.

35. Okafuji, A.; Schnegg, A.; Schleicher, E.; Moebius, K.; Weber, S., G-Tensors of the Flavin Adenine Dinucleotide Radicals in Glucose Oxidase: A Comparative Multifrequency Electron Paramagnetic Resonance and Electron-Nuclear Double Resonance Study. *Journal of Physical Chemistry B* **2008,** *112* (11), 3568-3574.

36. Mcdowall, J. S.; Ntai, I.; Hake, J.; Whitley, P. R.; Mason, J. M.; Pudney, C. R.; Brown, D. R., Steady-State Kinetics of a-Synuclein Ferrireductase Activity Identifies the Catalytically Competent Species. *Biochemistry* **2017,** *56* (19), 2497-2505.

37. Oleinikovas, V.; Saladino, G.; Cossins, B. P.; Gervasio, F. L., Understanding Cryptic Pocket Formation in Protein Targets by Enhanced

Sampling Simulations. *Journal of the American Chemical Society* **2016,** *138* (43), 14257-14263.

38. Beglov, D.; Hall, D. R.; Wakefield, A. E.; Luo, L.; Allen, K. N.; Kozakov, D.; Whitty, A.; Vajda, S., Exploring the Structural Origins of Cryptic Sites on Proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2018,** *115* (15), E3416-E3425.

39. Stoll, S.; Schweiger, A., Easyspin, a Comprehensive Software Package for Spectral Simulation and Analysis in Epr. *Journal of Magnetic Resonance* **2006,** *178* (1), 42-55.

Appendix D – Borane Catalyst

The following is a summary of the research published in the paper by Bennett *et al.*, 2019[1] (see original publication for Supporting Information). Contributions here included acquisition and analysis of the EPR spectra, and composition of a portion of the text describing and explaining these data.

The reactivity of Lewis acidic boranes with H_2 is now almost exclusively explored via the chemistry of frustrated Lewis pairs (FLPs)[2-6], which was introduced in 2006[7]. Although certain mechanistic details are still disputed[8-9], H_2 cleavage by FLPs is generally reliant on the action of the two sterically hindered (or, 'frustrated') reactive centres within the Lewis acid–base pair. The Lewis acid, which provides the vacant acceptor orbital is typically an organoborane, whilst the Lewis base, providing the donor orbital with which to cleave the strong H-H bond[10-12], is usually a phosphine or amine. This bond cleavage by borane-based FLPs is generally accepted to be heterolytic and is likely controlled by the relative strengths of the Lewis acidic/basic components and the degree of steric encumbrance between them[3-6, 13]. In contrast, in the transition metal-based complexes and biological systems that have typically dominated hydrogenation catalysis[14-15], where the metal centre provides both the donor and acceptor orbitals and the reactivity is largely dictated by the redox chemistry of the metal centre, homolytic H_2 bond cleavage is common.

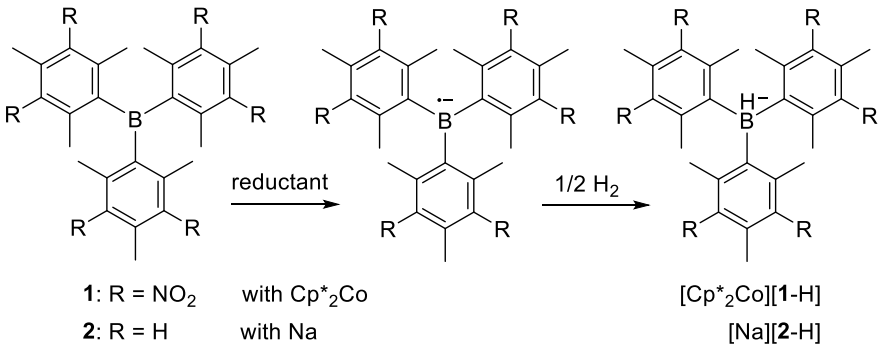
The heterolytic mechanism of activation of H_2 by FLPs agrees well with the general trends in reactivity and is supported by a number of computational studies[10-12], but experimental support remains to be seen. It is interesting to consider that observed patterns of FLP reactivity could be consistent with alternative H_2 activation pathways; that these trends may be consistent with radical mechanisms, wherein single-electron transfer (SET) from the Lewis base to the Lewis acid would generate highly reactive radical pairs capable of activating H_2. Whilst the thermodynamic and kinetic ability of an FLP to activate H_2 is well known to correlate with the hydride-ion affinity of the Lewis acid (consistent with heterolytic bond cleavage), these parameters also correlate well with the one-electron reduction potential of the Lewis acid (consistent with SET). Moreover, recent studies have implied that for some families of borane Lewis acids, reduction

potentials may even be a better indicator of reactivity towards H_2 than hydride-ion affinities[16-22].

A growing body of evidence shows the occurrence of radical mechanisms when small molecules, such as NO, Ph_3SnH, and peroxides, are employed in FLP reactions[23-29]. However, thus far these frustrated radical pair (FRP) mechanisms have not been observed with H_2 (nor has any FLP been shown to cleave H_2 via a radical mechanism). Previous work studying the electrochemistry of FLP components and the recent evidence for radical pathways in FLPs and FRPs raises an obvious question: can boranes react with H_2 outside of an FLP chemical framework, if they can operate via a hitherto unknown redox controlled, radical reaction pathway instead?

Herein it is explored whether tris(aryl)borane Lewis acids are capable of cleaving H_2 outside of the usual Lewis acid/base chemistry described by the concept of frustrated Lewis pairs (FLPs). Instead of a Lewis base a chemical reductant is used to generate stable radical anions of two highly hindered boranes. NMR spectroscopic characterisation reveals that the corresponding borane radical anions activate dihydrogen, whilst EPR spectroscopic characterisation, supported by computational analysis, reveals the intermediates along the hydrogen activation pathway. This radical-based, redox pathway involves the homolytic cleavage of H_2, in contrast to conventional models of FLP chemistry, which invoke a heterolytic cleavage pathway. This represents a new mode of chemical reactivity for hydrogen activation by borane Lewis acids.

