

Understanding movement and mechanism in essential mammalian membrane transporters

Jenny Hall

This thesis is submitted for the fulfilment of the degree of Doctor of
Philosophy

School of Chemistry
University of East Anglia
Norwich, UK

Submitted February 2022

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

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.

Jenny Hall,

February 2022

Acknowledgements

My first thank you goes to my supervisor Dr Fraser MacMillan, for the guidance you have provided, and the opportunities you encouraged me to take, throughout the course of this research experience. Many thanks also to Dr Jesus Angulo and Dr Justin Bradley for the additional support, and discussions that helped to inspire my research. A special thanks to Justin, for your hours of feedback on my writing.

Thanks to the various funding bodies who's aid allowed me to not only undertake this research, but also take part in a multitude of other research activities. Specifically, to the Norwich Research Park Doctoral Training Partnership, funded by the BBSRC, for funding over the course of my studies. Also to the EU-COST Action CM1306 and its organisers, for facilitating the various conferences and training school events that I was fortunate enough to attend.

Next, to those with whom I have collaborated in order to carry out this research. I'm grateful to Prof Thomas Stockner, Prof Raz Zarivach and Dr Shiran Barber-Zucker for providing essential guidance and protein and/or plasmids. Thanks also to Dr Jesus Angulo, Prof Eurico Cabrita and Dr Colin Macdonald for their guidance with NMR experiments.

I am hugely thankful to my friends, old and new, who have given me support, and provided no end of entertainment throughout the years. Specifically to Shona, once a project student, now an amazing PhD student and the best friend I made during this process. I thank my wonderful family for your constant love and belief in me - and yes Tom, I have 'found something new yet'! Finally, the warmest of thanks to Rory - you are my biggest cheerleader, and the best partner in crime I could wish for.

Abstract

The research presented within this thesis has utilised advanced magnetic resonance techniques to (a) probe fundamental unknowns of transport mechanisms in membrane transport proteins, and (b) test alternative approaches and magnetic resonance techniques for the study of these complex molecular machineries.

The cation diffusion facilitator (CDF) proteins contribute to maintenance of divalent cation homeostasis, and are conserved throughout all domains of life. Malfunction of CDF proteins in humans is linked to a number of serious diseases, including Type-II diabetes and neurodegenerative diseases. MamM is a (magnetosome associated) CDF protein of magnetotactic bacteria and has more recently been used as a system on which to model the structure and function of the human proteins. Research within this thesis has extended studies of MamM's C-terminal domain via electron paramagnetic resonance (EPR) spectroscopy. This has afforded insights into metal binding and metal type-dependent conformational change, metal binding stoichiometry and relevant affinities, and has aided in the demonstration (for the first time) that C-terminal domains of CDF proteins act as a first site for recognition and metal selectivity in CDF proteins.

A separate section of this thesis presents the study of membrane transporter GltPh, a both structurally and functionally well-characterised member of the solute carrier 1 (SLC1) family. Whilst GltPh typically represents a model system for understanding of mammalian excitatory amino acid transporters (EAATs), here it has been used as a system upon which to test alternative strategies to probe conformational change, and magnetic resonance approaches typically not used for the study of membrane proteins. This work describes the potential use of conformation 'locking' variants for deconvolution of complex EPR data, affording clearer insights into the solution state conformations adopted by this trimer under various conditions. A ligand-based NMR technique (known as STD NMR) has also been applied to demonstrate its use for K_D determination of two competing ligands, in different protein environments. This work has also provided evidence for potential cooperativity in GltPh trimers, a behaviour which is unique to protein reconstituted into proteoliposomes.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
Table of contents	v
List of figures	ix
List of tables	xiii
Glossary of terms	xiv
1 An introduction to magnetic resonance spectroscopy	1
1.1 Thesis layout and chapter summary	1
1.2 Fundamental EPR theory	2
1.2.1 The spin Hamiltonian	3
1.2.2 The electron Zeeman interaction and the resonance condition	3
1.2.3 The spectroscopic g factor	6
1.2.4 g value anisotropy	7
1.2.5 Hyperfine interactions	8
1.2.6 Electron dipole-dipole interactions	11
1.3 Continuous-wave EPR spectroscopy	12
1.3.1 Spectrum acquisition	12
1.3.2 Site-directed spin labelling of proteins	13
1.3.3 Rotational correlation times	14
1.3.4 The EPR spectrum of a nitroxide spin label	16
1.4 Pulsed EPR theory and applications	18
1.4.1 Magnetisation and the reference frames	18
1.4.2 Free induction decay and off-resonance effects	20

1.4.3	Electron spin relaxation processes	22
1.4.4	The double electron electron resonance (DEER) experiment	25
1.5	General NMR theory	33
1.6	Saturation Transfer Difference (STD) NMR	35
2	EPR as an integrative tool for studying membrane transporters	38
2.1	Summary	38
2.2	Membrane transport proteins	38
2.3	EPR of secondary active transporters	41
2.3.1	Symport	42
2.3.2	Antiport	47
2.4	Conclusions	51
3	Investigating conformational states and binding modes of MamM-CTD using advanced EPR techniques	53
3.1	Summary	53
3.2	The cation diffusion facilitator (CDF) protein family	54
3.2.1	The biological role of CDF proteins	54
3.2.2	CDF protein structures	55
3.3	Magnetotactic bacteria and magnetosome formation	59
3.3.1	From membrane invagination to protein sorting and activation	60
3.3.2	Alignment of magnetosomes into chains	63
3.3.3	Biomineralisation of the magnetic nanocrystal	64
3.3.4	Crystal size and shape	66
3.4	MamM from <i>Magnetospirillum gryphyswaldense</i> MSR-1	68
3.4.1	Structural and crystallographic studies of MamM-CTD	68
3.4.2	Biophysical characterisation and <i>in vivo</i> studies	73
3.4.3	The kinetic and conformational mechanism of MamM-CTD and other CDF proteins	78
3.5	Characterisation of MamM-CTD Zn ²⁺ binding using EPR spectroscopy	81
3.5.1	SDSL of MamM-CTD variants	81
3.5.2	Zn ²⁺ binding probed by cw- and pulsed EPR techniques	90
3.6	Metal selectivity of MamM-CTD	99
3.6.1	Initial indications of metal interaction probed by cw-EPR	100
3.6.2	DEER for probing metal binding and selectivity in MamM-CTD	105
3.7	Cu(II) - a special case	112
3.7.1	An introduction to Cu(II) EPR	112

3.7.2	Cu ²⁺ binding probed by DEER studies of MamM-CTD variants	113
3.7.3	Beyond pairwise distance determination in MamM-CTD	117
3.8	Discussion and conclusions	122
3.8.1	Conclusions	126
4	Advancing approaches for solving structural dynamics, modelled on GltPh	127
4.1	Summary	127
4.2	The SLC superfamily	128
4.3	Neurotransmission within the central nervous system	128
4.3.1	The chemical synapse	128
4.3.2	The glutamatergic synapse and glutamate transporters	129
4.3.3	Excitatory amino acid transporters	130
4.3.4	Stoichiometry of substrate transport and the alternate access model	130
4.4	GltPh - an archaeal homologue of the mammalian EAATs	132
4.4.1	General structure of GltPh	133
4.4.2	Substrate and ion binding sites	135
4.4.3	Crystallographic studies	136
4.4.4	Biophysical characterisation	140
4.4.5	Overall mechanism of substrate and ion transport	143
4.5	EPR Investigation of GltPh	145
4.5.1	SDSL and cw-EPR of a GltPh variant	145
4.5.2	DEER studies of GltPh _{SL}	149
4.6	Conformational locking of GltPh protomers	151
4.6.1	Ideas behind GltPh 'conformation-locking' variants	151
4.6.2	SDSL and cw-EPR measurements	153
4.6.3	DEER studies of conformation-locked variants	155
4.7	Discussion and conclusions (EPR)	158
4.8	STD NMR for probing binding events in GltPh	160
4.8.1	K _D determination by STD NMR and the effects of protein environment	160
4.9	Discussion and conclusions (STD NMR)	171
5	Materials and methods	174
5.1	MamM-CTD	174
5.1.1	Biological material	174
5.1.2	Site-directed spin labelling	174
5.1.3	Continuous wave X-band EPR spectroscopy	175
5.1.4	Pulsed X-band EPR spectroscopy	178

5.2	GltPh	182
5.2.1	Plasmids	182
5.2.2	Expression, purification and site-directed spin labelling of GltPh in de- tergent	182
5.2.3	Reconstitution into liposomes	183
5.2.4	Continuous wave X-band EPR spectroscopy	185
5.2.5	Pulsed X-band EPR spectroscopy	185
5.2.6	NMR spectroscopy	187
Bibliography		189
Appendices		205
I	Published research - Metal binding to MamM-CTD	205
II	Published research - Metal-type dependent binding modes and discriminatory behaviour of MamM-CTD	229
III	Published research - STD NMR investigation of substrate binding in GltPh . .	246

List of Figures

1.1	Spectroscopy schematic.	2
1.2	The electron Zeeman interaction.	5
1.3	Isotropic, axial and rhombic EPR spectra.	8
1.4	The nuclear Zeeman effect and hyperfine.	9
1.5	Nitroxide structure and the isotropic EPR spectrum.	15
1.6	The effect of T_c on nitroxide EPR spectra.	16
1.7	Multi-frequency nitroxide EPR.	17
1.8	Pulsed EPR off-resonance effects.	22
1.9	The effect of pulses on net magnetisation and spin populations.	23
1.10	The inversion recovery pulse sequence and its effect on magnetisation.	24
1.11	The Hahn echo pulse sequence and its effect on magnetisation.	25
1.12	The 4-pulse DEER sequence.	27
1.13	An example of DEER data analysis via Tikhonov regularisation.	30
1.14	An example of DEER data analysis using DeerNet.	31
1.15	The origin of the FID in NMR.	35
1.16	The STD NMR experiment.	37
2.1	Summary of membrane transport protein classes.	39
2.2	Predicted LeuT mechanism.	44
2.3	Predicted Mhp1 mechanism.	45
3.1	YiiP crystal structures.	56
3.2	CDF CTD structures, overlaid for comparison.	58
3.3	Magnetosomes within magnetotactic bacteria.	59
3.4	Schematic of vesicle formation.	61
3.5	Schematic of vesicle alignment.	63
3.6	Schematic of iron uptake into vesicles.	65
3.7	Schematic of magnetite formation in magnetotactic bacteria.	67
3.8	Overlay of apo and Cu^{2+} -bound MamM-CTD crystal structures.	71
3.9	Overlay of apo and Cd^{2+} -bound MamM-CTD crystal structures.	72

3.10 Suggested mechanism of Zn^{2+} binding by MamM-CTD.	79
3.11 MamM-CTD crystal structures.	82
3.12 Rotamer libraries of MamM-CTD labelled at positions 267 and 275 with MTSSL/IAP.	84
3.13 RT EPR spectra of spin labelled, apo MamM-CTD variants.	86
3.14 RT EPR spectra of MTSSL-labelled, apo MamM-CTD variants, highlighting broadened spectral features.	87
3.15 LT EPR spectra for MTSSL-labelled, apo MamM-CTD variants, with spectral fittings.	88
3.16 RT EPR spectra measured for MTSSL-labelled, apo MamM-CTD variants with spectral fittings.	89
3.17 RT EPR spectra measured for MTSSL-labelled MamM-CTD variants, in apo and Zn^{2+} -bound conditions, with spectral broadening highlighted.	91
3.18 LT EPR spectra measured for MTSSL-labelled, Zn^{2+} -bound MamM-CTD vari- ants with spectral fittings.	91
3.19 RT EPR spectra measured for MTSSL-labelled, Zn^{2+} -bound MamM-CTD variants, with spectral fittings.	92
3.20 The 2-pulse ESEEM pulse sequence, and the effects of cryo-protectants on 2P-ESEEM spectra.	94
3.21 The 4-pulse DEER pulse sequence, and field swept echo detected spectra for MamM-CTD variants (apo/ Zn^{2+} -bound).	95
3.22 Raw data: 4-pulse DEER measurements for MTSSL-labelled MamM-CTD variants, under apo and Zn^{2+} -bound conditions.	96
3.23 4-pulse DEER measurements for MTSSL-labelled MamM-CTD variants, un- der apo and Zn^{2+} -bound conditions.	97
3.24 4-pulse DEER measurements for MTSSL-labelled MamM-CTD variants fol- lowing Zn^{2+} titration.	99
3.25 RT EPR spectra measured for C267S _{SL} in apo and (various) metal-incubated conditions, in order of decreasing K_D	101
3.26 LT EPR spectra measured for C267S _{SL} in apo and (various) metal-incubated conditions, with spectral fittings.	102
3.27 RT EPR spectra measured for C267S _{SL} under apo and (various) metal- incubated conditions, with spectral fits.	104
3.28 MamM-CTD crystal structures with various divalent metal cations.	106
3.29 Field swept echo detected spectra for C267S _{SL} incubated with various metals.	108

3.30 4-pulse DEER experiments for C267S _{SL} following incubation with various metals.	111
3.31 Schematics of MamM-CTD variants and their available Cu ²⁺ binding sites, with possible inter-spin distances shown.	114
3.32 4-pulse DEER experiments, 1 and 2.	115
3.33 4-pulse DEER experiments measured for MamM-CTD variants with Cu ²⁺ , with differing pump and detect positions.	116
3.34 Pulse sequence for an inversion recovery experiment.	117
3.35 Inversion recovery versus field measurements for C267S _{SL} with Cu ²⁺	118
3.36 Inversion recovery-filtered field swept echos for C267S _{SL} with Cu ²⁺ for MTSSL and Cu ²⁺ suppression.	119
3.37 Pulse sequence for the <i>i</i> DEER experiment.	120
3.38 Overlaid 4-pulse DEER and <i>i</i> DEER experiments for C267 _{SL} with Cu ²⁺	121
3.39 <i>i</i> DEER experiments for C267S _{SL} with Cu _{SL} at a number of different filter times.	122
4.1 The glutamatergic synapse.	129
4.2 Transport stoichiometry in glutamate transporters.	131
4.3 Alternating-access models.	132
4.4 GltPh crystal structure and membrane topology.	134
4.5 Binding sites in GltPh.	135
4.6 Movement of the extracellular gate in GltPh with L-aspartate versus DL-TBOA.	137
4.7 Comparison of outward- and inward-facing GltPh structures.	138
4.8 Suggested transport mechanism by GltPh.	144
4.9 SDSL labelling position in GltPh.	146
4.10 EPR spectrum for spin labelled, apo GltPh _{SL}	147
4.11 EPR spectra for GltPh _{SL} in apo and multiple bound states.	148
4.12 Field-swept echo detected spectra for GltPh _{SL}	149
4.13 4-pulse DEER measurements for GltPh _{SL}	150
4.14 Movement of TM3 between outward- and inward-facing conformations in GltPh.	152
4.15 Highlighted residues involved in formation of a salt bridge for stabilisation of outward-facing GltPh.	153
4.16 EPR spectra measured for apo, spin labelled GltPh variants.	154
4.17 EPR spectra measured for GltPh variants in apo, and multiple bound states.	154
4.18 4-pulse DEER measurements for GltPh _{SL_{OF}}	156
4.19 4-pulse DEER measurements for GltPh _{SL_{IF}}	157
4.20 ¹ H NMR spectra, highlighting L-glutamate resonances.	162

4.21 T_1 analysis (plots and fittings) for L-glutamate resonances.	163
4.22 Reference and STD NMR spectra for WT GltPh (detergent-solubilised versus proteoliposomes) in the presence of NaCl.	164
4.23 Control reference and STD NMR spectra for WT GltPh.	166
4.24 Plots and fittings for L-glutamate K_D determination.	168
4.25 Plot of Na^+ dependency of the STD effect.	169
4.26 Plots and fittings for L-aspartate K_D determination.	170
4.27 Hill plots and fittings for L-aspartate binding.	171
5.1 Example script used in EasySpin, for simulated fits.	178

List of Tables

3.1	MMM analysis based on MamM-CTD crystal structures.	84
3.2	Spin labelling efficiencies for MamM-CTD variants.	85
3.3	Best fit simulation values for LT EPR spectra of MTSSL-labelled, apo MamM-CTD variants.	88
3.4	Best fit simulation values for RT EPR spectra of MTSSL-labelled, apo MamM-CTD variants.	89
3.5	Best fit simulation values for LT EPR spectra of MTSSL-labelled, Zn ²⁺ -bound MamM-CTD variants.	92
3.6	Best fit simulation values for RT EPR spectra of MTSSL-labelled, Zn ²⁺ -bound MamM-CTD variants.	93
3.7	Best fit simulation values for LT EPR spectra of C267S _{SL} , under apo and (various) metal-incubated conditions.	103
3.8	Best fit simulation values for RT EPR spectra of C267S _{SL} under apo and (various) metal-incubated conditions.	105
3.9	MMM analysis based on metal-bound MamM-CTD crystal structures.	108
4.1	MMM analysis based on multiple GltPh crystal structures.	147
4.2	Determined T ₁ values - GltPh NMR experiments.	163

Glossary of terms

ABC	ATP-binding cassette
AF	Amplification factor
B_0	Applied magnetic field
BCCT	Betaine/choline/carnitine transporter
B_{local}	Local magnetic field
B_{eff}	Effective magnetic field
CD	Circular dichroism
CDF	Cation diffuser facilitator
CM	Cell membrane
CNS	Central nervous system
CTD	C-terminal domain
cryo-EM	Cryogenic electron microscopy
CS	Central site
cw	Continuous wave
DDM	n-Dodecyl- β -D-Maltoside
DEER	Double Electron Electron Resonance
DL-TBOA	D,L-threo-beta-benzyloxyaspartate
DMSO	Dimethyl sulfoxide
DOPC	1,2-dioleoyl- <i>sn</i> -glycero-phosphocholine
DTT	Dithiothreitol
EAAT	Excitatory amino acid transporter
EEI	Electron-electron interaction
eL	External loop
EL	Extracellular loop
EPR	Electron Paramagnetic Resonance
ESEEM	Electron spin echo-envelope modulation
EZI	Electron Zeeman interaction
FID	Free induction decay
FSE	Field swept echo

HFI	Hyperfine interaction
HP	Helical hairpin
HS-AFM	High-speed atomic force microscopy
IAP	3-(2-iodoacetamido)-proxyl
<i>i</i> DEER	Inversion recovery-filtered DEER
IF	Inward facing
IRfFSE	Inversion recovery-filtered field swept echo
ITC	Isothermal calorimetry
LMV	Large multilamellar vesicles
LT	Low temperature
LUV	Large unilamellar vesicle
M_0	Bulk magnetisation
MAI	Magnetosome island
Mam	Magnetosome-associated membrane
MATE	Multidrug and toxic compound extrusion
MD	Molecular dynamics
MFS	Major facilitator superfamily
MM	Magnetosome membrane
MMM	Multiscale modelling of macromolecules
Mms	Membrane specific
MTB	Magnetotactic bacteria
MTSSL	S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl
...	methanesulfonylthioate spin label
sm-FRET	Single molecule fluorescence resonance energy transfer
NCS1	Nucleobase-cation-symport 1
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NQI	Nuclear quadrupole interaction
NSS	Neurotransmitter sodium symporter
NTD	N-terminal domain
NZI	Nuclear Zeeman interaction
OF	Outward facing
PL	Proteoliposome
PS	Peripheral site
REFINE	Relaxation-filtered hyperfine
RF	Radio frequency

RIDME	Relaxation-induced dipolar modulation enhancement
RT	Room temperature
SAXS	Small-angle X-ray scattering
SBDD	Structure-based drug design
SDSL	Site-directed spin labelling
SL	Spin label
SLC	Solute carrier
SMALPs	Styrene maleic acid lipid particles
ss-NMR	Solid-state nuclear magnetic resonance
SSSF	Sodium-solute symporter family
STD NMR	Saturation Transfer Difference Nuclear Magnetic Resonance
T_1	Spin-lattice (longitudinal) relaxation time
T_2	Spin-spin (transverse) relaxation time
T_c	Rotational correlation time
T_F	Filter time
TEM	Transmission electron microscopy
TMs	Transmembrane domains
WT	Wild type
ZnT	Zinc transporters

Chapter 1

An introduction to magnetic resonance spectroscopy

1.1 Thesis layout and chapter summary

The layout of this thesis includes firstly an introduction to the fundamental concepts of electron paramagnetic resonance (EPR) and nuclear magnetic resonance (NMR) spectroscopy, as well as key experiments used for acquiring and analysing the data presented in this thesis (chapter 1). Additionally, a short review of EPR studies of membrane transport machineries is given in chapter 2. Chapter 3 describes investigations into the mechanism of transition metal cation binding and transport by a membrane protein, MamM, utilising a variety of EPR methodologies for analysis of conformational change, binding modes and metal selectivity. Chapter 4 describes how approaches can be advanced for the solving of structural dynamics of large membrane transporters, modelled on a membrane protein known as GltPh. Preliminary results describing how (a) the introduction of conformational restraints, and (b) protein mutations to allow additional orthogonal labelling strategies, can be applied for deconvolution of complex spectral features are presented. Additionally, saturation transfer difference nuclear magnetic resonance (STD NMR) experiments have been employed for the first example of substrate dissociation constant determination for a membrane protein using this technique, also affording insight into the affects of protein environment on protein functionality. Chapter 5 includes all information relating to the materials and methods used for the work undertaken in this thesis.

The following chapter will introduce EPR spectroscopy and explain fundamental concepts of the EPR technique, as well as specific experimental procedures used for the acquiring of data within this thesis. These experiments include continuous wave- (cw-) EPR at both room and cryogenic temperatures, and pulsed EPR techniques including T_1 and T_2 measurements and double electron-electron resonance (DEER) experiments. In addition, a brief description of the STD NMR experiment, and its application, is given.

1.2 Fundamental EPR theory

Spectroscopy in general encompasses a vast number of different techniques which use electromagnetic radiation to induce transitions between discrete states in molecules or atoms, each of which has a corresponding energy. Spectroscopy affords information on the identity, structure and dynamics of samples under investigation through interpretation of the energy differences between these discrete states. Different spectroscopic techniques might monitor the absorption, emission or scattering of applied electromagnetic radiation by atoms or molecules. Methods have been developed to probe transitions induced throughout the entire width of the electromagnetic spectrum, and as such a vast number of atomic and molecular processes can now be probed. There is a direct relationship between the energy difference of transitions, ΔE , and the absorption of electromagnetic radiation which is described by the following equation 1.1,

$$\Delta E = h\nu \quad (1.1)$$

in which h is Planck's constant and ν is the frequency of the applied radiation. Transitions from lower to higher energetic states (as depicted in Figure 1.1A) occur only as a result of absorption by the atom/molecule under investigation of radiation corresponding to a specific frequency. Typically, this response is recorded as a function of frequency of the applied radiation, which is known as a spectrum (Figure 1.1B). Many spectroscopic tools are now routinely used in conjunction with other structural methodologies to afford understanding of structure, function and mechanisms of complex molecular machinery - this thesis presents such use of magnetic resonance spectroscopy techniques.

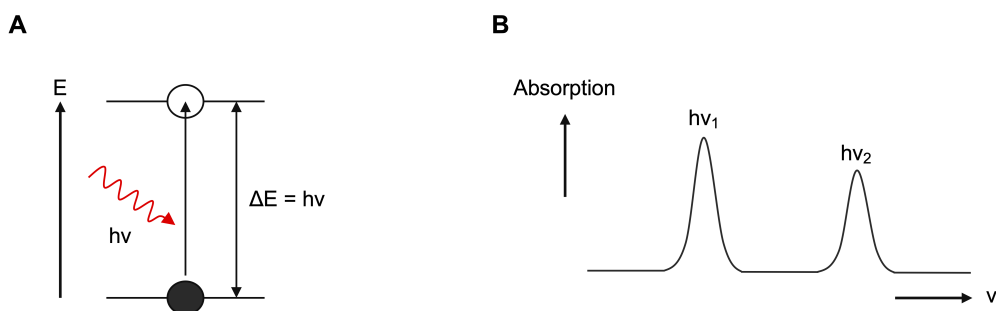


Figure 1.1: (A) Schematic representation of an energetic transition from lower to upper energy levels upon absorption of external electromagnetic radiation. (B) A spectroscopic spectrum representing absorption of two differing frequencies of electromagnetic resonances, resulting in two peaks.

1.2.1 The spin Hamiltonian

Electron paramagnetic resonance (EPR) spectroscopy probes transitions of unpaired electrons between discrete energy levels in the presence of an external magnetic field, induced by incident microwave radiation. Unpaired electrons are not typically found isolated within a vacuum, meaning multiple factors will affect the energetic states of the unpaired electron. The total energy of a spin system within a magnetic field is described by the spin Hamiltonian ($\hat{H}$) which includes operators for each different contribution and takes the following form:¹

$$\hat{H} = \hat{H}_{EZI} + \hat{H}_{NZI} + \hat{H}_{HFI} + \hat{H}_{EEI} + \hat{H}_{NQI} \quad (1.2)$$

where each contributing Hamiltonian term refers to specific interactions. These include the electron Zeeman interaction (EZI), the nuclear Zeeman interaction (NZI), the hyperfine interaction (HFI), the electron-electron interaction (EEI) and the nuclear quadrupole interaction (NQI). Those which are relevant to the systems perturbed as per this thesis will be discussed in more detail, but can be explained in short term as follows.

- EZI - interaction between the spin of an unpaired electron and an external magnetic field.
- NZI - interaction between the nuclear spin of a magnetic nucleus and an external magnetic field.
- HFI - interaction between the spin of an unpaired electron and the nuclear spins of nearby magnetic nuclei.
- EEI - interaction between spins of multiple unpaired electrons with one another, including those which are strongly coupled (described by zero-field splitting) and those which are weakly coupled (as in dipole-dipole interactions and superexchange).
- NQI - interaction between the electric quadrupole moment (in nuclei with nuclear spin $> \frac{1}{2}$) and the local electric field gradient at the nucleus.

1.2.2 The electron Zeeman interaction and the resonance condition

The phenomenon of EPR comes from the fact that an electron is a charged particle with an intrinsic property known as spin angular momentum (meaning that electrons effectively behave like rotating, charged particles); the combination of these gives rise to an intrinsic magnetic moment, μ_s , described by equation 1.3. Here, g_e is the electron spin g factor (a

dimensionless quantity equal to 2.0023 for a free electron),² $\mathbf{S}$ is the electron spin and μ_B is the Bohr magneton.

$$\mu_s = g_e \frac{\mu_B}{\hbar} \mathbf{S} \quad (1.3)$$

The orientation of electron spin (and hence the magnetic moment) is quantised, and is described by $m_s = +\frac{1}{2}$ and $m_s = -\frac{1}{2}$ states which are energetically degenerate in the absence of a magnetic field.^{3,4} However upon application of an external magnetic field, this degeneracy is lifted resulting in an energetic splitting between these states, which is a phenomenon known as the electron Zeeman interaction (EZI). Upon interaction of the external magnetic field with the electron's magnetic moment, the electron can be oriented either parallel ($m_s = -\frac{1}{2}$) or anti-parallel ($m_s = +\frac{1}{2}$) with respect to the direction of the external field. Orientation of the spin such that its magnetic moment aligns with the external magnetic field vector is more energetically favourable, and hence this represents the lower energy state. The magnitude of energy level splitting is directly proportional to the strength of the external magnetic field, B_0 .^{3,4} The contribution of the electron Zeeman interaction to the spin Hamiltonian is described by equation 1.4.

$$\hat{H}_{EZI} = g_e \cdot \mu_B \cdot B_0 \cdot \mathbf{S} \quad (1.4)$$

In multi-electron systems, the Pauli exclusion principle states that no two electrons can have the same quantum numbers, and as such electrons can only occupy the same orbital if they possess opposite values for m_s . In such a case, the spin angular momenta cancel one another out and so the overall magnetic moment of the system is equal to zero. Thus paired electrons are 'EPR silent' and produce no EPR response, meaning unpaired electrons are required in a sample for analysis by EPR.^{3,4}

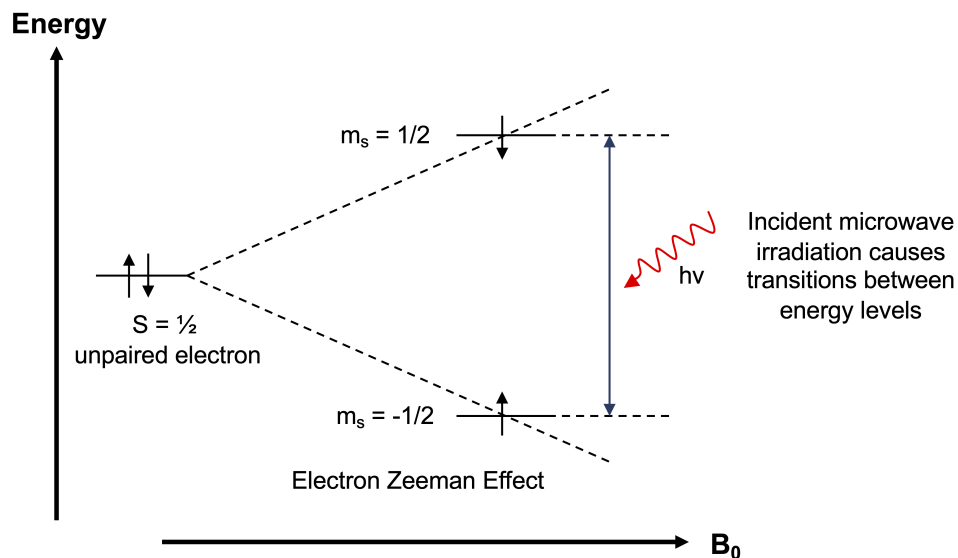


Figure 1.2: Image showing the electron Zeeman interaction within a spin $S = 1/2$ system containing one unpaired electron, in which application of an external magnetic field induces splitting of energy levels. The directions of alignment of spins with and against the external field are shown in lower and upper energy levels, respectively. Incident microwave irradiation (red arrow, of energy $h\nu$) of a specific frequency induces a transition between energy levels when the resonance condition is fulfilled.

At a particular magnetic field strength, thermal equilibrium results in a Boltzmann distribution of spins between the two states, described by equation 1.5, in which N_β and N_α correspond to the population of β ($m_s = -\frac{1}{2}$) and α ($m_s = +\frac{1}{2}$) states, respectively, k is the Boltzmann constant and T is the absolute temperature. In EPR, application of an external magnetic field causes divergence in energy levels, and the magnetic part of applied electromagnetic radiation of a specific energy induces transitions between these states; this is represented in Figure 1.2 and is known as the resonance condition.^{3,4}

$$\frac{N_\beta}{N_\alpha} = e^{(\Delta E/kT)} \quad (1.5)$$

This resonance condition (given by equation 1.6, where S is the total spin of the system) is typically fulfilled using electromagnetic radiation in the microwave region (GHz). However a large amount of radiation across the microwave frequency range is very effectively absorbed by water molecules present within the atmosphere. Thus specific 'bands' have been defined which correspond to specific frequencies of microwave radiation that are absorbed only minimally by atmospheric water molecules. EPR measurements are conducted at these frequencies, the most common of which is referred to as X-band, at approximately 9.5 GHz.⁵

$$\Delta E = h\nu = g_e \cdot \mu_B \cdot B_0 \cdot S \quad (1.6)$$

1.2.3 The spectroscopic g factor

The g factor is essentially a proportionality constant, which has a defined magnitude (~ 2.0023) for the 'free electron' that we have referred to thus far. In such a case, the electron is not constrained. However this situation changes when we consider an unpaired electron associated with an atom or molecule, in which the position of an electron is now constrained by its interaction with the nucleus, within spin orbitals.⁴ In this case we must now consider the phenomenon of orbital angular momentum, which affects electrons present in orbitals with a non-zero azimuthal quantum number (thus, all but s-orbitals). The orbital angular momentum of an electron is quantised (as was described for the electron's spin), and produces a magnetic moment with an associated energy which is also dependent on its orientation with respect to any external magnetic field. Now, the net magnetic moment associated with the electron has an additional orbital contribution, which in turn alters the energy required to fulfil the resonance condition. The orbital contribution (L) to the magnetic moment of the electron is now taken into account in our equation describing the electron Zeeman contribution to the spin Hamiltonian, shown in equation 1.7.⁴

$$\hat{H}_{EZI} = g_e \cdot \mu_B \cdot B_0 \cdot (L + S) \quad (1.7)$$

Unlike S , the magnitude of L is dependent on the orientation of the (non-spherical) molecule containing the unpaired electron, relative to the direction of the external magnetic field. Thus this contribution is termed anisotropic, and its orientation-dependence results in changes in the energy required (ΔE) to fulfil the resonance condition within the EPR experiment. As will be discussed further in the next section, EPR experiments are typically conducted using an incident microwave frequency that is fixed, whilst sweeping through the magnetic field to induce electronic transitions. Thus changes in L as a function of molecular orientation will manifest within the experiment as changes in the magnetic field strength required for transitions to occur. However, mathematical treatment of spin and orbital moments is complex and so typically the affects of S and L are combined, to give a fictitious spin, S' . Now the equation describing the resonance condition should be altered accordingly, as shown in equation 1.8.

$$\Delta E = h\nu = g_{eff} \cdot \mu_B \cdot B_0 \cdot S' \quad (1.8)$$

Note that a fictitious, and orientation dependent g factor - g_{eff} - is also assigned to this spin, which describes the rate of divergence of spin states as a function of the magnetic field; this

divergence is therefore also anisotropic. The magnitude of this fictitious g factor is indicative of the character of the orbital in which the unpaired electron resides, and hence can differ greatly depending on the species being measured.⁴

1.2.4 g value anisotropy

The g factor must be replaced by a 3x3 matrix in order to represent the orientational dependencies of these anisotropic interactions, relative to the cartesian axis, which is shown below:

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

In the principle axes (x,y,z) of this matrix, the magnitudes of the six off-diagonal elements (g_{xy} , g_{xz} , g_{yx} , g_{yz} , g_{zx} , g_{zy}) are equal to zero. Thus we are left with the diagonal components, consisting of g_{xx} , g_{yy} and g_{zz} . It is only ever necessary to define these three principle g values, since those at any other orientation can be determined using a weighted sum of g_{xx} , g_{yy} and g_{zz} .³

A system can be described as either isotropic, axial or rhombic, depending on the symmetry of the paramagnetic system. Conventionally, the molecular axis with the highest symmetry is defined as the z axis. For a system to be classed as isotropic in the solid (i.e. a frozen solution) state, it must possess symmetry such that all three principle axes (x,y,z of the molecule) are exactly equivalent.³ In this instance, the three principle g values will be equal and their magnitudes will be independent of the molecule orientation with respect to the orientation of an external magnetic field. Such symmetry is rare in the solid state, however g is also described as isotropic in a system which is undergoing rapid motion in the solution state; rapid tumbling and rotation of molecules relative to an external magnetic field means that information on orientation is lost, and the anisotropic g of the (equivalent) frozen solution (or powder) spectrum collapses to a single, averaged (isotropic) value.⁴

A system possessing axial symmetry is one in which the x and y axes are equivalent, whilst the z axis is not. This type of symmetry results in two principle g values which are equal, and one which is different; i.e. $g_{xx} = g_{yy} \neq g_{zz}$. In such cases, g_{xx} and g_{yy} might be described in one term as $g_{\perp}$, whilst g_{zz} might be termed $g_{\parallel}$. Now the resonance condition is fulfilled at different magnetic field strengths depending on the relative orientation of the

paramagnet with respect to that of the external magnetic field.³

Finally, a rhombic system is one in which all principle axes are inequivalent, i.e. $x \neq y \neq z$. This type of symmetry results in three inequivalent principle g values (meaning $g_{xx} \neq g_{yy} \neq g_{zz}$). Examples of isotropic, axial and rhombic EPR spectra are given in Figure 1.3, in which the relevant g values are labelled.³

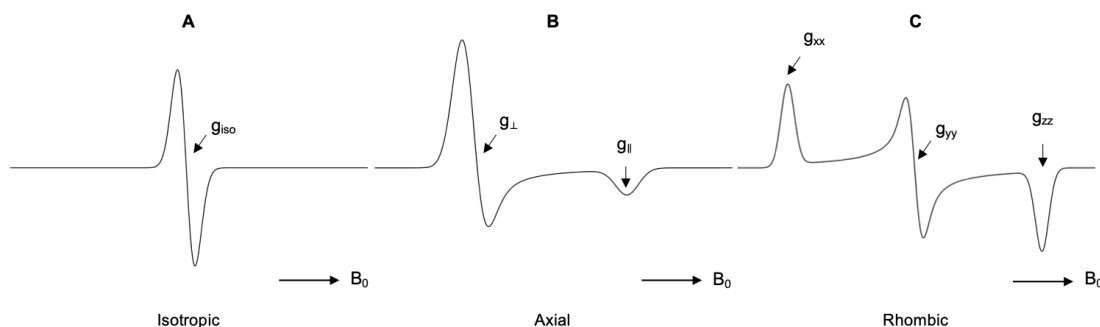


Figure 1.3: Simulated isotropic (A), axial (B) and rhombic (C) EPR spectra centred at $g = 2$ for an $S = 1/2$ species, with g factors indicated. The Easyspin package within Matlab was used for spectral simulations.

1.2.5 Hyperfine interactions

Thus far we have considered the interaction between an unpaired electron and the atomic nucleus it orbits in the context of spin-orbit coupling, and its affect on the g value. However the magnetic moment associated with an unpaired electron is also affected by nearby magnetic nuclei, resulting in characteristic splitting/s of the EPR signal.

In much the same way as was described for the electron, a charged nucleus which possesses nuclear spin (characterised by a non-zero nuclear spin quantum number, I) has an associated magnetic dipole moment; examples of such nuclei include protons ($I = \frac{1}{2}$) and nitrogen atoms (^{14}N , $I = 1$). The magnitude of nuclear spin moments are quantised, and can assume $2I + 1$ discrete orientations which are described by the magnetic quantum number, m_I . In the absence of a magnetic field, the energies of these discrete states are equivalent. However, this degeneracy is lifted in much the same way as that of the electron spin quantum states, under the influence of an external magnetic field; this is termed the nuclear Zeeman effect (or interaction) and is accounted for within the spin Hamiltonian as follows:^{3,4}

$$\hat{H}_{NZI} = -g_N \cdot \mu_N \cdot B_0 \cdot I \quad (1.9)$$

where g_N is the effective nuclear g-factor (which can have positive or negative values),

μ_N is the nuclear magneton and I is the nuclear spin quantum number. The magnitude of splitting of the nuclear energy levels (as a result of the nuclear Zeeman interaction) is far smaller than that of electron spin levels (as a result of the electron Zeeman interaction). Thus in the context of the EPR experiment, the combined interaction of the electron and nuclear magnetic moments with an external magnetic field results in further splitting of each of the non-degenerate electron Zeeman levels into $2I+1$ nuclear Zeeman levels.^{3,4} This effect is demonstrated in Figure 1.4 for an $S = \frac{1}{2}$, $I = \frac{1}{2}$ system. This results in a total of four energy levels in the example described, and selection rules mean that transitions may only occur between states where $\Delta m_s = \pm 1$, and $\Delta m_I = 0$, thus meaning there are two formally allowed transitions in this case.

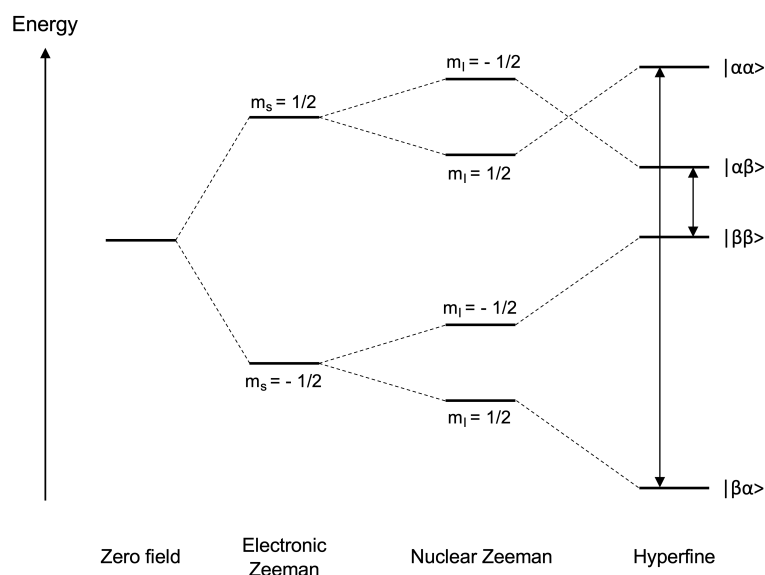


Figure 1.4: Schematic representation of energy level splitting caused by a combination of electron Zeeman, nuclear Zeeman and hyperfine interactions for an $S = 1/2$ paramagnetic species interacting with a proton, with a nuclear spin $I = 1/2$.

In the absence of any interaction between the electron and nearby nuclear magnetic moments, the energy required to induce each allowed transition within the EPR experiment would be equal, and thus the resulting spectrum would contain one single line, centred at g_{eff} . However these magnetic moments do indeed interact, in what is known as the Hyperfine interaction (described by $\hat{H}_{HFI}$ in equation 1.10, where $\mathbf{A}$ is the hyperfine interaction matrix). The hyperfine interaction causes a perturbation of the nuclear Zeeman energy levels towards higher or lower energy, as shown in Figure 1.4 (on the right hand side). Energy levels are designated $|\alpha\alpha\rangle$, $|\alpha\beta\rangle$, $|\beta\beta\rangle$ and $|\beta\alpha\rangle$, where $|\alpha\rangle$ and $|\beta\rangle$ refer to the spin states ($\frac{1}{2}$ and $-\frac{1}{2}$ respectively) in the external magnetic field. The extent of the perturbation depends on the magnitude of interaction, and is reflected by the hyperfine coupling constant, A .^{3,4} As a result, allowed transitions will now require different energies and therefore will fulfil the res-

onance condition with application of different magnetic field strengths, resulting in two lines in the EPR spectrum of an $S = \frac{1}{2}$, $I = \frac{1}{2}$ system.

$$\hat{H}_{HFI} = \mathbf{I} \cdot \mathbf{A} \cdot \mathbf{S} \quad (1.10)$$

There are two types of interaction which contribute to the magnitude of the total hyperfine interaction (which notably is field-independent) - namely the isotropic (a) and anisotropic (or dipolar) interactions (T). The isotropic contribution is also termed the Fermi contact hyperfine interaction, which relies on the finite probability of finding electron spin density at the (interacting) nucleus. Inside the nucleus, the magnitude of the local hyperfine field as a result of the nuclear magnetic moment is constant in all directions, and thus the extent of this hyperfine coupling is independent of orientation (hence isotropic). This situation is only ever the case for electrons residing in s-orbitals, or those with some s-orbital character (i.e. with no node at the nucleus). The strength of the isotropic hyperfine interaction is defined by the isotropic hyperfine coupling constant, A_{iso} .⁶

The anisotropic contribution is a result of dipole-dipole interaction between the electron and nuclear magnetic moments.¹ The energy of this interaction is dependent on both the distance and relative orientations of the electron and nuclear magnetic moments. The nuclear magnetic moments generate a local magnetic field (B_{local}) at the electron; thus the electron feels contributions from both external and local fields. Depending on the angle (θ) between B_0 and the vector joining the two dipoles, this local magnetic field will either add to or subtract from the external field, changing the effective field at the electron (B_{eff}).^{1,3,4} Since the electron is not localised at a fixed point, the energy of the dipolar coupling interaction (E_{dip}) must be averaged over the entire probability of the spin distribution. In a spherical electron cloud (i.e. and s-orbital), or in the case where rapid tumbling means that anisotropy relative to the external field direction is averaged, E_{dip} is averaged to zero. The hyperfine coupling in such cases is then described as A_{iso} .⁶ Meanwhile, for electrons present within axially symmetric orbitals, such as p-orbitals, E_{dip} becomes a finite value and the hyperfine coupling constant, A , describes contributions from both isotropic and anisotropic hyperfine coupling.

The hyperfine tensor, denoted $\mathbf{A}$, of the interacting nucleus comprises both an isotropic (A_{iso}) and purely anisotropic (or dipolar) part ($\mathbf{T}$), and is described by equation 1.11. The purely dipolar component is expressed by equation 1.12. $\mathbf{T}$ is a traceless, axially symmetric tensor which averages to zero when molecular tumbling is rapid. The hyperfine tensor can again

(much like for the g value) be described by just the three principle components, A_{xx} , A_{yy} and A_{zz} .⁴ Hyperfine coupling constants can be measured directly from EPR spectra, as will be demonstrated later.

$$\mathbf{A} = \begin{bmatrix} A_{xx} & A_{yx} & A_{zx} \\ A_{xy} & A_{yy} & A_{zy} \\ A_{xz} & A_{yz} & A_{zz} \end{bmatrix} = A_{iso} + \mathbf{T} \quad (1.11)$$

$$\mathbf{T} = \begin{bmatrix} T_{xx} & T_{yx} & T_{zx} \\ T_{xy} & T_{yy} & T_{zy} \\ T_{xz} & T_{yz} & T_{zz} \end{bmatrix} \quad (1.12)$$

1.2.6 Electron dipole-dipole interactions

The final contribution to the spin Hamiltonian to be expanded upon is the electron-electron interaction, which describes the interaction between spins of multiple unpaired electrons with one another. Specifically, we will consider those which are weakly-coupled via dipole-dipole interactions, since this is relevant to specific EPR methods used throughout the results chapters of this thesis. The energetic contribution from dipole-dipole interactions to the total spin Hamiltonian is described by equation 1.13,³

$$\hat{H}_{dip} = \mathbf{S}_1 \cdot \mathbf{D} \cdot \mathbf{S}_2 \quad (1.13)$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the electron spins of the two interacting electrons, and $\mathbf{D}$ is the traceless electron dipole-dipole coupling tensor (described as follows in equation 1.14).

$$\mathbf{D} = \frac{\mu_0}{4\pi\hbar} \cdot \frac{g_A g_B \mu_B^2}{r_{AB}^3} \cdot \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -\omega_{dd} & 0 & 0 \\ 0 & -\omega_{dd} & 0 \\ 0 & 0 & 2\omega_{dd} \end{bmatrix} \quad (1.14)$$

Here, ω_{dd} is the dipolar coupling frequency and can be expressed as follows:

$$\omega_{dd} = \frac{g_A g_B \mu_B^2 \mu_0}{4\pi\hbar} \cdot \frac{1}{r_{AB}^3} \cdot (1 - 3 \cos^2 \theta) \quad (1.15)$$

where μ_0 is the vacuum permeability, μ_B is the Bohr magneton, g_A and g_B are the g factors associated with the two unpaired electrons, r_{AB} is the distance vector between coupled spins (A and B) and θ is the angle between $\mathbf{B}_0$ and r_{AB} . The electron dipole-dipole interaction is anisotropic and depends on the distance between, and relative orientations of, the magnetic moments of interacting unpaired electrons - this is similar to the situation describing the interaction between electron and nuclear magnetic moments in the hyperfine interaction.

The distance-dependent dipole-dipole interaction between electrons affords the determination of inter-spin distances typically between 0.8 - 8 nm apart. Within this distance range, dipolar coupling means that the magnetic moment of each spin changes the effective magnetic field felt by the other, hence altering the Zeeman energies and the resonance condition. At inter-spin distances of approximately 0.8 - 2 nm, this typically manifests as a broadening of the EPR line shape, which becomes less discernable with increasing r_{AB} since the extent of broadening approaches the spectral linewidth. Advances in pulsed EPR methodologies have enabled longer-range inter-spin distance determination, through application of the double electron-electron resonance (DEER) experiment.

1.3 Continuous-wave EPR spectroscopy

1.3.1 Spectrum acquisition

There are many difficulties associated with the engineering of a microwave source capable of stably sampling a range of different frequencies. For this reason, it is more convenient in EPR spectroscopy to maintain a static microwave frequency and apply a sweeping magnetic field. Thus the transition of an electron between states occurs when the gap between energy levels (which is proportional to the external magnetic field) matches the energy of the incident microwave radiation. This manifests as an absorbance signal in the EPR spectrum plotted against the experimental magnetic field sweep width, in Gauss (G). With ΔE lying in the microwave region, the population difference between upper and lower energy levels is small - this translates to a net absorption signal which is also small. For example, the ratio of $N_\beta:N_\alpha$ at 293 K in a 0.34 T (3400 G) magnetic field is $\sim 1.0016:1$, therefore only slightly favouring the lower energy β level. To improve the signal:noise ratio of the continuous wave EPR signal, a field modulation technique is employed resulting in a first derivative spectrum of the absorption.^{3,4}

In the experimental set-up, the microwave field is established by a resonator (a rectangular or cylindrical cavity), into which the sample tube is introduced. The cavity is a metal box with either a rectangular or cylindrical shape which resonates with microwaves. The cavity stores the microwave energy and at the resonance frequency of the cavity no microwaves will be reflected back, but rather they will all remain inside the cavity. As a consequence of resonance, there will be a standing wave within the cavity. The electric and magnetic field components of a standing wave are at 90° to one another, meaning the point at which one is

at its maximum, the other is at its minimum. Within EPR, we are interested in the absorption of the magnetic component of microwave radiation, and so if the sample is placed in the electric field minimum (corresponding to the magnetic field maximum), the most intense signals and highest sensitivities are possible. Cavities are therefore designed for optimal placement of the sample.⁷

When a sample absorbs additional microwave radiation during resonance, detuning of the resonator occurs and the microwave power is reflected back to the bridge, which causes the EPR signal. The magnetic field strength that the sample 'sees' is modulated, at a fixed modulation frequency (in the order of kHz), over a width of magnetic field (in the order of Gauss, G). If the resonance condition is fulfilled and an EPR signal occurs, the field modulation sweeps through part of the signal, modulating the amplitude of microwaves reflected back from the cavity to the bridge. This controls the frequency range of the detected noise. The signal channel produces a signal proportional to the amplitude of the modulated EPR signal. The modulated signal is then compared with a reference signal, which has the same frequency as the magnetic field modulation, and signals that do not possess the same phase and frequency as the field modulation are filtered out. Field modulation results in the detection of dA/dB_0 (where A and B_0 are the absorbance and applied magnetic field, respectively), and so continuous wave EPR spectra are typically presented as first derivatives. The intensity of the detected signal increases with the difference between the absorption signals at either end of the modulation window.⁷

It is important to note that although the size of the detected EPR signal is proportional to the modulation amplitude, if the modulation amplitude is greater than the natural linewidth of the signal, line broadening will occur – the spectrum is therefore no longer a true representation of the system. Therefore, effective and practical spectrum acquisition involves being able to maintain a maximum signal to noise ratio, whilst preventing line broadening.⁸

1.3.2 Site-directed spin labelling of proteins

Many biological molecules of interest, including many proteins, do not naturally contain paramagnetic centres and as such their study via EPR was once considered limited. However a major turning point in the development of EPR methodologies was the introduction of site-directed spin labelling (SDSL), affording EPR to be used as a key technique for structural insights into biological processes. SDSL is possible due to the development of

(a) site-directed mutagenesis and recombinant protein expression, and (b) the synthesis of stable, paramagnetic spin probes with specific reactivity which can be attached to protein and nucleic acid molecules. As such, these spin probes (known as spin labels) can be attached to selected places of interest, rendering a once diamagnetic molecule now paramagnetic, for study using EPR in which the spin label acts as a reporter for the system.⁹ More specifically, the initial step in preparation of protein for SDSL and EPR involves mutagenesis of the protein sequence for incorporation of a cysteine residue/s at selected sites of interest. Sites are chosen with consideration of structural and functional implications of mutagenesis, and in regions dependent upon the aim of the studies. Recombinant DNA can then be transformed into a suitable host expression system, such as *Escherichia coli*, for gene expression and purification of the protein variant. Spin labelling can then be carried out for covalent attachment of stable radical molecules which react specifically with the thiol groups of introduced (or natural) cysteine residue/s; typical functional groups used for this reactivity include methanethiolsulfonate, maleimide and iodoacetamide moieties.¹⁰

Many of the most common spin labels contain paramagnetic nitroxide centres (N-O), usually incorporated into a heterocyclic ring, and the electron located in a π orbital with the most spin density located on N and O atoms. Modification of the nitroxide structure can have dramatic consequences for the reactivity of the spin label, and as a result many different spin labels have been designed. One such spin label - MTSSL - is widely used for protein studies and includes proximal dimethyl groups for steric shielding and increased stability of the radical species. MTSSL forms a disulphide bond to the thiol group of a cysteine residue, and this linkage allows flexibility of the spin label side chain, hence limiting any effects on the stability and structure of the protein to which it is attached. Additionally, this spin label can be easily cleaved using a reductant such as dithiothreitol (DTT) for control experiments.¹⁰

1.3.3 Rotational correlation times

An EPR spectrum reports on all orientations of molecules in a sample - in the solution state, this translates to the time-averaged ensemble of molecules. The rotational motion of molecules in solution means that anisotropic effects average to zero over time. Consider a nitroxide radical such as MTSSL (an $S = \frac{1}{2}$, $I = 1$ system). A typical solution state (isotropic) spectrum of MTSSL tumbling freely in water at room temperature and X-band frequency consists of three narrow lines, centred at an isotropic g factor (g_{iso}) and separated by an

isotropic hyperfine splitting (A_{iso}). The origin of this EPR signature with respect to Zeeman interactions and allowed transitions is demonstrated in Figure 1.5.¹¹

For the complete averaging of anisotropic contributions, the rate of molecular tumbling should be much faster than the difference between resonance frequencies for different molecular orientations. Whilst this might be the case for small paramagnetic molecules in solutions of low viscosity, the same is not true for larger molecules, or for molecules in more viscous solutions. Thus in the latter two examples, anisotropic contributions are only partially averaged which results in broadening and characteristic changes in the EPR spectral line shape.