Two boranes models were selected: tris(3,5-dinitromesityl)borane **1**, and tris(mesityl)borane **2** (Scheme D.1). Both have essentially identical steric shielding of the central boron atom by the six *ortho* methyl groups on the mesityl rings, leading to the formation of long-lived borane radical anions upon reduction[30-36]. Neither is currently known to be active for H_2 activation within an FLP. The addition of six electron-withdrawing nitro groups in **1** shifts the reduction potential in a positive direction to -1.57V vs. $Cp_2Fe^{0/+}$, making **1** as electrophilic and comparably facile to reduce as the archetypal electron-deficient borane $B(C_6F_5)_3$ used in FLP chemistry (-1.52 V vs. $Cp_2Fe^{0/+}$)[17, 20-22], and much easier to reduce than **2** (approximately -2.8 V vs. $Cp_2Fe^{0/+}$)[37]. The NO_2 groups in **1** also provide useful electron paramagnetic resonance spectroscopic ‘markers’ for the characterisation of reaction intermediates.



*Scheme D.1: Reduction of tris(3,5-dinitromesityl)borane **1**, and tris-(mesityl)borane **2**, and subsequent reaction with H_2 (taken from Bennett et al, 2019[1])*

To examine whether the radical anions of Lewis acidic boranes are capable of cleaving hydrogen, a solution of **1** in either CD_2Cl_2 or $[D_8]$THF was chemically reduced using decamethylcobaltocene (Cp^*_2Co, E^0= -1.94V vs. $Cp_2Fe^{0/+}$)[38], heated in the presence of H_2, and the reaction periodically monitored using multinuclear NMR spectroscopy (figure D.1(a) shows the resulting ^{11}B NMR spectra). The formation of the borohydride product $[Cp^*_2Co][$**1**-H] is clearly evident by the observation of a characteristic doublet at δ = -13.6 ppm ($^1J_{B,H}$ = 82 Hz) in the ^{11}B NMR spectrum.

In these reactions, the cleavage of H_2/D_2 must be homolytic as there is no apparent plausible mechanism to allow for the formation of H^+ (no counter anion). Whilst very strong acids are known to protonate Cp^*_2Co[39-40], there is no observable evidence for the formation of this in these reactions. To examine the proposed radical homolytic dihydrogen cleavage mechanism, **1** was again reduced with Cp^*_2Co under H_2 but this time in the presence of 1 equivalent of the radical spin-trap TEMPO ((2,2,6,6- tetramethylpiperidin-1-yl)oxyl). No H_2 cleavage was then observed, consistent with inhibition of a radical reaction by the TEMPO spin-trap. Additional control experiments confirm that Cp^*_2Co alone does not activate H_2 under these conditions and that THF/**1** mixtures do not result in the observable formation of $[$**1**-H$]^-$ via a solvent-FLP mechanism[41-43] in the absence of a reducing agent. Crucially, no evidence of reduction at the nitro groups is observed by NMR, EPR, nor IR spectroscopic characterisation of the reaction products.