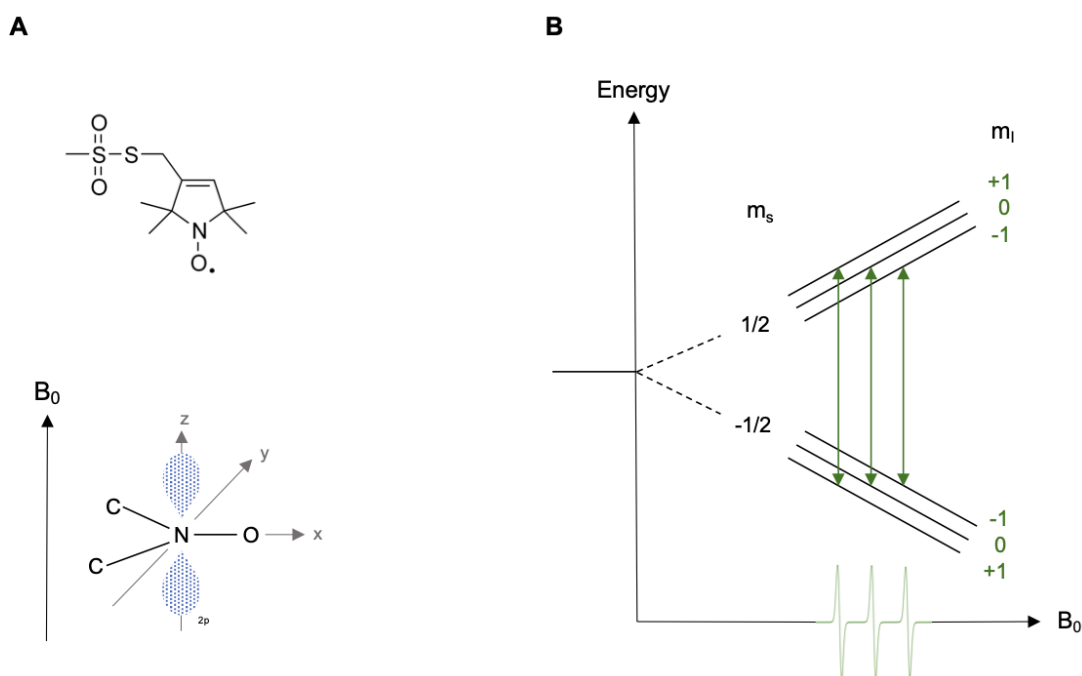


Figure 1.5: (A) The chemical structure of MTSSL (top), with the orientation of the molecular axis shown relative to the external magnetic field, B_0 (bottom). (B) The energy level diagram for a nitroxide radical species such as MTSSL, as a result of hyperfine coupling between the unpaired electron and the nitrogen nucleus. Transitions are shown which result in each line of the isotropic EPR spectrum below (green).

For MTSSL-labelled protein molecules, spectra will differ considerably based on the environment and rotational motion of the attached spin probe. The information that can be obtained from these spectra relates to the overall rotational motion of the spin label, defined by the rotational correlation time (T_c), and more generally referred to as the mobility of the attached probe. Multiple spectra of a nitroxide radical have been simulated and are presented in Figure 1.6 for varying rates of isotropic motion.

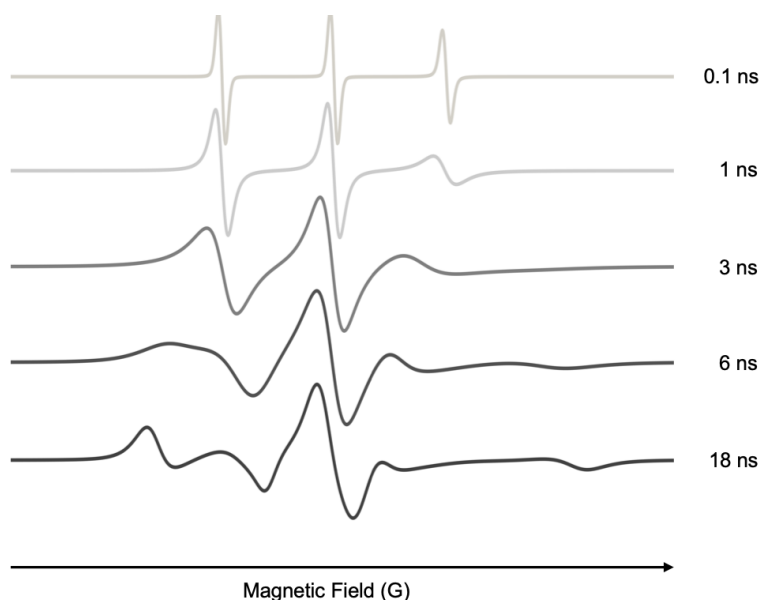


Figure 1.6: Simulated room temperature X-band nitroxide ($S = \frac{1}{2}$, $I = 1$) EPR spectra centred at $g = 2$, for a variety of rotational correlation times (T_c) which are indicated.

At the longest correlation time ($T_c = 18$ ns) the spectrum nearly resembles that of an immobilised powder spectrum. Gradually, as the rate of rotational motion increases (i.e. as T_c decreases) the splitting between high- and low-field resonances decreases and the overall line shape changes, firstly by loss of spectral anisotropy at the low-field ($m_I = 1$) manifold, followed by the high-field ($m_I = -1$) manifold. Eventually, all spectral anisotropy is lost, and the spectrum consists of three first derivative lines centred at g_{iso} , with an associated isotropic hyperfine coupling constant. Longer correlation times for spin labels attached to proteins might be indicative of a more buried labelling position, where steric factors mean that the probe is unable to undergo rotation on such a fast timescale as if it were attached at the solvent accessible protein surface, for example. Conformational changes might also occur on the timescale of a cw-EPR measurement, and as such the EPR spectrum will be the weighted sum of the spectra associated with each conformational state adopted by the protein in solution. Thus an EPR spectrum can report on the mobile conformational states of a protein exhibited in solution, as well as their relative populations.^{11,12}

1.3.4 The EPR spectrum of a nitroxide spin label

Since both the g tensor and hyperfine interactions (denoted by the hyperfine splitting constant A) are anisotropic, an important feature of powder spectra of nitroxide spin labels is the angle of the external magnetic field relative to the molecular axis defined in Figure 1.5A. Here, the z -axis points in the direction of the p -orbital which takes part in π -bonding across the N-O bond, and in which the unpaired electron resides. Meanwhile, the x -axis

lies along the N-O bond, and the y-axis is perpendicular to both z- and x-axes. Thus if B_0 is placed along the z-axis, we observe g_{zz} and A_{zz} components of the nitroxide EPR spectrum; the equivalent is true for g_{xx}/A_{xx} and g_{yy}/A_{yy} contributions when we place B_0 along either the x or y molecular axes.¹²

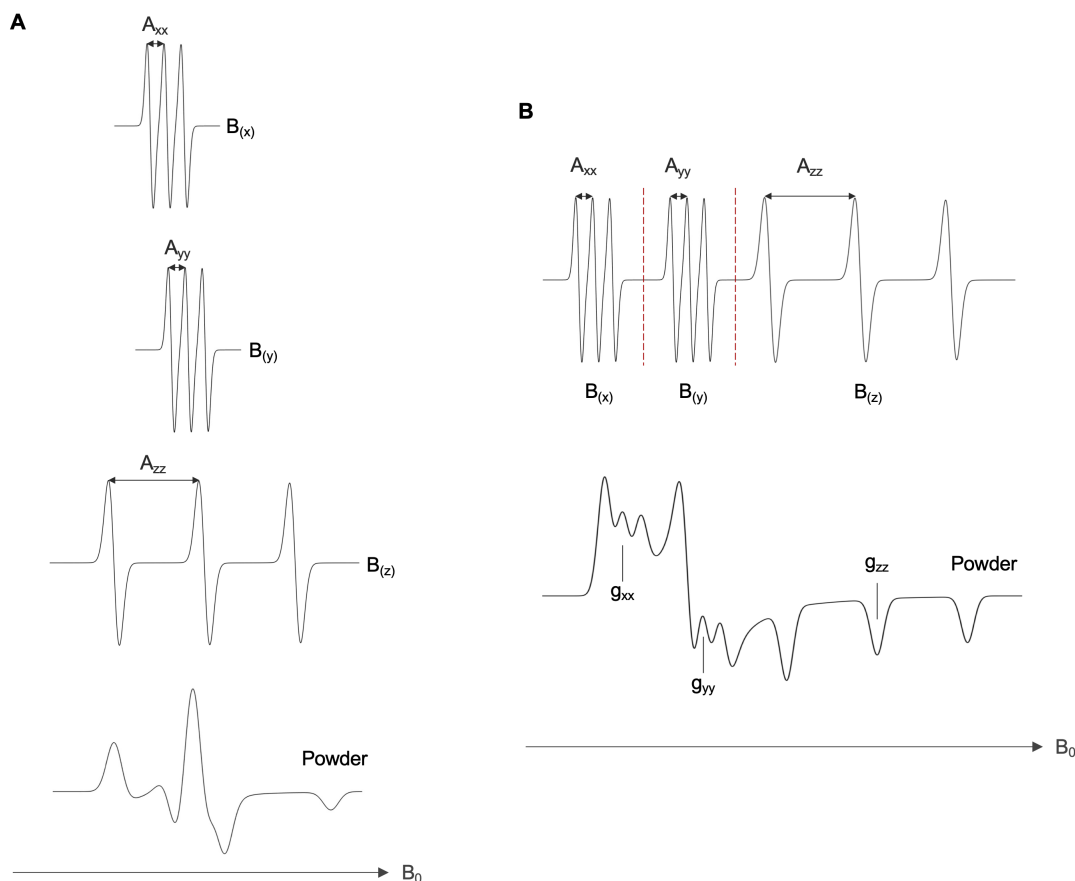


Figure 1.7: Simulated nitroxide EPR spectra at (A) X-band and (B) Q-band frequencies for both crystal and powder samples, showing the origin of the nitroxide powder spectrum and highlighting the increased resolution obtained at higher frequencies. Spectra have been simulated using Easyspin, within Matlab.

The effects of g value anisotropy and hyperfine interactions are shown in Figure 1.7, where parts A and B represent the situation at X-band and Q-band, respectively. Let us firstly consider a crystal sample in which all nitroxide spin label molecules are oriented in one direction. Contributions from each principle axis are shown as individual spectra, labelled according to the direction in which the external magnetic field is pointed ($B_{(z)}$, $B_{(y)}$ and $B_{(x)}$). At X-band frequency (Figure 1.7A, ~9.5 GHz) the effects of g value anisotropy are negligible compared to the effects of hyperfine interactions. The central position of each spectrum, defined as g_{xx} , g_{yy} and g_{zz} when the magnetic field points along the x,y and z axes, respectively, moves minimally. Meanwhile the hyperfine contribution along the molecular z-axis, A_{zz} , is much larger than either A_{xx} or A_{yy} . This is because the z-axis lies along the 2p orbital of the π -bond, hence giving rise to the strongest hyperfine

interaction as compared with interactions along x- and y-axes. Values for A_{xx} and A_{yy} are approximately equivalent, and thus the hyperfine interaction displays axial symmetry overall. The bottom spectrum in Figure 1.7A represents that of a powder sample measured at X-band, in which there are many orientations of the nitroxide spin label with respect to the external magnetic field direction. As such, the EPR spectrum gives a superposition of all overlapping contributions at each orientation; this spectrum no longer resembles the typical first derivative spectrum, since it is the derivative of the envelope formed by adding together all absorption packets across all possible orientations. The spectral width is dominated by A_{zz} which defines the outer wings of the spectrum - this is the only value we can resolve at X-band frequency, since all other contributions are overlapping in the spectrum.

At the much higher W-band frequency of ~ 94 GHz, (shown in Figure 1.7B) the powder spectrum (bottom) is now dominated by the effects of g value anisotropy; these effects increase linearly with the external magnetic field strength. Between each individual crystal spectrum, (again labelled $B_{(x)}$, $B_{(y)}$ and $B_{(z)}$ according to the direction of the external magnetic field), the central position of the spectrum varies more dramatically and these differences are now much larger than any of the hyperfine splittings. The effect of the hyperfine interaction is equivalent to that at X-band frequency, since it is not dependent on the magnetic field strength. The g tensor is not axial like the hyperfine interaction, and is instead rhombic since the nitroxide spin label molecule itself displays rhombic symmetry; for B_0 along the x and z molecular axes, the spectrum is centred to low and high field, respectively, as compared with the spectrum for B_0 along the y-axis.¹²

The polarity of the solvent also influences the nitroxide spectrum, causing changes in the A_{zz} and g_{xx} components of hyperfine and g tensors, respectively. A highly polar environment increases A_{zz} , whilst g_{xx} shifts to lower values. The increase in A_{zz} is reflected at X-band by an increased splitting between the outermost features of the spectrum, whilst changes to g_{xx} can only be observed by measuring at higher frequencies.

1.4 Pulsed EPR theory and applications

1.4.1 Magnetisation and the reference frames

Pulsed EPR techniques, as the name suggests, use pulses of electromagnetic microwave radiation within an external magnetic field in order to excite spin packets. Relaxation

of spin packets back to equilibrium is then monitored to give the EPR signal. However an understanding of the way in which bulk magnetisation associated with the perturbed electron spins is affected by these pulses of radiation, first requires an understanding of specific reference frames. Firstly let us consider the laboratory frame; this is a static frame described by a simple set of Cartesian axes (x , y , z which are mutually perpendicular) in relation to our EPR experimental setup. Here, the external magnetic field (B_0) lies parallel to the z -axis, and the microwave field (B_1) lies parallel to the x -axis. The y -axis is then orthogonal to both x - and z -axes.¹³

When an electron, which possesses spin angular momentum, is placed in a magnetic field along B_0 , it experiences a torque which causes precession of its magnetic moment about the z -axis at the Larmor frequency (ω_L). The Larmor frequency can be calculated using equation 1.16,

$$\omega_L = -\gamma \cdot B_0 \quad (1.16)$$

where γ is the gyromagnetic ratio (a constant of proportionality, in $\text{r.s}^{-1}.\text{T}^{-1}$). A free electron in a magnetic field of 3480 G precesses at a Larmor frequency of approximately 9.75 GHz - this corresponds to the EPR frequency at that magnetic field strength. As discussed previously, an electron spin within a magnetic field is characterised by two discrete, quantum mechanical states - where the associated magnetic moment aligns parallel or antiparallel relative to the orientation of the applied magnetic field. Parallel alignment is most energetically favourable, and so we have a net magnetisation along the z -axis; however spins will not favour any particular position in the x - y plane meaning their orientations within this plane will be random. The spins are still precessing about the z -axis, and so the various transverse components within the x - y plane effectively cancel one another out, resulting in stationary magnetisation (M_0) along B_0 .

Introduction of an additional 'rotating' frame¹³ is necessary to consider further principles of the pulsed EPR technique, since there are many rotating and precessing parts. EPR experiments are usually conducted with resonators which produce linearly polarised microwaves; these are designed to create a magnetic field (denoted B_1) which is perpendicular to the much stronger B_0 . A linearly polarised microwave is equivalent to the sum of two magnetic fields rotating in opposite directions at the microwave frequency. If the coordinate frame is rotated at this frequency, the result is a rotating frame of reference when $\omega_L = \omega_0$, where ω_0 is the microwave frequency. This makes one component of B_1 appear stationary, whilst the

other can be neglected. The rotating frame also makes the magnetisation precessing at the Larmor frequency appear stationary. In much the same way as was described for electron spins under the influence of B_0 , application of B_1 results in rotation about B_1 for the lifetime of the pulse, according to the following equation 1.17,

$$\omega_1 = \gamma \cdot B_1 \quad (1.17)$$

where ω_1 is known as the Rabi frequency. Within the rotating frame, we need only consider the affects of B_1 on M_0 ; this is explained by equation 1.16 since in the rotating frame there is no precession and hence B_0 can be ignored. For example, if B_1 is parallel to the x-axis, M_0 will rotate about the x-axis in the z-y plane for the duration that the pulse (which created B_1) is applied. The angle by which M_0 is displaced from the z-axis is described by the tip angle, α , according to equation 1.18 where t_p (s) is the pulse length.

$$\alpha = -\gamma |B_1| t_p \quad (1.18)$$

The pulses used in EPR experiments are labelled according to the tip angle that they induce, with a directionality according to the axis from which they originate. Hence a $+x \pi/2$ pulse is the application of field B_1 along the $+x$ direction, which will tip M_0 by $\pi/2$ (90°) into the $-y$ -axis. Detection will therefore be set up along the y -axis. Meanwhile, a $+x \pi$ pulse would tip M_0 by 180° to the $-z$ direction.¹³

1.4.2 Free induction decay and off-resonance effects

Emission of microwaves by a sample following excitation by means of a microwave pulse (which creates B_1) is most clearly understood within the static laboratory frame. Let us consider the application of a $+x \pi/2$ pulse; following the pulse precession of spins occurs again about the z-axis (in the x-y plane). Thus the static magnetisation along $-y$, when viewed from the rotating frame, becomes magnetisation which is rotating in the x-y plane at the Larmor frequency, when viewed from the laboratory frame. This generates currents and voltages in the resonator; the generated microwave signal is called the FID (free induction decay). This signal is at its maximum for magnetisation which is directly in the x-y plane.^{3,13}

The situation is not quite as simple as we have discussed thus far, since we have been working under the assumption that the Larmor frequency is exactly equal to the microwave

frequency (i.e. exactly on-resonance). In real systems, however, there will be many different frequencies which all contribute to the EPR spectrum, and these off-resonance effects must be considered. Let us again consider the behaviour of magnetisation following a $\pi/2$ pulse, after which point net M_0 lies along the -y-axis (in the rotating frame, this magnetisation appears to be static for spins which are exactly on-resonance). However if the frequency of precession by some spin packets (ω) is not equal to the microwave frequency (ω_0), then the magnetisation associated with these spin packets will no longer appear stationary, and will instead appear to rotate within the x-y plane (see Figure 1.8 A).^{1,3,13} This frequency difference ($\Delta\omega$) is then equal to:

$$\Delta\omega = \omega - \omega_0 \quad (1.19)$$

Magnetisation of off-resonance spins will rotate in either direction about the z-axis, depending on the sign of $\Delta\omega$. Thus all the individual frequency components of the EPR spectrum are represented by rotating magnetisation components in the x-y plane at specific frequencies, $\Delta\omega$. This is the means by which the EPR spectrum is encoded in the FID, since measurement of magnetisation in the x-y plane (transverse magnetisation) affords all frequency components to be extracted, and hence the construction of the EPR signal.

Since off-resonance effects result in transverse precession of magnetisation, the effect of B_0 is no longer zero and must again be considered. The consequence of this is that magnetisation is tipped by the vector sum of B_1 and B_0 , which is also known as the effective magnetic field (B_{eff} , see Figure 1.8B). As such, magnetisation cannot be tipped as effectively into the x-y plane since B_{eff} does not lie exactly along B_1 , and precession of the magnetisation takes on the motion of a cone (as oppose to moving in an arc as it does on-resonance). As $\Delta\omega$ deviates further from ω (in either 'direction'), the efficiency with which magnetisation is tipped into the x-y plane decreases, resulting in a decrease in the FID. For particularly broad EPR spectra in which there are many contributing frequencies ($\Delta\omega > \omega_1$), this poses a more significant problem. FID intensity is optimised by maximising ω_1 , or alternatively by minimising the pulse length.^{1,3,13}

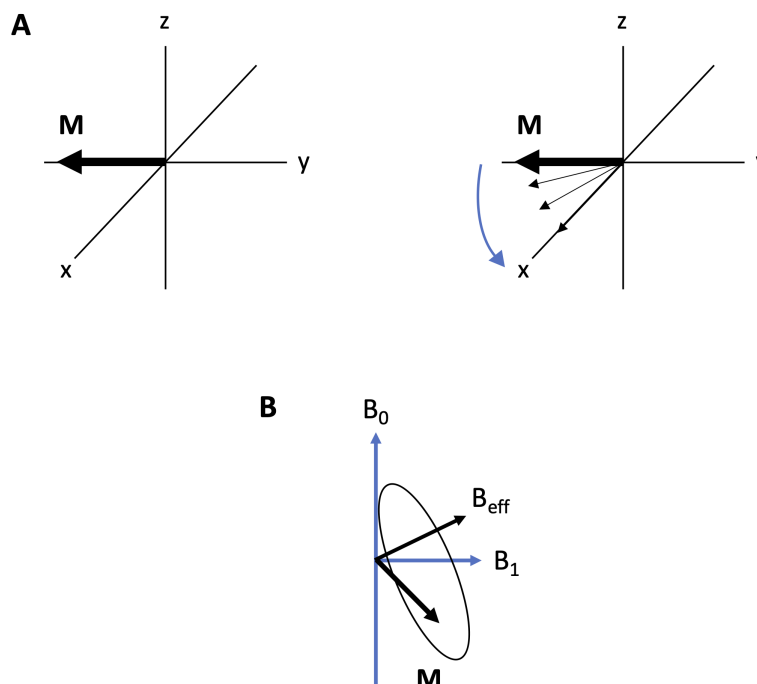


Figure 1.8: (A) Magnetisation in the rotating frame when exactly on-resonance (left) and when $\Delta\omega$ off-resonance. (B) The effective magnetic field produced by combination of B_0 and B_{eff} when off-resonance, and the tipping of magnetisation vector, M , about B_{eff} . The cone-like precession of magnetisation is also shown.

1.4.3 Electron spin relaxation processes

Thus far we have described excitation of spins via application of microwave pulses which results in tipping of magnetisation into the transverse (x-y) plane, but we have not considered the relaxation processes by which the net magnetisation returns (relaxes) to its equilibrium state. Relaxation occurs via different mechanisms since electrons are constantly interacting with their surroundings. Relaxation is characterised by two constants, namely the spin-lattice relaxation time (T_1) and the transverse (or spin-spin) relaxation time (T_2). The former defines the time it takes for net magnetisation to be restored along the z-axis, whilst the latter describes the time for transverse magnetisation (in the x-y plane) to disappear.

T_1 relaxation and the inversion recovery experiment

T_1 relaxation occurs between excited spins and surrounding molecules or atoms (the lattice) via non-radiative mechanisms including molecular rotations and thermal vibrations. We have already established that in the presence of an external magnetic field, electron spins are characterised by two discrete energy states.¹ A marginally larger population of associated

magnetic moments will align parallel to B_0 (along the +z direction) in the lower energy state as a result of the Boltzmann distribution at thermal equilibrium. Relative populations of each energy level can be perturbed (see Figure 1.9) such that spin populations (with magnetic moments in either +z and -z directions) equalise, and thus the net magnetisation (M) along the z-axis becomes zero. This condition in pulsed EPR is achieved through application of a $\pi/2$ pulse, tipping spins into the x-y plane. Spin-lattice relaxation then affords relaxation of spins and restoration of bulk magnetisation along +z. Alternatively, application of a π pulse results in the complete inversion of populations from the equilibrium state, and as the net magnetisation is tipped into the -z direction - this is what is also known as an inversion pulse.

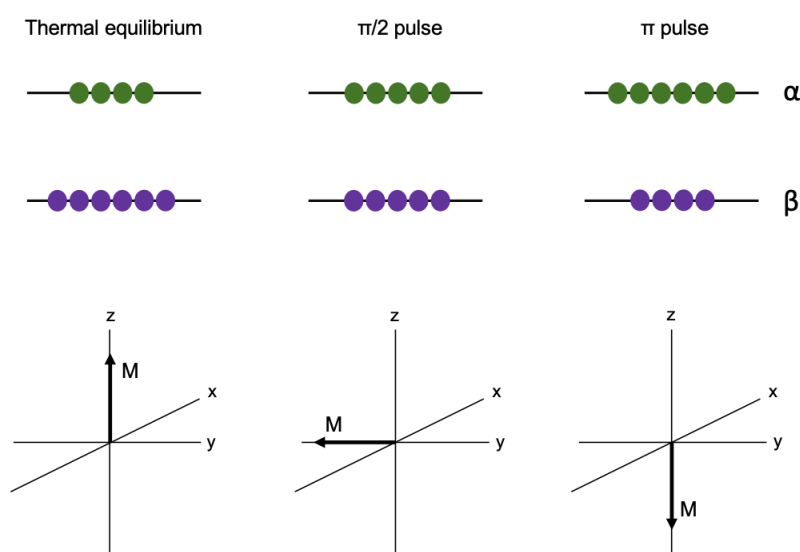


Figure 1.9: (Top) Comparison of spin populations in α and β states, for an $S = 1/2$ spin system in equilibrium (left), following a $\pi/2$ pulse (middle) and following a π inversion pulse (right). (Bottom) The direction of the bulk magnetic moment of the spins in an external field B_0 along the +z direction, for each condition mentioned above. Pulses are applied in the xy plane.

The T_1 relaxation time can be measured using a pulsed EPR technique known as inversion recovery;¹³ this experiment involves the three-pulse sequence shown in Figure 1.10. The first π pulse acts to flip the net magnetisation of spins by 180° from the +z direction to the -z direction (an inversion pulse). The second and third pulses make up the typical detection sequence used in any pulsed EPR measurement: the $\pi/2$ pulse flips spins into the xy plane, after which point gradual de-phasing of magnetisation in the xy plane occurs; the π pulse is then applied to flip the de-phasing spins a further 180° , which reverses the de-phasing and results in a refocused Hahn echo. The time T is the delay time between the inversion pulse and detection sequence, whilst τ is the time between the two pulses of the detection sequence. The spin-lattice relaxation behaviour can be monitored by measurement of the Hahn echo intensity as a function of increasing T , since during this time the net spin magnetisation begins to recover from the -z to the +z direction. The magnetisation vector

behaves according to equation 1.20 following a π inversion pulse, where $M_z(t)$ is the magnetisation along the z-axis at a given time, t . The rate constant at which M_z recovers to thermal equilibrium is the spin-lattice relaxation time, T_1 .¹³

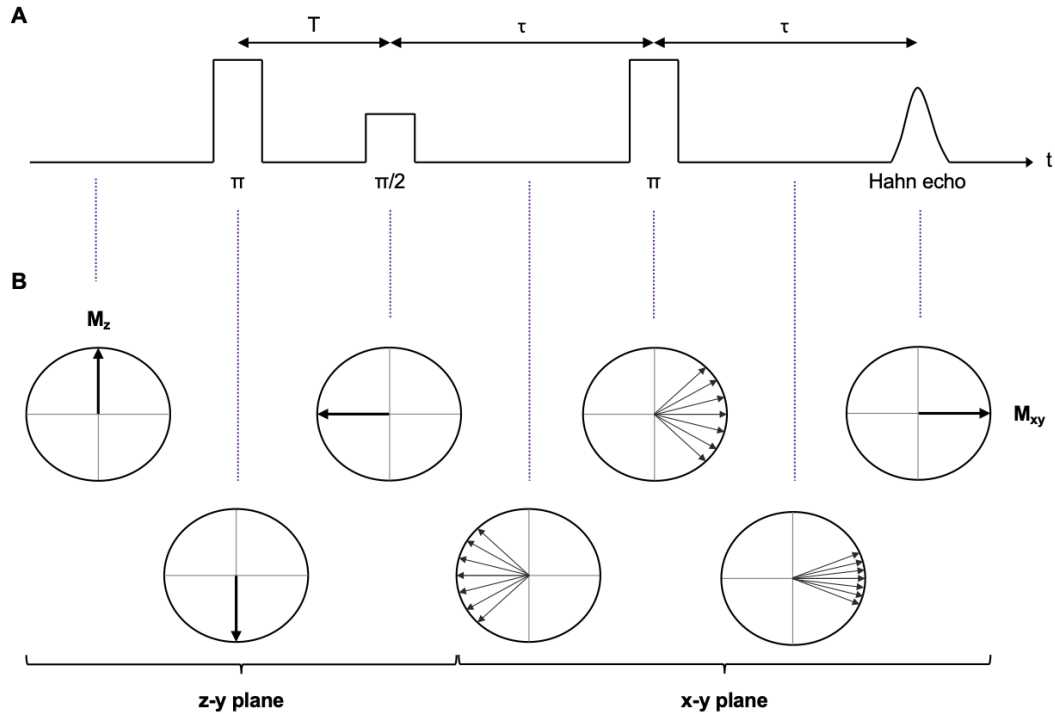


Figure 1.10: (A) The pulse sequence for an inversion recovery experiment, for T_1 relaxation time determination. (B) Schematic of the effect of the pulse sequence in (A) on the magnetisation vector of spins, with viewing planes (i.e. z-y vs x-y) highlighted.

$$M_z(t) = M_0 \cdot [1 - \exp^{-\frac{t}{T_1}}] \quad (1.20)$$

T_2 relaxation and the Hahn echo decay

There are two primary contributions to the process of transverse relaxation which are related to different broadening mechanisms, namely inhomogeneous and homogeneous broadening. Inhomogeneous broadening of spectra arises since spins feel slightly different magnetic fields as a result of unresolved hyperfine structure. The range of local field strengths corresponds to a distribution of $\Delta\omega$ in the rotating frame. Upon taking the sum of spin packet magnetisation within the rotating frame, many components will cancel one another out, decreasing the magnitude of the transverse magnetisation. The shape of this decay (an FID) is actually the shape of the EPR spectrum, and the associated time constant is denoted T_2^* .

Homogeneous broadening occurs when spins experience the same magnetic field, and interact with one another which causes mutual, random spin flip-flops. The result is that magnetisation de-phases in the x-y plane more quickly; the decay rate of transverse magnetisation (following a $+\pi/2$ pulse) due to homogeneous broadening is defined by T_2 (the spin-spin relaxation time) and takes on an exponential form,¹³ according to equation 1.21.

$$M_{-y} = e^{-\frac{t}{T_2}} \quad (1.21)$$

The spin-spin (transverse) relaxation rate can be determined in pulsed EPR via a simple Hahn echo sequence as shown in Figure 1.11. In this case, the Hahn echo intensity is monitored as a function of increasing τ , since the de-phasing described takes place within this time period. This is known as a 2-pulse electron spin echo envelope modulation (2P-ESEEM) experiment. As such we are able to determine the T_2 relaxation time, via fitting of the echo intensity as a function of time with an exponential decay.¹³

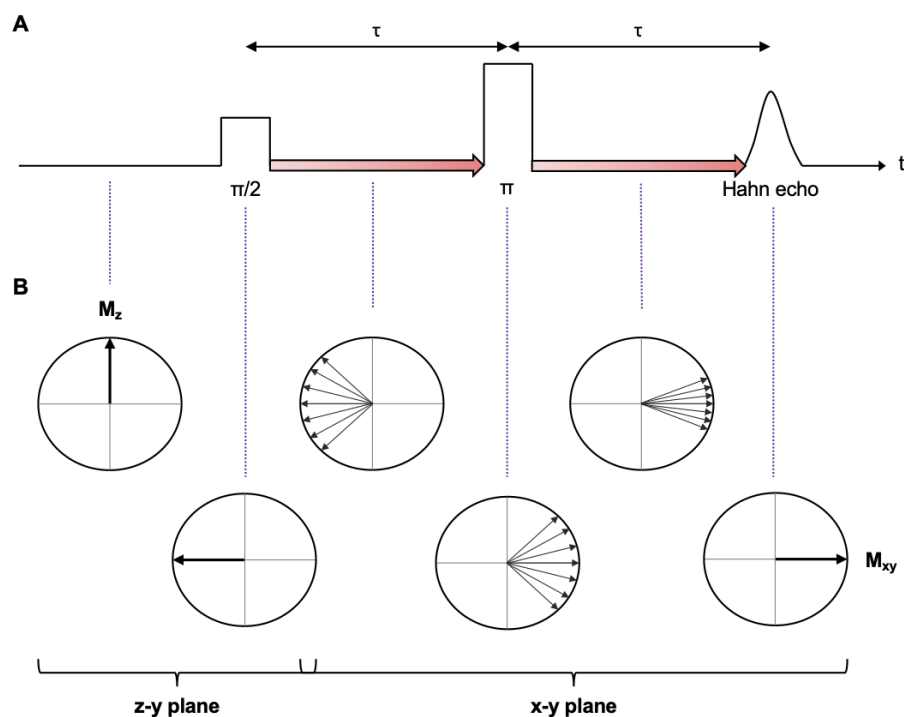


Figure 1.11: (A) The pulse sequence for a decaying Hahn echo, for T_2 relaxation time determination. (B) Schematic of the effect of the pulse sequence in (A) on the magnetisation vector of spins, with viewing planes (i.e. z - y vs x - y) highlighted.

1.4.4 The double electron electron resonance (DEER) experiment

The double electron electron resonance (DEER) experiment is most widely used for the determination of inter-spin distances within macromolecules. This methodology perturbs different spin packets by use of microwave pulses of differing frequencies, suppresses hy-

perfine couplings and selects for the electron-electron dipolar interaction between these spin packets, affording long range distance determination. Typically, distances between approximately 2 - 8 nm are accessible using this technique, however there are some examples of inter-spin distance determination using DEER of up to 17 nm.¹⁴

DEER for probing long range inter-spin distance measurements

The pulsed method of choice used for acquisition of data within this thesis is 4-pulse DEER, which will now be explained. The original DEER experiment was a 3-pulse experiment, however the inherent dead time associated with this pulse sequence lead to development of the dead time-free 4-pulse sequence which is now widely used.¹⁵

The DEER signal measured is the modulation of the refocused Hahn echo amplitude of one spin population (A, in resonance with microwave frequency ν_A), when a second spin population (B, in resonance with microwave frequency ν_B) is perturbed by a microwave pulse. This results in modulation of the echo intensity of A spins as a function of the timing of perturbation of the B spins. This modulation is a result of a local magnetic field from nearby, perturbed, dipolar coupled spins. We may also refer to our 'A' spins as observer spins, and the corresponding frequency as the observe frequency (ν_A), since it is the echo intensity associated with these spins that we are measuring during the DEER experiment. Meanwhile, our 'B' spins are also known as the pump spins, with the corresponding frequency ν_B known as the pump frequency.¹⁵ For the majority of DEER experiments presented and discussed in this thesis, the A and B spins are chemically identical nitroxide radical species - the nitroxide powder EPR spectrum has a width that is large enough to apply microwave pulses at two different frequencies in order to excite separate spin packets without overlap of the excitation profiles. It is this separation that describes our A and B spins.

The pulse sequence of the 4P-DEER experiment is shown in Figure 1.12. It begins by tipping the bulk magnetisation of A spins into the xy plane through application of a $\pi/2$ pulse along the y axis. Due to local field inhomogeneities, including dipolar interactions, spins precess about B_0 at slightly different frequencies and as such de-phasing of the bulk magnetisation in the xy plane begins. A π pulse is then applied to the A spins which flips them by 180° , effectively reversing the de-phasing, meaning spins re-converge resulting in a Hahn echo. This describes the typical Hahn echo pulse sequence. The aim is now to isolate the dipolar contribution of the total magnetic field that is experienced by the A spins. This is achieved through application of a π pulse (referred to as the pump or ELDOR pulse) at

the pump frequency, which causes flipping of B spins by 180° such that their magnetisation lies along the $-z$ axis. This pump is applied at a defined number of linearly spaced time points, relative to the Hahn echo position. Inversion of the pump spins leads to reversal of the dipolar contribution to the net angular rate experienced by the A spins. A second π pulse is then applied at the observe frequency, to once again refocus the de-phasing A spins following the first Hahn echo. However the change in net angular rate of the A spins caused by pumping of dipolar coupled B spins creates a dipole interaction-specific phase lag, meaning dipolar coupled A spins are not refocused. This therefore changes the intensity of the observed refocused Hahn echo measured following the second π pulse. The degree to which coherence is lost is dependent on both the strength of the dipolar interaction, as well as the time at which the pump pulse is applied. Variation of the time point (t) at which the pump pulse is applied relative to the Hahn echo causes changing of the de-phasing angle, resulting in changes to the refocused Hahn echo intensity. Thus the refocused Hahn echo intensity is measured for each time point (t) at which the pump pulse is applied, giving an oscillating decay in which the frequency of oscillation is equivalent to the dipolar interaction.^{15–17}

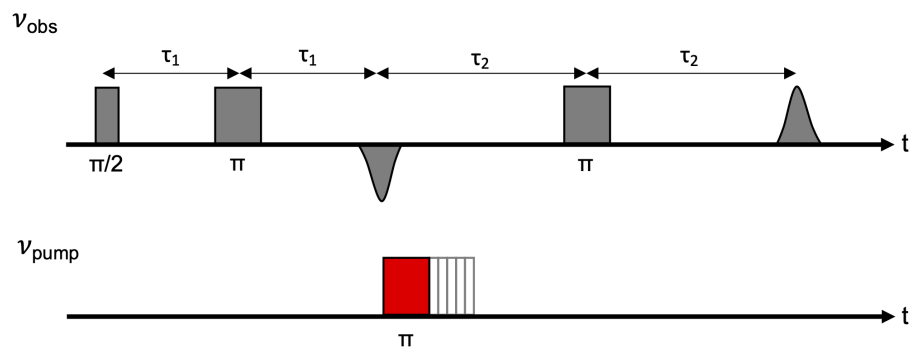


Figure 1.12: The 4-pulse DEER sequence, highlighting observe and detect frequencies and associated pulses.

According to equation 1.22 (repeated from section 1.2.6), the dipolar frequency is inversely proportional to the cube of the inter-spin distance. This means that faster oscillations are caused by shorter inter-spin distances, whilst slower oscillations are a result of longer inter-spin distances. In terms of the experimental set up, longer distances therefore require measurement of the refocused echo intensity over longer values of t (the spin evolution time). Since the re-focusing π pulse at the observe frequency must occur after the inversion pulse at the pump frequency, the maximum distance that can be determined using the DEER experiment is mainly limited by the T_2 relaxation time of the A spins; the Hahn echo cannot be re-focused once magnetisation in the xy plane has been lost due to de-phasing. In order to extend the number and length of time points (in t) over which the DEER trace

can be measured, hence extending the upper limit for distance measurements, experiments are typically undertaken at cryogenic temperatures. Additionally, the choice of solvent can have a pronounced effect on the length of T_2 , for example a fully deuterated sample at low temperature has a much longer transverse relaxation time than a fully protonated one.¹⁷

$$\omega_{dd} = \frac{g_A g_B \mu_B^2 \mu_0}{4\pi\hbar} \cdot \frac{1}{r_{AB}^3} \cdot (1 - 3 \cos^2 \theta) \quad (1.22)$$

As was described previously in section 1.3.4, the powder X-band EPR spectrum of a nitroxide radical like MTSSL is dominated by anisotropic hyperfine coupling due to interaction with the ^{14}N nucleus, which has a nuclear spin $I = 1$. Resolution of only the A_{zz} splitting at X-band frequency results in a spectrum of three lines, with a hyperfine splitting (A_{zz}) of typically around 100 MHz. The spins A and B in a DEER experiment are typically chosen such that their resonance positions are toward the low-field and centre of the nitroxide spectrum, respectively. This might be expected to introduce some orientation dependence to the DEER experiment, however the excitation bandwidths used act to counteract this. Excitation bandwidths of the observe and pump pulses used are small enough that spectral overlap is avoided, but still as large as possible within this limit to achieve the largest modulation depth. This ultimately means that a large range of orientations are actually excited by both the pump and observe pulses.^{17,18}

Interpreting the DEER distance distribution

All distance components contribute. The DEER signal obtained will not only reflect the average inter-spin distance/s within a sample, but also the distribution of distances around the average value/s. For a short and single inter-spin distance with little distribution of inter-spin distances about the average distance, the DEER signal will decay steeply and a single oscillatory frequency will be observed in which oscillations have a pronounced amplitude. The initial slope decay of the DEER signal becomes less steep for longer inter-spin distances, and fitting of this provides a good approximation for the average inter-spin distance within your sample. A DEER trace in which oscillations are unclear, or dampened, is representative of a wide distance distribution; if we imagine a doubly spin labelled macromolecule, wide distance distributions are indicative of conformationally heterogeneous, mobile molecules. In such cases, multiple oscillatory frequencies are in fact contributing to the DEER signal, each with an amplitude related to the relative population of interacting spins at each corresponding distance. This results in a DEER trace that

tends towards being featureless with increasing inhomogeneity, since the various frequency components destructively interfere with one another and cancel each other out. The larger the extent of different inter-spin distance relationships within a sample, the wider the distance distribution will be. For our doubly spin labelled system, the DEER signal contains both intermolecular and intramolecular contributions. However intermolecular interactions include a combination of randomly oriented intermolecular spin-spin relationships and so these result in a featureless exponential decay which can be subtracted for data analysis as a background contribution. This leaves only the intramolecular oscillating decay for calculation of this distance distribution.¹⁵

Determination of the DEER distribution. The most common method for analysis (through direct fitting) of DEER spectra is via Tikhonov regularisation. The time-domain DEER spectrum, also known as the dipolar evolution function, is a combination of what is known as the form factor, and a background factor. The form factor is the contribution from B spins interacting with A spins within the same molecule (the intramolecular contribution), whilst the background factor is the contribution from B spins interacting with A spins of different molecules (the intermolecular contribution). These factors can be separated for analysis of the time domain signal, based firstly on the fact that the form factor decays to a specific value (which is defined as the modulation depth) within the decay time; this decay time is determined by the maximum distance between spins as well as the distance distribution of inter-spin distances. The modulation depth is dependant on the number of interacting spins within the molecule and is calculated relative to the point of maximum echo intensity which is set to 1. The shape of the time-domain spectrum at times longer than the decay time of the form factor is governed by the background factor. In most cases, including those described within this thesis, paramagnetic centres are distributed homogeneously within solution, whilst the T_1 relaxation times of A and B spins can be disregarded, meaning the background factor can be described by a simple exponentially decaying function. In the dead-time free 4-pulse DEER experiment used throughout our experiments, the signal intensity at time 't' ($V(t)$) is normalised to $V(0) = 1$. Phase cycling is also carried out as part of the experiment such that receiver offsets are completely cancelled out, meaning the approximation $V(t \rightarrow \infty) = 0$ is applicable. Overall this knowledge allows fitting and subtraction of an appropriate background function in order to remove contributions of intermolecular interactions.¹⁹

Following this background correction, a simulation of the remaining time-domain form factor

is calculated via Tikhonov regularisation for conversion into the distance distribution, $P(r)$ (see Figure 1.13). This requires identification of an appropriate regularisation parameter that will result in a unique solution, without over-fitting the data. The regularisation parameter involves a weighting between a 'smoothing' parameter, and minimisation of the root mean squared deviation. Without smoothing, resulting distance distributions are very sensitive to signal:noise ratios and as such typically consist of multiple very sharp peaks.¹⁹

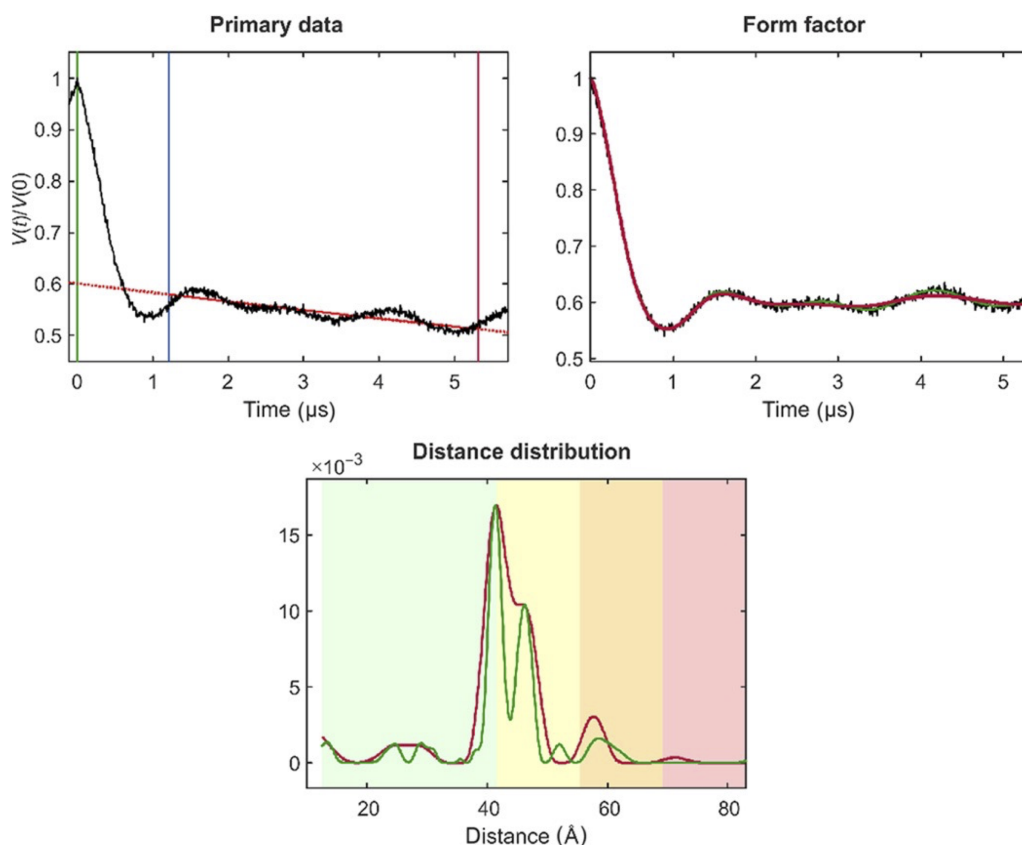


Figure 1.13: An example of DEER data processing via Tikhonov regularisation, showing the primary (raw) data (top left), background-corrected data (the form factor, top right) and calculated distance distributions (bottom) with colour-coded confidence bands. The colour-coded confidence bands represent the following: (a) distance ranges where the shape of the distribution should be reliable (green); (b) where mean distances and widths should be reliable (yellow); (c) where only mean distances should be reliable (orange); and (d) where data should not be interpreted (red). This figure was taken and adapted from Worswick2018.²⁰

A much more recent development in methodologies used for DEER data fitting - DeerNet - employs artificial neural networks to predict inter-spin distances from experimental datasets. These neural networks are trained on an expansive database of synthetic DEER traces computed from randomly generated distance distributions (within realistic limits), including variable baselines and amounts of noise. The ideal outlook is that once these networks have been trained on a database of sufficient size, they should be able to solve some or all of the problems associated with many other fitting routines, including Tikhonov regularisation.²⁰ The use of DeerNet will be demonstrated at points during the presentation and analysis of datasets within this thesis, not only for prediction of average inter-spin distances, but also

to demonstrate the level of uncertainty in these distances in particular cases, which is also predicted (see Figure 1.14 for an example).

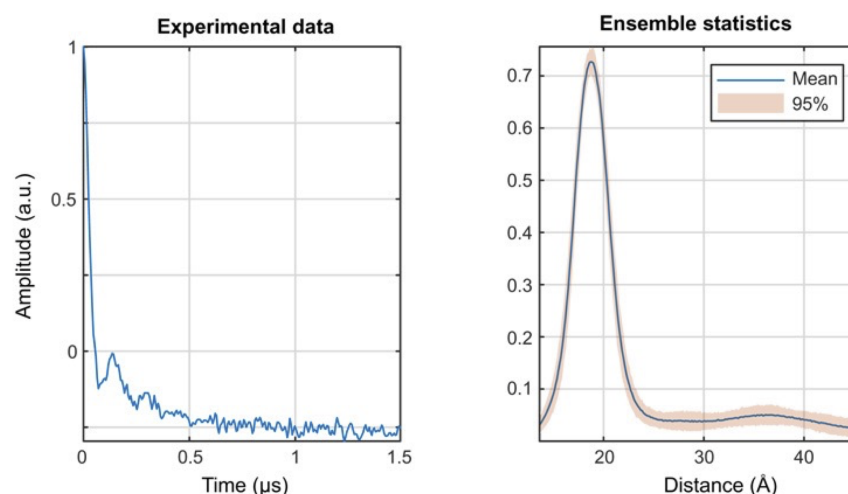


Figure 1.14: An example of DEER data processing using the DeerNet function within DeerAnalysis.¹⁹ (Left) Raw experimental data (blue line) is given alongside (right) the resulting distance distribution (blue line) and 95 % confidence bounds (pink). This figure was taken and adapted from Worswick2018.²⁰

The affects of spin label and protein dynamics on distribution widths. An intrinsic part of protein movements are equilibrium fluctuations which occur on multiple time scales, and with different amplitudes. These fluctuations reflect changes between local energy minima which might be a result of movements of side chains, flexible loop regions, or secondary structures and domains. Hence upon freezing of a sample, sampling of such conformations results in static disorder since it is unlikely that all proteins will be trapped in a conformation reflecting one single energy minimum. Aside from protein fluctuations, the intrinsic width of the distance distribution comes from flexibility of the spin label side chain. For example, MTSSL allows rotation about the linker region, consisting of 4 internal bonds between the nitroxide ring to the protein backbone. Thus, differences in the rotamers adopted by attached MTSSL across different protein molecules can alter the distance between intramolecular SL pairs. This manifests in a wider distance distribution about the average value.

As such, interpretation of the distance distribution requires some consideration of these points, in order to separate the intrinsic contribution from that of protein dynamics. The computer modelling programme MMM was developed for this aim, which utilises a simulated spin label rotamer library. The programme allows rotamers to be inserted into the protein at the position of interest (the labelling site) - these rotamers are evaluated for their interaction energy, taking into account repulsion and van-der-Waals interactions. Populations of each rotamer, and the overall spatial distribution of rotamers are computed. This is modelled by

a cloud of all contributing spin label rotamers in space, taking into account their relative populations. An additional output of this software is a predicted distance distribution from pairs of rotamers, weighted by their relative populations. Thus experimental results can be compared with such models for determination of protein conformations adopted in solution as compared with solved structural models.²¹

DEER for probing conformational change. Oftentimes proteins must undergo specific conformational changes in order to carry out their function. For example membrane proteins involved in the transport of molecules/ions/electrons across membranes tend to work via a number of well-defined mechanisms which involve specific movements of domains. Crystal structures represent a single conformation in space, and can be obtained in different conformations via locking of specific conformational states, however the affects of crystal packing and the conditions of crystallisation can result in protein conditions far from that of the native situation. Results of distance analysis via methods like DEER can thus be mapped onto structural models for enhanced insight into the biologically relevant conformations adopted and sampled within such systems.

Such motions are reflected in DEER datasets directly by changes in the average distance and/or shape and width of the distance distribution between samples. Different conformations might be monitored through addition of specific ions, substrates or inhibitors for determination of mechanistic changes in conformation. Differences in the average inter-spin distance/s report on the amplitude of movement exhibited between different states, whilst the widths of distributions reflect changes in the underlying conformational ensemble. In the most simple case, spin label rotamers will remain unchanged between different protein conformations, since changes in this can result in different widths of the distance distribution as explained previously. This can be checked using RT cw-EPR measurements for samples in different states and subsequent analysis of spectral shapes, relating to spin label mobility. Usually, labelling positions are chosen in highly accessible areas regardless of the protein conformation (again checked via MMM on different structural models of the protein), such that these effects are negligible. Thus changes in widths and average distances are reporting directly on protein conformational changes. This technique has been successfully implemented for a number of membrane proteins, as demonstrated in the following review chapter.