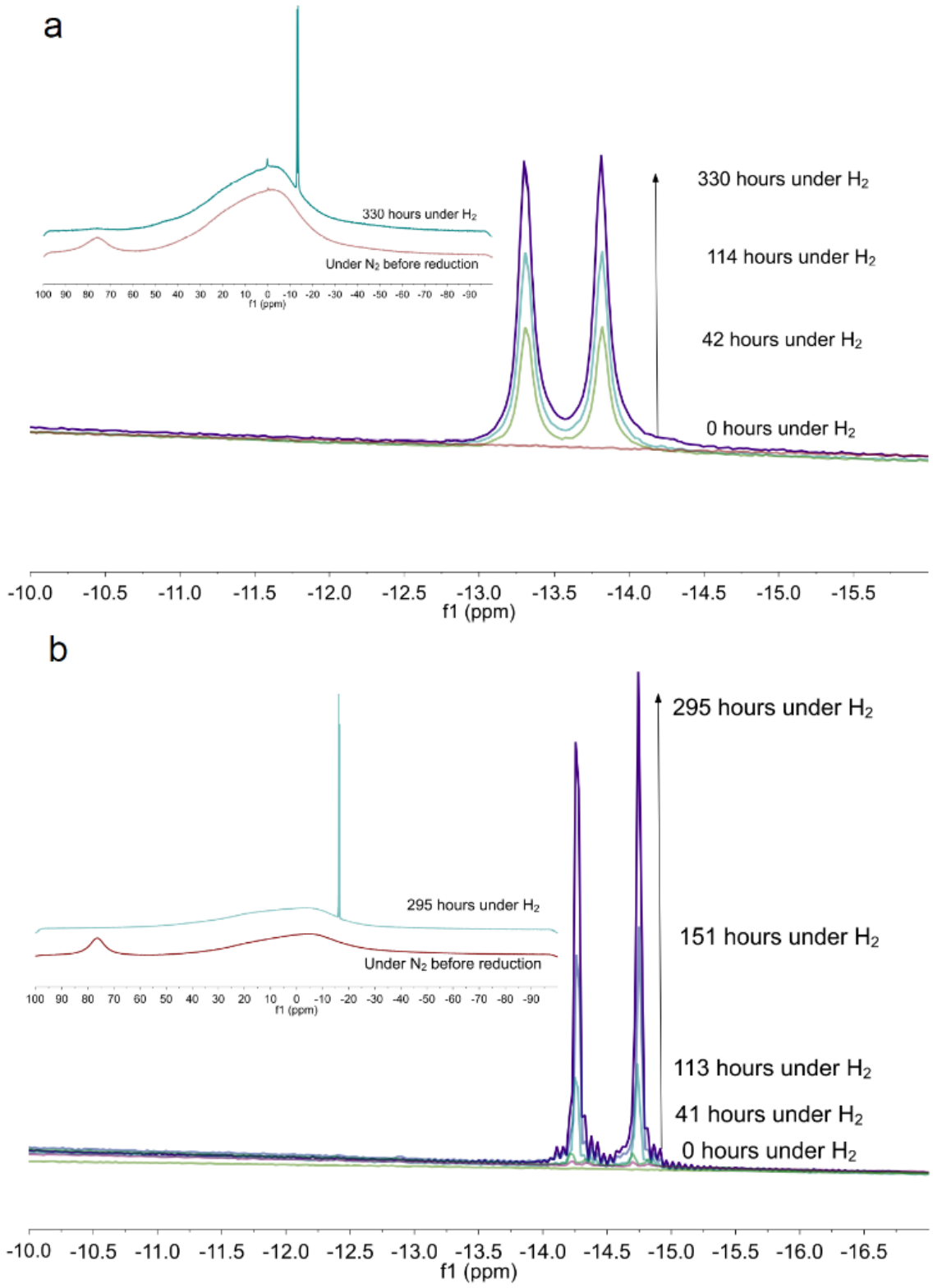


*Figure D.1: Overlaid ^{11}B NMR spectra expanded over the B-H bond region of interest, showing the progression of H_2 cleavage by chemical reduction of **1** in CD_2Cl_2 (a) and **2** in THF (b). Inset: The corresponding ^{11}B NMR spectra recorded at the start and end of the experiments showing the conversion of the parent borane starting material to the borohydride product upon reduction and exposure to H_2 (taken from Bennett et al, 2019[1]).*

The very negative redox potential of **2** necessitates the use of a stronger reducing agent. When a solution of **2** in [D_8]THF is reduced over sodium metal and heated in the presence of H_2 the appearance of a doublet in the ^{11}B NMR spectrum at δ = -14.5 ppm ($^1J_{B,H}$ = 78 Hz), and a corresponding 1:1:1:1 quartet in the ^{1}H NMR spectrum at δ = 3.75 ppm ($^1J^{H,B}$ = 77 Hz) is observed, characteristic of the formation of [Na]-[**2**-H] (Figure D.1 b). These experiments indicate that the borane radical anions **1**•$^-$ and **2**•$^-$ can cleave H_2 in the absence of any exogenous Lewis base. However, they are slow in comparison to typical FLP H_2 activation reactions. In the case of the **1**, this is advantageous, since it enables the reaction to be monitored in real time and reaction intermediates along the H_2 cleavage pathway to be observed using EPR spectroscopy.

Solutions of **1** dissolved in either CD_2Cl_2 or [D_8]THF were chemically reduced using Cp^*_2Co and the EPR spectra resulting from exposure to H_2 were recorded (Figure D.2a–d). Simulation of the spectra yields the isotropic hyperfine coupling constants for the various ^{1}H, ^{14}N, and ^{11}B nuclei, given in Table 1. These data, supported by DFT calculations (performed for the identifiable intermediates of both **1** and **2**), enable the structures of the intermediates to be characterised and gain valuable insights into the reaction mechanism (given schematically in Figure D.3) and the corresponding energetic profile by which organoborane radicals cleave H_2 homolytically (Figure D.4).

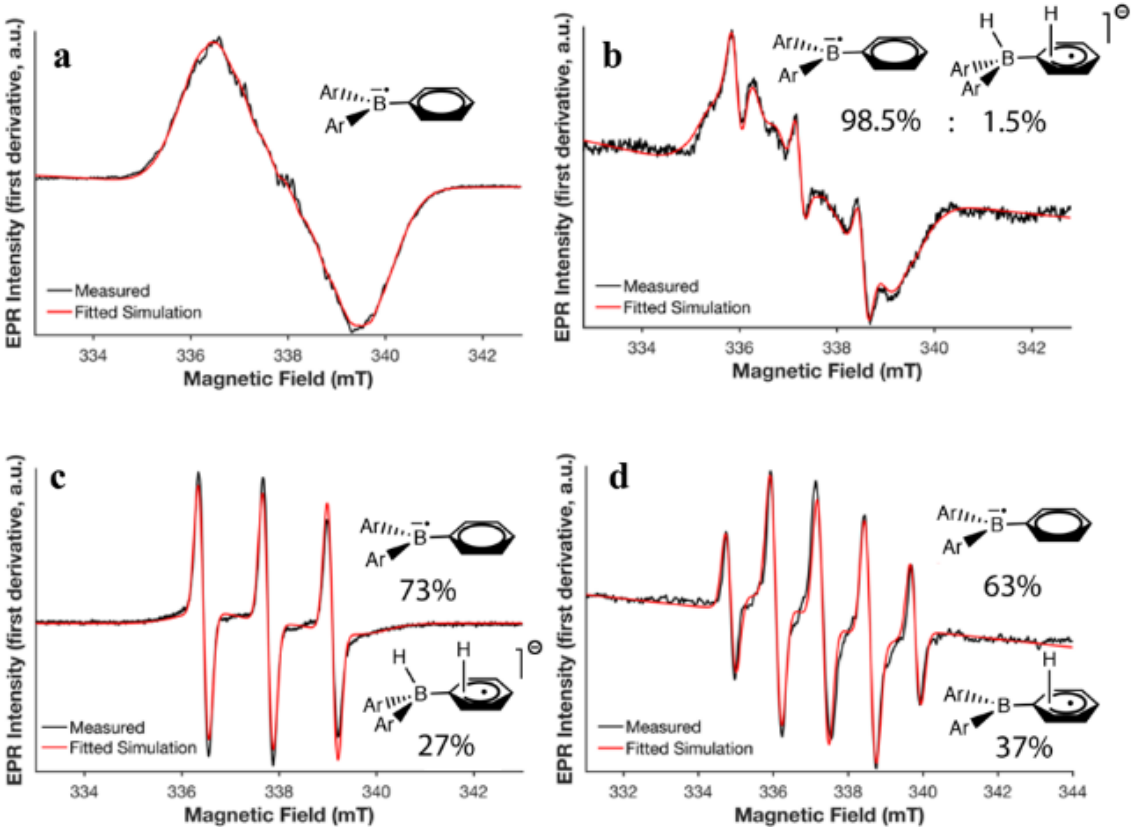


*Figure D.2: the room temperature X-band cw-EPR spectra of **1**•⁻ formed via chemical reduction of **1**, recorded under an atmosphere of N_2 (a), upon first exposure to H_2 but prior to heating (b), after heating under H_2 for 10 minutes (c), and after heating under H_2 for 48 hours (d). The structures of the paramagnetic species are shown with ring substituents removed for clarity; experimental data are shown in black with their respective fitted simulations in red (taken from Bennett et al, 2019[1]).*