1.5 General NMR theory

Nuclear magnetic resonance (NMR) spectroscopy is a commonly used magnetic resonance technique with a huge range of applications, from chemical analysis to biological studies. Not only does NMR allow for structural determination of proteins in close-to physiological conditions, its application can be extended to probe protein dynamics, kinetics, thermodynamics, and ligand-protein interactions. The technique is therefore routinely used in combination with X-ray crystallography and other structural determination methodologies, however one major limitation of macromolecular NMR is the size of the molecules that can be studied; studies are typically limited to proteins smaller than 35 kDa.²²

For NMR studies of membrane proteins specifically, sample preparation of high enough concentration is a common problem. Not only is expression of membrane proteins at high yield notoriously difficult, but stabilisation of functional protein in a suitable lipid environment for analysis can also lead to losses. The spectroscopic challenges faced when probing membrane proteins with NMR differ, based on whether the chosen method utilises solution NMR, or solid-state NMR (ssNMR). Larger molecules exhibit slower tumbling rates in solution and thus faster NMR signal relaxation rates. Specifically, faster transverse (spin-spin) relaxation results in decreases in signal intensity and increased line widths, hence limiting spectral resolution. For NMR experiments in which complex pulse sequences with long delays are required, the sensitivity of the experiment is then greatly reduced.^{22,23} Additionally, a large molecule equates to an increase in NMR-active nuclei, making the observed spectrum much more complex and difficult to analyse. Therefore, solution NMR experiments of membrane proteins are generally limited to those solubilised in detergent micelles and perhaps smaller bicelles and nanodiscs. In contrast, since ssNMR is not sensitive to the effects of protein tumbling due to size, proteins can be prepared in much more physiologically relevant environments - lipid bilayers. However, the issue of spectral crowding still remains. Regardless of these limitations, NMR is becoming an increasingly popular technique for studying membrane proteins, given its vast applicability even without the need for exogenous labels.^{22,23}

The combination of nuclear charge and spin gives rise to a magnetic moment associated with the nucleus. Nuclei with odd mass numbers have half-integer spin quantum numbers, whilst nuclei with an even mass number and an odd atomic number have integer spin quantum numbers; in both cases the nuclei are NMR active, which relies on the presence of nuclear spin. Nuclei with an even mass number and an even atomic number have spin

quantum numbers equal to zero, and are thus NMR inactive.^{24,25} Magnetic nuclei with nuclear spin quantum number (I) equal to $\frac{1}{2}$ most commonly used for NMR experiments include ^1H , ^{13}C , ^{15}N , ^{19}F and ^{31}P . $I = \frac{1}{2}$ nuclei have two independent spin states, with $m_I = +\frac{1}{2}$ and $m_I = -\frac{1}{2}$ which have degenerate energies in the absence of a magnetic field. This degeneracy, however, is lifted upon application of an external magnetic field (forming α and β states) as is the case with electron spin states in EPR. The splitting between these states is called the nuclear Zeeman splitting. During an NMR experiment, the NMR sample containing NMR active nuclei is placed in a strong, homogenous magnetic field (B_0). This results in alignment of the bulk magnetic moment of each set of NMR-active nuclei with the applied magnetic field along the z-axis. The spin polarisation moves around the applied magnetic field (z-axis) in a motion known as precession, at a frequency known as the Larmor frequency. A radio frequency (RF) pulse (B_1) of specified length and magnetic field strength is applied, perpendicular to B_0 ; this perturbs the nuclear spins such that the bulk magnetic moment flips into the xy (transverse) plane. The nuclei continue to precess about B_0 at characteristic frequencies that differ due to small changes in the local magnetic field as a result of local protein and solvent structure.^{24–26}

Maximum transverse magnetisation is achieved by applying a 90° pulse (Figure 1.15A), which acts to equalise the populations of nuclear spins in α and β states. The precessing magnetisation subsequently induces time-dependent current which is detected by a receiver coil. The signal measured is known as the Free Induction Decay (FID) and is equal to the sum of all nuclear FIDs for all spins within the sample. Due to the differences in precession frequencies (Figure 1.15B) exhibited by each spin, the FID gradually loses coherence and thus the signal intensity diminishes, leading to regained thermal equilibrium. Fourier transform of the FID then enables observation of the spectrum in the frequency domain (Figure 1.15C), affording analysis of three main characteristics: chemical shift (δ), intensity (I) and line width (γ). The position of a signal in the frequency domain spectrum is defined by its chemical shift, and is dependent on the local chemical environment of the nucleus.²⁶

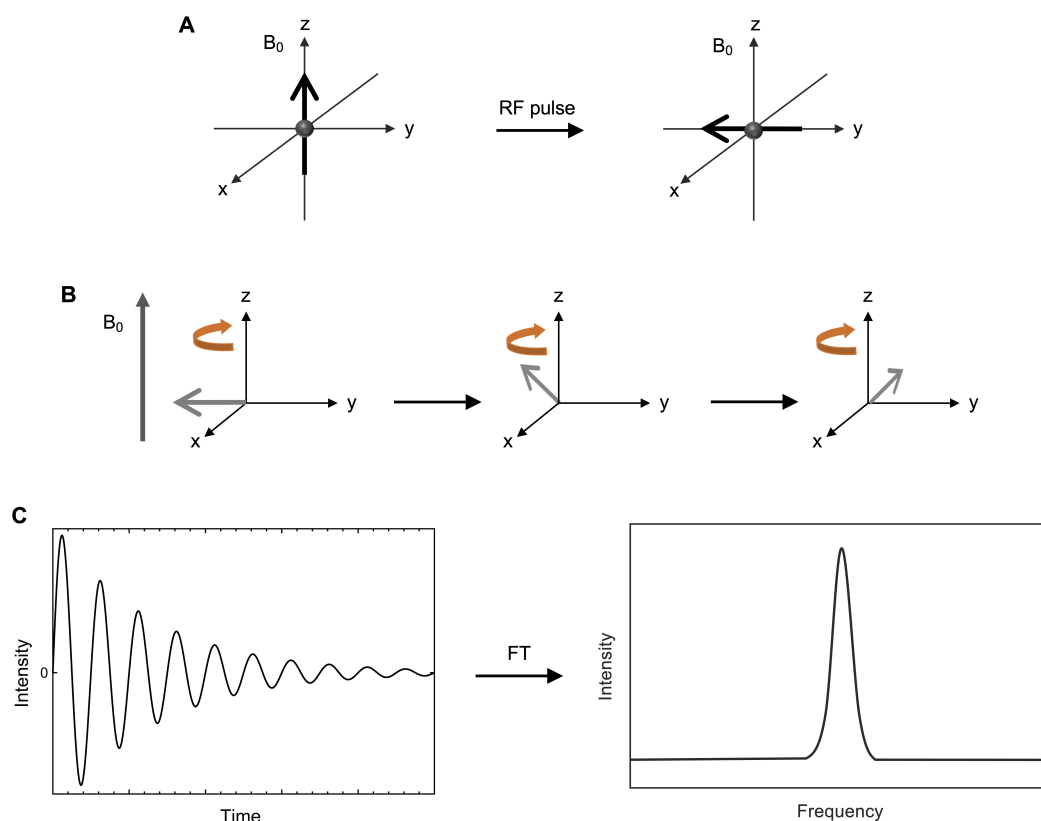


Figure 1.15: (A) The effect of application of a RF pulse at 90° to B_0 on nuclear spin net magnetisation, which flips to the xy plane, as shown by the direction of the bold black arrow. (B) Precession of the transverse magnetisation about B_0 following the RF pulse. (C) The FID recorded for the precessing magnetisation (left), and the frequency domain spectrum following Fourier transform of the FID (right).

1.6 Saturation Transfer Difference (STD) NMR

Recognition processes exhibited by macromolecules, including membrane transport proteins, are of huge scientific interest owing to the fact that such interactions enable and control many essential life processes. This includes the recognition processes of small ligands by macromolecular machines (receptors), which (where there are medical applications) are of great relevance to the rational design of compounds in pharmaceutical research. Thus, characterisation of such processes, both structurally and energetically, at a molecular level plays a key role in this rational design. A number of specific NMR methodologies have emerged as highly powerful and popular tools for structure-based drug discovery, including both protein-based and ligand-based approaches.²⁷ In the early stages of structure-based drug design (SBDD), ligand-based strategies are useful for the simple screening of 'hit' ligands - those which interact with the target molecule. Ligand-based approaches in fact have many favourable features including: (a) the requirement for more simple 1D NMR techniques; (b) the ability to undertake experiments without the need for isotopic labelling of the protein; (c) high sensitivity and fast data acquisition even with low protein concentrations (nM - low μ M); (d) no upper limit on protein size (which is particularly

beneficial for membrane transporters); and (e) protein sample purity is less important providing that any impurities present are not interacting with the protein in such a way that might affect its stability or function.

Saturation transfer difference (STD) NMR spectroscopy is one such ligand-based NMR technique which is routinely used for initial drug screening of many soluble proteins.²⁸ This experiment is based on the Nuclear Overhauser Effect (NOE), which arises from through-space dipole-dipole interactions between nuclei. Consider again an $I = \frac{1}{2}$ nucleus in the presence of a magnetic field, with populations of spins in α and β states corresponding to thermal equilibrium. These populations can of course be perturbed by a RF pulse at a frequency which corresponds to the transition frequency, after which point the spins will relax back to thermal equilibrium. When spins are perturbed by a RF pulse such that the population of spins in α and β states are equal, the condition known as saturation is reached; this condition results in a loss of signal since in order for a transition to occur, a population difference must be present. Dipolar interaction with neighbouring (typically within 0.5 nm) nuclear spins then results in a change in intensity of the resonance of these close-by nuclei, in the phenomenon of NOE.

The NOE builds up slowly in small molecules, and gives rise to an increase in intensity of neighbouring nuclei resonances; in such cases the effect is positive. In large molecules, however, the NOE builds up quickly and is negative, resulting in a decrease in resonance intensity of neighbouring nuclei. In large molecules exhibiting slower tumbling, the NOE can be transferred from one nucleus to another via spin diffusion - this does not occur with smaller molecules.²⁹ In STD experiments, a specific proton resonance (typically far from ligand resonances) of the target protein is selectively saturated by irradiation with a RF pulse, such that direct saturation of ligand resonances is avoided. This saturation then rapidly spreads through the macromolecule via spin diffusion. When a small molecule interacts and binds with the saturated macromolecules, the apparent molecular weight and rotational diffusion properties of the small molecule take on those which are representative of the protein; as such NOE results in saturation transfer to the ligand, from the protein (as shown in Figure 1.16). Additionally, intra-ligand NOE becomes much stronger, affording further saturation throughout the resonances of the ligand.²⁷ The STD experiment is only sensitive for medium-weak binding ligands, since ligand dissociation following saturation transfer is required for 1D measurements of the ligand resonance intensity. Following saturation transfer and ligand dissociation, its resonance intensities are much lower as compared with those in the absence of a saturating RF pulse (measured as a reference spectrum). The

difference in resonance intensities between reference and saturated spectra gives rise to the STD spectrum, in which only resonances of molecules which have interacted in such a way as to receive saturation from the ligand are observed. This experiment requires that the dissociation rate k_{off} is greater than the longitudinal relaxation rate of the ligand, hence it is not sensitive for very high affinity ligands.³⁰

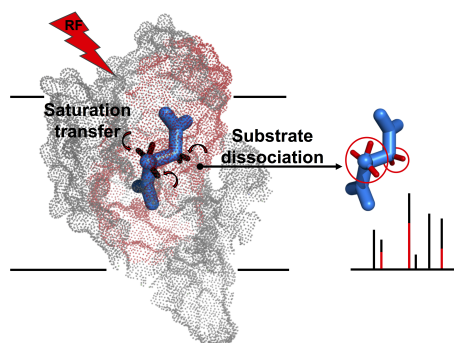


Figure 1.16: Representation of the STD effect, highlighting saturation transfer from protein to ligand protons (symbolised by proton colour change from black to red), dissociation of the ligand from the binding pocket and a resultant decrease in signal intensity.

The efficiency at which saturation is transferred to different resonances of the bound ligand is dependant upon the proximity of these protons to the saturated protein in the binding pocket.³¹ Thus protons in closest proximity receive the highest levels of saturation, resulting in the largest STD signals. This method is therefore useful for the general screening of ligands for protein interaction, and for determination at the atomic level of the ligand binding epitope. The latter is possible since protons of ligands which bind to the protein are easily identified from the STD NMR spectrum as they receive the most saturation, making their signals the most intense.³⁰ Additionally, the STD NMR experiment can be used for K_D determination via ligand titration, as is further explained and demonstrated in Chapter 4 of this thesis.³²

Chapter 2

EPR as an integrative tool for studying membrane transporters

2.1 Summary

In this review chapter, some examples of recent applications of both continuous wave (cw)- and pulsed electron paramagnetic resonance (EPR) methods to membrane transport proteins will be summarised. A particular focus has been placed on the use of specific techniques to probe (a) the local environment of exogenous labels attached to proteins, and (b) measure distances/distance changes as a function of state within doubly-labelled systems. The implications of these findings discussed, giving insights into the structure-function relationships of these membrane proteins. This section will highlight the use of EPR as an essential and integrative tool in understanding the functions and workings of such complex molecular machineries.

2.2 Membrane transport proteins

Biological membranes provide a means for the control of chemical exchange between cells or organelles and their surroundings. Transmembrane proteins specifically are responsible for the transport of many important compounds, including sugars, amino acids, nucleotides, inorganic ions and drugs, both into and out of cells or various cellular organelles (figure 2.1).³³ Transporters and channels allow the movement of hydrophilic compounds which ordinarily would be incapable of diffusion across lipid membranes.³⁴ In most cases, movement of specific molecules across the membrane via a transporter protein is associated with a fixed stoichiometry of molecules and/or ions, whilst the movement of ions or solutes through open channels is controlled by channel gating mechanisms and the single channel conductance.³³

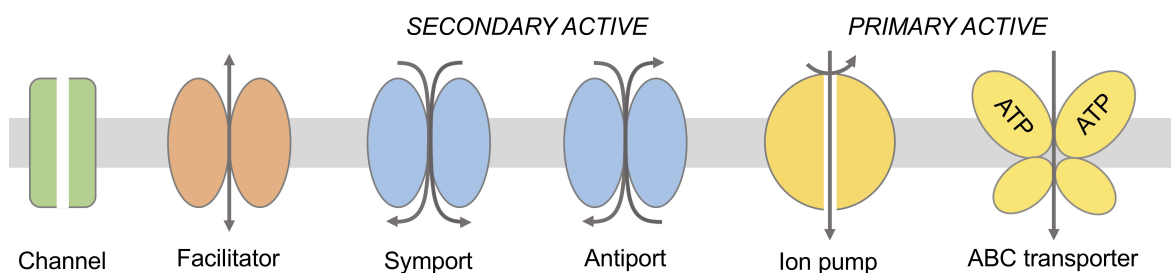


Figure 2.1: Schematic representation of membrane transport proteins including channels (green), facilitators (orange), secondary-active (blue) and primary-active (yellow) transporters. Arrows indicate the direction of substrate or ion movement.

Transporters are further divided based on the energy source used to drive transport, into uniporters (facilitators), primary-active and secondary-active transporters. Facilitators afford passive transport across membranes down the electrochemical gradient of specific solutes, including molecules such as glucose, amino acids and urea. Meanwhile active transporters create substrate gradients across the membrane via the use of various energy-coupling mechanisms. Primary-active transporters harness the energy from metabolic sources such as ATP to afford transport of solutes against their gradients. Examples of such transporters include ATP-binding cassette (ABC) transporters and ion pumps. Ion pumps allow movement of specific ions including Na^+ , K^+ , H^+ , Ca^{2+} and Cu^{2+} out of cells or into cellular organelles, and also generate electrochemical ion gradients. Such ion gradients are typically used during secondary-active transport, in which the movement of ions along their concentration gradient (passively) is coupled to the movement of a second solute against its own concentration gradient.³⁵

Since they afford the movement of a diverse array of solutes both into and out of cells and cellular organelles, membrane transport proteins play a vital role in every living cell and as such, their malfunction is often linked to severe health-related problems. This class of proteins is therefore of huge interest for fundamental research, and indeed approximately 60 % of therapeutic drug targets are membrane proteins.³⁶ Despite their importance and the growing interest in understanding their function, the study of membrane proteins outside of their natural lipid environment proves challenging. Of fundamental importance is being able to solve their three-dimensional structure when trying to elucidate the complex mechanisms of action which govern the structure-function relationships of these proteins. Regardless of the biophysical technique chosen for this purpose, large quantities of high quality, functioning protein are required. Whilst advances are being made in this field, production and subsequent purification of large quantities of membrane proteins is a difficult task, mainly due to (but not limited to) low expression yields, poor solubility and protein instability following extraction from the native lipid bilayer.

Despite the huge number of detergents that have been developed, there remains many draw backs associated with their use in stabilising membrane proteins. For example, they can be denaturing due to the lack of a physiologically-relevant lipid membrane environment. As a result, multiple other membrane mimetics are now used. Liposomes are one of the most commonly used alternatives, which form a spherical lipid bilayer and thus create an environment which is much closer to the native one. However, detergents (which are costly) are still required for reconstitution of proteins into liposomes, which are relatively unstable. Lipid nanodiscs provide another alternative, which consist of a lipid bilayer surrounded by a belt of membrane scaffold proteins. Preparation of nanodiscs typically yields a homogeneous sample, which remains fluid and intact even at very high concentrations. This is in contrast to sample preparation in liposomes, which often results in very turbid and viscous samples. However, once again detergents are required for sample preparation, size control is limited, and nanodiscs tend to give spectroscopic absorbance interference. Polymer nanodiscs (formed using styrene maleic acid (SMA)) have more recently been developed as a membrane mimetic. The SMA polymer is able to directly extract proteins from the native cell membrane environment, and the size of these nanodiscs can be controlled to some extent by changing the polymer:lipid ratio used. A limitation once again is the presence of strong absorbances by SMA in the UV region which is a hindrance for its use in other spectroscopic techniques. Additionally, SMA is unstable at a low pH making it unsuitable for use with proteins that require this.

The typical method of three-dimensional structure determination - X-ray crystallography - requires arrangement of the protein in a regular assembly, meaning only single conformations of these typically very dynamic proteins can be accessed at one time. This leaves many questions surrounding their biologically relevant conformations and dynamics, the answers to which must be provided by alternative means.^{37,38} Another common method used to determine structural and dynamic information is nuclear magnetic resonance (NMR) spectroscopy, however NMR methods are size limited and therefore restrictive.³⁹ Biophysical techniques providing both structural and functional information on such complex molecular machineries, specifically for the reliable and accurate determination of distance measurements and conformational restraints, are limited. Single molecule fluorescence resonance energy transfer (sm-FRET) and electron paramagnetic resonance (EPR) spectroscopies are most routinely used for this purpose, and often in conjunction with one another for direct comparison of results. Both are emerging as key tools in underpinning the mechanisms of complex systems, but for the purpose of this review chapter the main emphasis will be

on applications of EPR. The combination of advanced pulsed EPR techniques, with careful sample preparation in the presence and absence of known substrates/inhibitors/ions to trap complexes in specific states, can afford visualisation of the structure and dynamics of a membrane protein throughout its transport cycle. This can also be applied to protein stabilised in different membranous environments, including *in vivo*, reconstituted protein in proteoliposomes, protein incorporated into nanodiscs as well as styrene maleic acid lipid particles (SMALPs), affording information under conditions much closer to the natural membrane environment than those obtained with detergent.⁴⁰

Previously, EPR studies of proteins were limited to metalloproteins containing an inherent paramagnetic centre, or enzymes with paramagnetic cofactors. For EPR studies of proteins in which an inherent paramagnetic centre is not present, site-directed spin labelling (SDSL) is commonly employed.^{41,42} This process affords selective attachment of typically small, stable paramagnetic probes to specific amino acids of the protein molecule. This is often made possible by site-directed mutagenesis, meaning variants can be carefully designed depending on the scientific question being addressed. In contrast to X-ray crystallography or NMR studies, EPR does not currently allow for the determination of global structural information. However restraints - including accessibility restraints and distance restraints - are now highly sought after as unique contributions of EPR to structural biology.⁴³ The following short review focuses on more recent information gained via EPR techniques on secondary-active membrane transporters, relating to their structure-function relationships, with particular emphasis on the determination of inter-label distances, resolving relative orientations, and measuring distances under physiological conditions and in environments designed to mimic those which are native (i.e. membrane proteins stabilised within a lipid bilayer). This will demonstrate the powerful combination of EPR, and specifically in combination with SDSL, for providing a wealth of information which alongside other biological, biophysical and computational tools, helps to develop our understanding of these complex molecular machineries.

2.3 EPR of secondary active transporters

Secondary active transport involves the facilitated diffusion of ions down their concentration gradient, allowing the uphill movement of a second solute across the membrane, against its own concentration gradient. The coupled transport of ions and a solute can occur either together in the same direction (symport), or in opposite directions (antiport).⁴⁴

2.3.1 Symport

One well studied group of co-transporters is the sodium-solute symporter family (SSSF), which catalyse the uptake of a variety of solutes including sugars, amino acids, vitamins and ions. These proteins use a Na^+ gradient to drive the uphill transport of its solutes across the membrane. PutP, a Na^+ /proline symporter, is found within both bacteria and archaea in which it is responsible for the accumulation of proline as a nutrient or compatible solute under osmotic stress conditions.⁴⁵

The first EPR studies undertaken on PutP⁴⁶ utilised continuous wave (cw)-EPR measurements to probe the possibility of a then-recently proposed 13-helix model of PutP.⁴⁷ Site-directed labelling was employed on select cys-variants, followed by room temperature (RT) cw-EPR studies for information on label mobility and environment with respect to the membrane. The variants designed for study were focused on several transmembrane domains (TMs) and loops. Attached spin labels which were predicted to be buried in the TMs displayed more restriction and lower accessibility as compared to those attached to residues predicted to be exposed on the surface, strongly supporting the proposed 13-helix model. Changes in nitroxide mobility observed in 2 of the 16 tested positions suggested that conformational change induced by proline-binding affects only small, specific regions - specifically TM11 and loop 2 - of the protein. Following this, a combination of cw-and pulsed EPR methods were applied for the determination of inter spin distances in doubly spin labelled variants of PutP *E. coli*, affording insights into the unknown three-dimensional structure.⁴⁸ An inter-spin distance of 4.8 nm between residues in loops 4 and 7, measured using double electron-electron resonance (DEER) spectroscopy, strongly supported predictions that these loops are situated on different sides of the membrane. Shorter average inter spin distances were additionally found between labels attached to residues in loops 2 and 4 as well as loops 2 and 6, and broad distance distributions implied high flexibility in the loop regions studied. Finally, Na^+ binding was shown to induce conformational change by a change in distance between labels at positions 37 and 187. This study demonstrated for the first time distances between amino acids in PutP, and how these distances change in the presence of Na^+ within its tertiary structure. The use of DEER spectroscopy was later extended to *ab initio* structure determination, specifically the backbone structure of TM domain IX of PutP from *E. coli*.⁴⁹ In order to accomplish this, DEER measurements were undertaken on 16 site-directed doubly spin labelled PutP variants, reconstituted into

proteoliposomes. When modelled, the resulting distance distributions revealed a kink in this TM domain (recently proposed to be involved in the translocation pathway) close to residue 341 involved in substrate binding; this kink was predicted to act as a hinge, affording the opening and closing of an inwardly-facing cavity.

An example of the integrative nature of EPR and in particular DEER spectroscopy was demonstrated by the formation of a computational model of the helix bundle of PutP, determined using a combination of experimentally-determined distances (DEER) and template restraints.⁵⁰ cw-EPR experiments were also conducted on spin labelled variants for the analysis of spin label mobility and polarity of the surrounding environment, giving insight into secondary structure elements as well as rearrangements of the external loop (eL) 9 in the absence and presence of Na⁺ and/or L-proline. Results of this study supported the idea that eL9 is involved in controlling access to ion and substrate binding sites. This model was later advanced following further DEER experiments, utilising the resulting distance distributions between more site-directed pairs of labelled residues.⁵¹ The role of a second loop, eL4, was also probed and all 31 positions of eL4 in PutP - residues 294-324 - were targeted by cysteine-scanning mutagenesis.⁵² Following this, site-directed spin labelling and EPR experiments were conducted at each position. For inter-spin distance determination double cysteine variants were subjected to SDSL before DEER measurements were carried out. A combination of information on spin label mobility, polarity and accessibility of sites for paramagnetic quenching agents, along with inter-spin distances of doubly-labelled variants, afforded structural information about eL4 and provided evidence for structural changes upon addition of Na⁺ and/or L-proline. It was found that the structure of eL4 in PutP is composed of two helical segments - eL4a and eL4b - and results suggested that eL4 plays a role in the transmission of mechanical force during the transition between outward-open to inward-open conformations in this transporter, and those related.

A crystal structure of vSGLT (*Vibrio parahaemolyticus* sodium/galactose symporter) of the SSSF gave the first indication that SSSF proteins have a core architecture similar to that of LeuT, a leucine transporter from the unrelated neurotransmitter sodium symporter (NSS) family.⁵³ NSS membrane proteins are hugely important for transmitter uptake at the chemical synapse, utilising Na⁺ electrochemical gradients for the uphill transport of neurotransmitters. Dysfunction of the NSS proteins has been linked to diseases including depression, schizophrenia, epilepsy and Parkinson's disease. These proteins are both targets for specific therapeutic drugs, as well as addictive substances including cocaine and amphetamines.⁵⁴

EPR studies have been undertaken on LeuT by a single research group, and together with other crystallographic and biophysical studies have aided in the understanding of LeuT's conformational dynamics. Initially, 41 single and 6 double cysteine variants were produced and spin labelled, and EPR studies were carried out on protein reconstituted into proteoliposomes.⁵⁵ It was found that Na^+ -binding induced a dynamic outward-facing conformational state, largely mediated by extracellular loop (EL) 4. Leucine binding caused both local and global conformational changes of a different nature to those exhibited by inhibitor-bound LeuT. These results suggested that crystal structures available at that time were representative of inhibited conformations of LeuT. In a second study, distance measurements were used to understand the conformational changes that occur on the extracellular and intracellular sides of the protein, induced by Na^+ and leucine binding.⁵⁶ Motifs studied at both ends sampled open and closed conformations when the protein was in the apo state, whilst addition of Na^+ ions/substrate shifted the conformational equilibria indicating a key role of the studied motifs in substrate binding and translocation. A key finding of the study was the role of TM6 and TM7 specifically in the coupled movements of extracellular and intracellular sides of the protein, the nature of which ensures that at any one time only one side of the transporter is open. From this, a model of LeuT transport was predicted by the authors (Figure 2.2).

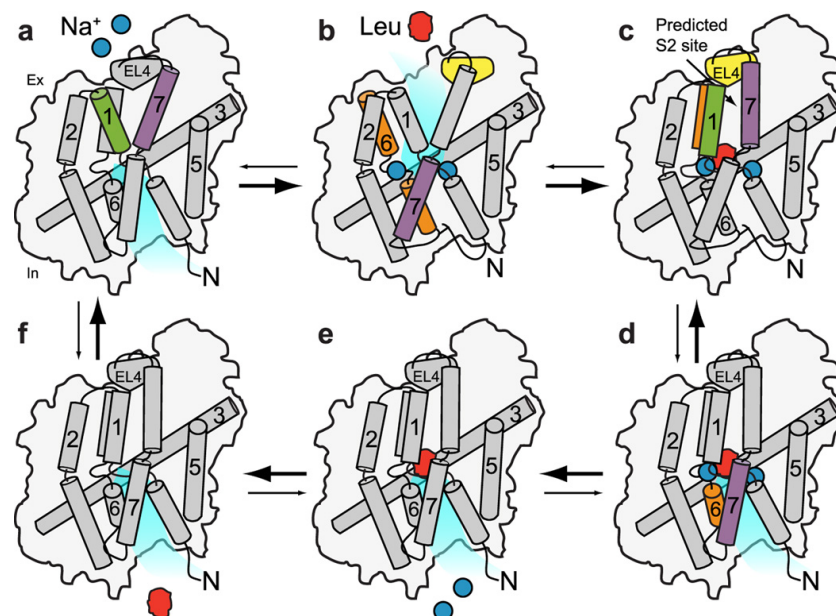


Figure 2.2: Predicted mechanism of transport by LeuT. This figure was taken from Kazmier2014.⁵⁶

Mhp1, a Na^+ /benzylhydantoin symporter from *Microbacterium liquefaciens*, is a secondary active transporter of the nucleobase-cation-symport-1 (NCS1) family. Within this family there are at least 800 known homologues across multiple domains, which between them transport

a multitude of substrates.^{57,58} The crystal structure of Mhp1 again displays similarities to members of the SSSF, and to LeuT, despite sharing relatively low sequence identity. Each of these proteins is believed to operate via an alternating access mechanism. There is a single EPR study published on Mhp1, which employed SDSL with DEER spectroscopy to probe the conformational changes associated with this alternating access mechanism. They found that in the apo state, the protein samples outward- and inward-facing conformations in an equilibrium. Whilst Na^+ binding was not found to affect this conformational equilibrium (unlike in the case of LeuT), it does increase the affinity of the protein for its substrate, the binding of which induces large-scale conformational movement.⁵⁹ A model for substrate/ion binding and transport was again predicted based on the results of these studies (Figure 2.3).

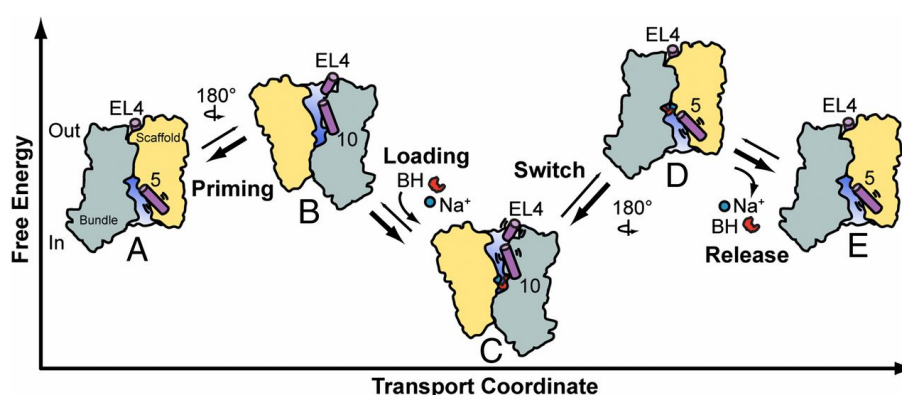


Figure 2.3: Predicted mechanism of transport by Mhp1. This figure was taken from Kazmier2014.⁵⁹

BetP of *Corynebacterium glutamicum* is a well characterised Na^+ /glycine betaine transporter belonging to the betaine/choline/carnitine transporter (BCCT) family; BetP also contains the LeuT-like fold at its core according to crystal structures.^{60,61} An initial EPR study of BetP employed SDSL, cw-EPR and DEER studies to probe C-terminal domain dynamics, aiding in the understanding of its importance in the osmosensory and osmoregulatory properties that BetP displays. Analysis of RT cw-EPR spectra for C-terminal cysteine variants led to the conclusion that reconstitution affords an increased interaction of the CTD with adjacent structures - these could be other domains of BetP itself, or alternatively the surrounding lipid bilayer. Activation of the transporter by osmotic stress appeared to have little consequence to the spin label mobility, whilst distance determination via DEER experiments suggested that the core structure of the trimer in liposomes is distinctly different when compared with that observed in solved crystal structures.⁶² More recently, distance restraints as a result of DEER measurements of double cysteine variants (2 cysteines per protomer) were used in combination with ensemble-based metadynamics simulations. This afforded the testing of the relevance of each conformational state of the protein (in proteoliposomes) as inferred from DEER studies, since some did not agree with the confor-

mations exhibited in crystal structures of BetP, as is often the case.⁶³ This is yet another example of how the results of EPR studies can be used in combination with other techniques and principles to underpin the structure, function and dynamics of these machineries.

GalP (galactose permease) couples proton and monosaccharide transport and is a member of the major facilitator superfamily (MFS). The MFS contains membrane proteins which transport a diverse array of substrates including ions, drugs, nucleosides, amino acids, neurotransmitters, peptides and sugars.⁶⁴ GalP itself is the closest homologue from *Escherichia coli* to the human glucose transporter GLUT1; GLUT proteins are essential for the maintenance of glucose homeostasis throughout the body, with their dysfunction being linked to a number of diseases including type 2 diabetes.³⁸ Another symporter belonging to this family, whose structure and function has also been investigated using EPR methods is LacY, the lactose permease of *Escherichia coli* which couples proton and galactopyranoside translocation across the membrane. This monomeric protein consists of twelve transmembrane α -helices arranged in two 6-helix bundles with an interior hydrophilic cavity,⁶⁵ and the combination of various studies indicates that LacY operates via an alternating access mechanism.⁶⁶

Solved three dimensional structures of GalP, in contrast to LacY, were not available at the time of published EPR studies on this system, however predicted models suggested a tightly packed assembly consisting of 12 membrane-spanning α -helices. The first EPR study of GalP utilised SDSL on protein overexpressed in plasma membrane vesicles, targeting residue cys³⁷⁴ which was already known to block activity of the protein upon any chemical modification, including labelling. Three differing spin labels were used to investigate the rotational diffusion rate of the protein in its natural membrane environment. Spin labelling results gave an indication that residue cys³⁷⁴ is close to residues of the sugar binding site, whilst label mobility and saturation transfer EPR techniques suggested that the label is highly constrained and therefore buried. Additionally it was predicted that the proposed 12-transmembrane helix assembly was in fact less tightly packed than models at the time suggested.⁶⁷ A second EPR study later probed lipid-protein interactions of the protein in vesicles of the inner membrane. In this case, spin labelling of the lipids (rather than the protein) was employed yielding GalP protein in a surrounding environment of labelled lipids. The resulting cw-EPR measurements revealed highly restricted spin label motion; this was in agreement with the prediction that GalP has a largely hydrophobic surface which will interact with the inner membrane.⁶⁸

A single EPR study was undertaken on LacY in which DEER measurements were used to probe the proposed alternating access model, and specifically inter-helical distance changes between apo and sugar-bound states. Cysteine pairs were introduced to trans-membrane helices on either the cytoplasmic or periplasmic sides, for both WT LacY as well as a conformationally locked variant. Distance changes of opposite sign between cytoplasmic vs periplasmic pairs for WT LacY inferred the simultaneous closure of the inward-facing cytoplasmic cavity and opening of a cleft on the periplasmic side. In the conformationally locked variant however, whilst there appeared to be closure on the cytoplasmic side, little conformational change was observed on the periplasmic side upon sugar-binding. These results were then utilised in computational simulations to develop a model for LacY sugar-induced conformational change.⁶⁹

Note that a review on the multiple published EPR studies of GltPh (a secondary active symporter) has been purposefully left out of this chapter, and is instead discussed in section 4.4.3, since this results chapter presents multiple spectroscopic studies of GltPh.

2.3.2 Antiport

The Na^+/H^+ antiporters are a family of membrane transport proteins responsible for the maintenance of both pH and Na^+ concentration within cells, and are found throughout all biological kingdoms^{70,71}. When the concentration of either of these ions becomes too high they become potent stressors, affecting the correct functioning of the cell and its proteins. NhaA of *Escherichia coli* is one of the most extensively studied Na^+/H^+ antiporters; NhaA was first shown to exist as a dimer, with monomers of 12 transmembrane segments, in 2D crystals. By utilising the electrochemical H^+ gradient across the membrane (i.e. in a pH-dependent manner), NhaA catalyses the exchange of H^+ for Na^+ across the cell membrane. The first EPR study on NhaA utilised cw-EPR and DEER measurements to investigate both the local environment of the attached label, as well as the biological assembly of the protein. Residues were chosen for spin labelling based on their positions in domains proposed to undergo pH-dependent structural changes, in the context of proteoliposomes.⁷² cw-EPR measurements revealed pH-dependent changes in spin label environment for one spin label position, indicating conformational change locally. Accessibility measurements using water or lipid soluble paramagnetic quenchers revealed that spin label side chains in both positions studied were in apolar environments, most likely interacting with the lipid bilayer. DEER measurements on each singly-labelled variant also gave a clear indication that in its

biologically active state NhaA does indeed exist as a dimer, corroborating the results of the 2D crystal data mentioned previously. DEER measurements were later extended to multiple other single cys-variants of NhaA in proteoliposomes, for an enhanced understanding of the physiologically relevant protein structure.⁷³ Results were compared with previously solved 2D and 3D crystal structures to further inform on NhaA's dimeric nature, including points of contact, pH dependence and any resulting conformational change.

The plasma membrane Na^+/H^+ exchanger isoform 1 (NHE1) is a human Na^+/H^+ antiporter similar in structure and function to NhaA.⁷⁴ A structural model of NHE1's transmembrane domain was produced based upon the solved crystal structure of NhaA, and EPR investigations were used to inform on the importance of TM4 and TM10-11 in inhibitor binding. Inter-spin distances between the labelled residues A173C (TMIV) and I461C (TMXI) were measured under apo vs inhibitor bound conditions, as well as in an environment of lower pH. Both the binding of cariporide, an inhibitor of NHE1, and acidic conditions rendered shorter inter-spin distances than in the apo case.

Antiporters of the MFS family whose structure and dynamics have been investigated using EPR spectroscopy include OxIT of *Oxalobacter formigenes* (an oxalate/formate exchanger)⁷⁵, LmrP of *Lactococcus lactis* (a multidrug transporter)⁷⁶ and EmrD of *Escherichia coli* (a drug/ H^+ antiporter).^{38,77,78} As described previously, members of the MFS share a common fold consisting of both N-terminal and C-terminal domains, each containing six transmembrane segments, with a cytoplasmic loop connecting the two. The gram-negative anaerobic bacterium *Oxalobacter formigenes* plays a role in oxalate degradation in the large intestine of humans. OxIT itself is present in the inner membrane of *O. formigenes*, in which it facilitates exchange of oxalate and formate. A study in 2014 utilised EPR spectroscopy to test specific hypotheses based on a previously published, experimentally tested GltP-based homology model of OxIT.⁷⁹ SDSL in combination with cw-EPR demonstrated that spin labels in sites S143C and S336C are separated by 8 Å under substrate-bound conditions. Both of these positions lie within GxxxG-like motifs which are known for their involvement in helix-helix interactions, and this experimentally-determined distance demonstrates that TM5 and TM11 come into close proximity during substrate translocation. DEER measurements for longer range inter-spin distance determination gave evidence for both inward-open and outward-open conformations in the presence of substrate.

LmrP utilises the downhill movement of protons for the expulsion of a variety of substrates,

including antibiotics, anticancer drugs and other cytotoxic compounds,⁷⁶ and for this reason LmrP has been implicated in multidrug resistance in cells. An initial EPR study involved the use of singly- and doubly-labelled variants for probing transport-induced conformational change. Results revealed distinct substrate-induced conformational rearrangements of helix VIII, implying that access of drugs to the binding pocket is likely controlled (at least in part) by the interface between N- and C-terminal halves of this MFS protein. Conformational change was also observed upon selective protonation of key residues.⁸⁰ DEER measurements between different TMs were later used to investigate the conformational states sampled by LmrP, any conformational equilibrium exhibited in the apo state and how changes in pH or addition of substrate might displace this equilibrium. A key finding was that LmrP exhibited and sampled at least two conformations, outward-open and outward-closed, and that the transition between these two states is controlled and driven by a change in the protonation state of specific residues. From this it was suggested that the conformational changes required for transport are afforded by the presence of a proton gradient which occurs via a proton relay between specific residues. A transport cycle was proposed based on these extensive DEER experiments in which substrate binding stabilises the outward-open conformation, through which the binding site is exposed to extracellular protons. Following this, protonation of key residues which coordinate the substrate allows substrate release to the extracellular side. This then results in further conformational changes, affording proton movement and release to the intracellular side.⁸¹

EmrD was the first drug/H⁺ antiporter of the MFS to be structurally characterised.^{38,77,78} Its crystal structure revealed a general structure similar to other members of the MFS, however in this case the large internal cavity was occluded to both the cytoplasmic and periplasmic sides,⁸² a state believed to be unstable and intermediary. SDSL in combination with EPR was employed to probe the conformations exhibited by EmrD in a native-like environment, reconstituted into proteoliposomes. 76 positions across six of the twelve transmembrane helices were identified for SDSL, and the mobility and accessibility of each attached label was analysed. Whilst much of the EPR data did in fact align with that of the crystal structure, some deviations in terms of the conformations in specific regions were apparent; specifically the orientation of TM5 and topologies of TMs 10 and 11. pH-dependent conformational change was also observed, in which a low pH resulted in opening of the cytoplasmic side based on an increase in mobility and accessibility of attached labels in the cleft formed by TMs 2-4. At the same time, a decrease in label mobility was observed at a residue facing the periplasm on TM6, indicative of simultaneous closure on the periplasmic side. Both *in vitro* and cell-based functional assays were attempted, however any significant multidrug

transporter activity was not reproducible, posing doubts as to the suitability of EmrD as a model for other MFS drug/H⁺ antiporters.⁷⁸

NorM is a membrane transporter belonging to the multidrug and toxic compound extrusion (MATE) family, and was the first of its kind to be classified.⁸³ MATE proteins are found throughout all kingdoms of life; in bacteria they play a key role in antimicrobial resistance, in plants they are important for detoxification of metabolic products, and in mammalian kidneys and livers they are essential for the excretion of toxic compounds.⁸⁴ Whilst the majority of MATE proteins act as Na⁺/drug antiporters, a few have been found to utilise H⁺ gradients in addition to, or instead of Na⁺ gradients. In terms of substrate specificity, the MATE proteins transport a wide variety of structurally unrelated molecules, with fluoroquinolones such as norflaxacin being recognised substrates of almost all MATE proteins.⁸⁵ Two X-ray crystal structures of Vc-NorM (of *Vibrio cholerae*) served as a basis for the structural components believed to be conserved throughout the entire class of proteins, which between them share approximately 40 % sequence similarity. These structures revealed an arrangement of two bundles, each containing six transmembrane helices, connected via a cytoplasmic loop, with a large extracellular-facing internal cavity.⁸⁶

An EPR study carried out by Steed *et al.* used a spin-labelled derivative of the NorM substrate daunorubicin, known as Ruboxyl, to investigate the position of substrate binding and translocation in Vc-NorM, with a particular focus on the effects of Na⁺ on this process. cw-EPR line shapes were analysed and compared for free and bound Ruboxyl, the latter of which gave rise to a highly immobilised species indicative of a sterically hindered substrate binding site. Evidence of inhibitors acting to block substrate binding was also apparent from cw-EPR measurements, as was an indication that Ruboxyl binds not only in the high-affinity, main binding site, but also elsewhere in a non-specific fashion. The extent of Ruboxyl binding in the presence and absence of Na⁺ was virtually identical. DEER measurements were conducted to probe this idea of non-specific binding by monitoring echo modulation depth (which is proportional to the number of interacting spins) as a function of protein to Ruboxyl excess. Addition of an inhibitor was shown to greatly reduce this modulation depth, presumably through substrate displacement. Distance measurements were also obtained for singly-labelled NorM with the bound paramagnetic substrate, with protein labelling positions chosen at the ends of helices proposed to surround the putative high-affinity substrate binding site. One single, specific binding site was inferred from these results when using an excess of protein:ligand, and the defined distances measured allowed the location of this binding site to be determined in the periplasmic cleft. Interestingly, the same

measurements conducted in the absence vs presence of Na^+ showed little difference, indicating no change in substrate position as a function of Na^+ . It was concluded that substrate translocation may be dependent on the presence of a Na^+ concentration gradient, which is cannot be achieved with detergent micelles, and thus further investigation was required.⁸⁷

A second EPR study was carried out more recently in which DEER and subsequent inter-spin distance determination on the periplasmic side of Vc-NorM highlighted specific Na^+ , H^+ and substrate-induced conformational states. In the apo state, distance distributions were convoluted with multiple components; such results are typically attributed to conformational fluctuations as was believed to be the case here. However addition of Na^+ appeared to stabilise an outward-facing state, whilst binding of doxorubicin suppressed the longer distance seen with Na^+ , and instead increased the population of shorter distances. It was suggested that Na^+ and doxorubicin therefore stabilise distinct, different conformational states of the transporter. pH-dependent conformational change was also observed, indicating a role for H^+ ions also in the mechanism of action exhibited by NorM, and distinct structural changes of TM1 induced by Na^+ and H^+ were experimentally observed for the first time. Site-directed mutagenesis allowed specific residues to be identified as having a role in H^+ - and Na^+ -driven conformational change, as well as stabilisation of the substrate-bound state. H^+ and doxorubicin binding were found to be mutually exclusive, indicating a role of protons in the release of substrate following translocation.⁸⁸

2.4 Conclusions

Membrane proteins, as demonstrated by the examples used in this review section, are vital for proper cellular function, with their dysfunction linked to a huge number and variety of sometimes life-threatening illnesses. Crucial to our understanding of their functions and mechanisms is knowledge of their structures at an atomic level. Crystallography remains the primary method used for 3-dimensional structure determination of proteins, nucleic acids, protein-nucleic acid complexes and complex molecules with large molecular weights. A huge advantage of this technique is that it is not size-limited, and the level of detail provided would not be possible using spectroscopic techniques. Indeed, the structures provided by crystallography are required as a starting point from which to base many of the spectroscopic experiments conducted on such molecules, as has been demonstrated in this review. However, practically speaking this process is difficult, and the results typically provide just a

snapshot of the possible conformational states exhibited by these proteins, and in conditions far from those which are physiologically true. Knowledge of the conformations sampled is vital if we are to fully and correctly understand how these proteins function, especially when working towards medicinal solutions to their dysfunction. EPR spectroscopy is a method very capable of supplying such information, via analysis of spin label environment, inter label distance determination, and overall protein dynamics. As with any technique, there are some disadvantages to be considered surrounding sample preparation and measurement of EPR data. For example site-directed mutagenesis for the deletion/insertion of amino acids to attach to exogenous spin labels should be carefully considered, in order to prevent alteration to the native 3-dimensional structure or function of the molecules under study. The fast relaxation of unpaired electrons also means that experiments must be conducted at extremely low temperatures, meaning specialist equipment must be used and sample preparation should be done extremely carefully such as not to affect protein function. Data analysis typically involves spectral simulation, for which there may be multiple solutions. However, the combined use of such biophysical studies and crystal structures is extremely informative, and EPR spectroscopy serves as one of the many powerful and integrative tools for the study of macromolecular machineries that govern life throughout all kingdoms.

Chapter 3

Investigating conformational states and binding modes of MamM-CTD using advanced EPR techniques

3.1 Summary

The following chapter presents firstly, an introduction to the cation diffusion facilitator family (CDF) of membrane proteins, including their biological role and general structure. The role of CDF (and other) proteins within magnetotactic bacteria, for the formation of magnetosomes, is then explained with reference to current literature. EPR investigations of the C-terminal domain of MamM - a CDF protein of the magnetotactic bacterium *Magnetospirillum gryphiswaldense* - are presented in which metal binding-induced conformational change is probed, in addition to metal selectivity, binding stoichiometry, relative metal binding affinities and allosteric effects. Results are finally discussed in a wider, biological context in which conclusions are made as to how these studies contribute to our knowledge of the structure-function mechanisms of CDF CTDs.

Here it is noted that specific sections of the results presented within this chapter also contribute to published manuscripts, in which I am a co-author. These manuscripts are given (in full) within the appendices of this thesis; experimental results undertaken by collaborators forming part of these publications, which are discussed in relation to my own EPR investigations, are referenced appropriately throughout the text. Meanwhile, any EPR data which has formed a part of these manuscripts has been edited and reformatted for inclusion within this thesis.

3.2 The cation diffusion facilitator (CDF) protein family

A large number of cellular functions within biological systems are reliant on the presence of heavy divalent metal cations at low cellular levels. Such processes include DNA synthesis, transcription, efficient enzyme activity, intracellular signalling and apoptosis, to name a few.^{89,90} However high cellular concentrations of ions including Zn^{2+} , Cu^{2+} , Ni^{2+} , Mn^{2+} and Fe^{2+} , can quickly lead to cytotoxic conditions, and hence maintenance of these levels is essential for cellular function. This is afforded (in part) by transport systems which have evolved within living organisms specifically for the movement of these cations.⁹¹

3.2.1 The biological role of CDF proteins

Cation diffusion facilitator (CDF) proteins are a family of transporters found in all domains of life, for example in the bacterial cell membrane and vacuolar membranes of plants and yeast, and in the Golgi apparatus of animal cells.⁹² These proteins play a role in the efflux of heavy-metal ions, thereby helping to regulate cellular metal ion homeostasis, affording tolerance to trace elements.⁸⁹ In most cases CDF proteins catalyse the transport of cytoplasmic divalent metal cations through membranes, either into extracellular space or intracellular compartments, by exploiting proton motive force.^{93,94} As a result of phylogenetic analysis, the CDFs were originally split into three major groups, which are dependent upon substrate specificity: (1) Zn^{2+} -CDFs, (2) $\text{Fe}^{2+}/\text{Zn}^{2+}$ transporters and (3) Mn^{2+} -CDFs⁹⁵. The most recent phylogeny however contains 18 clades, of which 13 contain at least one characterised CDF transporter, with the remaining 5 containing only uncharacterised CDFs.⁹⁶

CDF function has been shown in a number of bacteria to influence metal tolerance/resistance and cellular homeostasis through metal ion efflux. It is also believed that bacterial CDFs may be implicated in antibiotic resistance,⁹⁷ whilst MamM and MamB - CDFs found within magnetotactic bacteria (MTB) - contribute to magnetosome formation via the transport of Fe^{2+} .⁹⁸

Human CDF proteins are named the zinc transporters (ZnT), also known as the solute carrier family - SLC30, and in total 10 ZnT proteins have been identified.⁹⁹ These proteins allow and help to control the translocation of Zn^{2+} across cellular membranes for maintenance of appropriate and safe cytosolic Zn^{2+} concentrations, typically within the picomolar range.¹⁰⁰ Of the 10 known ZnT proteins, ZnT1 is proven to reside solely in the CM, where

it carries out Zn^{2+} translocation from the cytoplasm out of the cell into the extracellular space.¹⁰¹ Other ZnT proteins typically reside in the membrane of intracellular organelles for the binding and transport of Zn^{2+} into these specific compartments. Transport has been shown in multiple ZnT proteins to occur via a $\text{M}^{2+}/\text{H}^{+}$ antiport mechanism, utilising proton motive force.^{102–105} Currently there are no solved crystal structures of any mammalian CDF transporters - instead, bacterial homologues provide the bulk of structural information currently known about this protein class.

3.2.2 CDF protein structures

The CDF transporters are believed to share a conserved architecture, consisting of a transmembrane domain (TMD), made up of 6 transmembrane α -helices and a cytoplasmic N-terminus, and a relatively large cytoplasmic C-terminal domain (CTD).⁸⁹ The X-ray structure of the $\text{Zn}^{2+}/\text{Cd}^{2+}$ -transporting protein YiiP in complex with Zn^{2+} from *Escherichia coli*, described as exhibiting an outward-facing state, was the first solved structure of an entire CDF transporter in 2007.¹⁰⁶ This study reveals a homodimeric protein, in which the presence of two bands of aromatic side chains help to envisage the boundaries of the membrane-embedded region of the protein. This structure also shows two CTDs protruding into the cytoplasm, juxtaposed so as to form the major dimerisation interface.¹⁰⁶ The YiiP protein has been described as a Y-shaped protein, in which the TMDs are separated by approximately 20 Å forming the 'arms' of the homodimer, whilst the CTDs which are linked to their corresponding TMD via an intracellular loop form the 'stem'. In each protomer, the CTD adopts a metallochaperone-like fold, and each entire protomer binds four Zn^{2+} ions. An equivalent yet higher resolution structure was later obtained using X-ray crystallography by same group in 2009.¹⁰⁷

Cryoelectron microscopy (cryo-EM) in combination with molecular dynamics was later used to identify an inward-facing state of the closely related YiiP from *Shewanella oneidensis* in what was believed to be Zn^{2+} -free conditions, hence giving further insight into the mechanism of action of these transporters via visualisation of different conformational states.¹⁰⁸ This structure however is of very low resolution, which has more recently been extended from 13 Å to 4.1 Å once again utilising cryo-EM.¹⁰⁹ In this case significant electron density is found in three sites previously shown to house Zn^{2+} ions in X-ray crystal structures, despite the supposed Zn^{2+} -free conditions used. Considering this, Zn^{2+} ions were used in structural refinement revealing coordination sites very similar to those found previously

in X-ray crystal structures. When this structure is compared to that of Zn^{2+} -loaded YiiP in an outward-facing state (Figure 3.1), differences can be identified at an atomic level. These include an observed rocking and twisting of a 4-helix bundle in the TMD, which both houses a Zn^{2+} transport site and controls the accessibility of this site within each monomer. Additionally there is a distinct movement of the membrane-spanning domains; the inward-facing state shows the two in close proximity to one another, whilst they are more separated in the outward-facing state. The significance of this vast 'scissor-like' movement was tested via *in vitro* analysis of cysteine cross-linked variants during Zn^{2+} -transport assays, the results of which indicate that such a movement is not in fact necessary for substrate binding and transport. Instead the observed 'opening' of the TMD arms in the outward-facing state is attributed to the detergent conditions used for structural analysis, which does not limit movement of proteins in the same way a native surrounding membrane would. Site-directed mutagenesis was also carried out in select regions of the CTD, yielding no effect on dimer stability and therefore suggesting that the dimer must gain substantial stability from interactions between the TMDs.

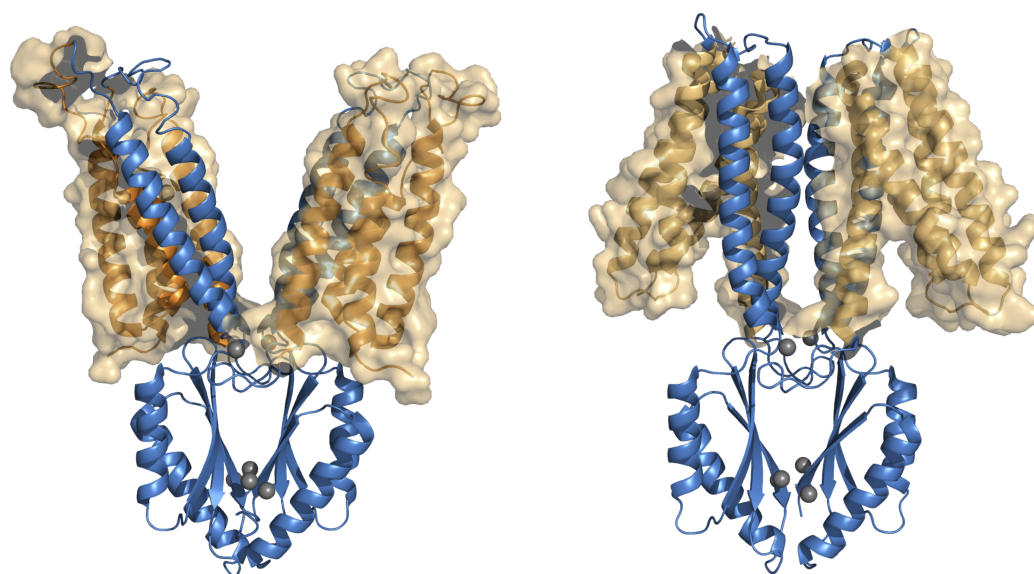


Figure 3.1: Cartoon representation of the outward-facing YiiP homolog from *E. coli* (left) and the inward-facing YiiP protein from *Shewanella oneidensis* (right), with TMs 1, 2, 4 and 5 shown in surface representation highlighting large-scale conformational changes. Grey spheres represent bound Zn^{2+} ions. Figures were produced using Pymol - The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Many other structural studies undertaken on the CDF proteins have been predominantly focused on the cytoplasmic CTDs, of which a number of crystal structures have been solved (see Figure 3.2). This focus is primarily due to the difficulties associated with transmembrane protein purification and crystallisation. Determined CTD crystal structures include the Zn^{2+} -bound and apo forms of CzcB from *Thermus thermophilus*,¹¹⁰ the apo form of TM0876 from *Thermotoga maritima*¹¹¹, MamM and MamB from *Magnetospirillum*

gryphiswaldense MSR-1 (MSR-1).^{112–116}

Apo and Zn²⁺-loaded CTD crystal structures of CzcB were solved by X-ray crystallographic modelling to resolutions of 1.7 Å and 1.8 Å respectively, yielding homodimeric structures. Each protomer of the apoprotein is highly structured, with a polar surface and a hydrophobic core, and made up in total of three β -strands and two α -helices. The dimer adopts a distinct V-shape with the membrane-connecting end splayed apart, similar to the CTDs of published YjiP structures, and with flattened β -sheets facing one another at the dimer interface. The Zn²⁺-bound structure displays only minor changes in protomer secondary and tertiary structure as compared to the apo state, the likes of which are assumed to be important for metal ion binding. The most dramatic of changes involves the arms of the V-shaped CTD, which appear to come closer in proximity along the flat surfaces of the β -sheets, via a hinge-like motion. Overall there is a 18.2 Å decrease in distance between the N-termini of the apo and Zn²⁺-bound forms. It is postulated that only 3 of the 4 Zn²⁺ ions shown in the bound structure are physiologically relevant, with two binding sites showing distinct similarities to two zinc states in the CTD of YjiP.¹¹⁰

Another comparative crystal structure belongs to the CTD of TM0876, from *Thermotoga maritima*. The monomeric structure of this protein is similar to that of the apo and Zn²⁺-bound forms of CzcB-CTD. This protein also adopts a V-shaped form within its dimeric structure, with highly conserved residues thought to be involved in Zn²⁺-binding in CzcB. Interestingly, the opening angle of the two N-terminal arms of the apo CzcB-CTD and TM0876-CTD structures are different, which could indicate that they may in fact be flexible in nature, as oppose to being confined to one specific conformation. This in turn is believed to aid in regulation of ion binding and transport.¹¹¹ Each of these solved CTD structures, including that of the YjiP protein and its homolog, includes the same metallochaperone-like fold, despite the overall distinct lack of sequence homology between the proteins. This type of fold is commonly involved in the trafficking and transport of metal ions from the cytoplasm. This often works through the formation of a metal donor:acceptor interface, with the acceptor being the relevant transport machinery for a particular system.¹¹⁷

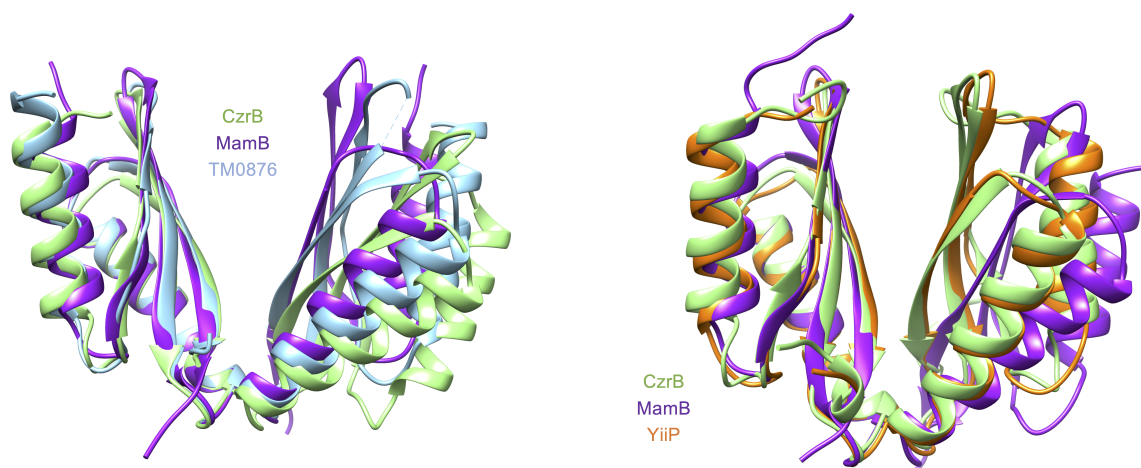


Figure 3.2: Overlay of (left) CzirB, MamB and TM0876 apo-form (PDB codes: 3BYP, 5HO5 and 2ZZT respectively) and (right) CzirB, MamB and YiiP metal-bound (PDB codes: 3BYR, 5HO1 and 3H90 respectively) CTD structures. Structures are shown in cartoon representation, and figures were made using UCSF Chimera (version 1.14).¹¹⁸

X-ray crystal structure determination of the CTD belonging to MamB, another membrane protein of magnetotactic bacteria, was more recently attempted for MSR-1 MamB-CTD with little success. And so instead, structural models based on X-ray diffraction studies for MamB-CTD from another *Alphaproteobacteria* MTB *Magnetospira* sp. QH-2 (QH-2), sharing significant sequence identity with MSR-1 MamB-CTD, were determined in both apo and Zn^{2+} -bound states.¹¹⁶ These demonstrate the presence of the same metallochaperone-like fold found in all other CTD crystal structures described here, with the dimeric protein adopting a V-shaped assembly with each monomer consisting of three β -strands and two α -helices. The apo MamB-CTD structure displays a degree of closure similar to that found for the closed metal-bound YiiP-CTD, which is more compact than in the apo CzirB-CTD structure; the Zn^{2+} -bound structure of MamB-CTD shows little structural and conformational differences to the apo state. Two Zn^{2+} ions are coordinated per monomer, one in a central location and one further towards the top of the dimer in a peripheral site, which considering its unique nature and lack of similarity with metal binding sites of any other CDF CTD, is assumed to be a consequence of non-specific binding.

The broad range of substrates of CDF proteins means that they play a role in a number of important biological pathways. Hence many severe human diseases, including cardiovascular diseases, hepatic cirrhosis, dystonia, Parkinsonism, chronic liver disease and Type-II diabetes have been directly associated with altered regulation of mutations within several human CDF transporters.¹¹⁴ This highlights the crucial role of the CDFs in cellular metal ion homeostasis, and the importance of undertaking research and gaining subsequent insight into their general structure and function. MamM, a protein isolated from *Magnetospirillum gryphiswaldense*, has of late been used as a model system for the study of CDF-related

Type-II diabetes.

3.3 Magnetotactic bacteria and magnetosome formation

Magnetotactic bacteria are a family of gram-negative bacteria which were independently discovered in 1963 and 1975 by Salvatore Bellini^{119,120} and Richard Blakemore¹²¹ respectively, and can be found in marine sediments and aquatic environments. The formation of internal organelle, known as magnetosomes (highlighted in Figure 3.3), within MTB is a unique system which utilises CDF transporters for iron transport and subsequent biomineralisation. It is the formation and chain-like arrangement of these organelle which allows MTB to passively align with and move along the earth's magnetic field to find optimum living and growth conditions.^{122,123} Magnetosomes are MTB-specific organelles, each of which is composed of a magnetic nanocrystal (either mixed-valence magnetite, Fe_3O_4 , or greigite, Fe_3S_4) surrounded by a protein-rich membrane; their sizes and shapes vary between different MTB strains but are conserved within each strain.

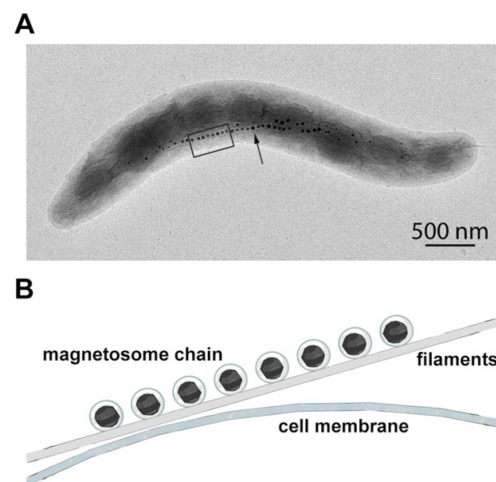


Figure 3.3: (A) Transmission electron microscope (TEM) image of *Magnetospirillum gryphiswaldense* MSR-1, with the magnetosome chain highlighted by the arrow. (B) A zoomed illustration of the black box highlighted in (A), showing magnetosomes surrounded by lipid membranes, aligned as a chain along filaments. This figure is taken from Barber-Zucker2016.¹²⁴

Published research suggests that firstly vesicles of the magnetosome membrane (MM) are formed via invagination of the cytoplasmic membrane, creating a distinct environment, separate to the cytoplasm and its surrounding membrane, for formation of the magnetosome particle. Protein sorting in parallel ensures that the correct, specific membrane proteins required for nanocrystal biomineralisation are present in the MM. Following this, alignment of magnetosomes into chains (requiring actin-like filaments) and magnetic nanocrystal

biomineralisation takes place, and finally magnetic nanocrystal growth results in strain-dependent sized nanocrystals.^{124–126} It is the alignment of several magnetosomes into either one single chain or multiple chains which generates a permanent magnetic dipole moment hence affording rotation of bacterium in the direction of the geomagnetic field and subsequent movement using its flagella, in a behaviour termed magnetotaxis.¹²⁷

There are a number of different strains of MTB - those more widely used as models for understanding magnetosome formation and MTB include the alpha-proteobacteria *Magnetospirillum gryphiswaldense* MSR-1 and *Magnetospirillum magneticum* AMB-1 (AMB-1) strains, with additional information relating to magnetosome diversity and gene organisation coming from studies of the *Magnetococcus marinus* MC-1 and *Magnetovibrio blakemorei* MV-1 strains.¹²⁶ Magnetosome formation is a highly controlled process, in which each identified and individual step is dependent on the presence of a specific set of proteins, mostly encoded for in a conserved set of genes within a segment known as the magnetosome island (MAI). Operons within the MAI include *mamAB*, which is conserved and found in all characterised MTB strains, *mamCDFG*, *mamXY* and *mms6* of MSR-1. Two different types of protein are encoded for by these operons, namely the magnetosome membrane-associated (Mam) and magnetic particle membrane-specific (Mms) proteins. These can be further categorised into individual groups related to the specific role they play in magnetosome formation within MTB. The following overview of magnetosome formation is based upon the proteins and processes known to occur in these aforementioned well characterised strains of MTB.

3.3.1 From membrane invagination to protein sorting and activation

Multiple proteins are believed to be essential for the first step of magnetosome formation, which is invagination of the magnetosome membrane from the cell membrane. It is highly important that the proper chemical environment is achieved, an environment distinct from the cytoplasm, by the formation of this MM. Not only does this provide the correct conditions for nucleation and magnetic nanoparticle formation (including a basic pH and high local iron concentration), but it also protects the remainder of the cell from any harmful by-products of this process.

The MM is a protein rich environment, containing proteins specifically for formation of the correct environment for magnetic nanoparticle growth. However empty magnetosomes

have been found prior to magnetic nanocrystal formation with bacteria cultured under Fe^{2+} -free conditions, and in non-magnetic MTB strains, indicating that MM formation is not dependent on magnetic nanocrystal biomineralisation.¹²⁸ Once Fe^{2+} is introduced to these bacteria, nucleation and nanocrystal biomineralisation then begins within these pre-existing membranes until they reach their final size. The absolute growth of the MM and its final size is therefore likely dependent on the biomineralisation process. Multiple studies on strains including AMB-1 and MSR-1 suggest that the MM originates from the inner cell membrane (CM). The presence of at least a proportion of magnetosomes - both empty and nanocrystal-containing - still attached to and invaginated from the CM suggests that magnetosomes can in fact remain connected to the CM throughout magnetosome formation and afterwards.^{129,130} This attachment to the CM would allow for more efficient alignment of the entire bacterium with respect to the earth's geomagnetic field. That this is not the case in all species, however, suggests some variation in membrane biogenesis.¹³⁰ It has also been suggested that magnetosomes which remain attached to the CM contain a channel which is open to the periplasmic space, hence affording easy exchange of the specific compounds required for biomineralisation. It is though unlikely that such a channel would be completely open, but rather partially occluded, to prevent the influx/efflux of potentially harmful substances between the magnetosome lumen (which could affect nanocrystal formation) and the periplasmic space.¹²⁵



Figure 3.4: Schematic representation of invagination of the inner membrane for vesicle formation in magnetotactic bacteria. This figure has been taken and adapted from Arakaki2015.¹³¹

Studies have probed not only the physical nature of the MM, but also the molecular mechanisms that result in its formation (see schematic in Figure 3.4). Bending of the lipid bilayer in such a way is afforded by the presence of specific proteins which provide the energy required for the process. However, genes encoded for by the MAI show no homology to proteins of eukaryotes previously found to allow membrane bending/movement. A number of studies have demonstrated that numerous Mam proteins are important for membrane invagination - in fact upon deletion of the entire *mamAB* operon (encoding for 17 proteins in total), only the reintroduction of seven genes together including *mamL*, *Q*, *B*, *I*, *E*, *M* and *O* affords proper MM invagination. Specifically, MamL, MamQ, MamM and MamB are the

only proteins found to play important roles in regular MM formation, since deletion of any one of these genes alone affects MM formation in some way. However the reintroduction of just these four proteins upon deletion of the *mamAB* operon without others was insufficient to restore MM formation. Of all the Mam proteins, MamB plays the most important part in MM formation specifically. MamB is a member of the CDF protein family, and so in addition to this role MamB is believed to afford magnetosomal iron uptake. It is important to note that whilst each of the aforementioned proteins is implicated in MM formation, this process occurs only when all are present together, rather than upon reintroduction of just any one of these proteins.¹³²

The requirement of very specific proteins, not only for MM invagination but also for nanoparticle biomineralisation, means that the protein composition of the MM is vastly different and unique compared with the CM. Therefore, protein sorting and activation are necessary via exclusive mechanisms, to ensure the correct proteins are in the necessary positions for optimal formation of magnetosomes.^{124,125,133} Studies of proteomics have previously suggested that whilst Mam and Mms proteins are present in the MM, the extent to which each protein is found varies not only between strains, but in some cases between individual magnetosomes depending on what stage of growth they are at. A recent study utilised mass spectrometry in combination with algorithmic calculations for the determination of the relative abundances of each protein in *Magnetospirillum gryphiswaldense*,¹³⁴ revealing what might be an unusually crowded protein composition within the MM. A number of proteins have thus far been associated with protein sorting and proper protein localisation at the MM; MamA, MamC, MamF and MamI were shown in AMB-1 cells to become mislocalized upon deletion of MamE. MamE itself contains a specific domain commonly believed to be involved in protein-protein interactions - this further supports its suggested functionality in protein sorting.¹³⁵ MamA has also been implicated in protein sorting, mainly due to its proven interaction with Mms6,¹³⁶ as well as the observation that upon its deletion, MamC - one of the most abundant proteins to reside in the MM in MSR-1 - was mislocated.¹³⁷ In general though the exact mechanisms by which the appropriate proteins are targeted to the correct places, both for initial MM invagination and subsequent particle biomineralisation, remain unclear.

3.3.2 Alignment of magnetosomes into chains

Each individual magnetic nanocrystal formed within one magnetosome is limited in size to that of a single magnetic domain, and hence individually magnetosomes are not sensitive to any effects of an external magnetic field. A larger magnetic dipole moment is achieved by the alignment of multiple magnetosomes in a linear chain or indeed multiple chains (see Figure 3.5), thereby allowing the passive alignment of MTB with the earth's geomagnetic field. This chain formation is mediated by a combination of magnetic behaviours with protein-activated assembly mechanisms. MamK is one such protein believed to play a key role in this process, of which there are multiple homologues found throughout different strains of magnetotactic bacteria.¹³³ Multiple studies have revealed the presence of long, linear filaments extending from one cell-pole to the other in MTB.¹²⁵ This actin-like filament is made of polymerised MamK, runs along the cell's inner curvature and is aligned with and attached to the magnetosomes.¹³⁸ Indeed deletion of the *mamK* gene has been shown to affect the efficient and proper formation of long magnetosome chains.¹³⁰

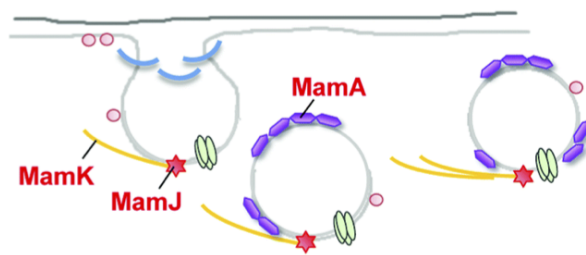


Figure 3.5: Schematic representation of the alignment of vesicles within magnetotactic bacteria. This figure has been taken and adapted from Arakaki2015.¹³¹

Assembly of MamK filaments is believed to be in part controlled by cellular salt concentration due to the largely mixed hydrophobic/neutral nature of the protein surface; polymerisation efficiency of MamK decreases with increasing concentrations of KCl, whilst concentrations of Ca^{2+} and Mg^{2+} higher than just 1 mM inhibit polymerisation altogether.¹³⁹ Studies suggest that following invagination of the MM, short chains of magnetosomes can form independent of the presence of MamK, however it is the formation of much longer chains that is facilitated by the active joining together of these shorter chains by MamK, which also ensures the correct cellular positioning of the magnetosome chain. This gap-closing process is suggested to happen by either (a) moving misaligned magnetosomes into gaps present, or (b) simply pulling existing magnetosomes more closely together, and in doing so creating a much tighter arrangement.¹⁴⁰ MamJ is another protein which has been proven to play a part in magnetosome alignment; *mamJ* deletion mutants of MSR-1 result in detachment of magnetosomes from the filament of MamK.¹³³

3.3.3 Biomineralisation of the magnetic nanocrystal

As mentioned previously, biomineralisation of the magnetic nanocrystal requires a suitable chemical environment. The appropriate chemical environment for magnetic nanocrystal biomineralisation typically requires a high iron concentration and a basic pH. In order to really understand biomineralisation and how it occurs, current literature reports on the proteins involved and the mechanisms by which they carry out their roles, since these biological enzymes are key players in affording magnetosome formation. Specifically, the following nucleation-related steps have been the focus of such studies: movement of iron both into and around the magnetosome; pH control inside the magnetosome; redox reactions of iron; and formation of the precursor ferrihydrite and its transformation to magnetite.

The transfer of iron across the MM and into the magnetosome (see Figure 3.6) is the first step required for biomineralisation of the magnetic nanocrystal, with a vast amount of evidence suggesting that the principle method of iron transport into the magnetosome is via specific transport systems. The two magnetosomal CDF proteins, MamM and MamB, have been associated with iron translocation, and are both found predominantly in the MM. MamM shares 47 % sequence identity with MamB, and either complete suppression of biomineralisation or changes in crystal size and morphology have been demonstrated upon entire deletion of the *mamM* gene, or indeed simply single point mutations.⁹⁸ As previously stated, deletion of *mamB* results in no magnetosome formation, whilst TMD metal binding site mutations have resulted in loss of magnetite formation.⁹⁸ It has also been shown that in MSR-1, MamM is required for the stabilisation of MamB, giving an indication that they may together form not just homodimers but also heterodimers. The presence of just one of these proteins cannot compensate for the absence of the other, indicating that they have distinct roles from one another. As CDF proteins, both MamM and MamB utilise the proton motive force (M^{2+}/H^{+} antiport) and as such this key process is also assumed to have a distinct effect on the internal pH of magnetosomes. By increasing the magnetosome lumen pH through proton expulsion, it is probable that these proteins lend a hand in optimisation of conditions for nucleation and crystal biomineralisation. Two other proteins, MamN (a Na^{+}/H^{+} antiporter) and MamP have also been implicated in pH regulation in the magnetosome lumen.¹²⁵

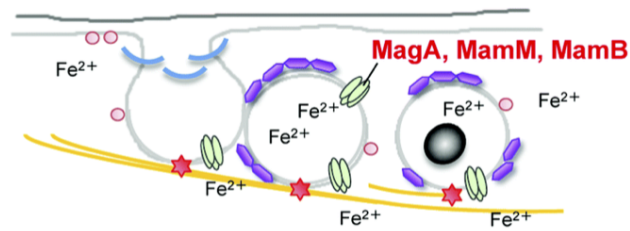


Figure 3.6: Schematic representation of iron uptake into vesicles during magnetosome formation. This figure has been taken and adapted from Arakaki2015.¹³¹

MamH and MamZ both contain the common domain of the major facilitator superfamily (MFS) and are thereby also predicted to play some part in iron transport for magnetosome formation. A structural model of MamH reveals a negatively charged cavity suggested to allow binding of metal cations and their transport. MamZ contains not only this MFS-like domain but also a conserved ferric reductase component, and so has been predicted to participate in electron transfer and redox reactions as well as iron transport.¹⁴¹ Another protein, MamO, whose entire function is not currently well understood, has also been implicated in iron transport, not necessarily through the MM but rather within the magnetosome, for nucleation.¹⁴²

Depending on the environment (i.e. anoxic/microoxic zones) in which specific MTB reside, Fe might be stored in ferrous or ferric forms within the cytoplasm. Accumulation of Fe^{2+} is typical within the magnetosome lumen due to the presence of ferric reductase and CDF proteins. However formation of magnetite and greigite requires oxidation of Fe^{2+} to Fe^{3+} , and the maintenance of an $\text{Fe}^{3+}:\text{Fe}^{2+}$ ratio of roughly 2:1 which is again controlled by specific proteins. As stated previously, MamZ (containing a ferric reductase transmembrane component) is believed to be one such protein. In addition, MamE, P, X and T all contain two or three CXXCH motifs, a c-type cytochrome motif for binding of heme, alluding to the role of these proteins also in the oxidation/reduction of iron. These MTB-specific motifs have been termed magnetochromes, and represent an entirely new type of cytochrome.^{141,143,144} Investigations involving depletion of the paired CXXCH motif of MamX have indicated its role in redox control for formation of mixed-valence Fe_3O_4 under conditions where oxidants (i.e. molecular oxygen and nitrate) are low.¹⁴¹ MamP displays some different properties when compared with other c-type cytochromes, including its ability to prevent the reduction of already-oxidised Fe^{3+} ; this property could be implicated in a role for the control of the $\text{Fe}^{3+}:\text{Fe}^{2+}$ ratio within magnetosomes, hence affording proper nanocrystal formation.¹⁴⁵ It has also been suggested that MamP oxidises ferrous iron (Fe^{2+}) to ferrihydrite, an Fe^{3+} oxide and precursor to magnetite - this is due to various inherent structural components,

including a specific cavity as well as a conserved H⁺ exit channel. The presence of more Fe²⁺ when MamP has been fully reduced should then allow the transformation of magnetite from ferrihydrite.¹⁴⁴ Maml is another protein whose role has been linked to the formation of the final magnetic nanocrystal at least in MSR-1, via the transformation of compact iron oxides to magnetite.¹²⁵

3.3.4 Crystal size and shape

A number of proteins play a vital part in controlling the size and shape of the magnetic nanocrystals formed in magnetosomes (see Figure 3.7), which is a strain-dependent process. Within strains however crystal size and morphology is highly uniform, with little size difference between nanocrystals, suggesting a highly controlled process. The most well characterised and widely studied of proteins known to be involved include MamC, D, F, G, MmsF and Mms6; all are integral membrane proteins found within *Alphaproteobacteria* specifically.¹²⁵ Indeed, individual deletion of either *mms6* or *mamCDFG* operons in *Alphaproteobacteria* strains results in nanocrystals smaller in size than in the equivalent WT strains. The four proteins encoded for by the *mamCDFG* operon appear to have a cumulative effect on nanocrystal growth, since it is only upon reintroduction of any three out of the four genes that nanocrystals reach the same size as found in WT strains.¹⁴⁶

A conserved hydrophobic N-terminal domain (NTD) found in each MamC, D, G and Mms6, all of which have been found to strongly bind magnetite, was hypothesised to allow co-localisation of these proteins for interaction with magnetite. Interestingly, these four proteins are not encoded for in other MB strains which produce bullet-shaped magnetite nanocrystals, and thus they are believed to play an important role in morphological control for cuboctahedral nanocrystal formation specifically.¹⁴⁷ Each of these proteins have either been shown or predicted in structural models to contain some form of acidic loop (MamC and MamG), or CTD (MamD and Mms6), extending into the magnetosome lumen and involved in iron binding. In MamC, this α -helical loop connects two integral transmembrane helices;¹⁴⁸ specifically two acidic residues at around 8 Å apart are believed to form the scaffold for cuboctahedral magnetite nucleation, in which iron atoms are 6 Å apart. MamC is predicted to bind specific magnetite faces, hence preventing growth in these directions and affording the correct, uniform morphology of nanocrystals.¹⁴⁹ This same magnetite interaction site has been predicted for MamG.¹⁵⁰ Both CTDs of MamD and Mms6 have also been implicated in control of crystal morphology, with deletion of *mamD* in MSR-1 leading

to a crystal face dissimilar to those in WT cells.¹⁴⁷ The CTD of Mms6, which is most well characterised, has been shown to bind iron with a preference for Fe^{2+} over Fe^{3+} , and remains a functional 6 KDa unit upon proteolytic cleavage, a process which has been found to readily occur in Mms6.¹⁵¹

Magnetite binding is also predicted for the homologous MmsF and MamF proteins, whose co-deletion results in crystals of decreased size and number - a phenomenon also observed but to a lesser extent upon deletion of *mmsF* alone.¹³⁷ An acidic C-terminal tail extending into the magnetosome lumen is found in MmsF and predicted for MamF, and similarly expected to be the site of magnetite binding.¹⁵²

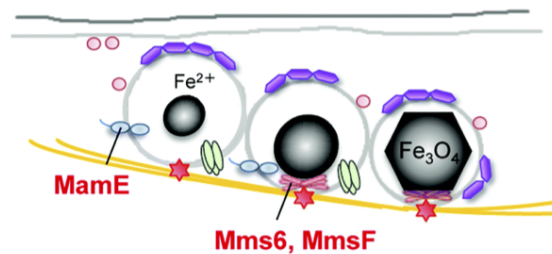


Figure 3.7: Schematic representation of magnetite formation within magnetosomes of magnetotactic bacteria. This figure has been taken and adapted from Arakaki2015.¹³¹

Altogether, the synthesis of magnetosomes within MTB is a highly controlled process which is regulated primarily by proteins including those mentioned here, which constitute the majority of magnetosome membrane and associated proteins encoded for in *Alphabacteria*. The importance of such uniform, strain-specific magnetosome formation is clear, but for reasons which are as yet unknown. However, a number of these proteins belong to large superfamilies, including the MFS and CDF families for example, whose human homologues have in some cases been implicated in serious medical conditions. In many cases, when trying to elucidate the functions and their associated mechanisms in human membrane transporters it is necessary to first utilise their homologues from other domains of life; this process can give huge amounts of information without the often difficult task of isolating the human protein for research directly. Indeed a large volume of work described in this thesis has been carried out on the MTB protein MamM, of the CDF family, for which the literature on its structure, function and mechanism of action as understood thus far will be concisely reviewed next.

3.4 MamM from *Magnetospirillum gryphswaldense* MSR-1

MamM has in recent years been used as a model system for the study of human disease-causing CDF proteins. This was first demonstrated when introduction of mutations to the CTD of MamM at homologous positions to disease-causing mutations in the human CDF protein, ZnT-8, resulted in altered magnetic responses in resulting magnetosomes, and altered crystal structures of MamM-CTD.¹¹⁴ Research has included crystallographic studies as well as biophysical methods for determination of structure, function and mechanism of action both *in vivo* and *in vitro*. Most well characterised currently is the CTD of MamM (MamM-CTD) - this is the cytoplasmic C-terminal unit most commonly believed to play a role in locating and capturing its substrate in the cytoplasm for the subsequent translocation across the membrane, although this function remains the topic of some debate across other bacterial homologues.

3.4.1 Structural and crystallographic studies of MamM-CTD

Structural studies of MamM have thus far been focused on, and limited to, its CTD, for which a number of crystal structures have been solved in a variety of states. The first major breakthrough in structural research of MamM was undertaken by N. Zeytuni *et al.* in 2011¹¹² in which the truncated MamM-CTD unit was successfully crystallised. Initial purification via size-exclusion chromatography revealed what was believed to be dimeric MamM-CTD with an observed molecular mass of approximately 25 kDa, as compared to the predicted 11.9 kDa mass of the 108 residue monomeric unit. This approximate molecular mass observed is comparative to that of the CzcB-CTD, which also exists as a dimer in solution.¹¹⁰ The preliminary validation of the crystal structure published here suggested only a single MamM-CTD monomer in the asymmetric unit. Nonetheless, the group predicted a V-shaped packing of MamM-CTD similar to that previously seen for the *Thermotoga maritima* TM0876-CTD.¹¹¹ A more in-depth analysis involving crystal structure determination was carried out at a later date by the same group.¹¹³ Purified MamM-CTD was crystallised and its structure solved in the apo state, other than the C-terminal tail composed of residues 294-318 whose structure was disordered and as such, undetected. Results revealed firstly a metallochaperone-like fold in each monomer typical of other CDF CTDs; truncation of MamM-CTD to eliminate the disordered C-terminal tail gives comparable results. Solved crystal structures show MamM-CTD as a stable dimeric unit in its apo state, exhibiting a V-shaped assembly. Some flexibility is seen towards the top of the dimer, whilst a stabilising dimerisation interface exists at the bottom. For further analysis

of this dimerisation interface, a single point mutation - V260P - was introduced which was suspected to alter the backbone hydrogen bonds between MamM-CTD monomers. Size-exclusion chromatography of this purified mutant however revealed the protein is still in its dimeric form, whilst molecular dynamics simulations of a V260P structural model shows a stable dimer fold.

A separate study - this time testing the use of *Magnetospirillum gryphyswaldense* as an effective model system for probing human CDF-related diseases - utilised this same position for structural determination.¹¹⁴ The X-ray crystal structure solved for the V260R variant of MamM-CTD displays monomeric units with folds that resemble those of the WT protein, whilst a portion of the C-terminal tail which was not previously solved adopts a rigid fourth β -sheet parallel to the three existing ones. Most interestingly, this variant provides the first structural evidence for an altered dimeric packing of MamM-CTD, in which a tighter and twisted conformation is observed relative to the WT structure, stabilised by a number of salt bridges, polar and hydrophobic interactions both at the top and bottom of the dimer.

On the theme of disease-causing mutations in the human transporters, position V260 was chosen specifically based on its homology to position 325 in the human transporter ZnT-8. In the more common risk allele for acquiring Type-II diabetes, an arginine is found at this position, whereas in the less common allele which provides some resistance to the disease, this position is instead occupied by a tryptophan. Out of the two MamM-CTD variants for which crystallisation trials were attempted - V260R and V260W - the former provided the only crystals for structural determination. Based on putative metal binding sites previously suggested for MamM-CTD¹¹³, the twisted dimeric fold observed prohibits the formation of binding sites predicted on the peripheral sides of the protein. Instead, two alternative binding sites specific to this structural arrangement were suggested.¹¹⁴

Investigation of another disease-causing mutation - L349P in ZnT-10 - was also carried out using MamM as a model system.¹⁵³ The crystal structure obtained for MamM-CTD M250L (the homologous position) is highly similar to that of the WT protein, insinuating that the introduced leucine maintains the dimeric structure of WT MamM-CTD. Free energy calculations suggest little free energy difference between the WT and M250L variant, whilst introduction of a proline at this position predicts a much higher free energy. This was further investigated with molecular dynamics simulations which overall give a clear indication of a destabilised protein state. Size-exclusion chromatography resulted in protein elution at a different volume than the WT, and in combination with multi-angle light scattering a

molecular weight equal to that of monomeric MamM-CTD was determined. Combined, these results imply that the protein exists as a disordered monomer with the M250P mutation. Small angle X-ray scattering (SAXS) was also employed for low resolution structural determination of this variant, giving rise to an elongated protein envelope relative to that of the WT, suggestive of an unfolded monomeric structure.

Of course in order to gain a deeper understanding of the structural elements which are closely linked to the proper functioning of complex proteins, information is required not only in the apo state but additionally in the presence of substrates. First attempts at MamM-CTD crystallisation in the presence of M^{2+} ions were attempted with metals including zinc, iron, cadmium, manganese plus others, with a variety of different crystallisation techniques.¹¹³ However cation-bound MamM-CTD crystals were not obtained and instead, vast precipitation was observed. Instead, a closed-state model (based on the CrzB zinc-bound state) of MamM-CTD was prepared and analysed, revealing three putative metal binding sites: one central site (CS) involving His285 and Asp249 of both monomers; and two symmetrical peripheral sites (PS), each involving His264 and Glu289 of different monomers. Molecular dynamics (MD) simulations were carried out, with four zinc ions docked to the closed-state model (two bound at the central site, and one at each peripheral site) revealing a highly stable zinc-bound state, whilst MD simulations of the same model without docked zinc ions reveal flexibility of the monomeric 'arms' after a period of just 9 ns. SAXS-calculated protein envelopes also give an indication of zinc binding-induced conformational change in MamM-CTD. MD simulations of binding site variants, including single and double variants, as well as crystal structures solved for all of these variants in their apo state display no changes as compared with the WT MamM-CTD structure.

Further structural studies on both central and peripheral binding site double mutants - specifically D249A-H285A MamM-CTD (termed CSM) and H264A-E289A MamM-CTD (termed PSM) - have been carried out more recently for additional elucidation of both apo and metal-bound conformational states.¹¹⁵ Based on a previous suggestion that MamM-CTD monomers might undergo a swivelling motion,¹¹³ in addition to laterally moving towards and away from one another, MD simulations were carried out using WT and CSM crystal structures. These simulations give an indication that in both cases, an open-close movement in addition to a swivelling motion, accounts for the majority of motions exhibited by the protein, whilst the amplitude of the swivelling motion is largest for CSM MamM-CTD. Additionally, CSM displays more flexible arms than the WT protein; overall these results indicate that removal of the central binding site residues affects the

apo MamM conformational states, affording larger and more dynamic swivelling movements.

The first successful attempt at crystallisation and structural determination for metal-bound MamM-CTD is documented in a recent publication,¹⁵⁴ and also represents for the first time structural evidence of any CDF CTD to bind not one, but multiple metals. MamM-CTD crystal structure models were successfully obtained following co-crystallisation with Cd^{2+} , Ni^{2+} and Cu^{2+} . Whilst structures of Cd^{2+} - and Ni^{2+} -bound MamM-CTD show no evidence of conformational change relative to that of the apoprotein, a clear conformational shift to a tighter dimeric interface in which the monomer 'arms' approach one another is observed with Cu^{2+} (see Figure 3.8). In all structures however, there is evidence of metal-coordination by at least one residue previously proposed to form the central and peripheral binding sites; indeed the Cu^{2+} -bound structure reveals one metal ion coordinated at each of these proposed sites, giving three bound Cu^{2+} ions in total. Additional water molecules also play a role in ion stabilisation at both peripheral sites.

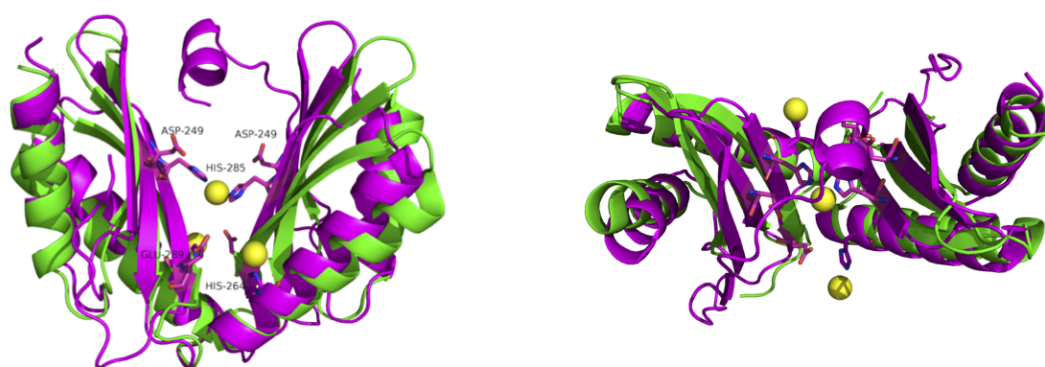


Figure 3.8: Both apo (PDB code: 3W5Y) and Cu^{2+} -bound (PDB code: 6GP6) MamM-CTD crystal structures, overlaid. Cu^{2+} binding site residues are highlighted (left).

The Cd^{2+} -bound structure (see Figure 3.9) reveals two bound metal ions per monomer giving four in total, with each ion in each monomer ligated by either D249 or H285 of the proposed central binding site, whilst the half occupancy shown for each ion suggests alternate binding. Just one Ni^{2+} ion was found to bind per MamM-CTD monomer, at the proposed peripheral binding site and coordinated by H264 and water molecules. A non-specific Ni^{2+} interaction is suggested, hence causing no conformational change, however the observed positions of bound ions and specifically their coordination to His264 further supports the importance of this residue in MamM-CTD function, possibly for the attraction of metal ions for binding and translocation. Out of all other CDF CTD crystal structures previously solved, only those of CzrB displayed metal binding-dependent conformational change, showing closure of the 'arms' to a distinct V-shape similar to

that observed for MamM-CTD with Cu^{2+} . The degree of closure with CzrB however is larger than that of MamM-CTD. Contrastingly, both apo and Zn^{2+} -bound MamB structures reveal a closed state, both of which are equivalent to one another but with tighter packing than apo MamM-CTD. Similarly, solved YiiP structures (including EcYiiP and SoYiiP), in which the CTD is always in the metal-bound state, show a greater degree of closure than in MamM-CTD, whilst relative to the unsolved apo states no comparisons can be made.

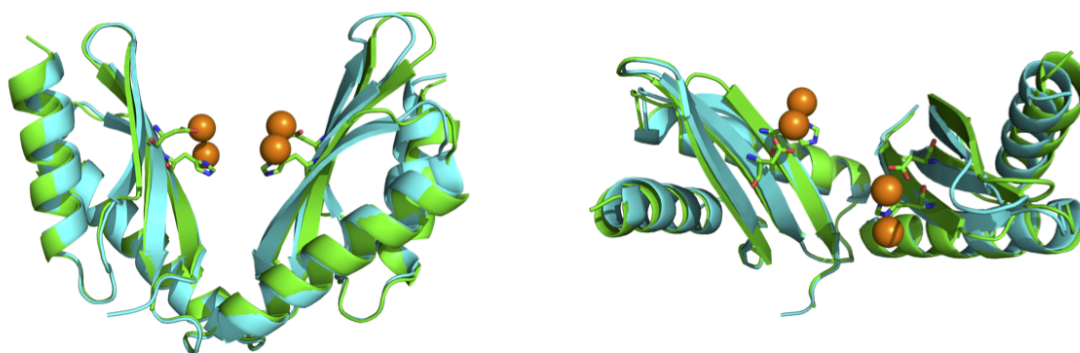


Figure 3.9: Both apo (PDB code: 3W5Y) and Cd^{2+} -bound (PDB code: 6GMT) MamM-CTD crystal structures, overlaid.

A yet more recent study has introduced binding site point mutations to investigate how the nature of the metal binding sites affects the CTD response to different divalent metal cations.¹⁵⁵ X-ray crystallography was used in an attempt to crystallise and obtain structural models for each variant in the presence of various divalent metal cations, of which 22 structures were successfully solved (all are available on the PDB). Discussion is primarily focused on solved structures of multiple variants with Zn^{2+} , Ni^{2+} and Mn^{2+} , the comparison of binding sites observed between variants and therefore how specific residues influence the binding capabilities and conformational changes of MamM-CTD with different metals. Firstly, crystal structures with Zn^{2+} were solved for each H264E, D249E and D249N variants, all of which exhibit a tighter dimeric fold than that of the apo WT protein. Previously, the Cu^{2+} -bound structure of WT MamM-CTD was assumed to represent that of the Zn^{2+} -bound state also, based on biophysical methods. In this case all three variants appear to share a symmetrical binding site composed of H285 of one monomer, and H236 and E289 of the second monomer. Based on the previously solved Cu^{2+} -bound structure, this solved Zn^{2+} binding site is suggested to be an unnatural binding site, despite it not containing any of the mutated residues. A second symmetrical binding site, coordinating two zinc ions, was found for the D249E and D249N variants, where in both cases H264 coordinates the Zn^{2+} ions along with C267 and E268 of the same monomer, with water molecules also involved. This again highlights the importance of H264 for metal binding in MamM-CTD. Whilst not

believed to cause metal binding-induced conformational change directly, binding of metals to this residue is suggested to be vital for metal attraction to the CTD for subsequent translocation.

Solved crystal structures with Ni^{2+} , for comparison with the previously solved Ni^{2+} -bound WT structure, were obtained for both E289H and D249N MamM-CTD variants. Both crystal structures exhibit conformations comparable to that of both WT apo and Ni^{2+} -bound structures, with Ni^{2+} ions coordinated on both peripheral sides of the dimer by H264 of each monomer and water molecules. Hence no conclusions could be made as to whether or not these mutations affect the biologically relevant conformational states exhibited. Multiple crystal structures were also solved for MamM-CTD variants in the presence of Mn^{2+} , however Mn^{2+} was undetected in all but one electron density map. Each solved crystal structure for MamM-CTD H285D, D249H-H285D, H264E-E289H variants show no conformational change relative to the WT apo structure, whilst that of both E289D and D249N MamM-CTD variants display a much tighter dimeric conformation. A weak Mn^{2+} density was observed in the E289D Mn^{2+} -bound structure, coordinated by H264 and a water molecule. However since both variants adopt the same conformational state, this Mn^{2+} binding was suggested to be non-specific. Instead, the conformational changes observed are attributed to different stabilisation effects not possible in the WT case. D289 of the E289D variant forms an additional hydrogen bond relative to the WT protein, afforded by the different size of the introduced residue. Similarly, the tighter dimerisation interface of D249N is a results of additional and stabilising hydrogen bonds between N249 residues of each monomer, which have replaced the negatively charged and therefore repulsive D249 residues present in the WT protein.

3.4.2 Biophysical characterisation and *in vivo* studies

Whilst crystal structures can provide a wealth of information in the study of protein function, these solved models only represent a single snapshot in space and time of what could be a highly mobile system. Additionally, as is expected to be the case in multiple MamM-CTD crystallographic studies, the conformations of the protein seen can be dominated by crystal contacts and therefore are not necessarily representative of those within the physiological environment. *In vivo* studies and further biophysical characterisations allow further information on the mechanisms of function and conformational states exhibited by these proteins. These can be linked back to structural data, which is typically gained under far from native

protein conditions. *In vivo* analysis of magnetite formation and magnetic behaviour of magnetosomes upon introduction of specific MamM-CTD mutations provides a means of determining which mutations lead to a decrease of MamM functioning. Biophysical techniques including Trp-fluorescence studies, isothermal calorimetry (ITC), solid-state nuclear magnetic resonance (ss-NMR) spectroscopy and EPR have also been routinely used for characterisation of binding stoichiometry, affinity and conformational change, for example. Such characterisations published in literature will be presented here, other than the EPR data, since these published studies will form a substantial portion of the results to follow in this thesis chapter.

Previously, alterations to residues believed to stabilise the dimeric interface of MamM-CTD were discussed in terms of solved crystallographic models, with significant attention in the literature being paid to V260 at the base of the dimer. In one such study¹¹³ *in vivo* studies of MSR-1 $\Delta mamM$ cells trans-complemented with the *mamMV260P* mutant gene reveals that proper functioning of MamM-CTD V260P is not possible, despite the protein maintaining its stable dimeric packing. Additionally, ITC measurements reveal fewer metal binding sites as compared to the WT protein, again implying decreased functioning of this variant. In a separate study¹¹⁴ this same position was utilised for site-directed mutagenesis and similarly *in vivo* analyses of magnetic nanoparticle formation and magnetic behaviour of trans-complemented MSR-1 cells were undertaken. Trans-complementation of MSR-1 $\Delta mamM$ cells with V260G and V260W mutant alleles restores magnetic response and hence affords magnetic nanoparticle formation, whilst V260R and V260D mutant introduction does not. Both latter variants however, according to ITC measurements, are still able to bind Zn^{2+} , although the number of binding sites is significantly reduced for the V260W mutant.

The M250P MamM-CTD mutation - for which the structural studies were discussed previously, highlighting the observed destabilised unfolded state - was additionally subjected to further *in vivo* and biophysical analysis relating to overall protein function.¹⁵³ Trans-complementation of MSR-1 $\Delta mamM$ with *mamMM250P* shows significantly decreased magnetism, and far smaller and fewer magnetic nanoparticles in the host MTB. Additionally a large proportion of cells lack magnetosomes entirely, providing evidence that this homologous disease-related mutation strongly decreases proper functioning of the protein. It was suggested in this study that the elongated envelope observed by SAXS for M250P might be as a result of loop breakage or secondary structure loss; this possibility was further investigated via circular dichroism (CD) spectroscopy. The CD spectrum obtained reveals obvious and significant differences for the M250P variants when compared with the WT,

and in fact resembles something more characteristic of random-coil proteins. This gives clear evidence of a loss of secondary structure. 1D ^1H -NMR was also utilised, given the differences in spectra typically observed in structured versus disordered proteins. Whilst specific features in the ^1H -NMR spectrum of WT MamM point towards a structured protein, that of the M250P variants again indicates a loss of structure.

Whilst X-ray crystal structures were not successfully solved for MamM-CTD in combination with metal ions until very recently, numerous other techniques have been routinely used to give an insight into protein-metal interaction, binding and associated conformational change. As mentioned previously, SAXS data in combination with MD simulations gave one of the first indications of Zn^{2+} binding-induced conformational change in MamM-CTD.¹¹³ Trp-fluorescence measurements, using the native Trp247 residue of each monomer, further informed on the conformational states exhibited by MamM-CTD with and without metal ions. Evidence for a change in local environment of the Trp247 residue, and hence conformational change, is observed upon increasing Zn^{2+} concentrations, in which a specific fluorescence peak shift is associated with increased hydrophobicity. Additionally, fluorescence signal quenching upon increasing concentrations of iron ions indicates residues close to Trp247 are involved in metal binding. ssNMR investigations undertaken as part of this same study give further evidence for conformational change associated with metal binding, specifically the conformations adopted by histidine residues. In turn, these said histidines are implicated in metal coordination. Restricted mobility caused by formation of intermolecular contacts between monomers in the hinge region between residues Thr259 and Glu257, as well as between Glu289 and His264 was found for Zn^{2+} -bound MamM-CTD but not for the apoprotein, again implying conformational change towards an overall more tightly packed dimeric structure.¹¹³

In vivo studies were also performed on MSR-1 ΔmamM trans-complemented with single point variant alleles in which residues of both the putative central and peripheral binding sites were altered.¹¹³ Interestingly, single point mutations of either histidines involved in central or peripheral sites do not affect magnetic response, the size, shape or number of magnetic nanocrystals compared with the WT gene. The E289A mutation of the peripheral site leads to much fewer nanocrystals per cell and ultimately a lower magnetic response, whilst the D249A central site variant has much the same effect. ITC measurements of each single point mutant with Zn^{2+} lead to the hypothesis that one metal ion binds at each peripheral site, and that the proper functioning of the CTD is afforded by metal binding to either central or peripheral sites but not necessarily both at once.

WT and binding site variants of MamM-CTD were further investigated more recently, with a focus on Zn^{2+} binding stoichiometry, affinity and the associated conformational changes that take place.¹¹⁵ The WT, PSM and CSM protein variants (as previously described) were subjected to CD measurements, both in apo and Zn^{2+} -bound states, for analysis of secondary structural elements. Results suggest little change in secondary structure for WT or variant proteins upon metal binding, whilst the majority of conformational change is suggested to take place about the dimerisation interface. DEER studies were used to further probe the conformational states exhibited in apo versus Zn^{2+} -bound MamM-CTD variants, the results of which will be discussed in detail later in this chapter. Zn^{2+} binding stoichiometry was determined via ITC measurements, however Zn^{2+} binding to PSM did not initiate enough heat capacity change for measurement and so results are of the WT and CSM variant alone. In total three Zn^{2+} ions are shown to bind WT MamM-CTD, and two ions bind CSM, indicating only one Zn^{2+} binds at the central site. Binding affinity of CSM for Zn^{2+} is higher than that of WT protein, which is attributed to a more stable Zn^{2+} -induced conformational change. Trp-fluorescence emission measurements were also obtained for all constructs, in which the reporting W247 residue lies in close proximity to the central binding site, hence informing on conformational changes around this site as a function of increasing Zn^{2+} concentration. All MamM-CTD constructs exhibit a Zn^{2+} -dependent blue shift in the emission spectra, as well as an increase in peak intensity relative to the WT protein, suggesting movement of W247 towards a more hydrophobic environment. The amplitude of this shift and the increase in peak intensity is larger for WT protein relative to CSM, attributed to both the alanine substituted at position 249 in CSM which is close to W247, as well as a different Zn^{2+} -bound conformational state. However, Zn^{2+} addition to PSM induces only a slight blue shift, with the largest of all peak intensities, indicating a different effect of Zn^{2+} binding on the protein altogether.¹¹⁵

This same W247 native residue was utilised for *in vitro* investigation of MamM-CTD binding to multiple metal species, following the first successful structural determination of metal-bound MamM-CTD models.¹⁵⁴ Quantification of metal binding firstly using Trp-fluorescence emission was attempted for states in which X-ray diffraction based structural models had been determined (with Ni^{2+} , Cd^{2+} and Cu^{2+}), or not (with Mn^{2+}), with the knowledge that upon binding of any paramagnetic metal ion close to W247, the fluorescence signal would be quenched. Results with Zn^{2+} and Fe^{2+} were described previously and hence not repeated in this study. Despite its paramagnetism, addition of Mn^{2+} makes minimal difference to either the fluorescence intensity or peak position. Contrastingly, addition of

Ni^{2+} and Cu^{2+} results in both quenching and slight spectral shifts, both of which are more prominent with Cu^{2+} . Cd^{2+} addition leads to a peak shift similar to that seen previously with Zn^{2+} and indicative of a change in tryptophan environment most likely due to local conformational change. Altogether these results give an indication that MamM-CTD can bind all metals tested besides Mn^{2+} . ITC results gave an initial estimation of K_D values, which vary from 6-48 μM with the following order of increasing affinity: Cd^{2+} , Zn^{2+} , Ni^{2+} then Cu^{2+} . No heat change was observed with Mn^{2+} , providing further evidence that this protein discriminates against Mn^{2+} . The binding stoichiometry for each metal ion was estimated, with the results in some cases differing from those of crystallographic studies. Taken together with additional DEER studies (which will again form part of this results chapter) it was concluded that Zn^{2+} , Ni^{2+} , Cu^{2+} and Fe^{2+} all bind in such a way that results in a distinctive conformational locking to a single, tighter V-shaped conformation than is exhibited in the WT apo crystal structure. Binding of Cd^{2+} is different, most likely due to its larger radius as compared to the other tested metals. DEER studies revealed for the first time, structural observation of Fe^{2+} induced conformational change in MamM-CTD.¹⁵⁴

Trp-fluorescence measurements were additionally utilised to further probe the affects of binding site mutations on the ability of MamM-CTD to bind different metals, with a particular focus on Zn^{2+} , Ni^{2+} and Mn^{2+} , the crystallographic results of which were discussed previously.¹⁵⁵ This again provides a clear example of how in specific cases, biophysical evidence clearly contradicts that of structural models, emphasising the need for multifaceted research approaches. Trp-fluorescence results for binding site variants with Zn^{2+} show smaller blue-shifts than that of Zn^{2+} -bound WT protein, relative to the apo state; results indicate that Zn^{2+} binding to these variants gives a less tight conformation than is likely for Zn^{2+} -bound WT protein. The peripheral H264E variant displays the smallest shift of all, suggesting that the inability of Zn^{2+} ions to bind this residue might lead to an even less stable closed state. In this case results for each the H264E, D249E and D249N are comparable to their Zn^{2+} -bound structural models.

Whilst crystal structures solved with Ni^{2+} for E289H and D249N display no conformational change relative to the apo WT protein, with Ni^{2+} ions bound by H264 of each monomer, Trp-fluorescence results indicate different binding capabilities between multiple variants. Fluorescence quenching implicates a role for D249 of the central binding site in Ni^{2+} binding, and substitution of glutamate at position 289 for histidine results in less quenching as compared to the WT protein, suggesting a negative charge at this position is required for metal binding at the central site. Substitution of H264, the only residue shown to ligate

Ni^{2+} in crystal structures, shows much the same results as for the WT protein. In this case significant quenching (caused by Ni^{2+} ligation at the central site) is not apparent, and so it was suggested that similarly to Zn^{2+} , H264 is crucial for correct binding and associated conformational changes in the peripheral sites.