*Table D.1: parameters obtained by simulation of the room temperature X-band cw-EPR spectra of **1**•⁻ formed via chemical reduction of **1**, recorded under an atmosphere of N_2 (Figure D.2a), upon first exposure to H_2 but prior to heating (Figure D.2b), after heating under H_2 for 10 minutes (Figure D.2c), and after heating under H_2 for 48 hours (Figure D.2d) (taken from Bennett et al, 2019[1]).*

Parameter	**Figure D.2a**	**Figure D.2b**		**Figure D.2c**		**Figure D.2d**	
	$1^{\bullet-}$	**$1^{\bullet-}$**	**$[1–\{H_2\}]^{\bullet-}$**	**$1^{\bullet-}$**	**$[1–\{H_2\}]^{\bullet-}$**	**$1^{\bullet-}$**	**$[1–H]^{\bullet}$**
g-value	2.00475	2.00473	2.00619	2.00473	2.0064	2.00473	2.00404
A (^{11}B) / MHz	23.2	23.3	-	23.3	–	23.3	35.4
A (^{14}N, *meta*–NO_2) / MHz	3.6	3.4	36.5	3.4	37.3	3.5	0.7
A (^{1}H, *ortho*–CH_3) / MHz	4.8	4.2	–	4.2	–	4.2	0.5
A (^{1}H, *para*–CH_3) / MHz	7.9	7.8	–	7.8	–	7.9	1.8
A (^{1}H) / MHz	–	–	–	–	–	–	32.2
Linewidth (Gaussian) / mT	0.15	0.25	0.22	0.25	0.22	0.2	0.26
Weighting	–	98.50%	1.50%	73.00%	27.00%	63.00%	37.00%
RMSD	0.022532	0.038452		0.034853		0.060845	

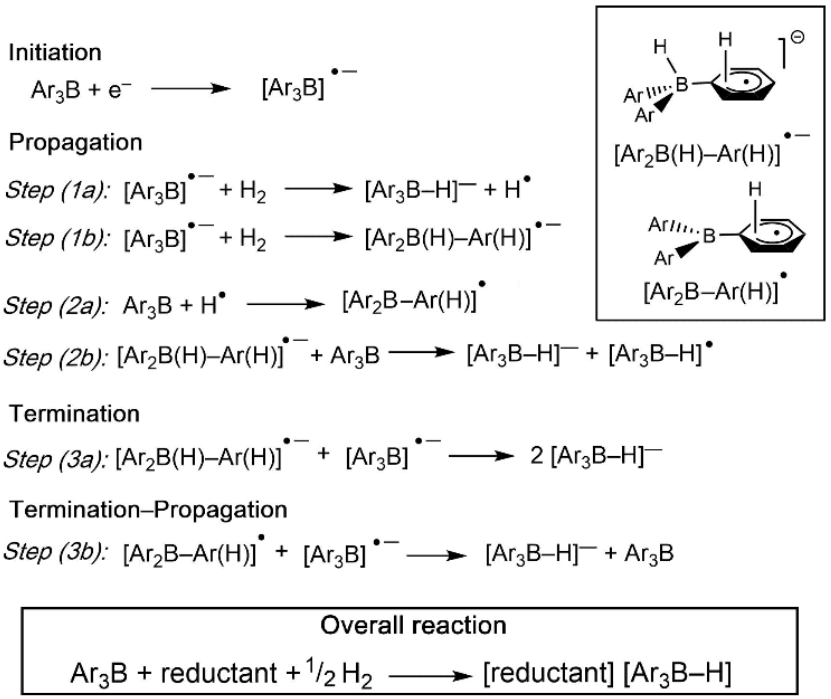


Figure D.3: The proposed radical chain-propagation mechanism for the homolytic cleavage of H_2 upon reduction of organoborane Lewis acids. Inset: the chemical structures corresponding to the $[Ar_2B(H)\text{-}Ar(H)]^{\bullet -}$ and $[Ar_2B\text{-}Ar(H)]^{\bullet}$ intermediates (substituents on the aryl rings have been omitted for clarity) (taken from Bennett et al, 2019[1]).