In the case of Mn^{2+} , for which specific binding was not demonstrated in any solved crystal structures of binding site variants, no Trp-fluorescence quenching or blue-shift is observed in any case, providing yet more evidence for MamM-CTD selectively not binding Mn^{2+} . The authors have suggested that MTB have evolved to prohibit the overload of Mn^{2+} into magnetosomes, however the specific mechanism by which this occurs in MamM-CTD could not yet be predicted, since studies thus far have not implicated any key residues in this selectivity against Mn^{2+} .¹⁵⁵

3.4.3 The kinetic and conformational mechanism of MamM-CTD and other CDF proteins

The overall aim of research into complex membrane transport systems like MamM is of course to gain insight into the biologically relevant pathways exhibited by proteins, and of primary importance are their kinetic and conformational mechanisms when considering ways in which we can target them, or alter their function, for medicinal purposes. Taking MamM-CTD specifically, recent work involving fluorescence based studies has given information on the kinetics associated with MamM-CTD metal binding, in which determination of the order of binding events has been afforded by binding site mutants, PSM and CSM MamM-CTD.¹¹⁵ Firstly, although no time-dependent changes in tryptophan fluorescence is detected for PSM, a Zn^{2+} -dependent increase in zero-time fluorescence intensity indicates an initial fast binding event within the dead-time (1.5 ms) of the experiment. Additionally the observed single phase fluorescence intensity shown by the WT data, which is unexpected due to the presence of multiple different metal binding sites, may also be attributed to binding of metal ion/s during the experimental dead-time. This gives the first insight into the order of binding events in MamM-CTD - since the PSM variant only contains the central binding site, it is suggested that for the WT protein it is rapid binding at this site which takes place first causing a slight closure of the dimer (WT^C) from an entirely open form (WT^O). The cause of the measured Zn^{2+} -dependent fluorescence intensity observed in both WT and CSM MamM-CTD has been defined as rapid binding of Zn^{2+} to the peripheral sites, which is the second binding event to take place in the case of WT protein; binding to the

peripheral sites (although rapid) is slower than initial binding to the central site. Observation of a high fluorescence state reveals an isomerisation event causing further closure of the CTD to a more compact form (WT^{CC}), a process which takes place over a longer period of time than either binding event. This suggested mechanism (see Figure 3.10, top panel) results in a total of 3 bound metal ions to the WT protein, as inferred from ITC results also.

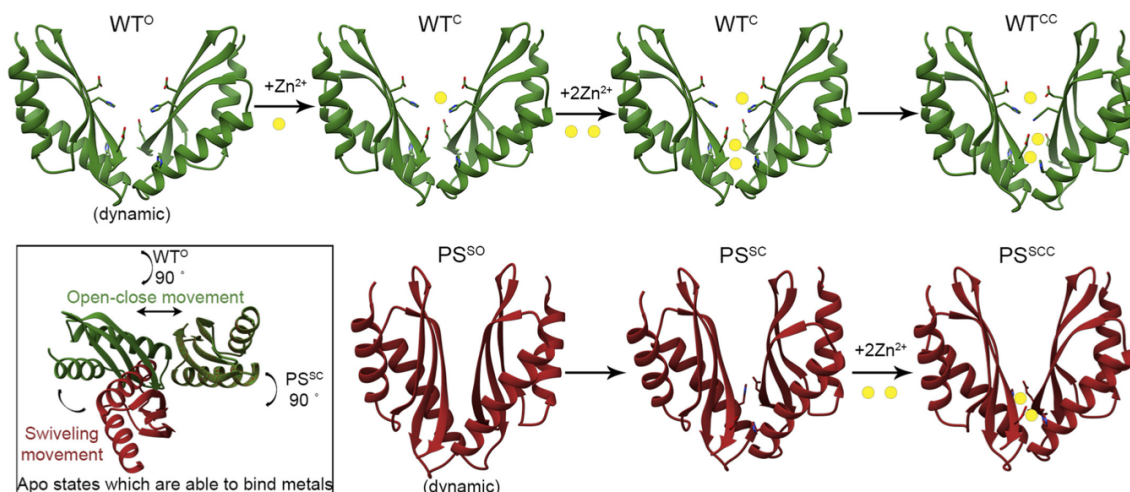


Figure 3.10: Suggested mechanism of Zn^{2+} binding and associated conformational changes, for each WT (top panel) and CSM (bottom panel) MamM-CTD. The lower left box gives an overlapped representation of WT^0 and CSM^{SC} to highlight the difference between the swivelling movement described, and the regular CTD closure. This figure has been taken from a publication of which I am a co-author.¹¹⁵

Differences in dissociation rates between WT and CSM MamM-CTD were suggested to be the result of different conformations adopted between the two upon Zn^{2+} binding at peripheral sites. As explained previously, additional swivelling motions have been observed for CSM MamM-CTD, and the results of pulsed EPR experiments to determine conformational change have also lead to the understanding that upon metal binding, CSM can adopt a more closed conformation than the metal-bound WT protein, due to the removal of bulky and charged residues from the central site. Additionally, the difference in trends of the k_{obs} dependence on Zn^{2+} concentration between WT and CSM is attributed to a different mechanism of Zn^{2+} binding altogether in which an additional isomerisation step precedes metal binding. The suggested mechanism (see Figure 3.10, bottom panel) includes two apo conformations of CSM MamM-CTD - the swivel-open (CSM^{SO}) and swivel-closed (CSM^{SC}) forms - of which the latter is the only conformation to afford metal ion binding. Subsequent binding of Zn^{2+} to CSM^{SC} MamM-CTD and binding-induced closure of the dimer to form CSM^{SCC} is then a much more rapid step, meaning k_{obs} is always dominated by the rate of isomerisation between swivel-open and swivel-closed states; in other words, this is the rate limiting step of this binding mechanism.

Overall, binding of a single metal ion at the central site (in the case of WT MamM-CTD) appears to facilitate binding of two further ions, one at each peripheral binding site, leading to a larger energetic cost and the most distinct conformational change of the dimer. The importance of metal binding in peripheral sites has also been demonstrated previously in both CzrB and YjiP metal-bound structures,^{110,156} whilst all known structures of CDF CTDs strengthen the evidence for binding at the central site in MamM. However considering that not all metal ion binding sites are universally conserved between the known structures, and neither is the nature of metal binding-induced conformation change, it is possible that these CTDs function using distinct and versatile regulation mechanisms.

In terms of the overall mechanism of action by entire CDF proteins (i.e. involving both the TMD and CTD), multiple suggestions have been made in the literature. Some proposals have been based upon solved crystal structures of YjiP, for which structural models derived from either X-ray diffraction or cryo-EM studies present both the TMD and CTD. One such study suggests a disruption of the dimerisation interface of the CTD in the absence of bound metal ions. This is based upon the electronegativity and side chain arrangements of ligands in a particular metal binding site, which upon release of metal ions is predicted to cause the splaying apart of CTD monomers due to charge repulsions and loss of inter-subunit interactions. Additional FRET measurements also indicate a decrease in distance between residues towards the bottom of YjiP-CTD upon increasing Zn^{2+} concentrations.¹⁰⁷ However, considering the extent of research into the stability of MamM-CTD about its dimerisation interface (as described previously), which altogether infers a very stable dimeric fold but with a high degree of flexibility towards the top of the dimer 'arms', an alternative suggested mechanism exists. This flexibility has been attributed to a combination of charge repulsion between monomers, and hydrophobic attraction at the centre of the dimerisation interface, with the suggestion that such opposite forces induce a spring-like flexibility allowing the CTD to search for its substrate in the cytoplasm. For MamM-CTD it is predicted that binding of metal ions and subsequent conformational locking into a more closed and rigid state affords translocation of substrate via initiation of the transport mechanism of the TMD. The suggested alternating-access model of the TMD of CDF proteins is based solely on YjiP solved structures, since this is the only CDF protein for which the TMD has been solved in multiple conformations, including both outward- and inward-facing states.

3.5 Characterisation of MamM-CTD Zn²⁺ binding using EPR spectroscopy

The introductory chapters of this thesis present a survey of the application of EPR methodologies, in conjunction with other techniques, to elucidate conformations adopted by proteins during their biological functions. Thus far, crystallographic, *in vivo* and other *in vitro* biophysical techniques have been used to probe the structure and function of MamM-CTD as a means to further understand how abnormal functionality of the human CDFs leads to disease. This investigation has now been extended through the use of advanced EPR techniques to assess the structure-function relationship of MamM-CTD.

3.5.1 SDSL of MamM-CTD variants

As previously mentioned in section 1.3, EPR investigations of many proteins require the introduction of a paramagnetic label, which is typically afforded by SDSL. In a homodimeric protein such as MamM-CTD, probing of function-related conformational change using dipolar coupling-based EPR techniques (for example the DEER experiment) may be afforded by labelling a single site on each monomer of the protein. Each MamM-CTD monomer contains two intrinsic cysteine residues; cysteines contain a thiol group which is the reactive partner in spin label attachment when using typical labels including MTSSL and IAP with thiol-reactive functionalities. These cysteines are present at position 267 and 275 in MamM-CTD, approximately midway along and towards the end of the same α -helix, respectively. In terms of their positioning with respect to the entire CTD structure, position 275 is located further from the dimeric interface in a peripheral position which is solvent exposed, whereas position 267 is closer to the dimerisation interface and faces the centre of the dimeric structure. Considering the positioning and apparent flexibility of MamM-CTD towards the tips of its 'arms' and close to the TMD, as inferred by a number of biophysical studies, labelling at position 275 might be expected to report on longer-range conformational changes (Figure 3.11A). The presence of inherent cysteines is useful since, assuming these positions are suitable for SDSL and subsequent EPR studies, genetic mutation for cys-introduction is not necessary. However the presence of multiple potential labelling sites within a monomer of a dimeric protein can lead to confusion when it comes to analysis of DEER data, for example, when probing conformational change; this is due to the possibility of multiple inter-spin distances which can result in datasets which are more difficult to analyse for specific distances (Figure 3.11B). Spin labelling attempts and

cw-EPR measurements were initially conducted using WT MamM-CTD (residues 215-318 of the full MamM gene) and will be presented within the next sections of the thesis. As will be discussed, results with the WT protein ultimately lead to the use of a C267S variant (simply termed C267S). C267S was then utilised for the further analytical EPR experiments discussed.

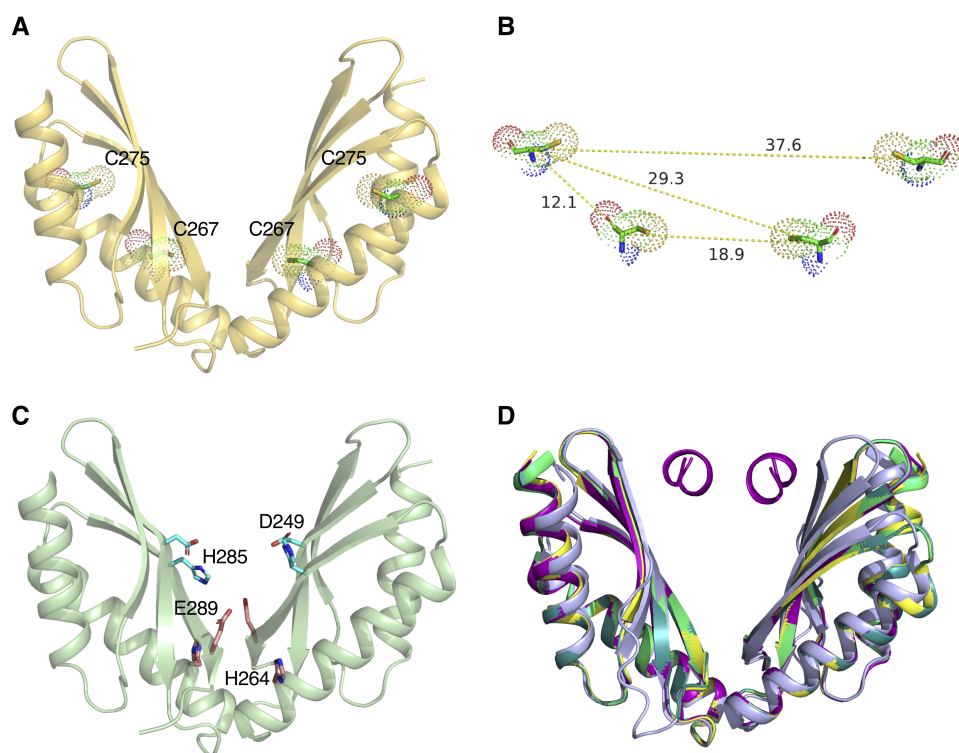


Figure 3.11: MamM-CTD crystal structures. All figures and measurements have been carried out using PyMol - The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. (A) The solved crystal structure of dimeric WT MamM-CTD (PDB ID: 3W5X) (cartoon representation), with native cysteines labelled and highlighted (sticks/dots). (B) $C\alpha$ - $C\alpha$ distance measurements between native cysteines. (C) Postulated binding site residues of the central site (D249 and H285) and peripheral sites (H264 and E289) are indicated (sticks). (D) Crystal structures of C267S (magenta, PDB ID: 6G55), PSM (deep teal, PDB ID: 6G64), PSM-C267S (lime green, PDB ID: 6G5E) and CSM (lilac, PDB ID: 3W8P) overlaid with the WT structure (yellow, PDB ID: 3W5X).

Three putative metal binding sites have already been described for MamM-CTD; these include two symmetrical peripheral sites (PSs) comprising histidine 264 and glutamate 289 of different monomers, and a central site (CS) comprising aspartate 249 and histidine 285 of both monomers (Figure 3.11C). Based on these predictions, two additional MamM-CTD variants have been utilised in the work described in this chapter for probing of binding-induced conformational change, binding stoichiometry and allosteric effects. These variants will be termed PSM-C267S (C267S with H264A and E289A mutations) and CSM-C267S (C267S with D249A and H285A mutations), which abolish the peripheral and central binding sites respectively. X-ray crystallography and structural determination has been attempted for each of the following variants in addition to WT and C267S MamM-CTD: PSM-C267S, PSM (lacking the C267S mutation), CSM-C267S and CSM (lacking the C267S mutation).

Structures were successfully solved for all variants in their apo states besides CSM-C267S MamM-CTD, for which a structure could not be obtained (Figure 3.11D).¹¹⁵ All variants for which structural models have been determined (based on X-ray diffraction experiments) display the same metallochaperone-like fold found in both WT MamM-CTD and other CDF CTDs, as well as the same dimeric packing and conformation as in the WT protein.¹¹⁵

MMM analysis

To assess the suitability of cysteine sites for spin labelling in proteins, *in silico* modelling of spin label attachment has been carried out using MMM (Multiscale Modelling of Macromolecules)²¹ - an open-source toolbox implemented in Matlab® (The MathsWorks Inc., Natick, MA). A rotamer library is calculated for a specific spin label attached at a chosen position in the protein; this is done by utilising a pre-calculated set of rotamers for the sterically unrestricted label, followed by *in silico* attachment of these rotamers at a defined position based on the solved three-dimensional model of the protein in question. The interaction energy for each rotamer with the surrounding protein is calculated, which affords a result in which only rotamers without substantial steric interaction with the surrounding protein are allowed. A higher number of calculated rotamers corresponds to a more accessible labelling site and hence the attached label would be mobile here, whereas a low number of rotamers or a low value for the partition function (another output of this analysis) indicates difficulty in labelling a site due to steric hindrance. An estimation of inter-spin distance can then also be calculated for multimeric proteins/proteins labelled at more than one site based on the average rotamer positions, which gives an idea as to whether or not the selected site/s are suitable for inter-spin distance determination using EPR methods. As already mentioned, the most common nitroxide-based spin labels used for labelling proteins include (MTSSL) and 3-(2-iodoacetamido)-proxyl (IAP), both of which can be used within MMM for *in silico* label attachment. Table 3.1 gives the results of MMM analysis including the number of rotamers, partition function and predicted inter-spin distances (the spin label position chosen for each distance calculation is given under 'Label position') for each WT and C267S MamM-CTD structural models labelled with MTSSL and IAP. Results have not been included for PSM or CSM structural models since their three-dimensional structures are so similar to those of WT and C267S MamM-CTD.

These results suggest that labelling of cys267 would likely be less successful than labelling of cys275; this is indicated by both the lower number of rotamers and a lower value for the partition function which would suggest a more sterically restricted labelling site.

PDB code	Variant	Label position	Number rotamers		Partition function		r (nm)	
			MTSSL	IAP	MTSSL	IAP	MTSSL	IAP
3W5Y	WT	267	21	13	0.3	0.42	1.03	1.68
3W5Y	WT	275	60	38	0.4	0.57	4.00	4.32
6G55	C267S	275	65	38	0.43	0.65	4.02	4.40

Table 3.1: Results of MMM analysis using structural models for WT and C267S MamM-CTD, including the predicted number of rotamers, partition functions and average inter-spin distances.

When considering these labelling sites for probing conformational change using inter-spin distance determination, position cys275 is again more favourable. This is because the DEER experiment is only accurate for distances between approximately 2-8 nm. Therefore an average predicted inter-spin distance of 1.03 nm or 1.68 nm between MTSSL- or IAP-labelled cys267 positions, respectively, is too short for accurate distances based on DEER experiments. Figure 3.12 gives the output of MMM analysis with predicted rotamers of attached MTSSL/IAP shown for each site of the WT MamM-CTD structural model.

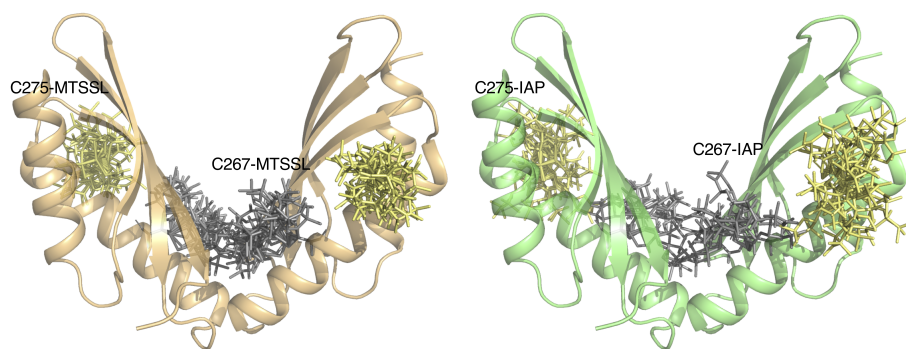


Figure 3.12: WT MamM-CTD structural models built using the published crystal structure (PDB ID: 3W5Y), made using PyMol. Dimeric apo MamM-CTD is shown in cartoon representation with calculated rotamer libraries of MTSSL (A) and IAP (B) attached at positions 267 and 275 shown as grey and yellow sticks, respectively. *In silico* modelling of rotamer libraries was performed using MMM.²¹

SDSL and labelling efficiencies

Preliminary investigations were focused upon identifying the most effective spin label for attachment to MamM-CTD prior to further EPR investigations. Whilst MMM analysis gives an initial indication for the likelihood of labelling at a particular site, this *in silico* model is based upon the singular conformation captured within whichever crystallographic structure is used, and cannot predict how labelling might change with conformational fluctuations such as those previously predicted for MamM-CTD in its apo state.¹¹³ In the WT protein, spin label is assumed to react with and bind to both cys275 and cys267 of each monomer; in all other variants containing the C267S background label attachment is only possible at cys275.

Solution state continuous wave- (cw-)EPR experiments (measured at room temperature) following SDSL provide a means for determination of the labelling efficiency - therefore labelling with MTSSL and IAP can be somewhat quantitatively compared. The approximate concentration of spin label in a sample can be calculated via comparison of its signal intensity with that of known standards. By estimating the expected concentration of spin label at 100 % labelling, based on the protein concentration and number of labelling sites, the ratio of actual label concentration versus expected label concentration is calculated, giving the labelling efficiency. Table 3.2 gives the (approximate) calculated labelling efficiencies which were routinely found for MamM-CTD variants with both MTSSL and IAP.

MamM-CTD variant	Labelling efficiency (%)	
	MTSSL	IAP
WT	25	15
C267S	50	40
PSM-C267S	50	40
CSM-C267S	40	15

Table 3.2: Results of SDSL labelling efficiencies for MamM-CTD variants using both MTSSL and IAP spin labels.

Labelling attempts with IAP gave consistently lower efficiencies whilst few labelling attempts have exceeded 50% efficiency with either label. Interestingly, labelling of WT is always approximately half as efficient as that of C267S - this could indicate that labelling at the additional cys267 position is very poor, or alternatively that labelling at one position in fact prohibits labelling at the other. In either case labelling of WT would make further EPR experiments and subsequent analyses more complex since it is not possible to definitively say which position is labelled. Labelling efficiencies for CSM-C267S (containing only the peripheral binding sites) are also continually lower than for C267S and PSM-C267S. This might suggest the presence/favouring of different conformations in its apo state which could limit site accessibility - this was also inferred recently by MD simulations and fluorescence measurements.¹¹⁵ Here it was suggested that removal of bulky and charged residues of the central binding site allows the monomeric arms to move closer to one another in conformations that are not possible in the case of WT or other constructs. Additionally, CSM-C267S is the only variant of those described here whose three-dimensional structure could not be solved, perhaps hinting towards a different apo conformation.

cw-EPR data collected at X-band frequency and room temperature for all four variants following SDSL with MTSSL/IAP is shown in Figure 3.13, in which the signal intensity is plotted as a first derivative. Solution state spectra also give information on the degree of movement exhibited by the attached label, afforded by the flexibility of linker regions within

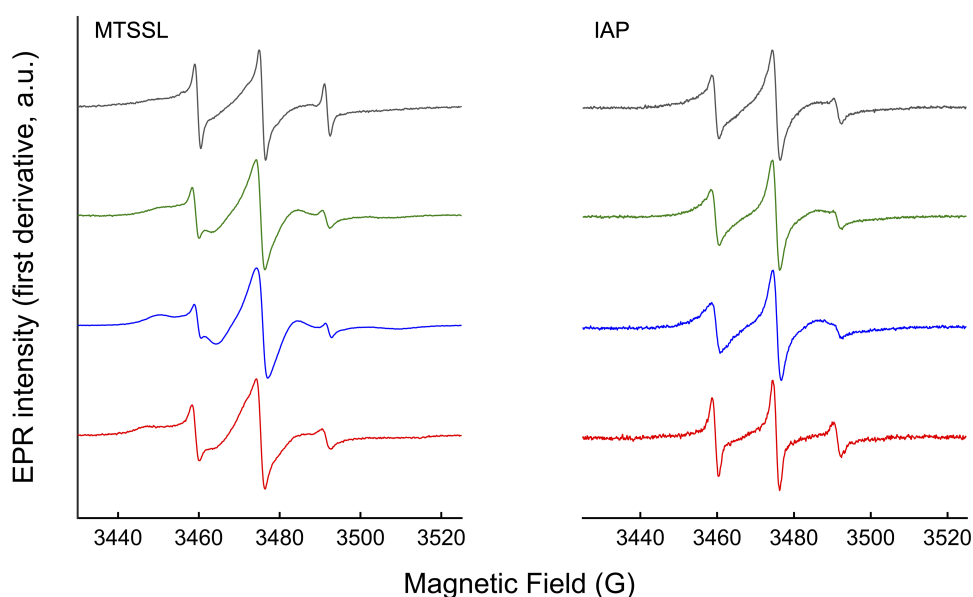


Figure 3.13: Room temperature continuous wave EPR measurements for WT (grey), C267S (green), PSM-C267S (blue) and CSM-C267S (red) MamM-CTD labelled with MTSSL (left) and IAP (right). Spectra have been normalised to equal intensities and figures were prepared using MatLab®. Experimental parameters: (IAP) microwave frequency = 9.757 GHz, centre field = 3475 G, sweep width = 100 G, modulation frequency = 100 KHz, modulation amplitude = 1 G, microwave power = 0.126 mW, number averages = 16-32. (MTSSL) microwave frequency = 9.760 GHz, centre field = 3478 G, sweep width = 110 G, modulation frequency = 100 KHz, modulation amplitude = 1 G, microwave power = 0.3955 mW, number averages = 36-56.

both MTSSL and IAP. As mentioned in section 1.3, broadening of spectral features in EPR spectra can be a result of an increase in anisotropy; this corresponds to a decrease in rotational movement of the attached spin label. Typically this is quantified by a parameter known as the rotational correlation time (T_c) which can be extracted via spectral simulation and fitting of room temperature spectra. The overall mobility of the spin label is reflective not only of the local environment in which it resides, but also the overall motions of the protein. Hence, the rotational correlation time is made up of three factors: (1) the rotational correlation time of the protein as a macromolecule in solution; (2) the backbone motions of the protein at the position of spin label attachment; and (3) the rotational motion of the spin label side chain about its point of attachment. The cw-EPR line shapes of large (>50 kDa) soluble proteins and membrane-bound proteins are dominated by the latter two.¹⁵⁷ Additionally, broadening can be a result of an increase in the hyperfine coupling constant, A , the magnitude of which is easily obtained by spectral simulation and fitting of LT spectra.

Since labelled C267S, PSM-C267S and CSM-C267S MamM-CTD are reporting on the same site, any spectral differences might indicate alternative conformational states and mobility exhibited in solution. Focusing on MTSSL-labelled variants, the spectral line shapes exhibited by each labelled variant are all slightly different (Figure 3.14A). The broader line shapes (as shown by the dotted lines) of both PSM-C267S and CSM-C267S spectra might

be indicative of a more restricted labelling site, or possibly different apo conformations compared with those exhibited by C267S, without binding site mutations. Qualitatively it can be seen that multiple spin label species contribute to the spectrum of each of these variants; an example of a spectrum comprised of multiple contributing species exhibiting different rotational correlation times (T_c) has been simulated in Figure 3.14B.

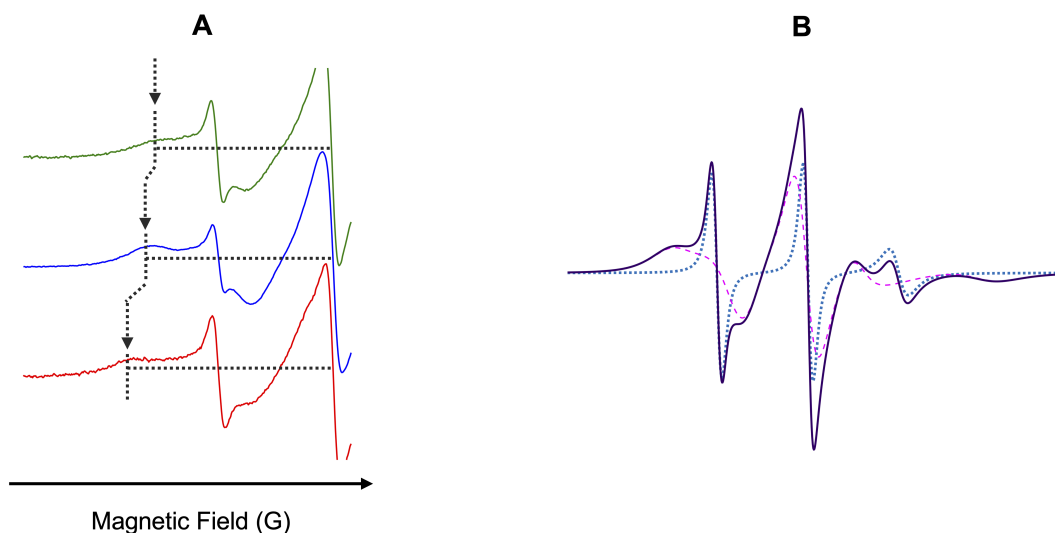


Figure 3.14: (A) Room temperature continuous wave EPR measurements carried out on C267S (green), PSM-C267S (blue) and CSM-C267S (red) MamM-CTD are presented between 3440-3480 G, highlighting broadened features in line shapes. (B) Spectral simulation and addition of nitroxide species with rotational correlation times of 1 ns (blue) and 6 ns (magenta), forming a multicomponent spectrum (purple). Individual spectra were simulated using EasySpin, an open-source Matlab[®] toolbox (microwave frequency = 9.75 GHz).

For quantitative analysis of SL mobility (reflected by the value for T_c) it is ideal to first obtain accurate values for the hyperfine coupling constant (A) along the defined z-axis through spectral simulation and fitting of spectra measured at cryogenic temperatures. At X-band frequency spectral anisotropy is dominated by the hyperfine coupling constant A_{zz} . As was explained in the section 1.3, whilst all hyperfine splittings are present within the LT (anisotropic) X-band spectrum, it is only the contribution along the z-axis which can be resolved, since the value for $2A_{zz}$ defines the spectral width between the centre of the outer-most peaks. Knowledge of this value can then be used in spectral simulation and fittings of RT spectra and kept constant, such that any further broadening can only be solved by fitting for T_c . X-band, low temperature (LT) cw-EPR spectra and associated simulated fits for apo, MTSSL-labelled C267S (C267S_{SL}), PSM-C267S (PSM-C267S_{SL}) and CSM-C267S (CSM-C267S_{SL}) are presented in Figure 3.15, along with the corresponding fitted parameters (Table 3.3). The simulations were fitted to the experimental data, through iterations of slight modifications to the input values, to reach a final fitting and extraction of these final best-fit parameters. It is noted here that well-documented, typical values for nitroxide g-tensors¹⁵⁸

were applied, and were not allowed to vary as part of any fitting in this results section. This is based upon the fact that g_{xx} , g_{yy} and g_{zz} cannot be resolved at X-band frequency, however it is reasonable to assume that the intrinsic g-tensor values in each case should not vary significantly from those typical of nitroxides. In cases where the central simulated resonance positions did not align with those of the experimental data, the centre field used for simulation was adjusted by ± 1 G; this accounts for the difference in field strength felt by the sample relative to the Hall probe (which monitors the experimental magnetic field strength). This solution was chosen over allowing variance in g-tensors, since these should not deviate between samples. All fittings and simulations of cw-EPR data in this chapter have been carried out using the EasySpin package¹⁵⁹ for Matlab®. In each case one single nitroxide component was used for spectral fitting (as is the case for all LT spectra fitted in this chapter).

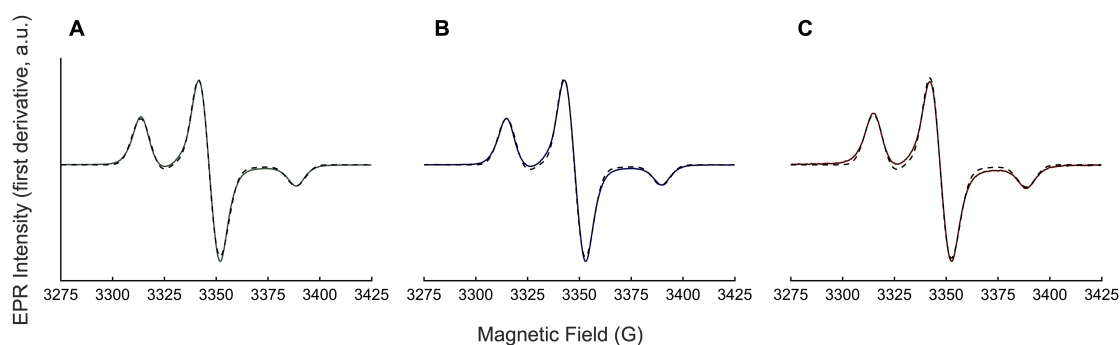


Figure 3.15: Low temperature (50 K) continuous wave X-band EPR spectra measured for MTSSL-labelled C267S (A), PSM-C267S (B) and CSM-C267S (C) MamM-CTD are plotted between 3275-3425 G, with associated spectral fits overlaid as black dashed lines. Experimental parameters: microwave frequency = 9.394 GHz, centre field = 3350 G, sweep width = 200 G, modulation frequency = 100 KHz, modulation amplitude = 4 G, microwave power = 0.01006 mW, number averages = 1.

MamM-CTD variant	g_{xx}	g_{yy}	g_{zz}	A_{xx} (MHz)	A_{yy} (MHz)	A_{zz} (MHz)	Linewidth (peak-peak, mT)	RMSD
C267S _{SL}	2.008	2.006	2.002	11	18	105	0.90	0.0148
PSM-C267S _{SL}	2.008	2.006	2.002	14	15	104	0.90	0.0179
CSM-C267S _{SL}	2.008	2.006	2.002	10	19	103	0.90	0.0255

Table 3.3: Best fit simulation values of the x,y,z components of both g- and A-tensors, as well as the linewidth, for low temperature experimental X-band EPR data measured on MamM-CTD variants labelled with MTSSL. The RMSD of each simulated fit is also given in the table.

Resolved A_{zz} values have been applied to multicomponent simulated fittings of the equivalent room temperature spectra. The resulting simulated fits are shown in Figure 3.16 plotted on top of the raw experimental data. In each case the raw data was taken and fitted with a two-component simulation script describing two differing nitroxide environments - one much less mobile (i.e. exhibiting a longer T_c) and more predominant than the other. Associated values of each fitting are highlighted in Table 3.4.

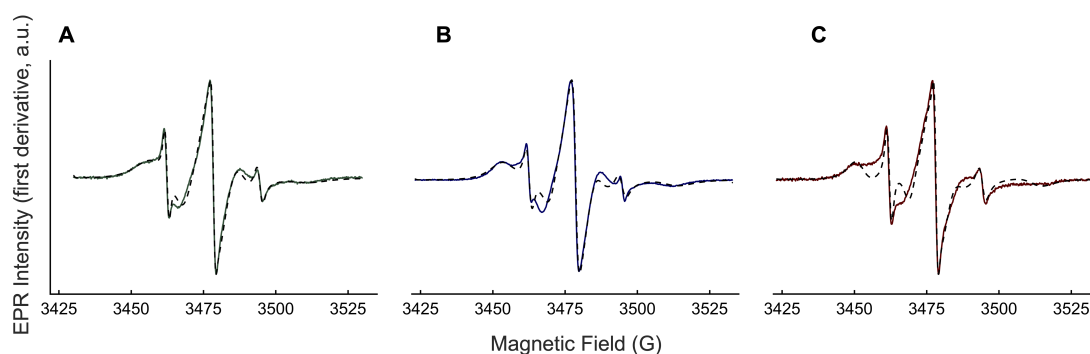


Figure 3.16: Room temperature continuous wave X-band EPR spectra measured for MTSSL-labelled C267S (A), PSM-C267S (B) and CSM-C267S (C) MamM-CTD (apo states) are plotted between 3424-3529 G, with associated spectral fits overlaid as black dashed lines. Experimental parameters were given previously.

C267S_{SL} (RMSD = 0.0330)

Component	g_{xx}	g_{yy}	g_{zz}	A_{xx} (MHz)	A_{yy} (MHz)	A_{zz} (MHz)	Linewidth (peak- peak, mT)	T_c (ns)	Relative weighting
1	2.008	2.006	2.002	11	18	105	0.08	4.90	0.94
2	2.008	2.006	2.002	14	17	104	0.10	0.67	0.06

PSM_{SL} (RMSD = 0.0409)

1	2.008	2.006	2.002	14	15	104	0.09	7.47	0.93
2	2.008	2.006	2.002	14	17	104	0.12	0.79	0.07

CSM_{SL} (RMSD = 0.0638)

1	2.008	2.006	2.002	11	18	103	0.10	10.80	0.94
2	2.008	2.006	2.002	14	17	105	0.12	0.78	0.06

Table 3.4: Best fit simulation values are given for apo, MTSSL-labelled C267S (C267S_{SL}), PSM-C267S (PSM-C267S_{SL}) and CSM-C267S (CSM-C267S_{SL}) spectra, measured at X-band and room temperature. Values include the x,y,z components of both g- and A-tensors, as well as linewidths and averaged rotational correlation times for the two spin label species simulated and contributing to each spectrum. The relative weightings of each species are also shown. The RMSD of each simulated fit is given above each set of results.

In each case the more mobile component constitutes less than 10% of the entire spectrum, with sub-nanosecond T_c values attributable to non-specifically bound or entirely free spin label tumbling in solution. Notably each spectrum in Figure 3.16 can be satisfactorily reproduced using very similar A_{zz} values (i.e. within $\sim 2\%$ of each other) for the less mobile component (which represents bound label). Thus we can conclude that the broadened spectral features seen for the binding site variants are explained by varying of the rotational averaging (T_c) between samples. A values were allowed to vary in the final iteration of simulated fits to find the local minima of fittings, and clearly do not change significantly from values deduced from LT spectra.

3.5.2 Zn^{2+} binding probed by cw- and pulsed EPR techniques

Within the 18 clades of CDF proteins recognised, a number of different divalent metal cations are known to be transported. Each CDF protein binds and transports specific metal cations typically by utilising proton motive force and undergoing conformational changes. The natural substrate of MamM is iron, whilst biophysical studies have implicated metal binding-induced conformational change upon binding of Fe^{2+} and Zn^{2+} , with a decrease in function observed both *in vitro* and *in vivo* upon mutation of either proposed central or peripheral binding sites.¹¹³ *In vitro* studies of MamM-CTD metal binding were only successful at basic pH's due to the histidine residues participating in metal ion binding; such conditions proved non-ideal for Fe^{2+} studies with MamM-CTD considering the extent of $\text{Fe}(\text{OH})_2$ precipitation at basic pH. Additionally, oxidation to Fe^{3+} occurs readily under aerobic conditions. In such cases, exact determination of Fe^{2+} concentration is not feasible which limits the application of biophysical techniques; for example analysis of binding stoichiometry and kinetics would not be possible. Thus our initial studies of MamM-CTD metal binding were conducted with Zn^{2+} , which was previously shown to bind MamM-CTD. Both cw-EPR and DEER measurements were conducted with C267S_{SL}, PSM-C267S_{SL} and CSM-C267S_{SL} for probing of changes in local spin label environment, binding-induced conformational change as well as binding stoichiometry. Zn^{2+} is a diamagnetic ion and hence does not have an EPR signature - for the time being this greatly simplifies things.

cw-EPR measurements

Cw-EPR datasets for C267S_{SL}, PSM-C267S_{SL} and CSM-C267S_{SL}, measured at RT using an X-band microwave frequency, are shown in Figure 3.17 both for the apo state and in the presence of a 5-fold excess of Zn^{2+} . As indicated by the dotted lines between spectra, clear broadening of the outer features (corresponding to the less mobile species, i.e. bound spin label (SL)) of the EPR signal is evident in all cases with Zn^{2+} addition. A broadened spectrum might be explained by an increase in hyperfine coupling (described by the hyperfine coupling constant A), or alternatively as a result of decreased SL mobility - this would be indicative of a change in the local SL environment causing restriction of its side chain mobility. This might occur as a result of large scale conformational change within the protein, hence placing the attached SL in a more buried and sterically hindered site. However, considering the position of the SL on the periphery of the protein (i.e. being solvent exposed) this seems unlikely. Alternatively, more restricted backbone movements at the position of SL attachment as a result of metal ion binding might contribute to this

broadening. It might be possible that the position of bound metal ions sterically restricts SL motion, however this again seems unlikely considering their solvent exposed positions. In any case, such a significant change strongly indicates conformational change of MamM-CTD as a function of Zn^{2+} binding, and serves as our first EPR 'marker' for metal binding.

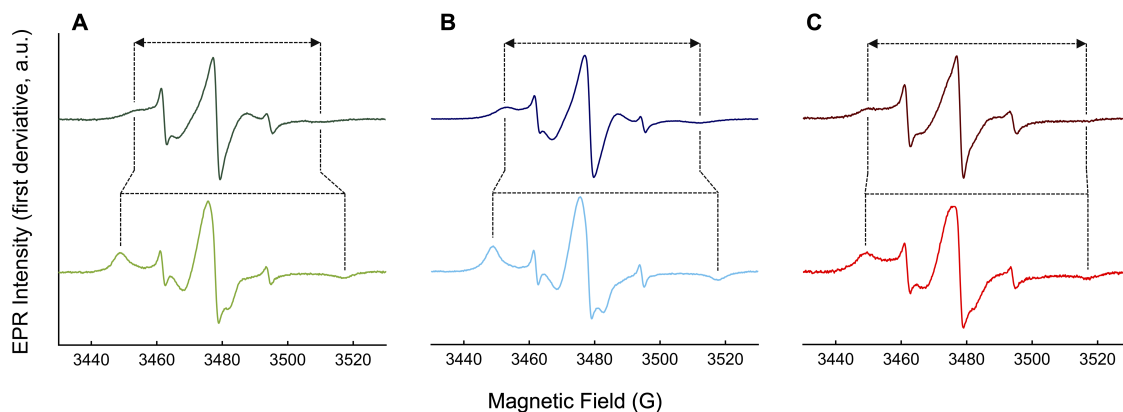


Figure 3.17: Room temperature continuous wave EPR measurements performed on (A) C267S_{SL}, (B) PSM-C267S_{SL} and (C) CSM-C267S_{SL} MamM-CTD variants under apo (top) and excess Zn^{2+} (5-fold) conditions. Dashed lines highlight spectral broadening as a result of Zn^{2+} addition. Experimental parameters (Zn^{2+}): microwave frequency = 9.760 GHz, centre field = 3478 G, sweep width = 110 G, modulation frequency = 100 KHz, modulation amplitude = 1 G, microwave power = 0.3955 mW, number averages = 36-56.

Again, as a pre-requisite for accurate spectral fitting and parameter extraction, equivalent spectra measured at cryogenic temperatures were simulated and fitted in order to obtain values for A_{zz} . These experimental datasets along with the associated simulated fits are shown in Figure 3.18, with the corresponding fitted parameters highlighted in Table 3.5. Upon comparison to the values obtained from spectral fittings in the absence of Zn^{2+} , it is evident that the magnitude of A_{zz} coupling does not dramatically change between states.

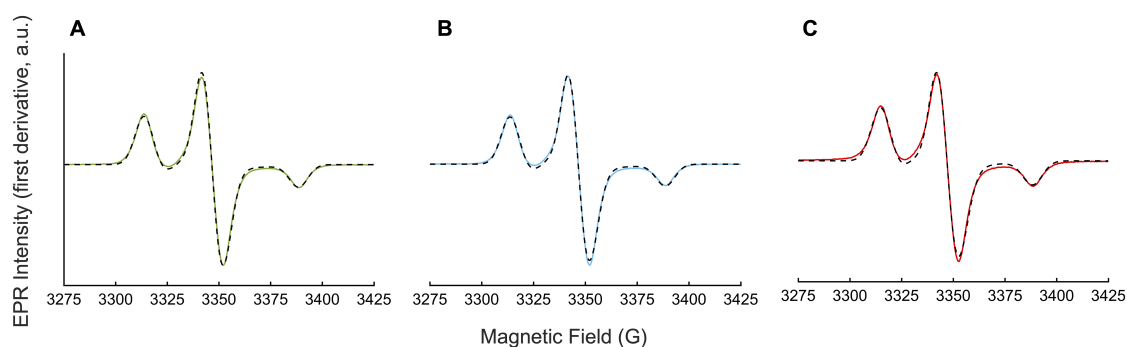


Figure 3.18: Low temperature (50 K) continuous wave X-band EPR spectra measured for MTSSL-labelled C267S (A), PSM-C267S (B) and CSM-C267S (C) MamM-CTD incubated with Zn^{2+} are plotted between 3275-3425 G, with associated spectral fits overlaid as black dashed lines. Experimental parameters: microwave frequency = 9.390 GHz, centre field = 3350 G, sweep width = 200 G, modulation frequency = 100 KHz, modulation amplitude = 3G, microwave power = 0.003157 mW, number averages = 4.

MamM-CTD variant	g_{xx}	g_{yy}	g_{zz}	A_{xx} (MHz)	A_{yy} (MHz)	A_{zz} (MHz)	Linewidth (peak-peak, mT)	RMSD
C267S _{SL} + Zn ²⁺	2.008	2.006	2.002	13	15	104	0.90	0.0177
PSM-C267S _{SL} + Zn ²⁺	2.008	2.006	2.002	11	17	105	0.90	0.0149
CSM-C267S _{SL} + Zn ²⁺	2.008	2.006	2.002	9	20	102	0.95	0.0227

Table 3.5: Best fit simulation values of the x,y,z components of both g- and A-tensors, as well as the linewidth for low temperature experimental X-band EPR data measured on MamM-CTD variants labelled with MTSSL, following incubation with Zn²⁺. The RMSD of each simulated fit is also given in the table.

In the same way as was previously shown for the apo spectra, resolved A_{zz} values obtained via spectral fitting of LT Zn²⁺-bound spectra have been applied to multicomponent simulated fittings of equivalent room temperature spectra. Resulting simulated fits are shown in Figure 3.17 plotted on top of the raw experimental data. In each case the raw data was taken and fitted with a two-component simulation script describing two differing nitroxide environments - one much less mobile (i.e. exhibiting a longer T_c) and more predominant than the other. The associated values of each fitting are highlighted in Table 3.6. It has already been established that the value for A_{zz} does not change significantly between apo and Zn²⁺-bound states, and so the extent Zn²⁺-induced broadening of the most prevalent nitroxide species should be reflected in the value for T_c alone. Indeed upon comparison of the average T_c values extracted between apo and Zn²⁺-bound states, there is a clear increase in rotational correlation time of the predominant species, indicating more restricted motion of the attached spin label species and thus, most likely, some conformational change.

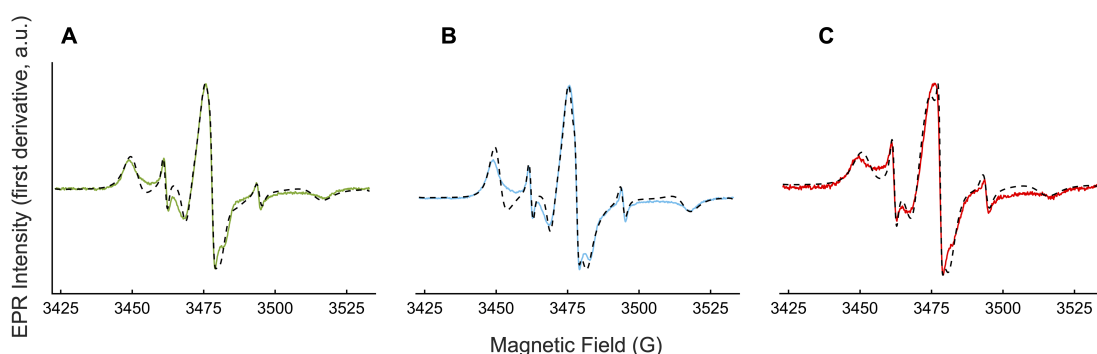


Figure 3.19: Room temperature continuous wave X-band EPR spectra measured for C267S_{SL} (A), PSM-C267S_{SL} (B) and CSM-C267S_{SL} (C) MamM-CTD following incubation with Zn²⁺ are plotted between 3424-3529 G, with associated spectral fits overlaid as black dashed lines. Experimental parameters were given previously.

C267S_{SL} + Zn²⁺ (RMSD = 0.0471)

Component	g_{xx}	g_{yy}	g_{zz}	A_{xx} (MHz)	A_{yy} (MHz)	A_{zz} (MHz)	Linewidth (peak- peak, mT)	T_c (ns)	Relative weighting
1	2.008	2.006	2.002	11	18	105	0.13	17.11	0.95
2	2.008	2.006	2.002	14	20	105	0.11	0.60	0.05

PSM_{SL} + Zn²⁺ (RMSD = 0.0557)

1	2.008	2.006	2.002	14	15	104	0.12	18.48	0.98
2	2.008	2.006	2.002	15	17	105	0.10	0.40	0.02

CSM_{SL} + Zn²⁺ (RMSD = 0.0635)

1	2.008	2.006	2.002	11	18	103	0.25	25.10	0.94
2	2.008	2.006	2.002	14	17	105	0.08	0.71	0.06

Table 3.6: Best fit simulation values are given for C267S_{SL}, PSM-C267S_{SL} and CSM-C267S_{SL} spectra following incubation with Zn²⁺, measured at X-band and room temperature. Values include the x,y,z components of both g- and A-tensors, as well as linewidths and averaged rotational correlation times for the two spin label species simulated and contributing to each spectrum. The relative weightings of each species are also shown. The RMSD of each simulated fit is given above each set of results.

DEER studies for distance determination

For a quantitative measure of Zn²⁺ binding-induced conformational change in MamM-CTD variants, DEER experiments were employed. As mentioned in section 1.4, such experiments are typically carried out at cryogenic temperatures in order to increase the spin-spin relaxation time (T_2) and hence lengthen the amount of time over which the Hahn echo can be probed and observed, ultimately allowing longer range dipole-dipole interactions to be measured. Low temperatures also minimise the spin-lattice relaxation time (T_1); this altogether affords longer measurements in the time domain with high repetition rates.¹⁶⁰ At cryogenic temperatures T_2 relaxation is largely due to nuclear hyperfine interactions with neighbouring protons, causing spin diffusion, and hence deuterated solvents are typically utilised. Additionally, cryo-protectants such as glycerol are routinely used in sample preparation of proteins for LT experiments to prevent ice crystal formation and protein aggregation, both of which severely minimise T_2 .¹⁶¹ Such measures also allow data collection of sufficient quality using a lower concentration of protein, which is extremely useful when working with proteins that express poorly (as is often the case with membrane transporters) or are unstable, for example. Initial pulsed investigations of MamM-CTD variants therefore involved optimisation of sample preparation based on 2-pulse electron spin echo envelope modulation (2P-ESEEM) experiments and T_2 analysis.

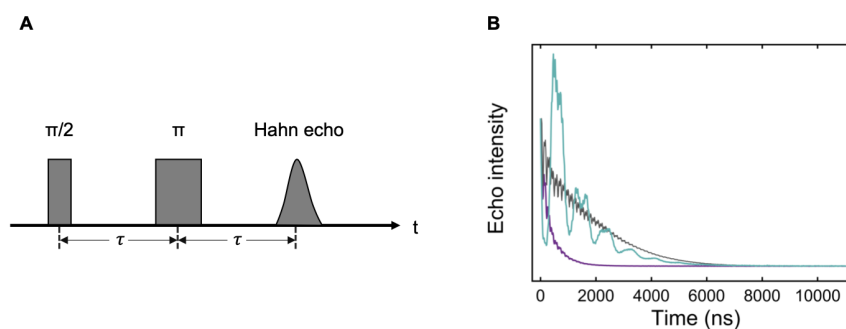


Figure 3.20: (A) The pulse sequence for 2P-ESEEM measurements. (B) 2P-ESEEM experiments (50 K) conducted on C267S_{SL} samples in protonated solvent and without any cryo-protectant (purple), in protonated solvent using glycerol as a cryo-protectant (grey) and in deuterated solvent using d8-glycerol as a cryo-protectant (teal). Experimental parameters: field position = 3444 G, number of points = 700, shots per point = 50, shot repetition time = 2 ms, number averages = 4.

For investigation of optimal spin labelled MamM-CTD sample conditions for maximum T_2 , samples were prepared in the following ways: (1) in protonated solvent without cryo-protectants; (2) in protonated solvent with glycerol as a cryo-protectant; and (3) in deuterated solvent with d8-glycerol as a cryo-protectant. The results of 2P-ESEEM experiments carried out for each sample can be seen in Figure 3.20B. As mentioned in section 1.4, 2P-ESEEM is a pulsed EPR technique in which the electron spin echo amplitude is measured as the time interval between the two pulses is increased (see pulse sequence in Figure 3.20A), the decay of which is dominated by T_2 relaxation effects. This decay is a superposition of the decaying Hahn echo amplitude, which is unmodulated, and coherence transfer echo's caused by interactions with nearby magnetic nuclei, which is demonstrated by modulations in the decay at specific frequencies. The sum and difference combinations of the nuclear frequencies are also present in the 2P-ESEEM spectrum. The T_2 for a spin can then be quantitatively determined by fitting the unmodulated part of the decay. As can be seen in Figure 3.20B, the additive effects of glycerol and deuteriation are substantial with respect to the T_2 relaxation time. What is also evident is the difference in modulation frequencies between samples with and without deuterated glycerol/solvent. This corresponds to electron-proton versus electron-deuteron modulation, the frequency of which is dependent upon the Larmor frequency of the interacting nuclei. From these initial studies it was proven that optimal sample preparation for DEER experiments would require the use of both d8-glycerol as a cryo-protectant, and deuterated solvents.