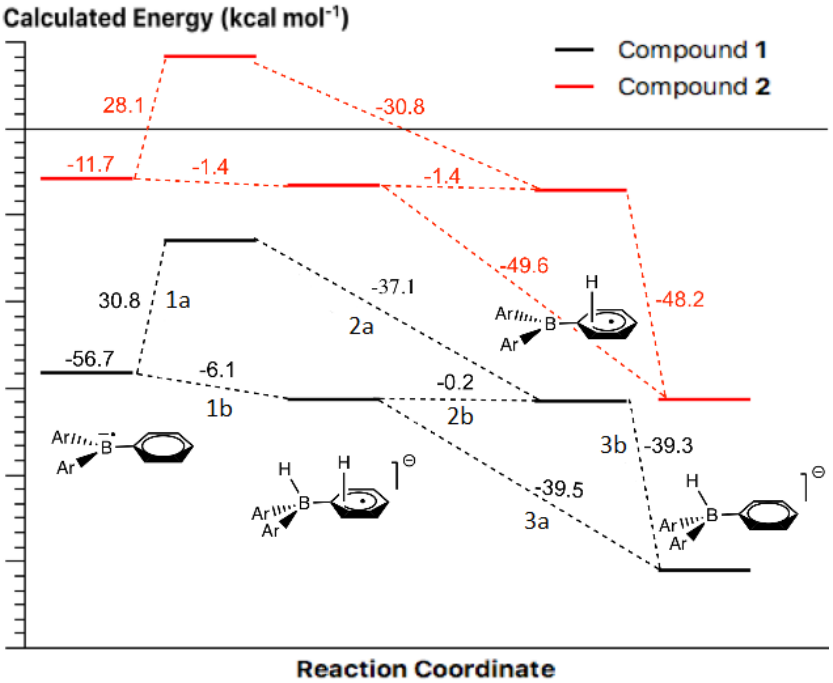


Figure D.4: Postulated reaction profile showing the relevant reaction intermediates involved in each step (ring substituents removed for clarity, steps labelled as in Figure D.3) together with the associated change in energy values along each reaction step obtained from DFT calculations (taken from Bennett et al, 2019[1]).

Upon reduction of **1** under N_2, the EPR spectrum shown in Figure D.2a is observed, which is characteristic of **1**•$^-$ with hyperfine coupling of the unpaired electron spin density to the boron nucleus as well as the methyl and nitro substituents on the aromatic rings (Table 1)[44-46]. The initiation step is calculated to be exothermic for both compounds (-56.7 and -11.7 kcal mol^{-1} for **1** and **2**, respectively) and reflects the relative LUMO energy and reduction potential of each borane.

Figure D.2b shows the resulting spectrum recorded upon first exposing the reaction to H_2 and before heating. An immediate change is evident with the appearance of a sharp 1:1:1 three-line signal superimposed on the original signal of the **1**•$^-$ parent. After heating the reaction for a further 10 minutes this three-line signal dominates the EPR spectral response (Figure D.2c) for the next 48 hours. The only change to the system is the addition of H_2 and computational modelling of the possible interactions between **1**•$^-$ and H_2 reveal two propagation pathways. Propagation 1a produces the diamagnetic borohydride product and is endothermic (+30.8 and +28.1 kcal mol^{-1} for **1** and **2**), albeit to a lesser extent than homolytic H_2 splitting itself (+107.1 kcal mol^{-1} at this level of theory). The alternative pathway, Propagation 1b avoids the release of free H-atom radicals and is slightly exothermic (-6.1 and -1.4 kcal mol^{-1} for **1** and **2**). This reaction produces a radical species consistent with that observed in Figures D.2b and D.2c. Computation reveals the structure of this intermediate to be [(Ar_2B(H)-Ar(H)]•$^-$ with hydride attached at a four-coordinate boron centre, and H• carried on one of the aromatic rings (denoted as [**1**-{H2}]•$^-$ with specific reference to borane **1**). DFT models indicate that there is little energetic discrimination for the H• to be attached to one or other carbon positions around the aromatic ring. Spin-density calculations confirm, however, that the isomer with the H• predominantly located at a *meta* carbon on the ring, *ipso* to one of the nitro groups, is consistent with the observed EPR spectra (Figures D.2b,c). Here the unpaired electron is coupled only to one of the nitrogen nuclei in the nitro groups of the aryl ring system and is not coupled to the boron nucleus at all (Table D.1).

After 48 hours of heating, the EPR spectrum changes once again (Figure D.2d) to reveal a 1:2:2:2:1 five-line hyperfine coupling pattern of a new persistent paramagnetic species. This does not fit the expected coupling pattern from two nitro groups which would give rise to a 1:2:3:2:1 splitting pattern. Instead, it arises from near coincident hyperfine coupling with both an additional single hydrogen

atom and the boron nucleus (similar to DFT calculations of a hydrogen–boron adduct)[47]. This is a neutral [**1**-H]• intermediate resulting from cleavage of the H_2 molecule.

Once again there are two possible pathways that result in the formation of the [**1**-H]• intermediate: Propagations 2a and 2b. Propagation 2a is exothermic by -37.1 kcal mol^{-1} and -30.8 kcal mol^{-1} for **1** and **2**, respectively. Interestingly, computation suggests that if [**1**-H]• is formed with the hydrogen atom at boron, as one might expect, the hydrogen atom immediately hops from the boron atom onto the aromatic ring system, until it arrives at the *para* carbon atom which is the most stable isomer in the case of **1** (whereas the *meta* position is most stable in **2**).This is supported by what is observed experimentally during the EPR spectroscopic monitoring of hydrogen splitting by **1** where the magnitude of the resulting H• atom hyperfine coupling fits well with coupling to spin density on the ring system in the para position located between the two nitro groups (Figure D.2d).

If the parent borane is present in excess of the radical anion (Propagation 2b), the hydrogen atom produced in step 1a (considered as $[Ar_2B(H)\text{-}Ar(H)]^{\bullet -}$) may be transferred, and the borohydride product and the neutral $[Ar_3B\text{-}H]^{\bullet}$ radical intermediate formed. Using the values calculated for propagation steps 1a and 2a, step 2b is energetically neutral. In the system reported herein, it is unlikely that the parent borane is present in excess of the radical anion initially, but as the reaction proceeds through step 3b and the consumption of the $[Ar_2B(H)\text{-}Ar(H)]^{\bullet -}$ progresses, this stabilisation may become more relevant towards the end of the reaction. This situation may also have relevance to potential radical-FLP hydrogen cleavage mechanisms, where the parent borane is most likely present in excess of any potential radical anion intermediates throughout.

The final step in the reaction, which cannot be observed by EPR spectroscopy, is the formation of the diamagnetic $[\mathbf{1}\text{-}H]^-$ product, which is detected by ^{11}B and ^{1}H NMR spectroscopic analysis of the reaction mixture at the end of the experiment. Aside from the obvious recombination of 2H• to form H_2 (the reverse of step 1), there are two termination pathways: Termination 3a (-39.5 and -49.6 kcal mol^{-1} for **1** and **2**, respectively), and Termination-Propagation 3b (-39.3 and -48.2 kcal mol^{-1} for **1** and **2**, respectively). Step 3a may also be written $[Ar_3B]^{\bullet -} + H^{\bullet} \rightarrow [Ar_3B\text{-}H]^-$ for consistency with the rest of the Scheme, or as a termolecular reaction: $2\ [Ar3B]^{\bullet -} + H_2 \rightarrow 2[Ar_3B\text{-}H]^-$. Step 3b yields both the terminal borohydride product and

regenerates the parent neutral borane for further reaction in propagation step 2a. Note that whilst it would appear from Figures D.2c and D.2d that the EPR spectra are dominated by the [**1**-{H_2}]$^{\bullet-}$ and [**1**-H]$^{\bullet}$ species, respectively, simulation of the spectral data reveals that these spectra are each superimposed over the parent **1**$^{\bullet-}$ radical anion species. As the reaction proceeds with heating the weighting between the systems changes (**1**$^{\bullet-}$: [**1**- {H_2}]$^{\bullet-}$ = 98.5 :1.5 in Figure D.2 b; 73.0 :27.0 in Figure D.2c, and **1**$^{\bullet-}$: [**1**-H]$^{\bullet}$ = 63.0 :37.0 in Figure D.2d. The rate of consumption of **1**$^{\bullet-}$ as measured by EPR (Figures D.2a–d) correlates with the rate of conversion to borohydride as measured by NMR spectroscopy (Figure D.1a).

In summary, using two model boranes, which produce stable radical anions upon one-electron reduction, this study has successfully demonstrated homolytic dihydrogen cleavage in the absence of a Lewis base. This represents a new mode of chemical reactivity by Lewis acidic boranes towards H_2 that opens up new borane, and potentially other main group chemistries, beyond the framework of conventional FLPs. The reaction between the model borane radical anions and H_2 is slow, and the intermediates are sufficiently stabilised so that several distinct intermediates along the homolytic dihydrogen cleavage pathway can be observed using EPR spectroscopy and the energetics of the reaction pathway can be modelled computationally. The application of boryl radical H_2 activation is currently being explored as a convenient route to more active borane hydride species, which may have applications in catalysis and energy materials.

EPR spectra were collected on a Bruker eleXsys E500 spectrometer using a standard rectangular Bruker EPR cavity (4102ST). Samples were measured in standard 4mm quartz EPR tubes (Wilmad) fitted with J Young valves to maintain the desired atmospheres. Experimental parameters: microwave power, 20 mW (figures D.2a, D.2b, D.2d) and 0.8 mW (figure D.2c); field modulation frequency, 100 kHz; field modulation amplitude, 0.1 mT (figures D.2a-c) and 0.2 mT (figure D.2d); all performed at ambient room temperature. All spectral simulations were performed using the MATLAB-based EasySpin package[48]. Inside the glovebox a 5 μM solution of 1 in either CD_2Cl_2 or THF–d_8 was reduced by adding an excess (3 molar equivalents) of Cp^*_2Co in a glass vial. The reaction mixture was briefly stirred before the solution was filtered (by pipetting through a glass wool plug to remove any residual reducing agent/solid) into a 4mm quartz EPR tube fitted with a J Young valve. EPR spectra were recorded for each sample before H_2 was admitted via a sequence of three freeze–pump–thaw cycles. The EPR spectra

were recorded again after addition of H_2 to the reaction mixture, and then the reaction mixtures were heated in the dark at 50°C (CD_2Cl_2) or 65°C (THF–d_8) with periodic monitoring by EPR spectroscopy.

References

1. Bennett, E. L.; Lawrence, E. J.; Blagg, R. J.; Mullen, A. S.; Macmillan, F.; Ehlers, A. W.; Scott, D. J.; Sapsford, J. S.; Ashley, A. E.; Wildgoose, G. G.; Slootweg, J. C., A New Mode of Chemical Reactivity for Metal-Free Hydrogen Activation by Lewis Acidic Boranes. *Angewandte Chemie-International Edition* **2019,** *58* (25), 8362-8366.

2. Lawson, J. R.; Melen, R. L., Tris(Pentafluorophenyl)Borane and Beyond: Modern Advances in Borylation Chemistry. *Inorganic Chemistry* **2017,** *56* (15), 8627-8643.

3. Stephan, D. W., Frustrated Lewis Pairs: From Concept to Catalysis. *Accounts of Chemical Research* **2015,** *48* (2), 306-316.

4. Stephan, D. W., Frustrated Lewis Pairs. *Journal of the American Chemical Society* **2015,** *137* (32), 10018-10032.

5. Stephan, D. W., The Broadening Reach of Frustrated Lewis Pair Chemistry. *Science* **2016,** *354* (6317).

6. Stephan, D. W.; Erker, G., Frustrated Lewis Pair Chemistry: Development and Perspectives. *Angewandte Chemie-International Edition* **2015,** *54* (22), 6400-6441.

7. Welch, G. C.; Juan, R. R. S.; Masuda, J. D.; Stephan, D. W., Reversible, Metal-Free Hydrogen Activation. *Science* **2006,** *314* (5802), 1124-1126.

8. Paradies, J., Mechanisms in Frustrated Lewis Pair-Catalyzed Reactions. *European Journal of Organic Chemistry* **2019,** (2-3), 283-294.

9. Scott, D. J.; Fuchter, M. J.; Ashley, A. E., Designing Effective 'Frustrated Lewis Pair' Hydrogenation Catalysts. *Chemical Society Reviews* **2017,** *46* (19), 5689-5700.