The pulse sequence for the 4P-DEER experiment used throughout this work is shown in Figure 3.21A, in which pulses are applied at two different microwave frequencies. These frequencies are dependent upon where you wish to pump and observe within your nitroxide spectrum, and so the first step in preparation for the DEER experiment is to measure a

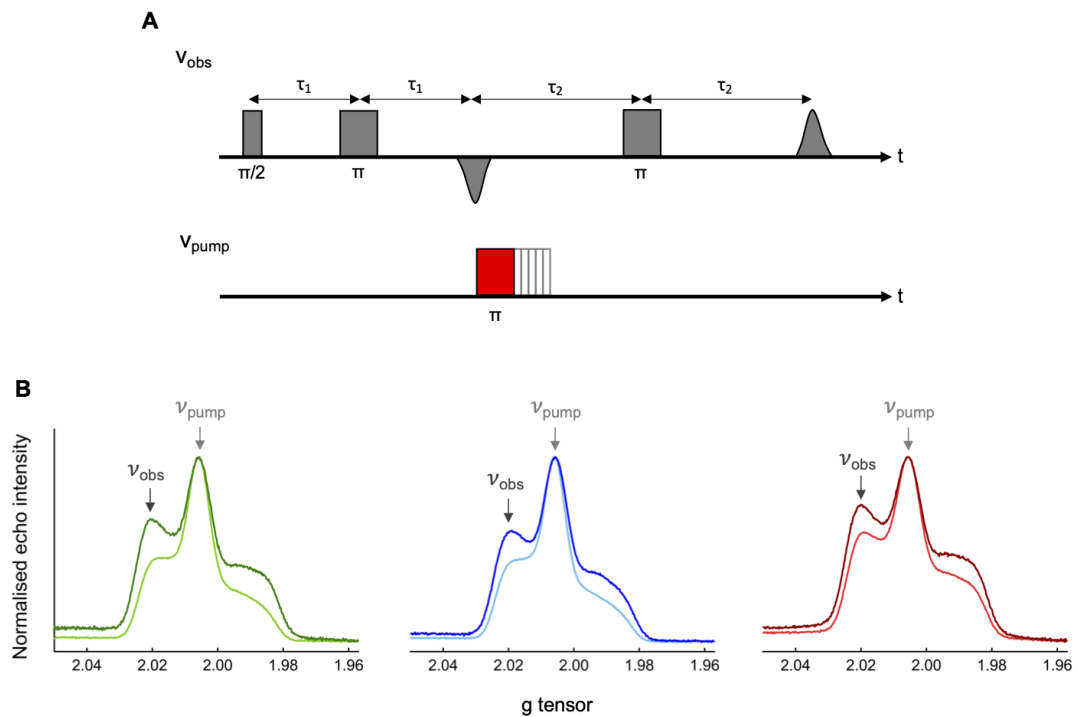


Figure 3.21: (A) The 4-pulse DEER sequence used for DEER measurements presented in this section. (B) Field swept echo spectra ($T = 50$ K) measured for C267S_{SL} (green, left), PSM-C267S_{SL} (blue, middle) and CSM-C267S_{SL} (red, right) MamM-CTD variants in apo and Zn²⁺-bound (5-fold excess) states, represented by light and dark colours respectively. Light and dark grey arrows represent pump and observe frequencies, respectively. Experimental parameters: centre field = 3440-3448 G, sweep width = 150-200 G, delay time (τ) = 120 ns, number of points = 500-600, shots per point = 50, shot repetition time = 2 ms, number of averages = 4.

field swept echo (FSE) of your sample. The relationship between the pump and observe frequencies is given by equation 3.1,

$$\nu_{pump} = \nu_{observe} \pm \Delta B * 2.83 \text{ MHz/G} \quad (3.1)$$

in which the magnetic field difference between the desired pump and observe positions is substituted for calculation of the pump frequency. Depending on whether you are pumping up- or down-field of your observe position, ν_{pump} will be lower (in the former case) or higher (in the latter case) than $\nu_{observe}$. Frequency offsets between 70-90 MHz were used in all DEER experiments described in this section. FSE data plotted for each variant in apo and Zn²⁺-bound states can be seen in Figure 3.21B, with arrows highlighting pump and observe positions for subsequent DEER measurements. Following FSE measurement, 2P-ESEEM measurements were undertaken on every sample to determine appropriate τ_1 and τ_2 values to be used in the DEER experiment - τ_2 determination also requires consideration of the experimental time required to obtain sufficient signal:noise for analysis, and in some cases a compromise must be made.

DEER measurements were recorded for each spin labelled variant, under both apo and Zn²⁺ excess conditions (5-fold) - the raw datasets are presented in Figure 3.22 as an exam-

ple of DEER data prior to any background-correction/fitting. Going forwards only processed (i.e. spectral fittings and resulting distance distributions) will be presented, for conciseness. DEER data processing was undertaken within DEERAnalysis¹⁹, a Matlab-based toolbox, using both Tikhonov regularisation and DEERNet.²⁰ The latter method has been more recently developed, and employs artificial neural networks for prediction of the inter-spin label distance distribution. Additionally the result includes a measure for the uncertainty of given distance distributions in the form of 95 % confidence bounds; results will be presented using this method for a clear indication of relevant distances where necessary.

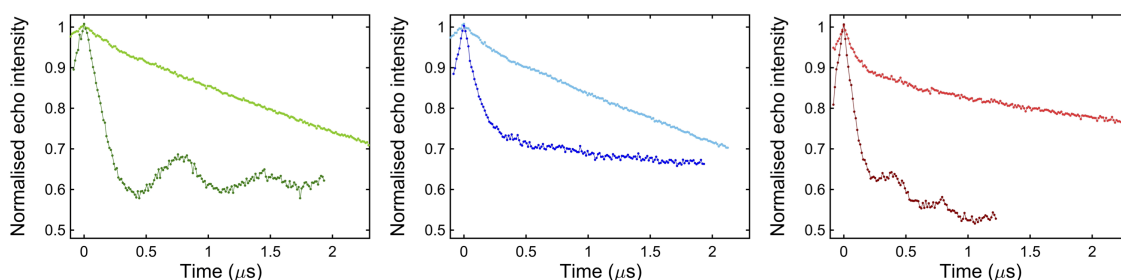


Figure 3.22: 4-pulse DEER experiments (50 K) for MamM-CTD spin labelled variants in their apo states, and following incubation with excess Zn^{2+} . Green, blue and red traces are the normalised raw dipolar evolution of C267S_{SL}, PSM-C267S_{SL} and CSM-C267S_{SL} MamM-CTD variants respectively. The traces for both apo and Zn^{2+} -bound states are presented in light and dark colours, respectively. Experimental parameters: frequency offset = 70 MHz, τ_1 = 386 ns (PSM-C267S_{SL} apo), 352 ns (C267S_{SL} apo), 120 ns (all others), τ_2 = 2000-4000 ns, shots per point = 50, shot repetition time = 2 ms.

MMM analysis predicts an inter-spin distance (using MTSSL) for the apoprotein of approximately 4.0 nm, in each C267S and binding site variants of MamM-CTD, based on their solved crystal structure models (see Table 3.1). DEER data obtained for each variant in both apo and Zn^{2+} -bound states can be seen in Figure 3.23, in which the upper panels show normalised background-corrected dipolar evolution data, whilst the lower panels present the corresponding distance distributions with 95 % confidence bounds represented by light grey areas around the lines. Both apo forms of C267S_{SL} and PSM-C267S_{SL} reveal minimal inter-spin interaction, which is evident from the low modulation depth of the time domain data (approximately 5 % in each case). This, combined with the lack of any dipolar modulations means a defined distance distribution between the labelled cysteines of the dimeric protein is not evident in either case. This is suggestive of a dynamic apo state, as opposed to a single preferred conformation such as that represented in crystal structure models. Flash-freezing of samples in which a protein is undergoing constant conformational change leads to a mixture of conformational states, and the sum of dipolar interactions between spins within these conformations results in a broad distribution as is seen here. The background-corrected dipolar evolution of CSM-C267S_{SL} however shows a steeper initial decay and increased modulation depth relative to the other apo cases - both are

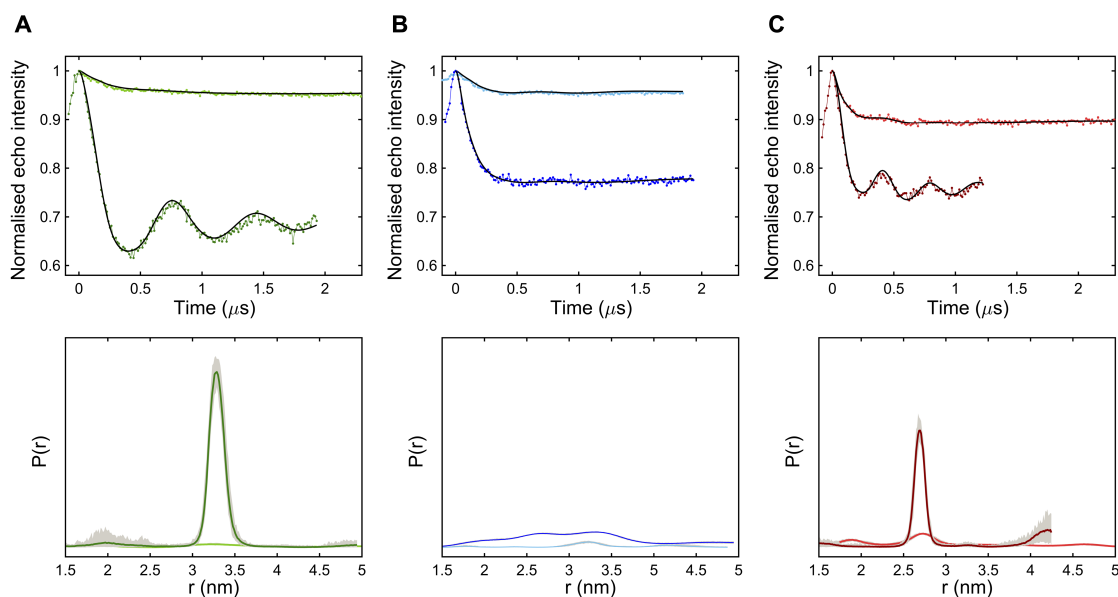


Figure 3.23: 4-pulse DEER experiments (50 K) for MamM-CTD spin labelled variants in their apo states, and following incubation with excess Zn^{2+} . (Top panel) Green, blue and red traces are the normalised background-corrected dipolar evolution of (A) C267S_{SL}, (B) PSM-C267S_{SL} and (C) CSM-C267S_{SL} MamM-CTD variants respectively. The traces for both apo and Zn^{2+} -bound states are presented in light and dark colours, respectively, with fittings shown as black lines. (Bottom panel) Distance distributions of apo and Zn^{2+} -bound MamM-CTD variants (A) C267S_{SL}, (B) PSM-C267S_{SL} and (C) CSM-C267S_{SL} shown in light and dark colours, respectively. Standard deviations in the form of 95 % confidence intervals are also given in each case and represented by light grey areas around the lines. Fitting of all dipolar evolutions and subsequent distance distributions were calculated with DEERNet.²⁰ Experimental parameters: frequency offset = 70 MHz, τ_1 = 386 ns (PSM-C267S_{SL} apo), 352 ns (C267S_{SL} apo), 120 ns (all others), τ_2 = 2000-4000 ns, shots per point = 50, shot repetition time = 2 ms.

indicative of increased interaction between spins over a shorter distance. The resulting distance distribution does give rise to a very small feature with an average distance of approximately 2.7 nm, however since the intensity of this distribution is low, it is likely that conformational equilibrium still remains.

Zn^{2+} addition (5-fold excess) results in large changes in the measured DEER traces and consequent distance distributions obtained for all variants. For C267S_{SL}, a single distance distribution centred at 3.3 nm was determined; this indicates that the conformational flexibility observed in the apo state is lost upon addition of and binding to Zn^{2+} , and that instead a single conformational state is now adopted. For PSM-C267S_{SL} with Zn^{2+} there is a clear increase in modulation depth, however the calculated distance distribution remains similarly broad when compared with the apo case. This could suggest that Zn^{2+} interacts in such a way that it affects the conformational movement exhibited by this variant, without inducing one single conformation in solution. In the case of CSM-C267S_{SL}, Zn^{2+} addition gives a single and narrow distance distribution with an average inter-spin distance of 2.7 nm and of much higher intensity than that seen in the apoprotein. Metal binding again appears to induce conformational 'locking' to a single preferred conformation, which in this case results in the proximity of the arms being much closer to one another than in the apo crystal structure. The DEER experiments discussed here have provided us with the first structural

insight into the dimeric conformations adopted by MamM-CTD variants in the presence of a divalent metal cation.

DEER studies to probe binding stoichiometry and allosteric effects

The modulation depth of DEER dipolar evolution traces is a concept which has already been mentioned when comparing apo vs Zn^{2+} -bound DEER datasets, since it is directly related to the number of dipolar-interacting spins within a sample. Therefore, analysis of the modulation depth as a function of substrate concentration (in a system with a distinct substrate-bound conformation) can serve as a tool for the estimation of binding stoichiometry. Stoichiometric Zn^{2+} titrations, with a DEER experiment carried out following each subsequent addition, were therefore carried out with both C267S_{SL} and CSM-C267S_{SL}, since binding-induced conformational change was evident in both cases, and both are also predicted to have more than one metal binding site. WT MamM-CTD Zn^{2+} binding stoichiometry was initially predicted using ITC at 4 metal ions per dimer.¹¹³

Results were obtained for C267S_{SL} upon varying the metal:protein ratio from 1:1 - 10:1, the results of which can be seen in Figure 3.24A. As indicated by the arrows, an increase in modulation depth is evident up to a 3:1 - 4:1 metal:protein ratio, at which point saturation is reached. The eventual decrease in modulation depth observed is likely due to the high metal ion excess which causes protein aggregation in solution. Results for CSM-C267S_{SL} were obtained with metal:protein ratios of 1:1 - 5:1 (see Figure 3.24B), considering the predicted binding stoichiometry of 2 ions per dimer - 1 at each putative peripheral site. In this case an increase in modulation depth is observed only up until a 2:1 metal excess, at which point saturation is already reached and no further increase is observed.

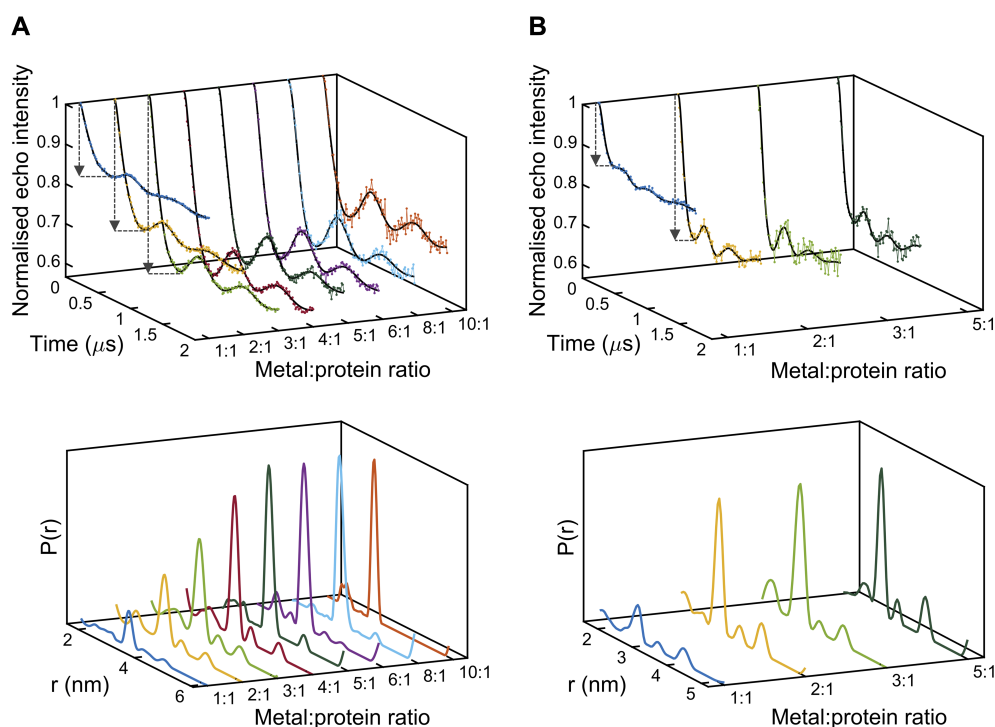


Figure 3.24: 4-pulse DEER experiments (50 K) for C267S_{SL} and CSM-C267S_{SL} following titration of Zn²⁺. (Top panel) The normalised background-corrected dipolar evolution of (A) C267S_{SL} and (B) CSM-C267S_{SL} following Zn²⁺ titration, with metal:protein ratios ranging from 1:1 - 10:1 and 1:1 - 5:1 excess, respectively (coloured lines). Spectral fittings are shown as black lines. Dashed dark grey lines/arrows indicate the modulation depth of their corresponding traces. (Bottom panel) Distance distributions of (A) C267S_{SL} and (B) CSM-C267S_{SL} corresponding to each titre of Zn²⁺. Fitting of all dipolar evolutions and subsequent distance distributions were calculated using Tikhonov regularisation. Experimental parameters were described previously.

ITC measurements have given negative values for the change in Gibb's free energy (ΔG) associated with Zn²⁺ binding to both WT and CSM MamM-CTD, with values equal to -25.71 and -27.11 kJ.mol⁻¹ respectively. This means binding should be spontaneous in solution. Additionally, calculated K_D 's indicate a slightly higher affinity for Zn²⁺ in the CSM variant as compared with the WT.¹¹⁵ Assuming similar thermodynamic properties for the C267S and CSM-C267S variants of MamM-CTD, binding of Zn²⁺ should be close to stoichiometric, and the concentration of apo protein within samples is below the detection limit of this technique.

3.6 Metal selectivity of MamM-CTD

Considering the diverse array of divalent metal cations known to be transported by CDF proteins, and phylogenetic studies which reveal substrate specificity within the CDFs, metal ion selectivity must be crucial for the proper functioning of these proteins. However, in addition to the exact transport mechanism employed by these transporters, factors which control metal selectivity in CDFs are not currently well understood. A specific question that remains is whether or not the CTD of CDF proteins plays a direct role in metal selectivity. Such knowledge is fundamental in building a full and clear understanding of the structure-function relationship exhibited in CDF proteins. The work presented in this section is part

of a collaborative effort towards extending our understanding of metal selectivity in CDF proteins by probing the role of MamM-CTD in metal binding and selectivity. Our EPR data will later be discussed with reference to other structural and biophysical methods employed in order to more fully assess such behaviour in MamM-CTD; these results present for the first time structural evidence of binding by any CDF CTD to multiple divalent metal cations.

3.6.1 Initial indications of metal interaction probed by cw-EPR

The metal selectivity experiments described in the following sections are limited to the C267S variant in order to gain a clear insight into such mechanisms exhibited in the WT protein, without any alterations to potential metal binding sites. Previously the results of both LT and RT cw-EPR measurements and simulated experimental fits were described for each C267S_{SL}, CSM-C267S_{SL} and PSM-C267S_{SL} in both apo and Zn²⁺-bound states. In each case, broadened outer spectral features were rationalised by a change in the rotational correlation time of the spin label, serving as our first indication of metal-protein interaction. Equivalent RT spectra are now presented for C267S_{SL} incubated with an excess of multiple other divalent metal cations, including Ni²⁺, Co²⁺, Cu²⁺ and Cd²⁺, the results of which are displayed in Figure 3.25A. Quite clearly the addition of each different metal ion has a quantifiable effect on the measured spin label EPR signature (reflected by changes in the outermost spectral features), which is representative of a change in the environment in which the spin label resides. This again is most likely due to the structural states of the protein in solution, as was also the case upon addition of Zn²⁺. These changes are highlighted in Figure 3.25A by the dashed lines running through the outermost features of the experimental plots. Spectra are stacked below the apo spectrum in order of decreasing K_D (increasing affinity) as inferred by recent ITC measurements,¹⁵⁴ and in each subsequent case broadening occurs to a different extent, making alternate metal-bound spectra distinct from one another. This effect can be quantified and described by simulated spectral fits, which will be presented shortly. Results with both Fe²⁺ and Mn²⁺ are not presented due to difficulties in room temperature sample preparation and measurement (with Fe²⁺), and spectral acquisition and analysis with paramagnetic Mn²⁺, which itself has an EPR signature at RT.

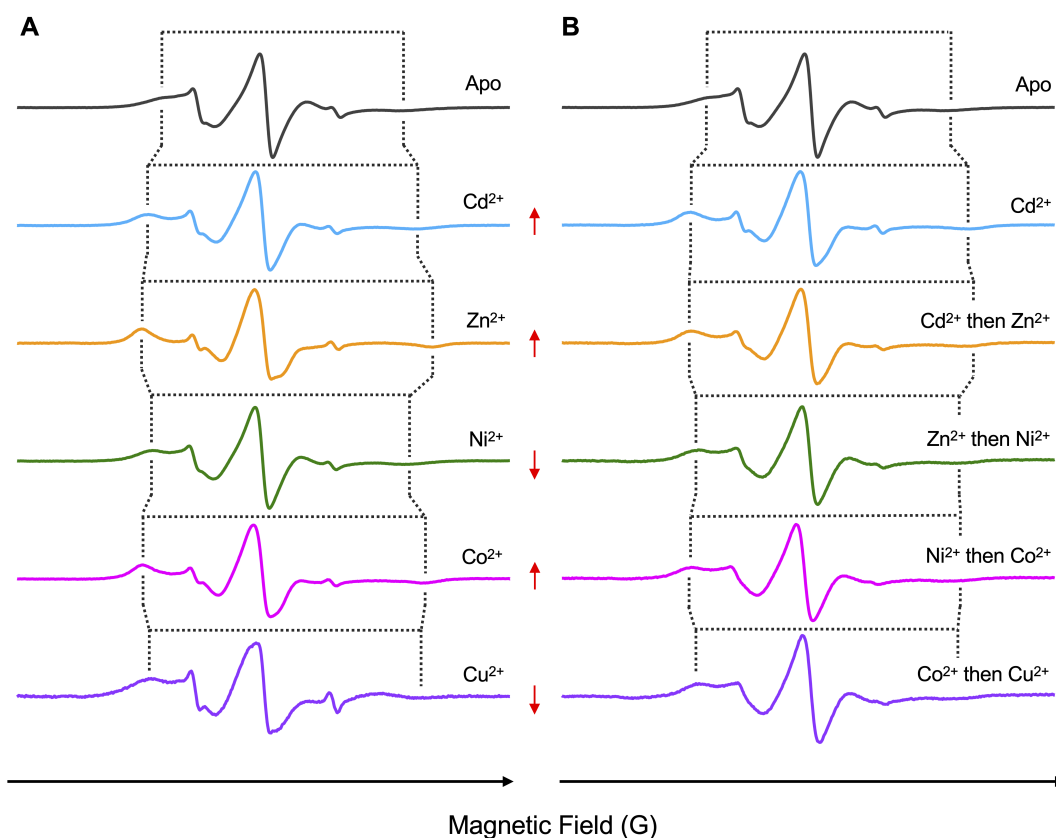


Figure 3.25: (A) Room temperature continuous wave X-band EPR spectra measured for C267S_{SL}, in its apo state, and following incubation with divalent metal cations, with central peaks aligned. (B) Room temperature continuous wave X-band EPR spectra for C267S_{SL}, in its apo state, and following a displacement titration of various divalent metal ions in order of decreasing K_D , with central peaks aligned. Arrows between spectra highlight the expected change in value of T_c with each successive metal ion. Experimental parameters: microwave frequency = 9.758 GHz, centre field = 3478 G, sweep width = 120 G, modulation frequency = 100 KHz, modulation amplitude = 2 G, microwave power = 0.3955 mW, number averages = 49.

A separate set of experiments was also undertaken in order to test these proposed metal binding affinities, based upon the observable change in broadening as a function of the K_D , which in theory should be replicated following competitive displacement of each metal. During this experiment metals were added in order of increasing affinity, with each subsequent addition and incubation followed by a wash step to remove excess displaced metal, prior to cw-EPR measurement. As a final addition, Cd^{2+} was re-added to the supposed Cu^{2+} -bound sample to ensure that spectral changes were a result of metal displacement as oppose to sample washing. The raw experimental results of this displacement experiment are shown in Figure 3.25B, in which the dashed lines highlight the observed pattern in spectral broadening.

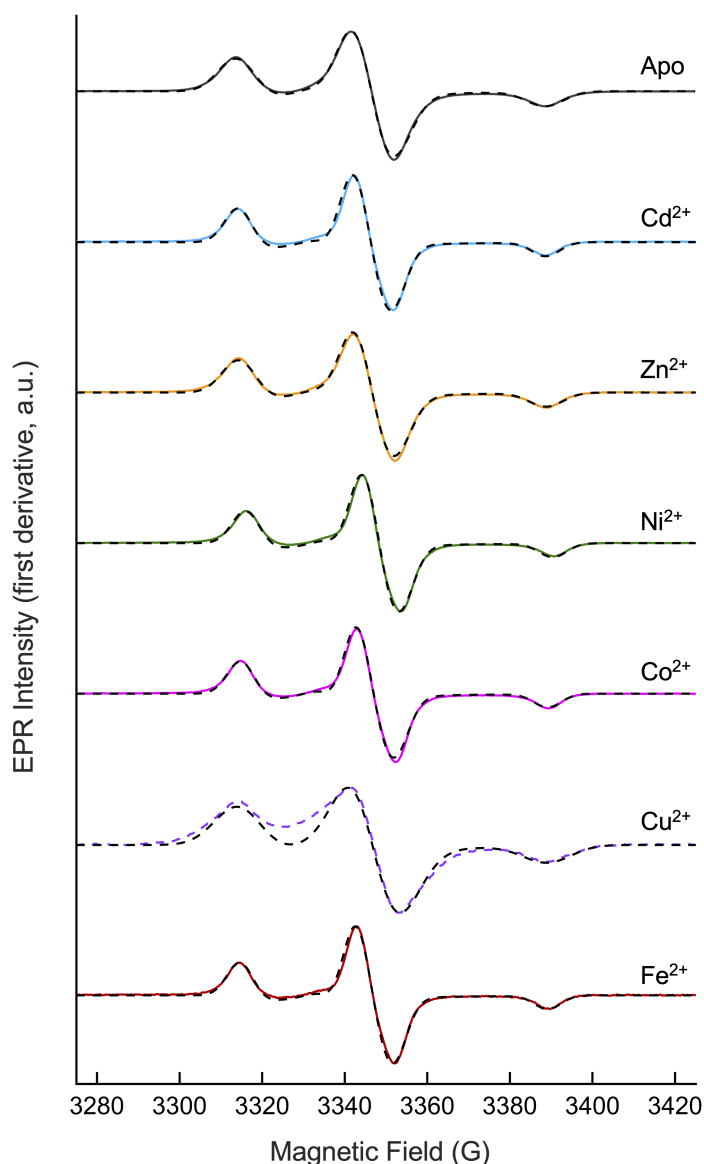


Figure 3.26: Low temperature (50 K) continuous wave X-band EPR spectra measured for C267S_{SL} in apo and various metal-incubated states are plotted between 3275-3425 G, with associated spectral fits overlaid as black dashed lines. The metal used for incubation is highlighted above each respective plot. Experimental parameters: microwave frequency = 9.390 GHz, centre field = 3350 G, sweep width = 200 G, modulation frequency = 100 KHz, modulation amplitude = 3G, microwave power = 0.003157 mW, number averages = 4.

Prior to spectral simulation and fitting of these RT spectra, the equivalent samples were also prepared for LT experiments at X-band frequency, the results of which can be seen in Figure 3.26, as well as each corresponding spectral simulated fitting (both apo and Zn²⁺-bound spectra are the same as those previously presented). This analysis afforded accurate A-value determination (highlighted in Table 3.7) from which it is apparent that the addition of metals has little effect on the broadness of these anisotropic spectra, indicated by the lack of variance in the value for A_{zz} in each case. Notably the Cu²⁺-bound spectrum was most difficult to fit since there is clear evidence of broadening which is not present in any other case - since Cu²⁺ is paramagnetic this might serve as an initial indication of interaction between the SL unpaired electron and the Cu²⁺ ion.

Protein state	g_{xx}	g_{yy}	g_{zz}	A_{xx} (MHz)	A_{yy} (MHz)	A_{zz} (MHz)	Linewidth (peak- peak, mT)	RMSD
Apo	2.008	2.006	2.002	12	18	105	0.90	0.0148
Cd ²⁺	2.008	2.006	2.002	12	14	104	0.68	0.0230
Zn ²⁺	2.008	2.006	2.002	13	16	104	0.82	0.0161
Ni ²⁺	2.008	2.006	2.002	13	13	104	0.67	0.0195
Co ²⁺	2.008	2.006	2.002	13	15	104	0.66	0.0181
Cu ²⁺	2.008	2.006	2.002	12	19	104	1.15	0.0732
Fe ²⁺	2.008	2.006	2.002	12	14	104	0.63	0.0218

Table 3.7: Best fit simulation values of the x,y,z components of both g- and A-tensors, as well as the linewidth for low temperature experimental X-band EPR data measured on C267S_{SL}, both in apo and metal-incubated states. The RMSD of each simulated fit is also given in the table.

This in turn allows us to rationalise the broadening effect observed in RT spectra with incubated metals, which considering the similarities in A_{zz} , must again be a result of different SL rotational correlation times. It is expected that the broader the spectrum, the slower (longer) the rotational correlation time, and so the expected length of T_c 's relative to the previous experiment in each stack of RT plots (see Figure 3.25) is represented by upwards (for longer) and downwards (for shorter) arrows. The A_{zz} parameters obtained from LT experimental fittings were utilised in RT simulated fits for regular metal addition, which are shown in Figure 3.27. The magnitude of T_c values obtained from simulated fittings, as shown in Table 3.8, fits the expected trend for the regular excess metal conditions as well as the metal displacement experiment. This result is one example of how simple, continuous wave EPR experiments can be utilised to not only qualitatively identify binding-induced changes in protein state and/or conformation, but also to investigate relative binding affinities and competition between substrates for the same binding sites.

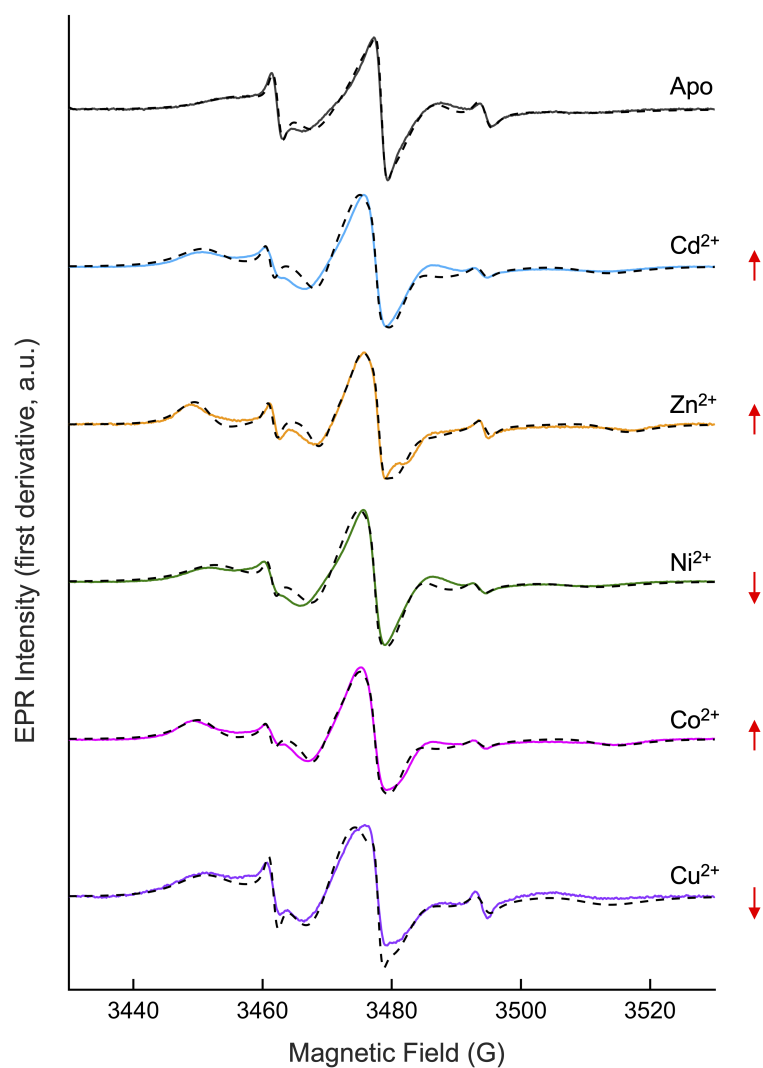


Figure 3.27: Room temperature continuous wave X-band EPR spectra measured for C267S_{SL}, in its apo state, and following incubation with various divalent metals cations (identities are highlighted), between 3430-3530 G, with associated spectral fits overlaid as black dashed lines. Experimental parameters were given previously.

C267S_{SL} apo (RMSD = 0.0330)

Component	g_{xx}	g_{yy}	g_{zz}	A_{xx}	A_{yy}	A_{zz}	Linewidth (peak-peak, mT)	T_c (ns)	Relative weighting
1	2.008	2.006	2.002	11	18	105	0.08	4.90	0.94
2	2.008	2.006	2.002	14	17	104	0.10	0.67	0.06

C267S_{SL} + Cd²⁺ (RMSD = 0.0583)

1	2.008	2.006	2.002	12	14	104	0.08	7.36	0.94
2	2.008	2.006	2.002	14	20	105	0.09	0.70	0.06

C267S_{SL} + Zn²⁺ (RMSD = 0.0471)

1	2.008	2.006	2.002	11	18	105	0.13	17.11	0.95
2	2.008	2.006	2.002	14	20	104	0.11	0.60	0.05

C267S_{SL} + Ni²⁺ (RMSD = 0.0511)

1	2.0076	2.0053	2.0017	13	13	104	0.08	5.67	0.95
2	2.0073	2.0057	2.0017	15	17	104	0.09	0.70	0.05

C267S_{SL} + Co²⁺ (RMSD = 0.0372)*A values FL

1	2.0076	2.0056	2.0014	15	16	104	0.09	10.99	0.96
2	2.0074	2.0056	2.0017	14	17	105	0.08	0.78	0.04

C267S_{SL} + Cu²⁺ (RMSD = 0.0710)

1	2.0075	2.0055	2.0013	12	19	104	0.31	5.84	0.93
2	2.0073	2.0053	2.0012	14	17	104	0.08	0.73	0.07

Table 3.8: Best fit simulation values are given for C267S_{SL}, in its apo state, and following incubation with divalent metal cations, measured at X-band and room temperature. Values include the x,y,z components of both g- and A-tensors, as well as linewidths and averaged rotational correlation times for the two spin label species simulated and contributing to each spectrum. The relative weightings of each species are also shown. The RMSD of each simulated fit is given above each set of results.

3.6.2 DEER for probing metal binding and selectivity in MamM-CTD

Recent solved crystal structures of WT MamM-CTD with Cu²⁺, Cd²⁺ and Ni²⁺ were published alongside the DEER results presented next, and were described in the introductory section of this chapter. These crystal structures are the first to show CDF CTD binding and conformational change with a metal ion other than Zn²⁺. Thus far a mixture of cw-EPR investigations have been described which have given us our first spectroscopic insights and indication of metal binding by MamM-CTD not only to these 4 metal ions, but additionally to Co²⁺. Further DEER studies will now be presented and later discussed with reference

to these crystal structure models, as well as ITC and fluorescence measurements also conducted as part of this collaborative project.

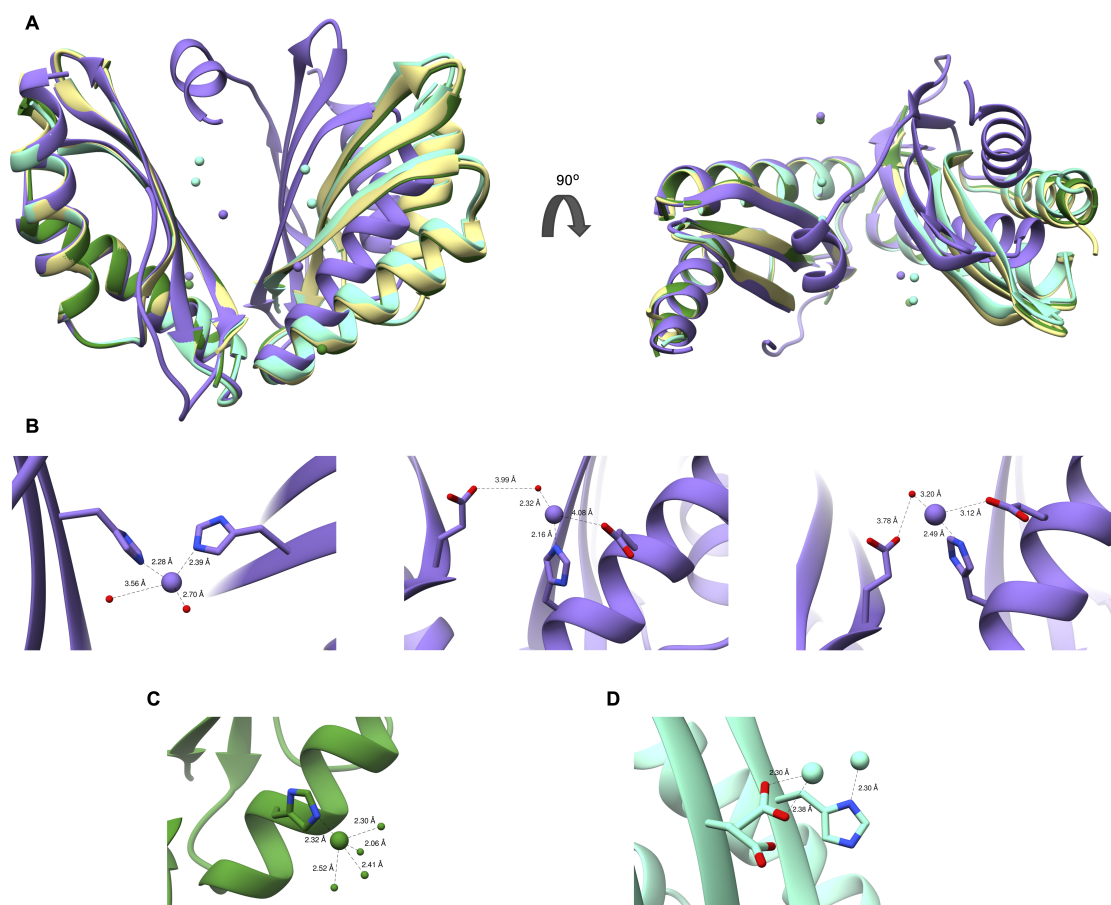


Figure 3.28: (A) MamM-CTD crystal structures with different divalent metal cations. The three metal-bound MamM-CTD crystal structures (Cd²⁺ (PDB ID: 6GMT) - cyan, Cu²⁺ (PDB ID: 6GP6) - violet and Ni²⁺ (PDB ID: 6GMV) - green) are overlaid with the WT apo crystal structure (PDB ID: 3W5X) - yellow. Structures are represented from two angles to display the conformational change exhibited by the Cu²⁺-bound structure. Metal binding sites are highlighted for (B) Cu²⁺, in central (LHS) and peripheral (middle and RHS) sites, (C) Ni²⁺, in the binding site of one monomer and (D) Cd²⁺, in the binding sites of one monomer.

As a recap, solved crystal structures with Cu²⁺, Cd²⁺ and Ni²⁺ (see Figure 3.28) gave direct evidence for metal binding at each of residues of the proposed metal binding sites (one CS and two PSs). In the Cu²⁺-bound structural model 3 metal ions were identified, one at the CS ligated by residues H285 of both monomers in addition to water molecules, and one at each PS ligated by residues H264 of one monomer. Additionally, a bridging water molecule was identified between the metal ion and E289 of the adjacent monomer, whilst direct ligation by E268 was observed in just one PS. This Cu²⁺-bound form exhibits a conformational change in its dimeric packing to a more closed conformation than is seen in the WT apoprotein, much like that observed via DEER spectroscopy with Zn²⁺ in the C267S variant as well as in other CDF CTDs with Zn²⁺. Interestingly, C-terminal residues 292-314 formed an additional β -strand followed by an α -helical segment, both of which are structural

elements which were not resolved in the apo case. The additional β -strand extends the existing 3-stranded β -sheet to 4 strands, and exhibits a stabilising bonding network with the residues of the adjacent monomers in a non-symmetric manner. Moving on to Cd^{2+} , in total 4 metal binding sites were observed in which residues D249 and H285 (both of the proposed CS) each bind one cadmium ion in each monomer. The bound metal ions possess half occupancy in the electron density map, perhaps suggesting alternate binding at these sites. The Cd^{2+} -bound MamM-CTD crystal structure revealed no conformational change relative to the WT apoprotein, which might be expected given that no metal ions were found at the PSs. Finally, the Ni^{2+} -bound structure revealed in total 2 metal ions bound per MamM-CTD dimer, one at each PS and ligated only by H264 and water molecules, and again exhibited no conformational change relative to the apo case. Taken together this could imply a non-specific binding event, whilst highlighting the importance of H264 specifically for metal ion attraction in MamM-CTD.

MMM analysis was carried out using the solved structural models for Cu^{2+} -, Cd^{2+} - and Ni^{2+} -bound MamM-CTD for *in silico* prediction of inter-spin distances compared with the apo C267S case. Table 3.9 gives the output of these analyses, including the number of rotamers, partition function and inter-spin distance for each structure. The lack of dimeric conformational change seen with Cd^{2+} and Ni^{2+} is reflected in the average inter-spin distances predicted for both, which are close to that predicted for apo C267S at approximately 4 nm. However closure of the dimeric interface to a more compact conformation with Cu^{2+} results in a predicted decrease in this distance to 3.23 nm - this closely matches the SL-SL distance found for Zn^{2+} -bound C267S_{SL} in our previously mentioned DEER studies. In a similar fashion as was previously described for the Zn^{2+} binding studies, DEER experiments conducted on metal-incubated C267S_{SL} have been undertaken using a 5:1 metal:protein ratio, other than for Cu^{2+} , in which a 3:1 ratio was used due to the quenching effect observed under RT cw-EPR conditions. Experiments have been conducted on C267S_{SL} incubated with multiple metals including Cd^{2+} , Cu^{2+} , Ni^{2+} , Co^{2+} , Mn^{2+} and Fe^{2+} (the natural substrate of MamM). The positions of pump and observe pulses for each experiment are highlighted in Figure 3.29.

PDB code	Protein state	Number rotamers	Partition function	r (nm)
6G55	Apo	66	0.36	4.00
6GP6	Cu ²⁺ -bound	64	0.30	3.23
6GMT	Cd ²⁺ -bound	43	0.24	4.14
6GMV	Ni ²⁺ -bound	39	0.28	4.11

Table 3.9: Results of MMM analysis using structural models for C267S and WT metal-bound MamM-CTD, including the predicted number of rotamers, partition functions and average inter-spin distances.

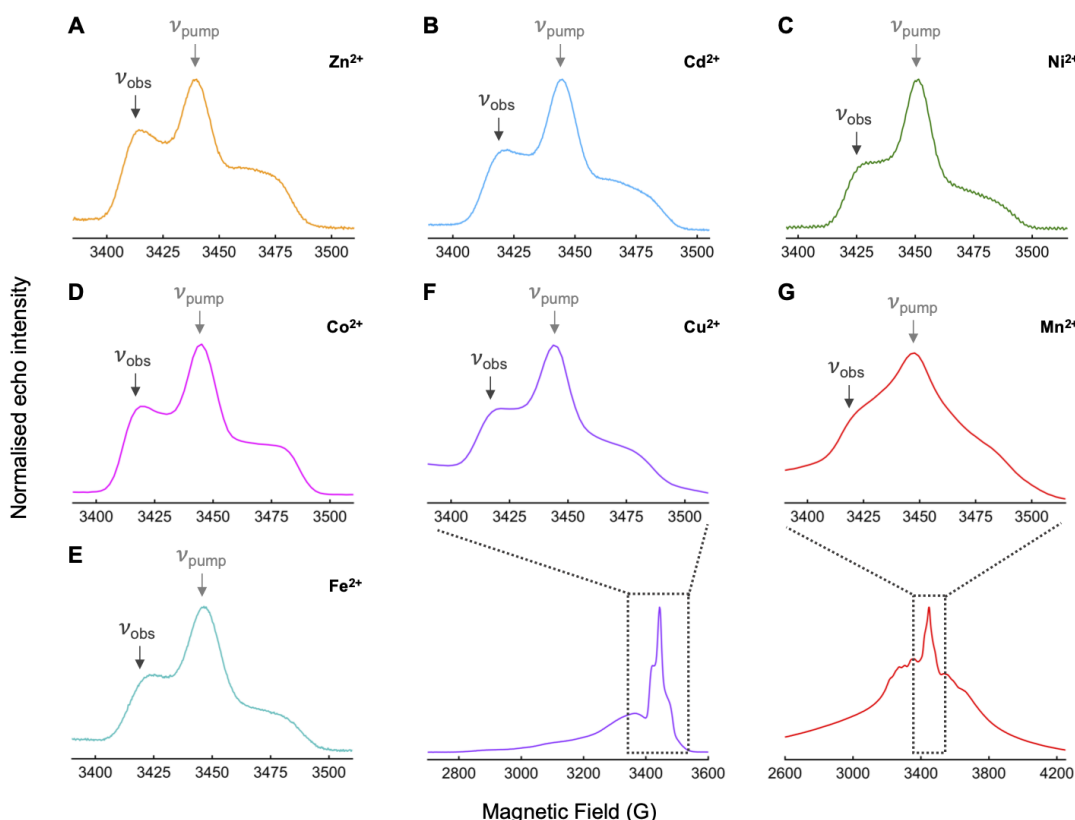


Figure 3.29: Field swept echo-detected spectra (50 K) for C267S_{SL} incubated with (A) Zn²⁺, (B) Cd²⁺, (C) Ni²⁺, (D) Co²⁺, (E) Fe²⁺, (F) Cu²⁺ and (G) Mn²⁺. Pump and observe positions for DEER experiments are highlighted with arrows. Wide-sweep plots can also be seen for the latter two metals in the lower panels. Experimental parameters: centre field = 3365-3473 G, sweep width = 120-1800 G, delay time (τ) = 120 ns, number of points = 300-1000, shots per point = 50, shot repetition time = 2 ms, number averages = 4.

Background corrected dipolar evolution datasets with their corresponding fittings (where appropriate) and the resulting distance distributions are presented in Figure 3.30 for each M²⁺ ion, including Zn²⁺ as a reference which is overlaid with the apo case. The addition of Cd²⁺, Cu²⁺, Ni²⁺, Fe²⁺ and Co²⁺ each results in an average inter-spin distance centred at 3.3 nm, matching that observed in the presence of Zn²⁺ and which was attributed to dimeric closure forming a single and stable conformation upon metal binding from a largely dynamic apo state. In addition, distance distributions of Cu²⁺- and Co²⁺-bound C267S_{SL} contain a second peak, centred at 2.4 and 2.1 nm respectively. Conversely, the Mn²⁺-incubated sample gave rise to no oscillatory effects in the time domain and consequently a distance distribution could not be determined in this case. The apparent

conformations adopted according to the distance distributions obtained in the presence of each metal can be rationalised when considered against crystal structure data, ITC and tryptophan fluorescence measurements, as will be discussed later.

Analysis of the Cd^{2+} -bound and Ni^{2+} -bound DEER datasets both give rise to a single and defined distance distribution at 3.3 nm, similar to that observed with Zn^{2+} . In each case the observed inter-spin distance contradicts that expected based on the solved crystal structure models, for which MMM analysis predicts an inter-spin distance of 4.14 nm and 4.11 nm with Cd^{2+} and Ni^{2+} respectively. Despite the similarity in conformational states adopted in each metal-bound crystal structure and that of the apoprotein, it is apparent from our spectroscopic data that addition of both Cd^{2+} and Ni^{2+} appears to cause conformational locking to a single conformation in solution, in which the dimeric arms are packed more tightly than in the apoprotein.

Despite the low solubility of $\text{Fe}(\text{OH})_2$ in basic conditions¹⁶² we were able to successfully prepare a sample of C267S_{SL} with Fe^{2+} , using argon for expulsion of oxygen, for DEER studies. The resulting dipolar evolution time trace is clearly different to that of the apoprotein (shown as raw data in Figure 3.30F to indicate these differences more clearly), and upon analysis an average inter-spin distance centred again at 3.3 nm is determined. Whilst this distribution is broader than those obtained in the cases of other divalent metal cations (including Ni^{2+} , Cd^{2+} , Cu^{2+} , Zn^{2+}), the difference as compared with the apo case alludes to a more restricted conformation in solution with Fe^{2+} , with the average inter-spin distance indicating a closed conformation similar to that shown in the Cu^{2+} -bound structure. The low modulation depth of the background corrected data is likely explained by the presence of only a very low concentration of Fe^{2+} in solution, or alternatively signal quenching by the paramagnetic metal ions in solution. Importantly, this is the first example of a quantifiable conformational change of MamM-CTD induced by binding of its natural substrate, Fe^{2+} , and the observation of a 3.3 nm inter-spin distance matching those in the cases of other divalent metals gives a good indication that these results do indeed represent biologically relevant states of MamM-CTD in solution.

Both incubation with Cu^{2+} and Co^{2+} gives rise to results clearly different from those with the other divalent metal cations already discussed here, in the form of bimodal distance distributions. For C267S_{SL} incubated with Cu^{2+} , and indication of SL-metal interaction was already displayed in the form of significant quenching of the RT X-band cw-EPR spectrum, in addition to broadened spectral features. Our DEER measurements and subsequent

analyses show one distance distribution centred again at 3.3 nm (see Figure 3.30D) in close agreement with the inter-spin label distance determined by MMM (with the Cu^{2+} -bound structure) of 3.23 nm. In addition, a second sharp distribution centred at 2.4 nm is present, the cause of which might be rationalised by a number of possibilities. Cu^{2+} is a paramagnetic species, which will therefore give rise to an EPR signal of its own when bound to MamM-CTD, in addition to that of the nitroxide SL of C267S_{SL}. The most simple explanation would be the presence of multiple protein conformations, resulting in two differing inter-spin label distances being present. However, the axial Cu^{2+} EPR spectrum at cryogenic temperatures overlaps with that of a nitroxide-based SL such as MTSSL (see Figure 3.29F, bottom panel), and so in this case our DEER experiment is not limited to excitation of the nitroxide spins, but also Cu^{2+} spins. This in turn means that we could be recording dipolar SL- Cu^{2+} and Cu^{2+} - Cu^{2+} interactions, in addition to the SL-SL interaction. Considering the similarity to SL-SL distances determined for other divalent metal cations, it is most likely that the 3.3 nm distance corresponds to a SL-SL interaction. The exact nature of the 2.4 nm distance however can only be determined with the application of further DEER and other pulsed techniques, which will be addressed later. Contrastingly, whilst Co^{2+} also exhibits its own EPR signature, in this case its resonance would not overlap with that on a nitroxide SL such as MTSSL, and so this second feature at 2.1 nm cannot be rationalised as either a metal-metal or metal-SL distance. In this case, a bimodal distribution would indicate two distinct but stable conformations of the protein in solution. Notably, both distances within the distribution are represented by much broader features than in the case of most other metals tested, perhaps indicating more fluctuation SL positions relative to one another in solution. It therefore seems as if both apparent conformations are less structurally stable than in other cases.

DEER measurements conducted on C267S_{SL} in the presence of Mn^{2+} were attempted numerous times under different conditions, all of which yielded dipolar evolution time traces without evidence for dipolar coupling between nitroxide species. In the presence of d8-glycerol (see Figure 3.30G), the decay is modulated by deuterium at a frequency equivalent to that seen in 2P-ESEEM experiments (see Figure 3.20B) at ~ 2 MHz. This modulation effect is not present under entirely protonated conditions. Whilst all DEER experiments presented have utilised a deuterium suppression cycle, the effectiveness of this cycle decreases with decreasing modulation depth.¹⁶³ Thus for a sample in which there is no apparent modulation in the DEER time-domain trace due to inter-(electron)spin interaction, this effect in the presence of deuterated glycerol is not surprising. Hence our results allude to discriminatory behaviour by MamM-CTD against Mn^{2+} binding.