10. Rokob, T. A.; Hamza, A.; Stirling, A.; Soos, T.; Papai, I., Turning Frustration into Bond Activation: A Theoretical Mechanistic Study on Heterolytic Hydrogen Splitting by Frustrated Lewis Pairs. *Angewandte Chemie-International Edition* **2008,** *47* (13), 2435-2438.

11. Rokob, T. A.; Papai, I., Hydrogen Activation by Frustrated Lewis Pairs: Insights from Computational Studies. In *Frustrated Lewis Pairs I: Uncovering and Understanding*, Erker, G.; Stephan, D. W., Eds. 2013; Vol. 332, pp 157-211.

12. Skara, G.; De Vleeschouwer, F.; Geerlings, P.; De Proft, F.; Pinter, B., Heterolytic Splitting of Molecular Hydrogen by Frustrated and Classical Lewis Pairs: A Unified Reactivity Concept. *Sci Rep* **2017,** *7*, 15.

13. Houghton, A. Y.; Autrey, T., Calorimetric Study of the Activation of Hydrogen by Tris(Pentafluorophenyl)Borane and Trimesitylphosphine. *Journal of Physical Chemistry A* **2017,** *121* (46), 8785-8790.

14. Blaser, H.-U., The Development and Application of Industrially Viable Catalysts for the Selective Hydrogenation of Complex Molecules. *Topics in Catalysis* **2010,** *53* (15-18), 997-1001.

15. Halpern, J., The Catalytic Activation of Hydrogen in Homogeneous, Heterogeneous, and Biological Systems. *Advances in Catalysis* **1959,** *11*, 301-370.

16. Blagg, R. J.; Lawrence, E. J.; Resner, K.; Oganesyan, V. S.; Herrington, T. J.; Ashley, A. E.; Wildgoose, G. G., Exploring Structural and Electronic Effects in Three Isomers of Tris{Bis(Trifluoromethyl)Phenyl)Borane: Towards the Combined Electrochemical-Frustrated Lewis Pair Activation of H-2. *Dalton Transactions* **2016,** *45* (14), 6023-6031.

17. Blagg, R. J.; Simmons, T. R.; Hatton, G. R.; Courtney, J. M.; Bennett, E. L.; Lawrence, E. J.; Wildgoose, G. G., Novel B(Ar ')(2)(Ar '') Hetero-Tri(Aryl)Boranes: A Systematic Study of Lewis Acidity. *Dalton Transactions* **2016,** *45* (14), 6032-6043.

18. Blagg, R. J.; Wildgoose, G. G., H-2 Activation Using the First 1: 1: 1 Hetero-Tri(Aryl) Borane. *Rsc Advances* **2016,** *6* (48), 42421-42427.

19. Lawrence, E. J.; Herrington, T. J.; Ashley, A. E.; Wildgoose, G. G., Metal-Free Dihydrogen Oxidation by a Borenium Cation: A Combined Electrochemical/Frustrated Lewis Pair Approach. *Angewandte Chemie-International Edition* **2014,** *53* (37), 9922-9925.

20. Lawrence, E. J.; Oganesyan, V. S.; Hughes, D. L.; Ashley, A. E.; Wildgoose, G. G., An Electrochemical Study of Frustrated Lewis Pairs: A Metal-Free Route to Hydrogen Oxidation. *Journal of the American Chemical Society* **2014,** *136* (16), 6031-6036.

21. Lawrence, E. J.; Oganesyan, V. S.; Wildgoose, G. G.; Ashley, A. E., Exploring the Fate of the Tris(Pentafluorophenyl)Borane Radical Anion in Weakly Coordinating Solvents. *Dalton Transactions* **2013,** *42* (3), 782-789.

22. Ashley, A. E.; Herrington, T. J.; Wildgoose, G. G.; Zaher, H.; Thompson, A. L.; Rees, N. H.; Kraemer, T.; O'hare, D., Separating Electrophilicity and Lewis Acidity: The Synthesis, Characterization, and Electrochemistry of the Electron Deficient Tris(Aryl)Boranes B(C6f5)(3-N)(C6cl5)(N) (N=1-3). *Journal of the American Chemical Society* **2011,** *133* (37), 14727-14740.

23. Merk, A.; Grossekappenberg, H.; Schmidtmann, M.; Luecke, M.-P.; Lorent, C.; Driess, M.; Oestreich, M.; Klare, H. F. T.; Mueller, T., Single-Electron Transfer Reactions in Frustrated and Conventional Silylium Ion/Phosphane Lewis Pairs. *Angewandte Chemie-International Edition* **2018,** *57* (46), 15267-15271.

24. Liu, L.; Cao, L. L.; Shao, Y.; Menard, G.; Stephan, D. W., A Radical Mechanism for Frustrated Lewis Pair Reactivity. *Chem* **2017,** *3* (2), 259-267.

25. Liu, L. L.; Cao, L. L.; Shao, Y.; Stephan, D. W., Single Electron Delivery to Lewis Pairs: An Avenue to Anions by Small Molecule Activation. *Journal of the American Chemical Society* **2017,** *139* (29), 10062-10071.

26. Liu, L. L.; Cao, L. L.; Zhu, D.; Zhou, J.; Stephan, D. W., Homolytic Cleavage of Peroxide Bonds Via a Single Electron Transfer of a Frustrated Lewis Pair. *Chemical Communications* **2018,** *54* (54), 7431-7434.

27. Longobardi, L. E.; Liu, L.; Grimme, S.; Stephan, D. W., Stable Borocyclic Radicals Via Frustrated Lewis Pair Hydrogenations. *Journal of the American Chemical Society* **2016,** *138* (8), 2500-2503.

28. Tao, X.; Kehr, G.; Wang, X.; Daniliuc, C. G.; Grimme, S.; Erker, G., Rapid Dihydrogen Cleavage by Persistent Nitroxide Radicals under Frustrated Lewis Pair Conditions. *Chemistry-a European Journal* **2016,** *22* (28), 9504-9507.

29. De Oliveira, M., Jr.; Wiegand, T.; Elmer, L.-M.; Sajid, M.; Kehr, G.; Erker, G.; Magon, C. J.; Eckert, H., Solid-State Epr Strategies for the Structural Characterization of Paramagnetic No Adducts of Frustrated Lewis Pairs (Flps). *J Chem Phys* **2015,** *142* (12).

30. Feng, R.; Zhang, L.; Chen, C.; Fang, Y.; Zhao, Y.; Tan, G.; Wang, X., Reversible Self-Assembling of Boryl Radical Anions to Their Diradicals with Tunable Singlet Ground States. *Chemistry-a European Journal* **2019,** *25* (16), 4031-4035.

31. Yuan, N.; Wang, W.; Wu, Z.; Chen, S.; Tan, G.; Sui, Y.; Wang, X.; Jiang, J.; Power, P. P., A Boron-Centered Radical: A Potassium-Crown Ether Stabilized Boryl Radical Anion. *Chemical Communications* **2016,** *52* (86), 12714-12716.

32. Liu, L. L.; Stephan, D. W., Radicals Derived from Lewis Acid/Base Pairs. *Chemical Society Reviews* **2019,** *48* (13), 3454-3463.

33. Kawamoto, T.; Uehara, S.; Hirao, H.; Fukuyama, T.; Matsubara, H.; Ryu, I., Borohydride-Mediated Radical Addition Reactions of Organic Iodides to Electron-Deficient Alkenes. *Journal of Organic Chemistry* **2014,** *79* (9), 3999-4007.

34. Kushida, T.; Yamaguchi, S., A Radical Anion of Structurally Constrained Triphenylborane. *Organometallics* **2013,** *32* (22), 6654-6657.

35. Olmstead, M. M.; Power, P. P., First Structural Characterization of a Boron-Centered Radical: X-Ray Crystal Structure of [Li(12-Crown-4)2]+ [Bmes3]-.Bul. *Journal of the American Chemical Society* **1986,** *108* (14), 4235-4236.

36. Power, P. P., Persistent and Stable Radicals of the Heavier Main Group Elements and Related Species. *Chemical Reviews* **2003,** *103* (3), 789-809.

37. Cummings, S. A.; Iimura, M.; Harlan, C. J.; Kwaan, R. J.; Trieu, I. V.; Norton, J. R.; Bridgewater, B. M.; Jakle, F.; Sundararaman, A.; Tilset, M., An Estimate of the Reduction Potential of B(C6f5)(3) from Electrochemical Measurements on Related Mesityl Boranes. *Organometallics* **2006,** *25* (7), 1565-1568.

38. Liu, Y.; Solari, E.; Scopelliti, R.; Tirani, F. F.; Severin, K., Lewis Acid-Mediated One-Electron Reduction of Nitrous Oxide. *Chemistry-a European Journal* **2018,** *24* (70), 18809-18815.

39. Chalkley, M. J.; Del Castillo, T. J.; Matson, B. D.; Peters, J. C., Fe-Mediated Nitrogen Fixation with a Metallocene Mediator: Exploring Pk(a) Effects and Demonstrating Electrocatalysis. *Journal of the American Chemical Society* **2018,** *140* (19), 6122-6129.

40. Chalkley, M. J.; Del Castillo, T. J.; Matson, B. D.; Roddy, J. P.; Peters, J. C., Catalytic N-2-to-Nh3 Conversion by Fe at Lower Driving Force: A Proposed Role for Metallocene-Mediated Pcet. *Acs Central Science* **2017,** *3* (3), 217-223.

41. Mahdi, T.; Stephan, D. W., Enabling Catalytic Ketone Hydrogenation by Frustrated Lewis Pairs. *Journal of the American Chemical Society* **2014,** *136* (45), 15809-15812.

42. Scott, D. J.; Fuchter, M. J.; Ashley, A. E., Metal-Free Hydrogenation Catalyzed by an Air-Stable Borane: Use of Solvent as a Frustrated Lewis Base. *Angewandte Chemie-International Edition* **2014,** *53* (38), 10218-10222.

43. Scott, D. J.; Fuchter, M. J.; Ashley, A. E., Nonmetal Catalyzed Hydrogenation of Carbonyl Compounds. *Journal of the American Chemical Society* **2014,** *136* (45), 15813-15816.

44. Kwaan, R. J.; Harlan, C. J.; Norton, J. R., Generation and Characterization of the Tris(Pentafluorophenyl)Borane Radical Anion. *Organometallics* **2001,** *20* (17), 3818-3820.

45. Elschenbroich, C.; Kuhlkamp, P.; Behrendt, A.; Harms, K., Metal Pi Complexes of Benzene Derivatives .49. Halfsandwich Complexes of Trimesilylborane Mes(3)B: Synthesis and Structure or Mes(2) B(Eta(6)-Mes)Cr(Co)(3) , Mesb (Eta(6)-Mes)Cr(Co)(3) (2), and B (Eta(5)-Mes)Cr(Co)(3) (2). Redox Behavior and Questions of Intramolecular Interaction (Vol 129, Pg 859, 1996). *Chem. Berichte* **1996,** *129* (11), 1407-1407.

46. Dupont, T. J.; Mills, J. L., Arylborane Anions - Electrochemical Study. *Journal of the American Chemical Society* **1975,** *97* (22), 6375-6382.

47. Walton, J. C.; Brahmi, M. M.; Monot, J.; Fensterbank, L.; Malacria, M.; Curran, D. P.; Lacote, E., Electron Paramagnetic Resonance and Computational Studies of Radicals Derived from Boron-Substituted N-Heterocyclic Carbene Boranes. *Journal of the American Chemical Society* **2011,** *133* (26), 10312-10321.

48. Stoll, S.; Schweiger, A., Easyspin, a Comprehensive Software Package for Spectral Simulation and Analysis in Epr. *Journal of Magnetic Resonance* **2006,** *178* (1), 42-55.