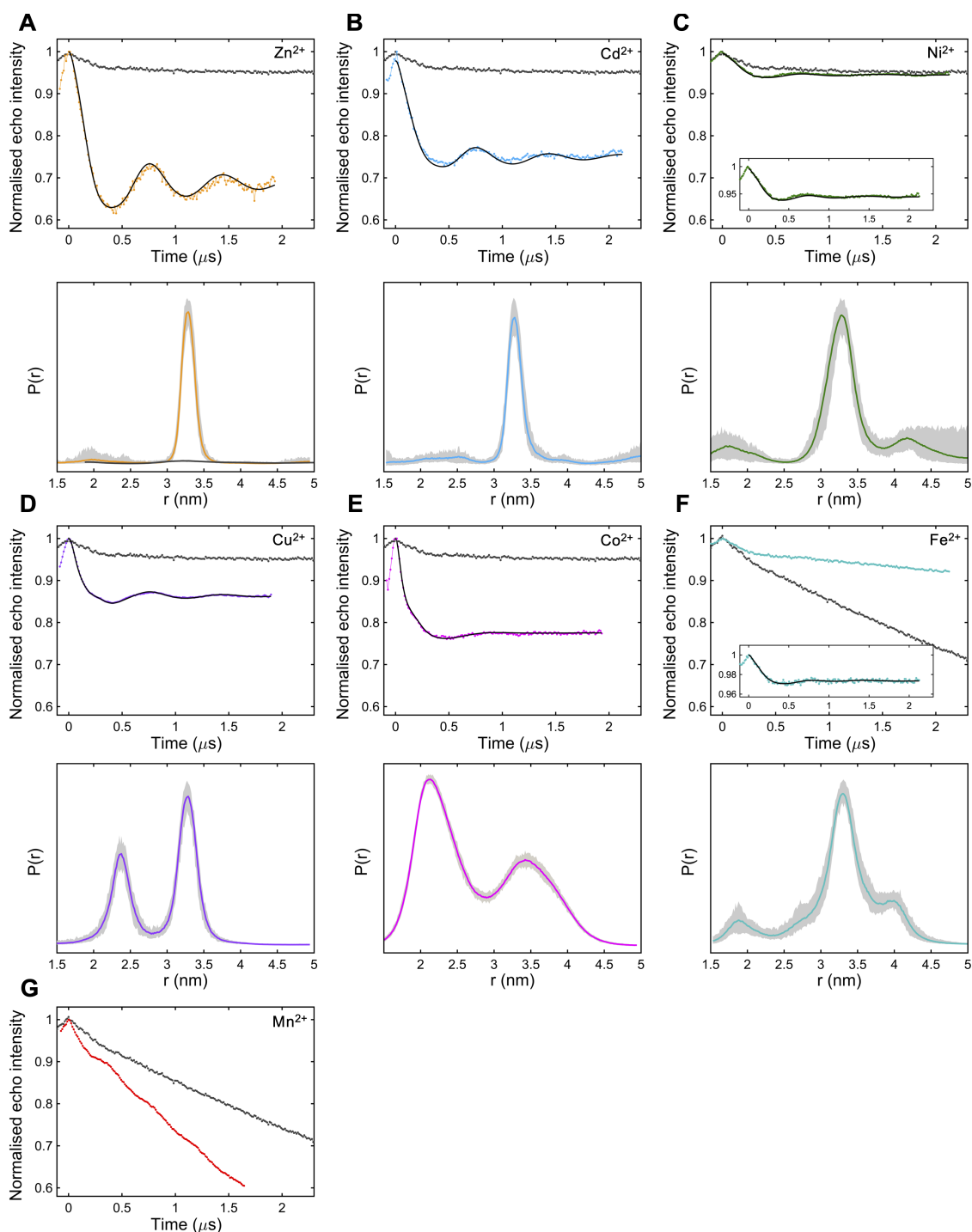


Figure 3.30: 4-pulse DEER experiments (50 K) of double spin labelled MamM-CTD C267S with a range of divalent transition metal ions, including (A) Zn^{2+} , (B) Cd^{2+} , (C) Ni^{2+} , (D) Cu^{2+} , (E) Co^{2+} , (F) Fe^{2+} and (G) Mn^{2+} . (Top panel) The normalised background-corrected dipolar evolution of metal-bound samples (bright colours) presented together with the apo spin labelled protein (grey). Spectral fittings are shown as black lines. Exceptions; For both Mn^{2+} and Fe^{2+} the non-baseline corrected time traces are shown to clearly demonstrate the change in decay caused by specific versus non-specific metal binding. (Bottom panel) Normalised distance distributions (where appropriate) of metal bound C267S_{SL}. The distance distribution of apo C267S_{SL} is also shown for reference, plotted against the Zn^{2+} -bound results and shown in dark grey. Standard deviations in the form of 95 % confidence intervals are also given in each case and represented by light grey areas around the lines. Fitting of all dipolar evolutions and subsequent distance distributions were calculated with DEERNet.²⁰ Experimental parameters: frequency offset = 80 MHz, τ_1 = 350 ns (Ni^{2+}), 120 ns (all others), τ_2 = 1300-2200 ns, shots per point = 50, shot repetition time = 2 ms.

3.7 Cu(II) - a special case

The following section describes work undertaken following the more complex DEER result previously presented with C267S_{SL} and Cu(II), in which a bimodal distance distribution was obtained, with overlapping Cu(II) and nitroxide label FSEs at the high field end of the Cu(II) spectrum. The work presented here aims to effectively pull apart the distance distribution obtained for C267S_{SL} with Cu(II) in order to define the exact origin of spin-spin interactions corresponding to each distance. Initially, regular DEER experiments over multiple observe and pump positions were utilised with both C267S MamM-CTD and binding site variants. Following this we have applied the inversion recovery-filtered *i*DEER technique, and further developed its application for elimination of the observation of specific interactions, and subsequent spectral filtering.

3.7.1 An introduction to Cu(II) EPR

Recently, the field of Cu(II)-nitroxide interactions studied by EPR methodologies has gained significant interest, with a primary focus surrounding two main areas. Firstly, on the potential of strongly exchange-coupled Cu(II)-nitroxide spin pairs as switchable magnetic materials,^{164–169} and secondly on the use of both native Cu(II) centres in proteins,^{170–174} and Cu(II)-based spin labels/labelling strategies,^{175,176} as targets and tools for EPR investigation throughout biological systems. The latter of these applications typically probes weakly coupled spins and thus requires the use of more sensitive pulsed dipolar EPR methodologies. Such experiments pose a number of challenges when applied to Cu(II) systems, including limited sensitivity due to the small fraction of spins perturbed when using a regular rectangular shaped microwave pulse over the wide Cu(II) EPR spectrum. This in turn means that dipolar modulation depths are limited (as will be seen shortly) when pumping Cu(II) spins, and signal-to-noise ratios suffer. This situation actually worsens at higher frequencies, since the width of Cu(II) spectra is dependant upon the (field-dependant) g-value anisotropy, which already dominates in the factors governing spectral width at X-band frequency. Development of new EPR setups and hardware, as well as new methodologies such as the RIDME (relaxation-induced dipolar modulation enhancement) experiment are providing a means of negating these limiting factors at both X-band and higher frequencies. Another issue that arises when investigating Cu(II)-nitroxide interactions is overlap of spectral features at lower frequencies; this can greatly complicate experimental data and is an issue which will be addressed through application of the *i*DEER experiment in the coming section.

The Cu(II) ion has a $3d^9$ electronic configuration, in which irrespective of its coordination geometry the electrons are always paired to give a total spin $S = 1/2$.¹⁷⁷ Ligands of Cu(II) can be found in an array of coordination geometries, with tetra-, penta- and hexa-coordination spheres; the nature of these geometries directly affects the electronic properties of the ion. The coordination geometry determines the order of d-orbital energy levels in which the electrons reside, and in turn the ground state of a particular complex.¹⁷⁸ Principle g-tensor values can be used in EPR to assign coordination geometries of complexes, and deviate significantly from the value of a free electron (g_e) as a result of spin-orbit coupling. The unpaired electron typically resides in either the $d_{x^2-y^2}$ or d_{z^2} orbital in the ground state of a Cu(II) species, resulting in an axial powder EPR spectrum in which the x- and y-axes are equivalent, with two principle g-values, namely $g_{||}$ (g_z) and $g_{\perp}$ (g_x and g_y). The nuclear spin of both naturally occurring isotopes ^{63}Cu and ^{65}Cu is equal to $3/2$, and thus the EPR spectrum will show hyperfine splitting of four lines at each principle g-value when fully resolved. The magnitude of $A_{||}$ is much larger than that of $A_{\perp}$, meaning that the nuclear hyperfine splitting at $g_{||}$ is resolved at X-band frequency (where g-value anisotropy already dominates the spectral width). Intrinsic linewidths of Cu(II) spectra mean that nuclear hyperfine splittings at $g_{\perp}$ are typically not resolved.¹⁷⁷ Additional hyperfine structure from ligands associated with the Cu(II) ion can also be observed, in the form of superhyperfine splitting.

3.7.2 Cu^{2+} binding probed by DEER studies of MamM-CTD variants

As was described previously, a recent X-ray crystal structure model of WT MamM-CTD with Cu^{2+} shows one central, and two peripheral Cu^{2+} binding sites, with Cu^{2+} binding clearly associated with conformational change involving closure of the dimeric arms. This gives a tighter dimeric packing than in the apo case. Considering the paramagnetic nature of Cu^{2+} ions, Cu^{2+} -bound spin labelled MamM-CTD will give rise to an EPR signature containing overlapping nitroxide and Cu^{2+} resonances. Hence the picture we can form of possible inter-spin interactions that might be measurable by DEER spectroscopy begins to look somewhat complex. Figure 3.31A depicts the possible pairwise distances between paramagnetic species in C267S_{SL} MamM-CTD, with approximate distances measured between (a) Cu^{2+} - Cu^{2+} , (b) Cu^{2+} -Cys and (c) Cys-Cys species; measurements were made in UCSF Chimera (version 1.14)¹¹⁸ and taken from C_{α} atoms of Cys residues.

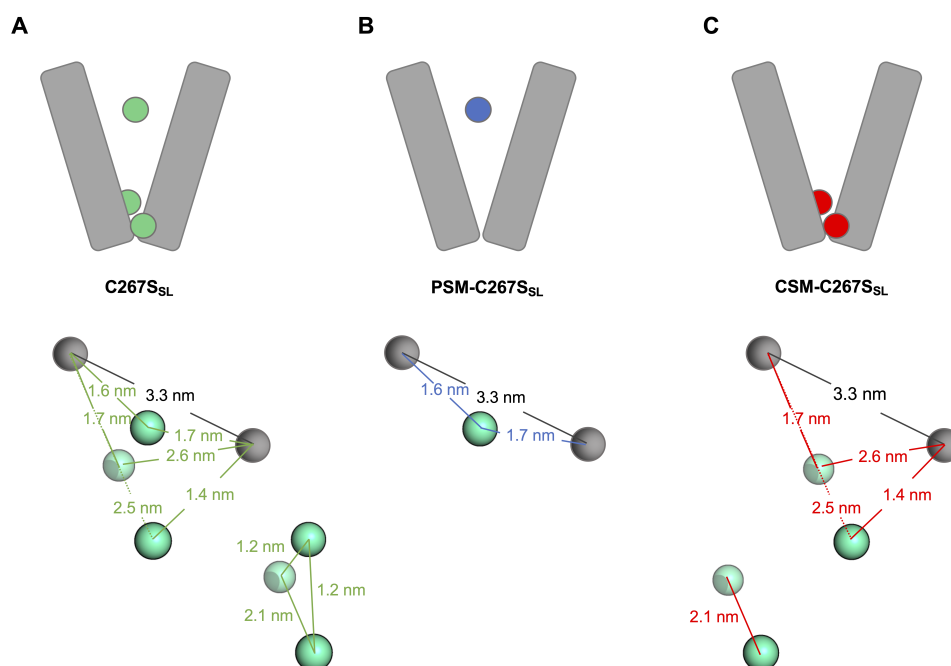


Figure 3.31: (Top) Schematic representations of C267S (A), PSM-C267S (B) and CSM-C267S (C) MamM-CTD, with potential metal binding sites available highlighted as coloured spheres. (Bottom) All possible inter-spin distances in each case for spin-labelled, Cu²⁺-bound MamM-CTD where all possible binding sites are occupied.

Previously DEER data was presented for C267S_{SL} with Cu²⁺, in which pump and observe positions were in the overlapping SL/Cu²⁺ region of the spectrum, resulting in a bimodal distance distribution. Interpretation of this data alone to determine the exact origin of each contributing distance is difficult, not only because of the diverse array of inter-spin couplings we might be measuring, but also due to the overlapping nature of the SL and Cu²⁺ resonances at the high field end of the Cu²⁺ spectrum. It is for these reasons that we have utilised (a) binding site variants of MamM-CTD and (b) the inversion recovery filtered *i*DEER technique to gain further insight into the origin of the bimodal distribution seen in our regular DEER experiment.

Variation of pump and observe positions in MamM-CTD variants

The distance distribution obtained for C267S_{SL} with Cu²⁺ (Figure 3.30D), in which pump and observe positions were localised at the SL region of the spectrum, gives two distributions centred at 2.4 and 3.3 nm. Indeed the 3.3 nm distance agrees nicely with inter-SL MMM predictions as well as the SL-SL distances determined for C267S_{SL} with other metal ions. However without further spectral filtering it is not possible to definitively say which inter-spin distance/s it is that we are observing. Figures 3.31B and C give an explanation for how, upon introduction of MamM-CTD binding site variants in combination with SDSL,

the maximum number of interactions we are probing can be greatly simplified. Additionally, since the Cu^{2+} spectrum is much more broad than that of the nitroxide SL, we can be selective of which species we pump/observe by simply moving our pump/observe positions (see Figure 3.32).

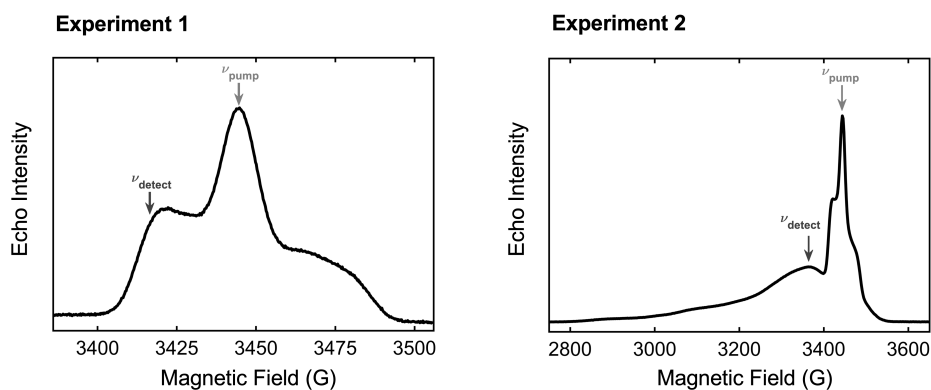


Figure 3.32: Simulated FSE spectra for a nitroxide species alone (left), and in the presence of Cu^{2+} (right), with pump and observe positions highlighted for the two types of DEER experiment.

The results of DEER experiments and subsequent data analysis in the form of distance distributions for C267S_{SL}, PSM-C267S_{SL} and CSM-C267S_{SL}, utilising different pump and observe positions, are shown in Figure 3.33. DEER experiments will be referred to as either 1 or 2, as described in the previous figure, depending on which pump and observe positions were utilised; results are plotted in light and dark colours for experiments 1 and 2 respectively. Starting with experiment 1, a bimodal distance distribution is determined for CSM-C267S_{SL} with Cu^{2+} (Figure 3.33B) in which the central distances and their distribution widths are highly similar to those observed in the absence of binding site mutations (at 2.4 and 3.3 nm). However in the case of PSM-C267S_{SL} (in which residues of the PS have been mutated) with Cu^{2+} only a single distance distribution is observed at 3.3 nm (Figure 3.33C). Introduction of binding site variants has therefore already given some insight into our bimodal distribution, since the additional 2.4 nm distance measured in the case of both C267S_{SL} and CSM-C267S_{SL} can now be linked to Cu^{2+} being present at the peripheral binding sites (which are present in these two variants alone).

Experiment 2 simplifies things further, as in this case we are observing at a position on the Cu^{2+} spectrum which does not overlap with that of the nitroxide SL. Meanwhile, we are pumping at a position which includes overlapping features of the SL and Cu^{2+} spectral signatures. Therefore this experiment can only perturb either Cu^{2+} - Cu^{2+} or Cu^{2+} -SL interactions, and not SL-SL interactions. Results for all three variants of spin labelled MamM-CTD (shown plotted in dark colours) are indeed different from those of experiment

1; specifically the 3.3 nm distance distribution in each case has been eliminated. This therefore confirms that the interaction resulting in the inter-spin distance of 3.3 nm in each case is occurring between SL species, which is in agreement with both MMM predictions based on the WT Cu^{2+} -bound structure, as well as SL-SL measurements of C267S_{SL} with other divalent metal cations. Additionally, the 2.4 nm distance distribution remains clearly present, with a peak of significant magnitude for both C267S_{SL} and CSM-C267S_{SL} . There is an additional feature in the case of C267S_{SL} , centred at 3 nm, however the 95 % confidence bounds would indicate that this feature is not relevant. The distance distribution obtained in the case of PSM-C267S_{SL} exhibits a small but sharp feature centred at 2.7 nm which is not present for experiment 1 - it is therefore likely that this is an artefact perhaps caused by electron-deuteron modulation effects for example. These initial results imply that the 2.4 nm distance is (a) caused by some interaction with Cu^{2+} bound at the peripheral sites, and (b) as a result of Cu^{2+} - Cu^{2+} or Cu^{2+} -SL interactions, as opposed to SL-SL interactions. All three variants appear to adopt similar conformations in solution upon addition of Cu^{2+} , considering the 3.3 nm SL-SL distance determined throughout the datasets.

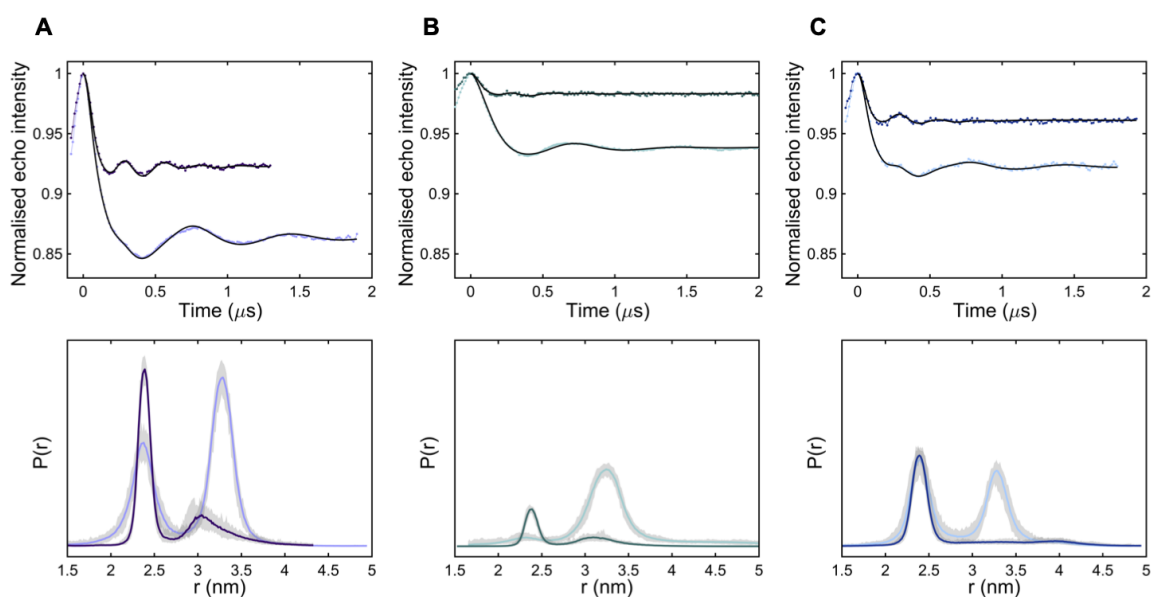


Figure 3.33: 4-pulse DEER experiments (50 K) for (A) C267S_{SL} , (B) PSM-C267S_{SL} and (C) CSM-C267S_{SL} MamM-CTD each following incubation with excess Cu^{2+} . (Top panel) The normalised background-corrected dipolar evolution of each metal-bound sample for experiments 1 and 2, overlaid and shown as light and dark colours respectively. Spectral fittings are shown as black lines. (Bottom panel) Distance distributions all plotted on equivalent axes such as to highlight the relative intensities of peaks throughout distributions. Standard deviations in the form of 95 % confidence intervals are also given in each case and represented by light grey areas around the lines. Fitting of all dipolar evolutions and subsequent distance distributions were calculated with DEERNet.²⁰ Experimental parameters: frequency offset = 80 MHz (exp 1) and 210 MHz (exp 2), τ_1 = 380 ns (PSM-C267S_{SL} exp 1), 498 ns (PSM-C267S_{SL} exp 2), 120 ns (all others), τ_2 = 1500-3000 ns, shots per point = 50, shot repetition time = 2 ms.

3.7.3 Beyond pairwise distance determination in MamM-CTD

The presence of overlapping spectral features, as we are faced with here, is in fact common when studying many biological systems, for example metalloproteins or paramagnetic co-factors. As mentioned previously, the use of higher microwave frequencies can be beneficial for EPR experiments in cases where contributing species exhibit different (field-dependent) g -values, thereby affording separation of spectral features. Alternatively, distinction can also be afforded according to differing electron spin relaxation times; examples in literature utilise specific pulse separations which are chosen for suppression of species with shorter T_1 or T_2 relaxation times (depending on the type of experiment). Short shot repetition times (i.e. how soon each successive pulse sequence can commence after the previous one) have also been used for suppression of species with a longer T_1 relaxation time. However these methods act only as low-or high-pass filters, and as such are not selective for very specific species.

The spin-lattice relaxation times (T_1) of paramagnetic species can be easily determined by measurement and subsequent fitting of a simple pulse sequence (see Figure 3.34), named the inversion recovery experiment. In the case of multiple overlapping species with different spin-lattice relaxation times, detection of one species can be 'turned off' by means of inversion recovery-filtered EPR techniques. Previous examples in literature primarily use inversion recovery-filtering in combination with hyperfine spectroscopy (REFINE)^{179–181}, but in theory this technique can be applied to many pulsed EPR methodologies, including the DEER experiment.¹⁸²

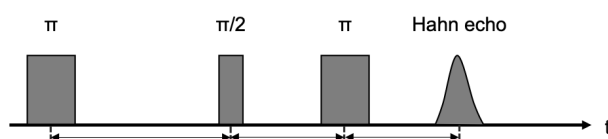


Figure 3.34: Pulse sequence for an inversion recovery experiment.

Figure 3.34 gives the pulse sequence for an inversion recovery experiment, which may be measured at one, or across multiple magnetic field values. The relaxation of spins following an initial π inversion pulse (optimised for full inversion) back to thermal equilibrium is monitored by the intensity of the Hahn echo produced by the following $\pi/2$ and π pulses; the time between the inversion pulse and this detection sequence is increased for a specified number of points at regular intervals. This recovery can be fitted using the following equation 3.2,

$$M_z(t) = M_0 \cdot [1 - 2ae^{-t/T_1}] \quad (3.2)$$

for a single spin species, in which M_z is the net magnetisation, M_0 is the equilibrium magnetisation, a is the inversion factor, t is time and T_1 is the spin-lattice relaxation time. This T_1 relaxation time will be specific for each individual species present within a perturbed system, and as such the magnetisation decay of each species will traverse its own zero-crossing point of the magnetisation M_z . It is at this specific time for each individual species that its own Hahn echo is not observed following the two-pulse detection sequence and hence this species is suppressed. This defines the filter time (T_F) which is utilised in inversion recovery-filtered experiments.

Inversion recovery-filtered FSE experiments for determination of T_1 relaxation profiles

Inversion recovery experiments were conducted for C267S_{SL} in the presence of Cu²⁺ at regular magnetic field increments across a sweep width which covered both Cu²⁺ and SL contributions (inversion recovery versus field measurements). Figure 3.35 gives both 2D (left hand panel) and 3D (right hand panel) representations of the experimental results. It can be clearly seen in the 3D representation that the inversion recovery curves measured at each magnetic field point vary considerably across the spectrum, with the MTSSL spin label exhibiting much slower T_1 relaxation than Cu²⁺, since restoration of maximum, positive echo intensity following the inversion pulse takes a longer period of time.

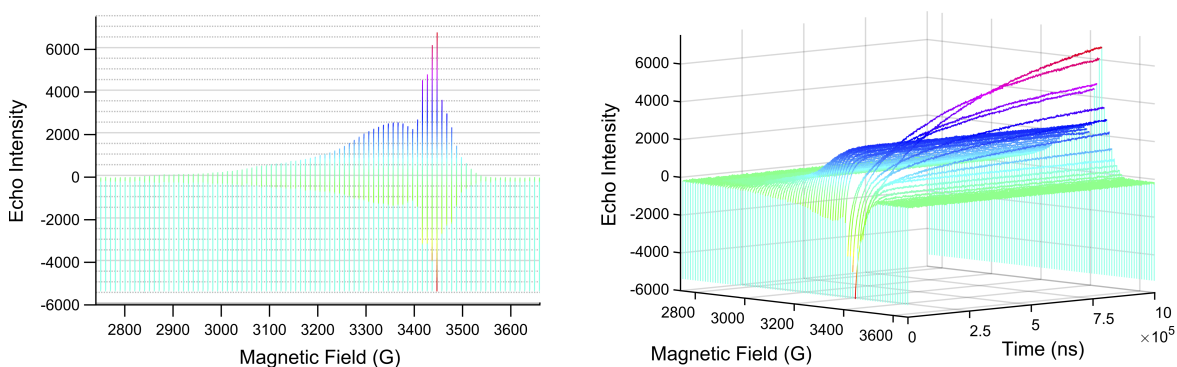


Figure 3.35: Results of the inversion recovery versus field measurements (50 K) conducted for C267S_{SL} with a 3x excess of Cu²⁺, giving both 2D (left) and 3D (right) representations. Slices are shown for inversion recovery measurements at each increment along the magnetic field axis. Experimental parameters: microwave frequency = 9.665 GHz, centre field = 3250 G, sweep width = 1000 G (100 points), pulse displacement increment = 1000 ns (1000 points), shots per point = 4, shot repetition time = 20 ms.

The following section will describe inversion recovery-filtered DEER (*i*DEER) experiments carried out on C267S_{SL} in the presence of Cu²⁺, in which the pump and observe positions used are equivalent to experiment 1 in Figure 3.32. In order to determine the appropriate filter times for such experiments, it is necessary to know the time at which the macroscopic

magnetisation (and thus echo intensity) of each species, at the detection field, traverses zero. It is this time which defines the filter time (T_F) required for filtering of the echo contribution of one specific species, in a sample containing two species. Thus, only the echo of one species should be detected if a field swept spectrum is recorded using an initial inversion pulse and a T_F corresponding to the zero crossing point of the second species. Figure 3.36 gives the individual field swept spectra of C267S_{SL} with Cu²⁺ measured in the absence of an initial inversion pulse (a regular FSE), and in the presence of an inversion pulse with two different T_F values which effectively filter out contributions from either component.

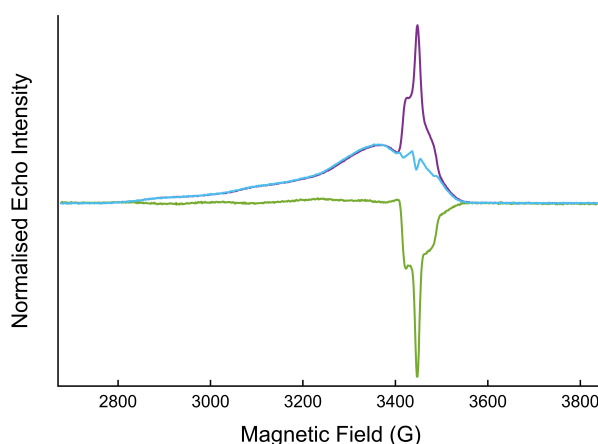


Figure 3.36: Inversion recovery-filtered FSE (IRfFSE) spectra (50 K) are plotted for C267S_{SL} with a 3x excess of Cu²⁺, with filter times (T_F) equal to 6000 ns (green) and 150000 ns (blue), and overlaid with that measured in the absence of an inversion pulse (purple, a regular FSE measurement). Experimental parameters: microwave frequency = 9.677 GHz, centre field (G) = 3375 G, sweep width = 1400 G, number of points = 1000, shots per point = 50, shot repetition time = 2 ms.

A much smaller T_F (6000 ns) is required for filtering out contributions from Cu²⁺, compared with that required for filtering out SL contributions (150000 ns) due to differences in their T_1 relaxation profiles. It is this knowledge which allows selective filtering of the distances observed (but not probed) in the iDEER experiment. The IRfFSE recorded at the shorter filter time of 6000 ns gives an inverted (negative) spin-label echo, since its magnetisation has been inverted by the initial inversion pulse and its relaxation back to equilibrium is longer than that of the Cu²⁺ present.

Inversion recovery-filtered DEER experiments for deconvolution of multiple distances

DEER spectroscopy, as mentioned previously, utilises two defined microwave frequencies to pump and probe different sets of spins within a sample according to their resonance positions. The DEER experiment perturbs one set of spins which are excited by the

observe frequency (the **A spins**), and one set of spins which are excited by the pump frequency (the **B spins**). In the *i*DEER experiment, an initial inversion pulse is applied at the observe frequency which will perturb the **A spins**, however the difference between pump and observe frequencies is large enough that a pulse of this length will not affect **B spins** at the pump frequency (pulse sequence as shown in Figure 3.37). Following this inversion pulse, T_1 relaxation means that the macroscopic magnetisation of **A spins** relaxes and at a specific time, T_F , passes through zero. It is at this time that no Hahn echo is observed by the detection sequence, and thus detection of the **A spins** has been suppressed. In our case we have two paramagnetic species present with overlapping spectral features at both pump and observe positions, each with distinctly different T_1 relaxation times. If the 4-pulse DEER sequence is applied at the T_F of one species, the associated **A spins** will be suppressed and will not give an echo, however the corresponding **B spins** will still be excited by the pump pulse at the pump frequency. So in the case of C267S_{SL} with Cu²⁺, if a T_F of around 6 μ s is used there will be no detected echo of the Cu²⁺ **A spins**, whilst the echo of SL **A spins** will still be observed. Cu²⁺ **B spins** are still excited at the second microwave frequency, as are the SL **B spins**. Thus we will detect the effects of (a) SL **B spins** on SL **A spins**, and (b) Cu²⁺ **B spins** on SL **A spins**, at our observe position of 3420 G. As such, this technique serves to filter out detection of dipolar interaction between **A** and **B** spins of the same species, whilst all other interactions will still be observed.

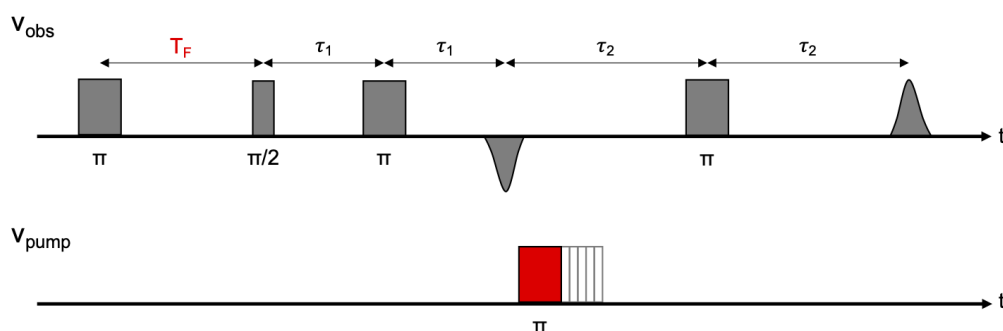


Figure 3.37: Pulse sequence for the *i*DEER experiment, with the filter time labelled T_F following the initial inversion pulse.

Filter times for C267S_{SL} of 6 μ s and 150 μ s were previously determined from inversion recovery filtered FSE experiments as optimal T_F values for filtering out detection of Cu²⁺ and SL, respectively. These were therefore used as starting values for attempts at filtering out detection of potential Cu²⁺-Cu²⁺ and SL-SL interactions in *i*DEER experiments. It has already been demonstrated that the 3.3 nm distance present in the regular DEER experiment 1 is reporting on SL-SL interactions, and this agrees well with MMM predictions. It is therefore expected that this distance should be completely filtered out at the correct

T_F , corresponding to the time at which the macroscopic magnetisation of SL spins passes through zero. The *i*DEER experiment carried out using T_F equal to 210 μs results in the most effective filtering of this 3.3 nm peak; this is relatively close in value to the T_F determined via IRfFSE experiments for SL suppression. The complete eradication of this peak gives further evidence still that the 3.3 nm distance is reporting only on SL-SL interactions, as expected.

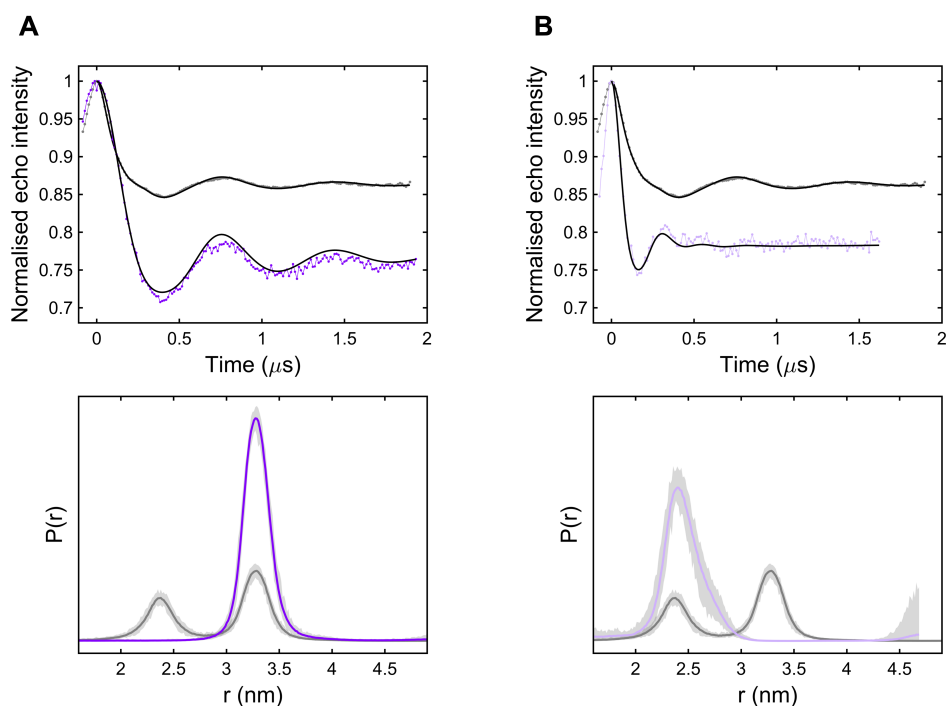


Figure 3.38: 4-pulse DEER experiment for C267S_{SL} following incubation with excess Cu²⁺ (grey), overlaid with the *i*DEER experiment (both 50 K) for the same sample (purple) using filter times (T_F) equal to (A) 24 μs and (B) 210 μs . (Top panel) The normalised background-corrected dipolar evolution for each experiment, overlaid. Spectral fittings are shown as black lines. (Bottom panel) The associated distance distributions, overlaid such as to highlight the relative intensities of peaks throughout distributions. Standard deviations in the form of 95 % confidence intervals are also given in each case and represented by light grey areas around the lines. Fitting of all dipolar evolutions and subsequent distance distributions were calculated with DEERNet.²⁰ Experimental parameters: frequency offset = 70 MHz, T_F = 24 μs (A) and 210 μs (B), τ_1 = 120 ns, τ_2 = 1800-2000 ns, shots per point = 50, shot repetition time = 2 ms.

Meanwhile, the exact origin of the distance distribution centred at 2.4 nm is less certain; whilst it has been shown to be linked with the peripheral binding site, it is still possible that this distance is reporting on either SL-Cu²⁺ or Cu²⁺-Cu²⁺ distances. The latter case would, with the relevant T_F , mean complete eradication of the 2.4 nm distance in the *i*DEER experiment, since only detection of interactions between spins of the same species are filtered out. *i*DEER experiments conducted at a T_F of 6 μs (as was determined for Cu²⁺ suppression in IRfFSE experiments) result in partial filtering of the second distance, centred at 2.4 nm, whilst complete filtering is achieved with a T_F of 24 μs . Since it is possible to filter out the entirety of this distance, we can conclude that 2.4 nm corresponds to the distance between Cu²⁺ species, most probably in the peripheral binding sites. Figure 3.38 gives

both the baseline corrected time evolutions, and distance distributions obtained at each T_F using the *i*DEER experiment, in each case plotted against the regular DEER experiment for demonstration of distance filtering.

Further *i*DEER experiments were undertaken at values either side of those determined as the most effective for spectral filtering, to demonstrate the gradual disappearance and reappearance of each distance as a function of T_F . The results of these experiments are shown in Figure 3.39 for C267S_{SL}.

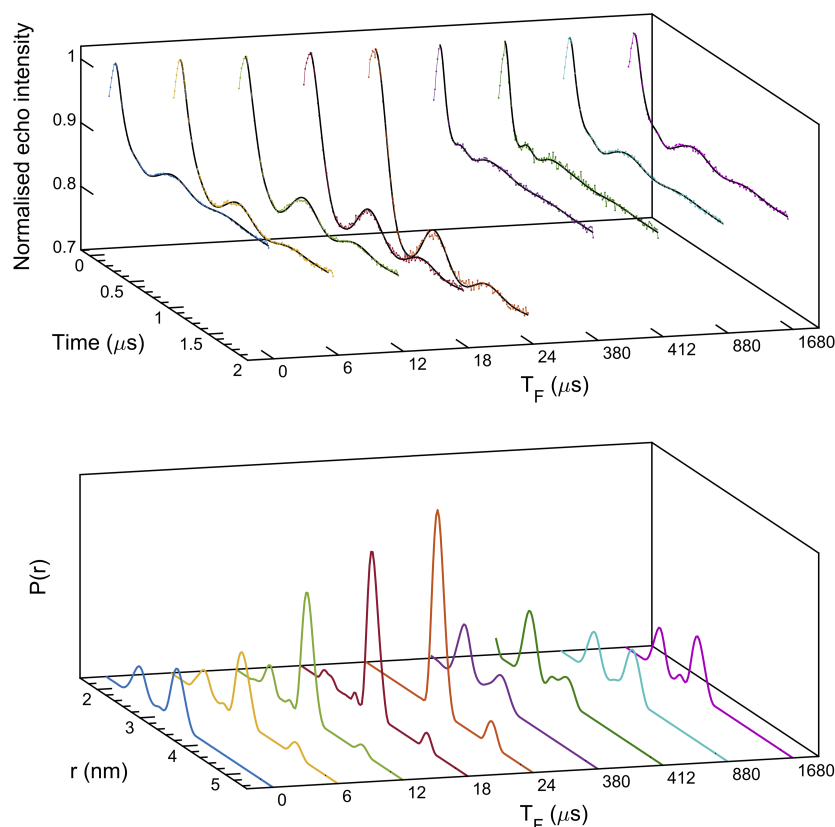


Figure 3.39: *i*DEER experiments(50 K) for C267S_{SL} following incubation with excess Cu²⁺ at a range of filter times (T_F). (Top panel) The normalised background-corrected dipolar evolution (coloured lines) for experiments at each filter time (T_F). Spectral fittings are shown as black lines. (Bottom panel) The associated distance distributions (coloured lines), with standard deviations in the form of 95 % confidence intervals also given in each case and represented by light grey areas around the lines. Fitting of all dipolar evolutions and subsequent distance distributions were calculated with DeerAnalysis.¹⁹ Experimental parameters are equivalent to those in the previous figure.

3.8 Discussion and conclusions

A number of conclusions can be drawn from the data presented in this chapter in relation to the function of MamM-CTD, especially when considered in parallel with published results using alternative structural and biophysical techniques for protein characterisation. Indeed many of these ideas and conclusions have been discussed as part of the published manuscripts in which a large portion of the EPR data within this thesis is presented.^{115,154}

Altogether, the EPR data obtained for C267S_{SL} MamM-CTD in its apo state would suggest rapid fluctuations between different conformations in solution. This is consistent with MD simulations, from which both a native-like open-close movement and a swivelling movement are predicted.¹¹⁵ Solved crystal structures (via X-ray diffraction) for both the WT^{112,113} and C267S MamM-CTD,¹⁵⁴ however, show the dimeric protein in an 'open', V-shaped conformation; this state might be attributable to stabilisation of the open conformation by crystal packing interactions which are not present in solution. cw-EPR results for both CSM-C267S_{SL} and PSM-C267S_{SL} at RT indicate potential differences in apo conformations, or the flexibility exhibited by dimers, in solution, as a result of binding site mutations. Results of DEER experiments with CSM-C267S_{SL} give increasing evidence for sampling of a more tightly-packed dimeric structure than is possible for either C267S_{SL} or PSM-C267S_{SL}. This result can be rationalised when we consider the nature of central binding site residues; removal of bulky histidines and charged aspartates facilitates a tighter dimeric closure and thus has an affect on the apoprotein dynamics in solution. This result is again in agreement with those of MD simulations, as demonstrated in the published manuscript highlighting these findings.¹¹⁵

DEER data obtained following binding experiments for C267S_{SL} MamM-CTD with Cu²⁺ revealed a well-defined bimodal distance distribution, with distributions centred at 2.4 and 3.3 nm. The latter inter-spin separation indicated by these results fits well with the predicted SL-SL distance based on MMM analysis of the Cu²⁺-bound X-ray crystal structure, which represents a 'closed' conformation of the dimer. Meanwhile, additional DEER and *i*DEER experiments confirmed this 3.3 nm distribution to be due to the coupling of two nitroxides, whilst the shorter 2.4 nm separation was shown to be due to coupling between (paramagnetic) Cu²⁺ ions. According to Figure 3.31A (bottom panel), this Cu²⁺-Cu²⁺ separation fits well with the distance between Cu²⁺ ions bound to the peripheral sites of the closed conformation; data on both CSM-C267S_{SL} and PSM-C267S_{SL} are consistent with this interpretation. It appears that binding of Cu²⁺ at either the peripheral or central binding sites causes conformational stabilisation to this closed state. Considering all results it seems likely that the solved crystal structure with Cu²⁺ represents well the Cu²⁺ binding sites and conformation adopted in solution, the closed nature of which is shown by the DEER results; this is despite the low value for *N* as determined by ITC measurements.¹⁵⁴

EPR data obtained for C267S_{SL} MamM-CTD with Mn²⁺ suggests no binding of this metal ion, and thus a discriminatory behaviour of MamM-CTD against Mn²⁺. These results,

coupled to those measured by our collaborators using alternative techniques,¹⁵⁴ provide the first indication of metal selectivity in MamM-CTD. This is suggestive of a CDF CTD-related mechanism for discrimination against Mn^{2+} (in MamM), or potentially other metals in different transporters. This behaviour by MamM-CTD, when put in the context of the biological role of MamM in magnetosome formation, can be explained by the fact that Mn^{2+} exhibits a higher affinity to magnetite than other divalent metal cations and can in fact be incorporated into iron-based magnetic nanocrystals upon exposure of magnetotactic bacteria over a prolonged period of time.^{183,184} This in turn changes the magnetic and structural properties of the magnetic nanocrystals, hence affecting their function.¹⁸⁵ *In vivo* there is an additional TMD associated with CTD of MamM and other CDF proteins which is also required for metal ion transport, but given the presumed orientation of the protein in the MM (i.e. the CTD facing the cytoplasm and the TMD facing the magnetosome lumen), metal ions present within the cytoplasm will encounter the CTD first. Thus discriminatory behaviour of the CTD against Mn^{2+} might be an evolved mechanism to prevent contamination of the magnetic nanocrystals.

The EPR data obtained for C267S_{SL} following incubation with Zn^{2+} , Cd^{2+} , Ni^{2+} , Co^{2+} and Fe^{2+} also indicates a 3.3 nm separation between SL moieties, supporting the interpretation that MamM-CTD adopts a closed conformation upon metal binding. Experiments undertaken on binding site variants of MamM-CTD with Zn^{2+} reveal that conformational change in C267S_{SL} is likely induced by binding at the peripheral sites; this is in contrast to results with Cu^{2+} , in which the same conformation appears to be adopted regardless of which proposed binding sites are eradicated. This has led to conclusions to be drawn surrounding allostery in MamM-CTD, specifically for Zn^{2+} binding - we propose that binding at the CS occurs first, increasing allostery at both PSs, and it is binding at the PSs which causes the closed conformation seen in the Cu^{2+} -bound crystal structure.¹¹⁵ In addition, binding of Zn^{2+} by CSM-C267S_{SL} affords a more tightly packed, closed conformation than is exhibited by either C267S_{SL} or PSM-C267S_{SL}. This again is attributed to the removal of CS binding residues.

Additional differences can be seen in the sharpness of distance distribution features centred at 3.3 nm, most specifically for Ni^{2+} and Fe^{2+} . Low modulation depths were also observed in both cases. For Ni^{2+} , the latter might be a consequence of a Ni^{2+} -mediated quenching mechanism, or alternatively a smaller proportion of spins interacting; meanwhile for Fe^{2+} this is likely explained by a low concentration of Fe^{2+} (and therefore Fe^{2+} -bound MamM-CTD) in solution due to precipitation effects already discussed. In either case,

these differences cannot be explained in terms of binding affinity when considering those of other (lower affinity) metals which exhibit sharper peaks and larger modulation depths. If Ni^{2+} binding was non-specific, as could be inferred from the crystal structure, then binding-induced conformational change as is shown by this spectroscopic DEER data is still possible. However taken together with an N value of approximately 3, indicating binding of three ions per dimer (as is seen in the Cu^{2+} -bound crystal structure), and the tryptophan fluorescence quenching close to the central site, it is most likely that Ni^{2+} exhibits a binding mode similar to that of Cu^{2+} and Zn^{2+} in which one metal ion binds to the central site, and one binds at both peripheral sites, with conformational locking to the same closed conformation. Similarly, the observed 3.3 nm-centred distance affords us confidence in the relevance of studies undertaken with metals other than the natural substrate of MamM, owing to the fact that this same distance is observed throughout. It is therefore likely that the Fe^{2+} -bound conformation is similar to that observed in the Cu^{2+} -bound crystal structure, and that in general the conformations indicated by the DEER distance distributions discussed so far reflect biologically relevant states of the protein.

Results of the DEER experiment in the presence of Co^{2+} are intriguing; this is the second example of a bimodal distance distribution for C267S_{SL} as a result of metal interaction. The most prominent feature of the calculated distance distribution sitting at 2.1 nm is entirely unique when compared with all other cases. Meanwhile both distances within the distribution (centred at 2.1 and 3.5 nm) are represented by particularly broad features, indicating that the conformations reflected by these inter-spin distances are less stable than those induced by binding of other metals discussed. Since the Co^{2+} EPR signature does not overlap with that of MTSSL, any distances observed are a result of SL-SL interactions. It might be possible that binding of Co^{2+} at different sites results in alternative conformations, and thus multiple features in the distance distribution. This would indicate a unique binding mechanism of Co^{2+} with C267S_{SL} . Alternatively, interaction might be non-specific thus inducing somewhat random preferred dimer conformations in solution. Further structural and biophysical data describing the interaction of MamM-CTD with Co^{2+} is not available for comparison, but from this DEER result alone we can conclude that interaction of MamM-CTD with Co^{2+} is unique.

Displacement experiments conducted with C267_{SL} and multiple metal ions indicate that MamM-CTD does not bind multiple different metals at one time. This might indicate that metal binding sites for all metals tested are conserved, or simply that the metal binding requires the involvement of certain conserved ligands (i.e. histidines), meaning that a

high affinity metal will always displace a low affinity one. However, considering the nature and similarity in conformations adopted by MamM-CTD as a result of metal binding in all test cases other than Mn^{2+} , I would suggest that the former argument is most likely. Since contaminated magnetite particles are rarely found within MTB *in vivo*, it is likely that further discrimination against all metals shown here to bind to the CTD other than Fe^{2+} , occurs within the TMD at the transport site, which thus provides a second recognition site. The tighter packing induced by metal binding is likely to afford the conformational change required of the TMD for the transport of Fe^{2+} through this domain, in return for protons, and the resulting formation of pure magnetite.

3.8.1 Conclusions

Overall, the distance distributions obtained for MamM-CTD with Zn^{2+} , Cd^{2+} , Ni^{2+} , Fe^{2+} and Cu^{2+} indicate metal-bound conformations in solution which are equivalent to that seen in the Cu^{2+} -bound crystal structure model. A number of other CDF CTD crystal structures have been solved in either their apo, Zn^{2+} -bound, or both states; each was presented briefly in the introductory section of this chapter. However the DEER experiments conducted and discussed here give some of the first structural insights into binding by a CDF CTD to divalent metal cations other than Zn^{2+} , as well as discriminatory behaviour against a specific metal ion - Mn^{2+} . Additionally, the *i*DEER technique has been successfully utilised to filter out detection of interactions between either SL-SL or Cu^{2+} - Cu^{2+} pairs for the first time, demonstrating a method for probing multiple inter-spin distances within a sample containing multiple paramagnetic species. These studies provide a meaningful and complementary contribution to the understanding of the structure-function relationships of CDF CTDs, when considered in parallel with results of other biophysical and structural methods of characterisation. There is huge scope for the continued characterisation of metal binding by MamM-CTD using EPR methodologies. An as-yet untouched avenue is the study of the full length protein using EPR, for which no structure is currently available. Results could, however, initially be mapped onto structural models based on the full length YiiP structure in order to probe events that occur downstream of metal binding at the CTD.

Chapter 4

Advancing approaches for solving structural dynamics, modelled on GltPh

4.1 Summary

The following chapter presents, firstly, an introduction to the glutamatergic synapse as well as the specific membrane transport proteins involved in the control of glutamate levels at the synapse. Published structural and biophysical investigations of GltPh - a related glutamate and aspartate transporter - are presented to provide an overview of the general structure and function of GltPh, as well as similarities/differences when compared with its human counterparts. The first section of experimental results includes the application of conformation 'locking' variants in combination with EPR methodologies for the study of conformational change in GltPh - these are discussed with reference to published EPR studies in order to demonstrate the potential of such an approach in deconvolution of complex data. Next, results of STD NMR studies for K_D determination for multiple substrates of GltPh are presented, in different protein environments. These results are discussed in terms of the affects of protein environment on substrate affinity, as well as cooperativity between protomers of this trimeric protomer - evidence presented for the latter is not seen elsewhere in the literature. Overall, these studies demonstrate new approaches, and the application of techniques which are not routinely in use for the study of membrane proteins, on GltPh as an already well-characterised, model transport system.

Here it is noted that the STD NMR results presented within this chapter are also published within a manuscript of which I am the primary author. This manuscript is given (in full) within the appendices of this thesis; all data and text has been edited and reformatted for inclusion within this thesis.

4.2 The SLC superfamily

An overview of the different types of membrane transporter was given at the beginning of chapter 2. From this we already know that movement of specific molecules across the membrane via a transporter protein is typically associated with a fixed stoichiometry of molecules and/or ions. One major superfamily of membrane transporters is known as the solute carrier (SLC) superfamily, which is primarily involved in the specific uptake of small molecules into cells. The superfamily consists of 395 human membrane-spanning transporters divided into 52 families.³⁴ SLC transporters range in specificity, from those that interact with only a small number of substrates and are therefore highly specific, including amino acid transporters, to those with a very wide range of chemically diverse substrates. Recycling of neurotransmitter molecules from the synaptic cleft during neurotransmission at the chemical synapse is regulated by secondary-active transporters known as neurotransmitter transporters.¹⁸⁶ Human glutamate transporters are members of the SLC1 family of transporters.

4.3 Neurotransmission within the central nervous system

4.3.1 The chemical synapse

Communication between neurons in the human brain is controlled by the chemical synapse. An important feature of all chemical synapses is the presence of synaptic vesicles, holding one or more neurotransmitters, within the terminal of the presynaptic neuron. It is these chemicals that when released act as messengers for communication between neurons. An action potential at the terminal of the presynaptic neuron allows neurotransmitter release, thereby increasing the concentration of neurotransmitter in the synaptic cleft. Following exocytosis, the neurotransmitters diffuse across the synaptic cleft and bind to specific receptors on the membrane of the post-synaptic neuron. Binding causes channels in the postsynaptic membrane to open/close, affecting ion flow and depolarising the postsynaptic neuron, causing a postsynaptic receptor potential. Membrane transport proteins act to clear the neurotransmitter from the synaptic cleft, decreasing the concentration of transmitter to basal level, and leaving the synapse ready for the next cycle of activation.¹⁸⁷

4.3.2 The glutamatergic synapse and glutamate transporters

L-glutamate is an excitatory neurotransmitter of glutamatergic synapses (Figure 4.1) in the mammalian central nervous system. Glutamatergic synapses are responsible for fast excitatory neurotransmission, and are essential for normal development and function, including learning and memory. Various neurological diseases and problems are linked to dysfunction of such synapses, including amyotrophic lateral sclerosis, Alzheimer's disease, stroke, cerebral ischemia, traumatic brain injury, epilepsy and Huntington's disease. The clearance of glutamate from the synapse via sodium-dependent transporters needs to be fast and efficient for regular and usual excitatory neurotransmission and prevention of glutamate-induced excitotoxicity.¹⁸⁸

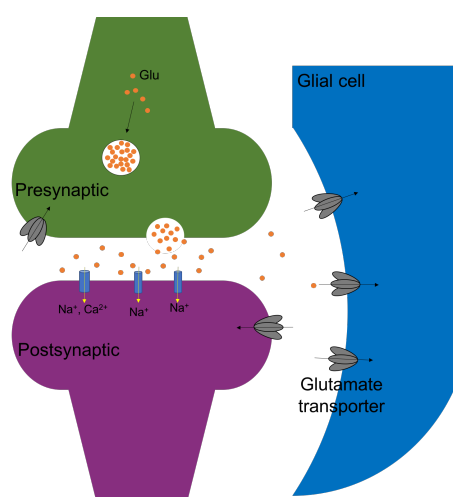


Figure 4.1: A schematic representation of the glutamatergic synapse, highlighting the role of glutamate transporters (EAATs) in the clearance of glutamate from the synaptic cleft.

EAATs (excitatory amino acid transporters) are membrane proteins within the mammalian central nervous system (CNS), and they play an important role in regulating the concentration of glutamate in the extracellular space by transporting it, as well as other ions, across cellular membranes. Following the release of glutamate as a result of an action potential along a neuron, glutamate transporters quickly remove the neurotransmitter from the extracellular space, maintaining low concentrations and thereby terminating synaptic transmission. This prevents the build-up of glutamate, leading to excitotoxicity, and also allows glutamate to be recycled for repeated release.^{189,190} More specifically, glutamate is present in presynaptic vesicles which fuse with the presynaptic membrane, allowing glutamate to enter the synaptic cleft. Glutamate molecules then bind to specific membrane-bound/spanning receptors located at pre- and post-synaptic neurons and glial cells. In ionotropic receptors, this binding leads to conformational change and opening of intrinsic ion channels. Influx of Na^+ (and/or Ca^{2+} , depending on the type of receptor) triggers membrane depolarisation

of the post-synaptic neuron which is sufficient to induce signal transmission. Metabotropic receptors are slower acting and work via a different mechanism, but in both types binding glutamate ultimately causes signal transmission. Glutamate transporters are then hugely important for clearance of excess glutamate in the extracellular space and prevention of over-excitation of the neuron. In humans, there are five subtypes of glutamate transporters - EAAT1-5, belonging to the solute carrier 1 (SLC1) family of transporters, and each have slightly different properties that regulate their ability and mechanism of glutamate transport. Examples of the characteristics of glutamate transporters that will affect synaptic neurotransmission include their concentrating capacity, binding affinity to glutamate molecules, the rate at which glutamate transport is carried out, intrinsic chloride conductance, expression levels of the various subtypes and the way in which a particular synapse is made up.¹⁹¹

4.3.3 Excitatory amino acid transporters

The first three glutamate transporter subtypes belonging to the SLC1 family were cloned and classified nearly simultaneously. The first was named GLAST1 (GLutamate ASpartate Transporter 1),¹⁹² and the second named GLT1 (GLutamate Transporter 1) [193], both isolated from the rat brain. The third, EAAC1 (Excitatory Amino Acid Carrier 1),¹⁹⁴ was isolated from the small intestine of a rabbit. A standardised nomenclature was later introduced following the cloning and classification of the human homologues of these three glutamate transporter subtypes, based on the acronym EAAT.¹⁹⁵ Based on the percentage of amino acid sequence identity shared between pairs, EAAT1 was defined as the human homologue of rat GLAST1 (sharing 95 % identity), EAAT2 the human homologue of GLT1 (sharing 95 % identity) and EAAT3 the human homologue of rabbit EAAC1 (sharing 92 % identity). Following this, EAAT4 was later cloned and classified as a human aspartate/glutamate transporter expressed predominantly in the cerebellum,¹⁹⁶ followed by EAAT5, the final human subtype of the EAATs classified, expressed primarily in the retina.¹⁹⁷ Cloning and classification of these five subtypes of the SLC1 family has provided insight into their locations, their function with respect to substrate and ion specificity, and ideas on their roles in various diseases of the nervous system.

4.3.4 Stoichiometry of substrate transport and the alternate access model

Excitatory amino acid transporters are electrogenic transport systems which undergo secondary active transport for the movement of glutamate across the membrane, using the free energy stored in the cross-membrane electrochemical gradients of the coupling

ions.³⁵ Soon after the cloning of glutamate transporters it was found that transport of substrates is coupled with the translocation of net charge across the membrane,¹⁹⁸ where it was also suggested that glutamate transport is associated with the co-transport of more than one Na^+ ion. More extensive studies for the understanding of substrate and ion coupling stoichiometry were carried out with EAATs 1-3, leading to the final conclusion that one glutamate anion is co-transported with one proton and three Na^+ ions, and the counter transport of one K^+ ion.^{199,200} This allows the transporters to maintain up to a 10^6 -fold glutamate concentration difference across the membrane of a neuron or astrocyte, or to keep extracellular levels at low nanomolar concentrations.²⁰⁰ A thermodynamically uncoupled, glutamate-gated chloride conductance within glutamate transporters was also recognised,¹⁹⁶ highlighting their ability to act as both secondary active transporters as well as ligand-gated ion channels (Figure 4.2). This uncoupled chloride conductance would limit the amount of glutamate available for synaptic transmission, and offset the depolarisation caused by electrogenic glutamate transport, hence limiting neuronal excitability.

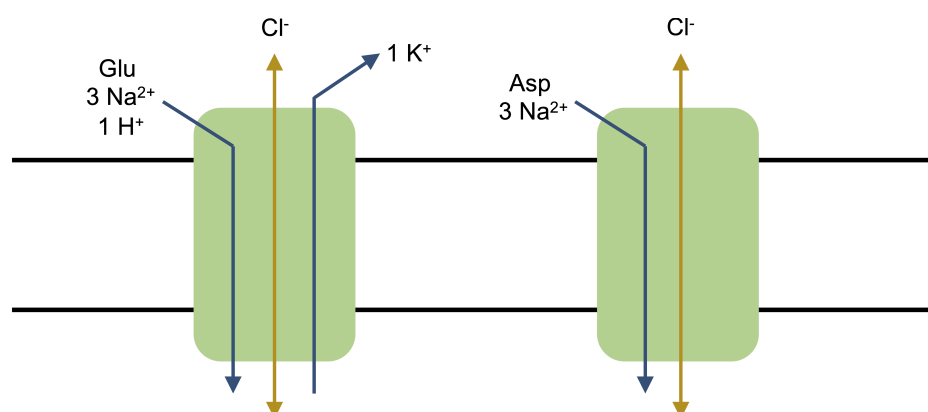


Figure 4.2: Comparison of the transport stoichiometry in mammalian EAATs (left) versus archaeal glutamate transporters, including GltPh (right). Black parallel lines represent either side of a membrane bilayer, in which the transporters (green) are stabilised. Arrows depict the direction of substrate and ion transport.

The most common model for secondary active transporters involve 'alternating access', first proposed in 1957.²⁰¹ According to this model, the binding sites for substrates and ions are exposed to either one side of the membrane or the other, but never both at once, during transport. Multiple versions of the model have been suggested. In the 'rocker-switch' mechanism,^{202,203} the transporter is made up of two domains which take part in a rigid-body rocking motion relative to one another. In this case, whilst access to the binding site by one side of the membrane is permitted, access by the other side is prohibited (see Figure 4.3A). The 'two-gated pore' mechanism²⁰⁴ describes a pore, spanning the membrane, which is terminated by a gate at either end. Access to binding sites by either the extracellular or

intracellular sides is therefore controlled by the opening and closing of these gates, and only one will open at any one time. Binding of substrate and ions from one side causes closure of the gate on that side - the state in which both gates are closed and the substrate is trapped within the protein is termed the occluded state. Meanwhile, the elevator mechanism involves translocation of molecules, in which a transport domain (housing binding sites) travels across the membrane against a rigid scaffold domain (see Figure 4.3B).

Upon release from the presynaptic terminal, glutamate molecules diffuse across the synapse and bind to glutamate receptors, generating a postsynaptic signal. Understanding the ability of EAATs in controlling glutamate concentrations is greatly aided by the alternating access model. The various EAAT proteins differ in their functional capacity, and therefore to what extent they may influence glutamate concentration within the synapse. The main events that are likely to influence this when considering an alternating access mechanism are: (a) the rate at which the substrate binds the transporter; (b) the likelihood of a transport event; (c) how quickly the glutamate is transported; and (d) the amount of glutamate transporters close to the synapse.¹⁹¹

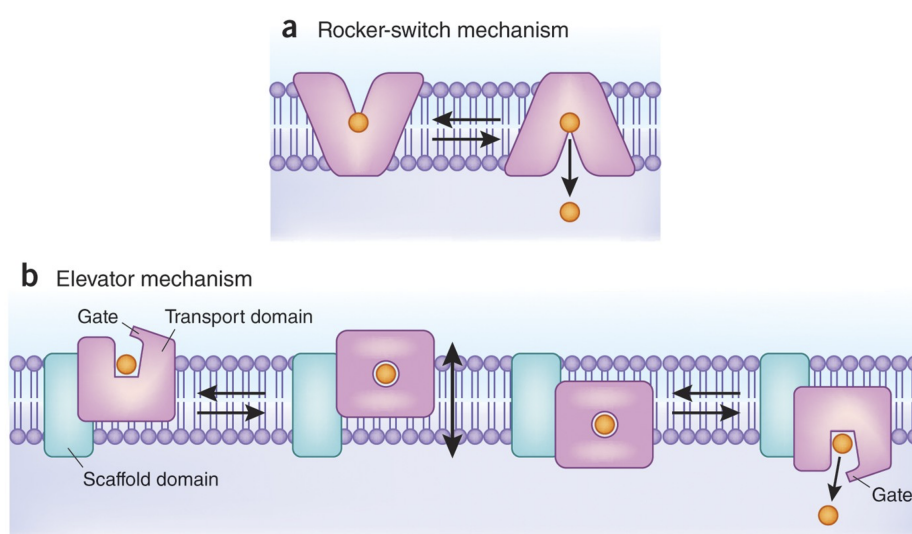


Figure 4.3: Examples of alternating-access models in membrane transporters, including (a) the rocker-switch mechanism and (b) the elevator mechanism. Glutamate transporters are known to function via the latter of these mechanisms. Figure taken from Slotboom 2014.²⁰⁵

4.4 GltPh - an archaeal homologue of the mammalian EAATs

A major breakthrough in the understanding of glutamate transporter structure was made in 2004, when the first crystal structure of an archaeal homologue from *Pyrococcus horikoshii* (GltPh) was solved, and since this time GltPh has been studied extensively via various

methods.²⁰⁶ Prokaryotic and eukaryotic glutamate and neutral amino acid transporters share significant amino acid sequence identities throughout their entire polypeptides.²⁰⁷ The carboxy-terminal half of both eukaryotic and prokaryotic transporters has been found to be responsible and essential for substrate binding, transport and ion coupling,²⁰⁸ whilst residues in the amino-terminal half have been found in eukaryotic transporters to be involved in the uncoupled chloride conductance.²⁰⁹ GltPh shares around 35% sequence identity with the mammalian EAAT proteins, and importantly many of the functionally important amino acids, involved in substrate and ion binding, are highly conserved between GltPh and its human homologues. GltPh has also has an intrinsic chloride conductance much like in the case of the EAATs;²¹⁰ these factors combined make GltPh an excellent model system for understanding mammalian glutamate transporters further. With further crystal structures of GltPh solved in different conformations, in the presence of various substrates and inhibitors,^{211–217} these studies have provided a structural framework for the study of the structure-function relationship of glutamate transporters, as well as allowing further insight into the molecular basis of function.

GltPh binds and transports both L-aspartate and L-glutamate, with increased selectivity for the former.²¹⁸ In contrast, the mammalian EAATs transport L-glutamate and L-aspartate with similar efficiency.²¹⁹ The binding affinity of GltPh for L-aspartate (reported K_D values in the nanomolar range)^{211,214,220–223} is much higher than in the EAAT proteins (reported K_D values in the micromolar range)^{211,222}. Another difference between GltPh and its human homologues is that neither co-transport of a proton with substrate, nor counter-transport of a K^+ ion are required for substrate binding and translocation.²¹⁸ Substrate turnover rates differ between GltPh and the EAATs, with one L-aspartate molecule transported approximately every three minutes at 30 °C in GltPh^{218,224} compared with every 80 ms in EAAT2.²²⁵ Despite these differences, GltPh is considered an excellent model system to elucidate structural and functional information related to the EAAT proteins - the following sections will provide some insight into the types of studies that have been used for this purpose thus far.

4.4.1 General structure of GltPh

The first solved crystal structure of GltPh represents an outward-facing state made up of three identical monomeric units.²⁰⁶ The bowl-shaped homotrimer contains an extracellular-facing, hydrophilic basin, and a pointed base which faces the cytoplasm. The basin, which

is the depth of approximately half the trimers height and protrudes well into the plane of the membrane, affords solvent exposure on the extracellular side to roughly midway through the membrane bilayer.

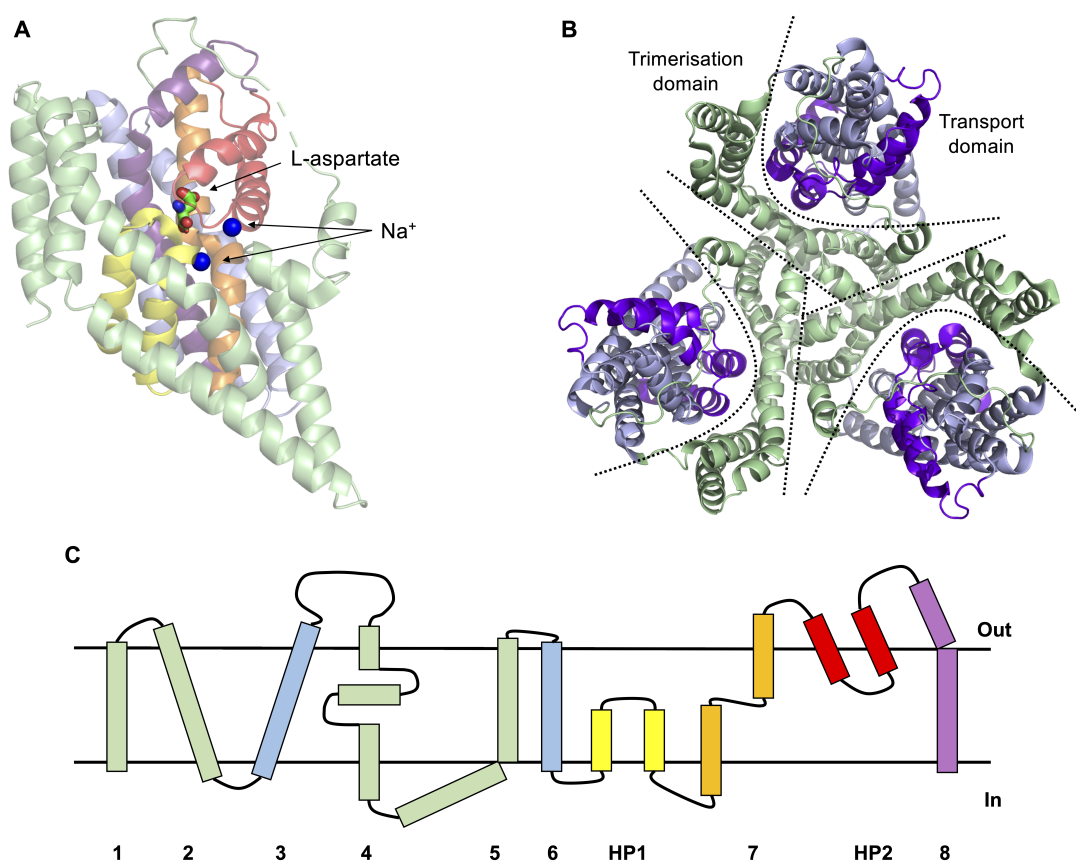


Figure 4.4: (A) Cartoon representation of a single GltPh protomer, bound to Na⁺ ions (blue spheres) and L-aspartate (sticks)(PDB code: 2NWX). The transport domain, consisting of HP1 (yellow), TM7 (orange), HP2 (red) and TM8 (violet) is highlighted. (B) Extracellular view of trimeric GltPh (PDB code: 1XFH) with trimerisation domains shown in green and transport domains shown in blue. Straight lines represent separation of different protomers, whilst curves lines represent separation of trimerisation and transport domains. (C) A schematic representation of the GltPh fold; TM helices of the trimerisation domain (TMs 1, 2, 4 and 5) are coloured green, whilst TM helices and HPs of the transport domain (TMs 3, 6, 7, 8 and HPS 1 and 2) are highlighted by different colours. Figures (A) and (B) were both made using PyMol.

Each wedge-shaped protomer consists of eight primarily α -helical transmembrane domains (TMs 1-8) and two helical hairpins (HPs 1 and 2). Protomers are described as containing both a trimerisation (or scaffolding) domain, and a transport domain.²¹² The former is made up of TM1, TM2, TM4 and TM5, it forms a rigid scaffold and mediates inter-subunit contacts; the latter comprises TM3, TM6, HP1, TM7, HP2 and TM8, and houses all substrate and ion binding sites. Each promoter therefore contains its own individual binding sites (see Figure 4.4). It was first suggested by Vernool *et. al.* that once substrate and ions are bound, a transport pathway through each protomer to the cytoplasm is formed via the movement of HP1 relative to other structural elements, thus acting as an intercellular gate. Meanwhile, HP2 which resides above the substrate binding site was suggested to act as an extracellular gate; the combination of both gates allows for an alternating access mechanism in which

substrates and ions can only be exposed to one side at any one time.²⁰⁶ HP1 and HP2 are therefore also termed re-entrant loops.

4.4.2 Substrate and ion binding sites

Co-crystallisation of GltPh with L-aspartate and thallium ions (TI^+) afforded identification of GltPh ion and substrate binding sites in 2007.²¹¹ The substrate L-aspartate binds within a buried, polar chamber which reaches halfway across the membrane bilayer; this binding site is formed by the tips of HP1 and HP2, the unwound region of TM7 and polar residues of TM8. Solvation of the highly charged L-aspartate molecule is afforded by a large number of polar contacts. The side-chain of D394 (TM8) and main-chain carbonyls of both R276 (HP1) and V355 (HP2) stabilise the amino-group of the substrate. Meanwhile interactions with the α -carboxyl of L-aspartate involve the amide nitrogen of N401 (TM8), the hydroxyl of T398 (TM8) and the main-chain nitrogen of S278. Finally, the guanidinium group of R397 (TM8), the hydroxyl of T314 (TM7) and the main-chain nitrogen of G359 (HP2) interact with the γ -carboxyl group (see Figure 4.5A).²¹¹ These residues are highly conserved between GltPh and EAAT substrate binding sites, making the selective preference for L-aspartate over L-glutamate in GltPh difficult to understand.

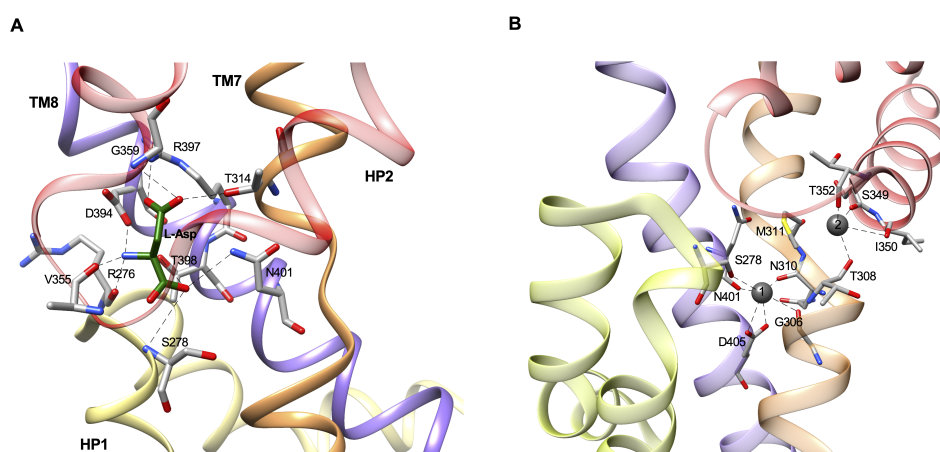


Figure 4.5: Cartoon representation of binding sites as seen in structural models for (A) L-aspartate and (B) Na^+ and Na^+ in GltPh (PDB code: 2NWX). Binding residues are labelled in each case, whilst interactions with specific groups are represented by dashed lines. Figures made using UCSF Chimera (version 1.14).¹¹⁸

Transport of substrate in both GltPh and the EAATs is coupled to the co-transport of three Na^+ ions; two sodium ion binding sites were solved in this same study by co-crystallisation with TI^+ (which has a large anomalous scattering signal in comparison to Na^+).²¹¹ Upon examination of anomalous difference Fourier maps of crystals soaked in thallium nitrate,

two strong peaks were observed per subunit, as a result of the presence of Ti^+ at these sites. Competition experiments using Li^+ , Na^+ and K^+ revealed that only Na^+ was capable of Ti^+ displacement, confirming that the Ti^+ -labelled sites are indeed genuine Na^+ binding sites. Both Na^+ binding sites found in each monomer are in close proximity to that of L-aspartate. It was inferred that Na1 (Na^+ bound at site 1) is coordinated by residues in TM7 and TM8, buried deep within the protein and binds below L-aspartate. Buried just under HP2, Na2 (Na^+ bound at site 2) is coordinated by four carbonyl oxygens of TM7 and HP2 (see Figure 4.5B). The location of the third binding site for Na^+ was not detected, however a number of computational studies have been conducted for elucidation of this site, named Na3.^{226–230} This predicted site was later confirmed in the solved crystal structure of GltTk, in which all three Na^+ binding sites were resolved using Na^+ rather than Ti^+ .²³¹ This third site is found between the unwound region of helix 7 (containing the conserved NMDGT motif) and TM3, with the Na^+ ion bound 8 Å from Na1.

4.4.3 Crystallographic studies

Structural models of GltPh, solved using with X-ray crystallography or cryo-electron microscopy, have provided huge gains in the understanding of structure and structure-related function in glutamate transporters. The occluded nature of the substrate binding site solved with L-aspartate left the question of how exactly the substrate molecule reaches its binding site from the extracellular or intracellular solution.²¹¹ However crystallisation of GltPh with a non-transportable, competitive inhibitor - D,L-threo-beta-benzyloxyaspartate (DL-TBOA) - revealed HP2 in an 'open' conformation (Figure 4.6). The positioning of HP2 is allowed by its movement by up to 10 Å from its position in the L-aspartate-bound state, and ultimately lead to the hypothesis that HP2 serves as an extracellular gate. The structure was solved by placing DL-TBOA into unmodelled density observed in data from DL-TBOA-exposed crystals²¹¹ - the bulky benzyl group of this inhibitor appears to lodge against the tip of HP2, holding it open. This is afforded by contacts made with G359 in HP2, and interaction with M311 in TM7; the latter residue is highly conserved and thus gives a structural explanation for the inhibition by DL-TBOA observed throughout eukaryotic transporters as well as GltPh. The role of HP2 as an extracellular gate was further evidenced in the first ion-only-bound crystal structure of GltPh;²¹⁵ the observed open position of the hairpin tip is thought to facilitate L-aspartate access to its binding site and to stop the inward movement of the ion-only-bound transport domain.

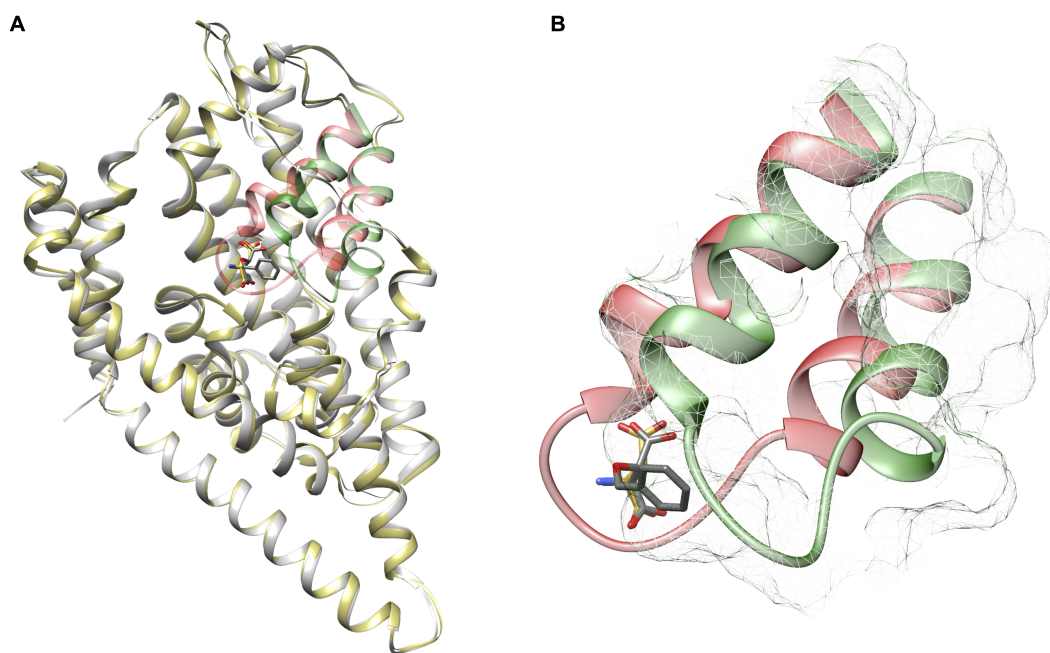


Figure 4.6: Relative position of HP2 (the extracellular gate) in the presence of L-aspartate versus DL-TBOA. (A) gives an L-aspartate-bound and DL-TBOA-bound protomer overlaid (PDB codes: 2NWX and 2NWW respectively), with transport domains shown as yellow and grey ribbon, respectively. HP2 is highlighted in red (L-aspartate) and green (DL-TBOA) in each case, whilst L-aspartate (yellow) and DL-TBOA (grey) molecules are shown as sticks. (B) shows HP2 and either substrate or inhibitor alone in each case, highlighting the open tip of HP2 in the DL-TBOA-bound state, relative to its closed position with L-aspartate.

A further step in the transport mechanism was elucidated in 2009, with the first solved crystal structure of GltPh in a substrate-bound 'inward-occluded' conformation.²¹² This was obtained via cysteine cross-linking of a double cysteine mutant with either copper 1,10-phenanthroline (CuPhen) or divalent mercury (Hg^{2+}) - point mutations were undertaken at positions K55 (TM2) and A364 (HP2) for cysteine introduction. The most striking difference in this structure compared with those previously solved (see Figure 4.7) is the substrate binding sites of each monomer, which have moved by 18 Å from the bottom of the bowl close to the extracellular solution, to a position much closer to the cytoplasm. Transmembrane segments making up the trimerisation domain are largely unchanged, whilst the remainder of the protein has undergone substantial movement. Further solved apo structures later revealed similar findings,²¹⁵ in that the entire transport domain shows identical structures when facing either outward or inward; this suggests movement of the transport domain as a rigid body relative to the trimerisation domain even in the apo state. This gives some insight into how the transport domain might translocate when free of substrate and ions to complete the transport cycle (i.e. from an empty inward-facing conformation to an empty outward-facing conformation, ready for re-binding and transport). For capturing of each apo, ions-only-bound and fully bound GltPh in outward-facing states, a single point mutant of GltPh was used (R397A) which displays a drastically decreased affinity for substrate. For inward-facing structure determination the same double cysteine

mutant was used as previously described²¹², cross-linked using mercury.

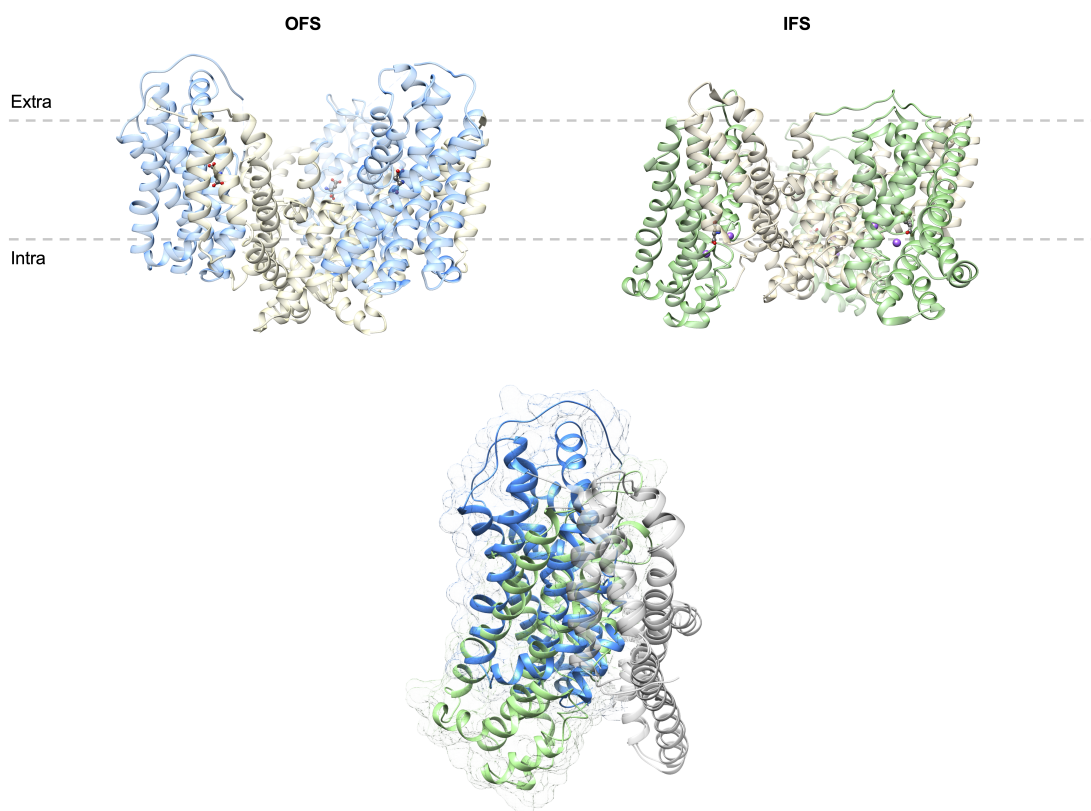


Figure 4.7: (Top) Trimeric GltPh in outward-facing substrate-bound (left, PDB code: 2NWL) and inward-facing substrate/ion-bound (right, PDB code: 3KBC) conformations, with substrate and ions shown as sticks and spheres respectively. Trimerisation domains are shown as grey/yellow ribbon, whilst OF and IF transport domains are blue and green respectively. Dashed lines represent either side of the membrane bilayer, affording visualisation of transport domain movement from OF to IF states. (Bottom) Overlaid single protomers of GltPh in OF and IF states, with trimerisation domains shown as grey ribbon, and transport domains shown as blue (OF) and green (IF) ribbon, with transparent mesh. Transport domain movement is clear between conformations, whilst the trimerisation domain remains stable.

This same cross-linked variant has been utilised much more recently for structural determination of GltPh reconstituted into nanodiscs,²¹⁷ to explore the motions of the transport domain in the IFS as a result of ion, substrate and inhibitor binding. Substrate-bound structures replicated nearly exactly those solved previously both with²¹² and without substrate²¹⁵ in inward-facing occluded conformations, in which HP2 occludes the substrate binding site and packing of both domains leaves no space for its opening. Notably, structures solved for apo-open, ions-only-bound, DL-TBOA-bound and TFB-TBOA-bound reveal that whilst regions of HP2 and TM8a which typically take part in domain-packing remain juxtaposed against the scaffold (trimerisation domain), a large proportion of the transport domain swings out and away from HP2 and the trimerisation domain. The extent of this movement is variable, depending on the protein state, with the largest splaying observed for IF GltPh bound to TFB-TBOA. This leaves the question of whether this 'swinging-out' of the transport domain occurs in order to allow subsequent opening of HP2 for substrate release, or whether this movement alone is sufficient for substrate binding site access on

the intracellular side.

The first solved crystal structure of GltTk from *Thermococcus kodakarensis*, a closely related homolog of GltPh, in an entirely apo state afforded further understanding of the mechanism by which the substrate-free transporter 'resets' following substrate transport, since somehow the transporter must reorient to the outward facing state.²³² This revealed an overall structure which is very similar to that of GltPh, with the transport domain adopting an outward-facing state and both helical hairpins occluding the substrate binding site. It was thus concluded that movement of the apo transport domain across the membrane will most likely occur in the same manner as that of the substrate-bound domain. The usual substrate binding site is partially filled by the overlap of a highly conserved Arg residue in this apo structure, stabilised by ionic interactions with a conserved Asp which, in substrate-bound GltPh, interacts with the amino group of the substrate molecule. Further rearrangements around this Arg create a pocket, proposed to be a potential potassium ion binding site in the mammalian transporters. Soaking of apo GltPh crystals with Ti^+ later revealed additional cation binding sites, with the protein remaining in an apo-like state;²¹⁵ one such site overlaps with that of the substrate and as such binding of cations here would compete with binding of both Na^+ ions and substrate. It was concluded that whilst the closed, empty state of GltPh and other prokaryotic homologs appears to be stable, this might not be the case in mammalian transporters. Thus binding of potassium cations at equivalent sites would afford coupling of K^+ counter-transport and completion of the transport cycle.²¹⁵

Binding of non-transportable inhibitors such as DL-TBOA and TFB-TBOA is believed to prevent closure of HP2 following binding, hence preventing transport. Both blockers, DL-TBOA and TFB-TBOA, share the same backbone structure as L-aspartate, but have additional benzyl ring/s which cannot fit within the substrate binding site; transport is believed to be blocked by their binding to the OFS of transporters, forcing HP2 into an open position with no room for closure. Solving of the OFS bound to both inhibitors in nanodiscs confirmed this same conformation is adopted in the context of a lipid bilayer, without the effects of crystal contacts.²¹⁷ The IFS crystal structure with bound TFB-TBOA²¹⁷ revealed a wide-open conformation of HP2 much the same as that seen in the OFS of DL-TBOA-bound GltPh.²¹¹ The IFS of DL-TBOA-bound GltPh, however, does not display this same HP2 conformation - in this case the extracellular gate remains mostly closed, with its N-terminal arm separated from its C-terminal arm. This movement appears to alter one Na^+ binding site, giving further evidence for a different Na^+ binding stoichiometry with this inhibitor as was already proposed.^{214,216} This ultimately opens up a small crevice into the cytoplasm;

whether or not this conformation has biological relevance is still unclear.

4.4.4 Biophysical characterisation

Key methodologies which have been used for biophysical characterisation of function-related conformational change and rate constants for GltPh include fluorescence-based techniques and EPR spectroscopy. Both techniques require the attachment of small organic radicals (for EPR), or fluorophores, to cysteine residues which have often been purposely placed in positions likely to report on conformational fluctuations. WT GltPh contains a single cysteine which, when not required, can be mutated to an alternative residue with little effect on the overall protein structure nor function.²²⁴ Such biophysical methodologies have been employed in the context of protein solubilised in detergent micelles, and reconstituted into proteoliposomes.

Multiple fluorescence based analyses have informed on binding constants for both aspartate and Na^+ ions. Measured dissociation constants (K_D s) for L-aspartate are in the nM range,²¹¹ exhibiting a Hill coefficient close to 1, indicative of independent binding in each protomer. Sodium-dependence of L-aspartate binding was also observed.^{211,214} Na^+ titrations have been conducted and monitored by fluorescence changes, giving apparent K_D s for Na^+ in the mM range.^{211,221} Studies indicate that at least two Na^+ ions are coupled to L-aspartate binding;^{211,214,221} binding of two Na^+ ions to the empty protomer was found to be cooperative even in the absence of substrate.^{214,221} The dependence of DL-TBOA binding on Na^+ was investigated and found to be much weaker, with binding of the inhibitor coupled to only one Na^+ ion. The affinity of the transporter for Na^+ is higher in the presence of substrate, with Na^+ binding being slow and likely rate-limiting, involving conformational changes that are essential for the subsequent binding of substrate. This ultimately leads to the conclusion that uptake of L-aspartate is only afforded in the presence and upon protomer binding to Na^+ ions.^{214,221} Additionally, the chemical potential of Na^+ ions is converted to substrate binding energy.

Despite the vast availability of crystal structures in multiple conformations, the question still remains as to which of these conformations are related to, and required for substrate transport, and whether the movement of individual subunits correlates with the transport event. To address such questions, multiple single molecule fluorescence resonance energy transfer (sm-FRET) studies have been undertaken, investigating conformational dynamics in

the context of GltPh reconstituted into proteoliposomes. In all such studies, the protein was found to undergo 'elevator-like' movements between OF, intermediate and IF conformations, depending on the protein state (i.e. empty, bound only to Na⁺ ions, and bound to both ions and substrate or inhibitor).^{220,224,233} An intermediate conformation appeared to be frequently adopted from either OF or IF conformations.²²⁴ It was suggested that this might serve as a 'barrier' between OF and IF extremes in the absence of L-aspartate, hence playing a role in prevention of Na⁺ leakage.²²⁴ Whilst a structural basis for this 'barrier' was suggested, no further experiments were conducted to test this hypothesis.

Protomers were found to move and function independently of one another, in corroboration with multiple other studies. In some cases, the kinetics of conformational fluctuations were found to be heterogeneous, sampling long quiescent (dormant) periods with occasional dynamic periods, the latter of which are believed to reflect transmembrane movements of transport domains.^{220,233} These dynamic states are observed in both the empty transporter, and when bound to both ions and substrate, with the lifetime of the quiescent periods (representing stable states) being longer in the latter case. It was thus suggested that there is allosteric coupling between substrate binding and interactions at the domain interface.²³³ Upon addition of DL-TBOA, the occasion of rapid dynamics decreases by more than 80 %, giving a strong indication that these dynamics are as a result of substrate translocation across the membrane. Additionally, the short-lived states adopted as a result of these rapid periods of activity are thought to be structurally different from conformations seen in crystal structures, exhibiting lower substrate binding affinities.

The use of gain-of-function 'humanising' variants (which result in higher rates of substrate uptake) in combination with physiological ion gradients revealed destabilisation of the previously observed quiescent 'locked' states.²²⁰ This translates to an increase in transporter dynamics paralleled to a decrease in substrate affinity, supporting previous findings. The absence of one FRET state seen in the WT but not in the 'humanising' variant was explained by a crystal structure, revealing the presence of all IF protomers but with transport domains exhibiting entirely different conformations than seen previously. Increased stability of the OFS in the presence of bound ions and substrate indicated that transport domain movements might be the rate-limiting step in the 'humanising' mutant; altogether introduction of such mutations affects both transporter dynamics and substrate affinity, despite the mutations not forming part of the substrate binding site. This provides further evidence for coupling between substrate binding and transport domain dynamics.²²⁰

High-speed atomic force microscopy (HS-AFM) has developed over recent years for real-time, real-space observation of proteins in native-like environments; for GltPh this has afforded analysis of transport dynamics of the membrane-embedded protein.²³⁴ Movements were observed initially of the substrate free state of the transporter, revealing recurring transitions between OF and IF conformations. Addition of Na^+ only resulted in a significant decrease in the number of moving protomers, whilst those still moving remained largely OF. Such a lack of movement again reveals a mechanism for prevention of Na^+ leakage. Dynamic motions largely returned under saturating concentrations of both Na^+ ions and substrate, although some remained quite dormant. This might correlate with long quiescent periods observed in previous SM-FRET studies.^{220,233} The motions of individual transport domains within each trimer were not found to correlate, hence indicating a lack in cooperativity.

The first EPR study on GltPh²³⁵ included SDSL of double cysteine variants in combination with cw-EPR to probe the conformational changes that occur in HP2, previously proposed to function as an extracellular gate for control of Na^+ /substrate translocation.²¹¹ Residues for mutagenesis and subsequent labelling were chosen based on previously solved crystal structures, showing variability in the position of HP2 in the presence of either DL-TBOA or L-aspartate, defined as open and closed states respectively.^{206,211} The strategy employed cw-EPR measurements of each labelled cysteine pair in maximally-labelled and under-labelled protein (equal to approximately one spin label per pair of residues), reconstituted into proteoliposomes. It was found that in the apo state, HP2 adopted a conformation similar to that of the L-aspartate-bound closed state, whilst upon Na^+ binding the conformation of HP2 resembled that of the DL-TBOA-bound open state. A proposed binding mechanism was suggested in which the closed state of HP2 in the apoprotein prohibits any binding of aspartate in the binding pocket before Na^+ is bound. Na^+ binding then affords opening of this gate to the extracellular side, and subsequently aspartate binds in its binding site. Then closure again of HP2 effectively closes off exposure of the substrate to the extracellular side.

In 2013, two EPR studies on GltPh were published almost simultaneously, both of which focused on the use of DEER for probing large-scale conformational change in both the scaffold and transporting domains of GltPh, in the context of both detergent and proteoliposome environments.^{236,237} In one study,²³⁷ DEER data gave a good indication that the scaffold domain forms the stable core of the protein, whilst the transporting domains were captured in multiple conformations, in all states (i.e. with/without coupled ions/substrates). Distance changes observed in proteoliposomes were smaller than those in detergent,

which led to the suggestion of more intermediate/asymmetric conformations sampled in proteoliposomes. They also suggested that equilibrium constants for each transporter conformation are of a similar magnitude, since multiple conformations in all protein states were observed. It was concluded that the transport domain is continually moving across the membrane, ready for binding/transport/expulsion of ions and substrates, and that local movements such as those shown for HP2²³⁵ afford ion and substrate loading and expulsion. The second study²³⁶ demonstrated much the same findings, describing the scaffold domain as the main structural core, as well as multiple conformations in all states, with the energies of each conformation being similar. In specific variants, the lipid environment lead to more pronounced single conformations as compared with other variants, caused by differing interaction of spin label side chains with the interface between scaffold and transporting domains. The distance changes observed in proteoliposomes were again smaller than those exhibited in the detergent environment; this was attributed to a more compact protein structure in the membrane. In both studies there are discrepancies as to which conformation is preferred in which state, depending on where the spin label pairs are positioned.

A more recent study demonstrated a novel method for the *in vitro* assembly of multimeric membrane transport proteins for (a) the study of cooperativity between protomers and (b) the deconvolution of complex and confusing spectroscopic data when dealing with multimeric proteins.²³⁸ GltPh heterotrimers were assembled from a mixture of WT GltPh and V216C-I294C-GltPh variant subunits and subsequently spin labelled. A comparison of DEER measurements from control V216C-I294C-GltPh homotrimers vs heterotrimers revealed sharper, more defined distance distributions for the heterotrimer, in which the number of possible distances measured is greatly reduced. Interestingly, these results revealed the spin labelled monomer of the heterotrimer quite clearly adopts an inward-facing conformation, in both apo and substrate-bound states.

4.4.5 Overall mechanism of substrate and ion transport

An amalgamation of results such as those described here has lead to what is currently understood as the transport mechanism in GltPh (see Figure 4.8). Starting with the apoprotein, the extracellular gate that is HP2 is able to adopt an open conformation, resulting in accessibility of substrate and ion binding sites to the extracellular environment in the OFS of the transporter. Upon binding of the first two Na⁺ ions to the OFS, the tip of HP2 remains open meaning uncoupled uptake of Na⁺ doesn't occur. The slow and weak binding of these two Na⁺ ions affords conformational movement to a state that can bind both the

substrate molecule, and the final Na^+ ion. This is followed by closure of HP2, occluding both substrate and ions from the extracellular side. The actual transport cycle involves binding and release of substrate at either side of the membrane; this process utilises an electrochemical gradient of Na^+ ions. Translocation of L-aspartate across the membrane involves elevator-like movements, in which the transport domain moves relative to the rigid scaffold (trimerisation) domain and traverses the membrane. During this movement, allowing translation between OF and IF states, the transport domain undergoes rotation of 37° , followed by translation of 15-18 Å. This results in binding sites now facing the cytoplasmic side.

When considering substrate and ion release on the intracellular side, Na2 (which is believed to be in rapid equilibrium with the solution) release is required for disassembly of the complex.²¹⁶ To this date, the identity of the intracellular gate remains unclear, as does the specific mechanism by which substrate release occurs. Both HP1 and HP2 have been suggested in different studies to play a role in substrate and ion release in the IFS, however further investigation is still required to solve this part of the mechanism. Molecular dynamics simulations coupled with free energy calculations have indicated that initial unbinding of Na2 is followed by intracellular gate opening, allowing L-aspartate release, followed by the release of Na1 and Na3 in succession, with the latter being rate limiting.²³⁹ Substrate release allows the transport domain to adopt an occluded conformation which can then move from the IF to OF state, completing the transport cycle. Chloride permeation and an uncoupled chloride conductance is known to occur between the scaffold and transport domains, however the conformational states and residues associated with this process remain unclear.

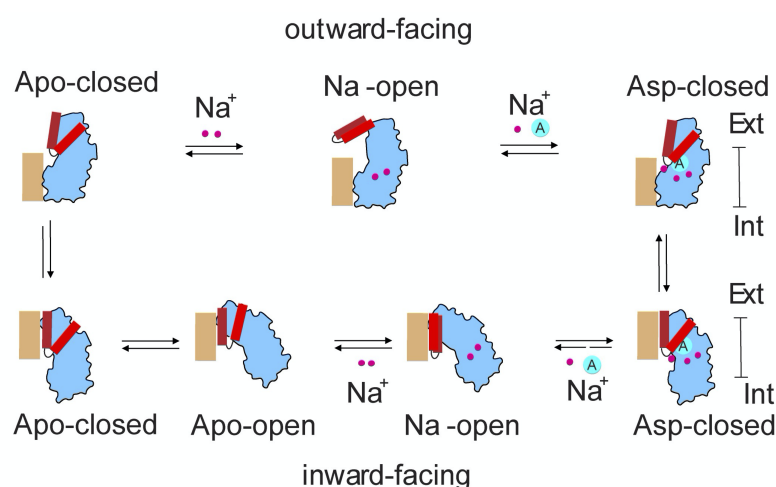


Figure 4.8: Schematic representation of the transport mechanism of GltPh, as suggested by (and taken from) Wang and Boudker 2020.²¹⁷ This highlights their proposed gating mechanism on both extracellular and intracellular sides of the membrane.

4.5 EPR Investigation of GltPh

As has been described in the introductory section of this chapter, a mixture of cw-EPR and DEER studies have been undertaken to probe conformational changes that occur during ion, substrate and inhibitor binding and/or transport. However, results of these studies have been contradictory and often not entirely clear, thus not necessarily providing clear insights into the mechanisms by which GltPh functions.^{235–237} The most recent study described the formation of heterotrimers to aid in such experiments, and indeed afforded far less complex, convoluted distance distributions.²³⁸ However the purification strategies required in heterotrimer formation mean that yields are even more limited. The aim of the EPR studies presented in these next sections was to consider and test alternative, less time-consuming and yield-limiting approaches to address the issue of convoluted DEER data due to multiple labelling sites in this trimeric protein, and thus give a clearer insight into the conformations adopted by GltPh.

4.5.1 SDSL and cw-EPR of a GltPh variant

GltPh is another protein in which there is no inherent paramagnetic centre, and as such SDSL is required in order for any EPR studies to be conducted. For the following studies, a single-cysteine variant of GltPh was utilised - C321A-A71C GltPh. The naturally occurring cysteine at position 321 in the WT protein is in a position that is buried and inaccessible for labelling; indeed MMM studies (based on crystal structures of the protein) predict only 1 or 2 rotamers in each protomer, with associated partition functions that suggest the site is too tight for successful labelling. The alternative labelling position chosen, at residue 71, is located at the base of the trimer, in the loop region between TM2 of the scaffold domain and TM3 of the transport domain (see Figure 4.9).

This position is exposed on the surface of the protein and as such forms an accessible labelling site. There is no evidence in the literature to suggest that mutation at this site should have any effect on the protein structure or function. cw-EPR experiments following efficient labelling will be expected to report on changes in the surrounding environment of, and/or rotational motions of the attached spin label, whilst DEER studies will afford inter-protomer distance determination.

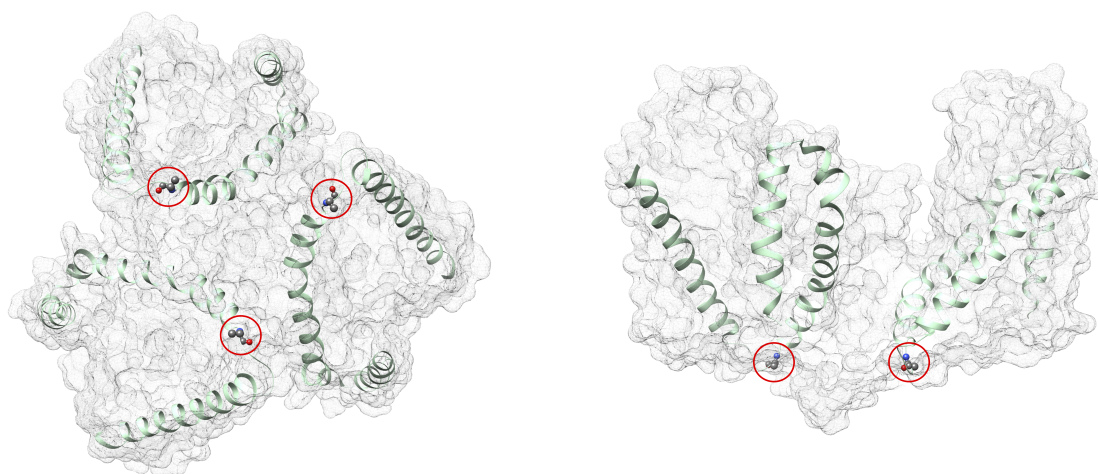


Figure 4.9: Outward-facing GltPh (PDB code: 2NWX) with position 71 highlighted in 'stick and dot' representation. The trimeric protein is shown from the bottom (left) to clearly show the relative position of this residue in each protomer. In the side-on representation (right), chain C has been omitted. TM1-3 in both representations are shown as ribbon, with the remainder of the protein shown as transparent mesh. Figures made using UCSF Chimera (version 1.14).¹¹⁸

MMM analysis, SDSL and cw-EPR measurements

MMM analysis was undertaken (as described previously with MamM in chapter 3.8.1) using MTSSL at position 71 on multiple solved structural models of GltPh, providing analysis of protomers in entirely inward- or outward-facing conformations, and a mixture of conformations. This ultimately allows the prediction of inter-spin distances for GltPh when apo, Na⁺/L-aspartate-bound and Na⁺/DL-TBOA-bound. Results are given in Table 4.1, including the predicted number of rotamers (averaged), partition functions (averaged) and inter-spin distances between specified protomers. Results suggest that position 71 is much more accessible when protomers adopt OF states, as seen in apo, Na⁺/L-aspartate and Na⁺/DL-TBOA-bound states. There is an expected change in inter-spin distance between all-OF and all-IF conformations of approximately 0.3 nm, with the OF structures tending towards longer distances. Since GltPh protomers are predicted to undergo continual conformational fluctuations in the absence of Na⁺ ions/substrate/inhibitor, labelling should be viable under the buffer conditions used.

PDB code	Description	Average number rotamers	Average partition function	r (nm)		
				Chains A/B	Chains A/C	Chains B/C
1XFH	Apo OF	57	0.89	2.63	2.74	2.69
3KBC	Apo IF	10	0.050	2.21	2.21	2.24
3V8G	Asymmetric	5	0.052	1.96	1.95	2.06
2NWX	Na ⁺ /L-asp OF	66	1.08	2.60	2.62	2.61
2NWW	Na ⁺ /DL-TBOA OF	61	0.46	2.63	2.77	2.73

Table 4.1: Results of MMM analysis using structural models for GltPh in multiple conformations and states, including the predicted number of rotamers (averaged between protomers), extinction coefficients (averaged between protomers) and inter-spin distances.

SDSL was attempted to begin with using detergent-solubilised, purified GltPh trimers in a similar manner as is described for MamM in chapter 5.2.6, in which Zebaspin desalting columns are used for removal of reductant and excess MTSSL post-labelling. However this approach caused large losses of purified protein during the multiple wash steps required and as such was deemed unsuitable for this system. The labelling strategy was thus adapted such that labelling occurs during the purification process, whilst the protein is bound to the metal affinity resin and prior to wash and elution steps. The protocol was tested without the use of the reductant dithiothreitol (DTT) since it is incompatible with the metal affinity resin used (due to interference with the cobalt metal ions binding to the resin), and resulted in yields of purified protein equivalent to those in the absence of labelling. The cw-EPR RT X-band spectrum of C321A/A71C GltPh following this SDSL strategy is shown in Figure 4.10, for which the estimated SDSL efficiency is close to 100 %. From now on, C321A/A71C GltPh which is spin labelled with MTSSL will be referred to as GltPh_{SL}.

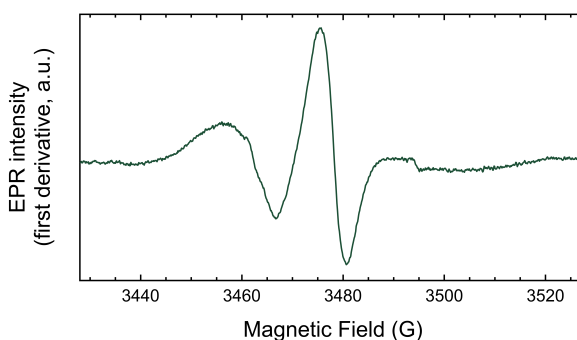


Figure 4.10: X-band RT cw-EPR spectrum of apo GltPh_{SL} following SDSL. Experimental parameters: microwave frequency = 9.75899 GHz, centre field = 3478 G, sweep width = 100 G, modulation frequency = 100 KHz, modulation amplitude = 2 G, microwave power = 0.3965 mW, number averages = 64.

Addition of Na⁺ ions, substrate and inhibitor

Binding and transport of L-aspartate (the high affinity substrate of GltPh) by GltPh is well known to be dependent on the binding of three Na⁺ ions by means of a transport mechanism such as that described in Figure 4.8. It has been suggested that binding of the first two Na⁺ ions affords conformational movement to form the L-aspartate binding site. It is therefore expected that some conformational movement is expected around the Na⁺ and substrate binding sites during this binding process. DL-TBOA is a known inhibitor of GltPh which appears to bind in the same binding pocket as L-aspartate, again only afforded by prior binding of two Na⁺ ions, with its bulky benzyl group prohibiting closure of HP2 and any further steps of the transport mechanism. Our labelling position sits at the base of each protomer of the trimer, and is not involved directly in substrate or inhibitor binding; thus it might not be expected to inform on such local conformational changes. However MMM analysis indicates that this site becomes much more restricted when protomers adopt the IF conformation (reflected by decreased numbers of rotamers and partition functions) - thus we might expect some increased rigidity if this conformation is adopted in solution. MMM also predicts some expected changes in inter-spin distances (see table 4.1) as a result of the movement of the transporter between OF and IF states which might be accessible via DEER experiments.

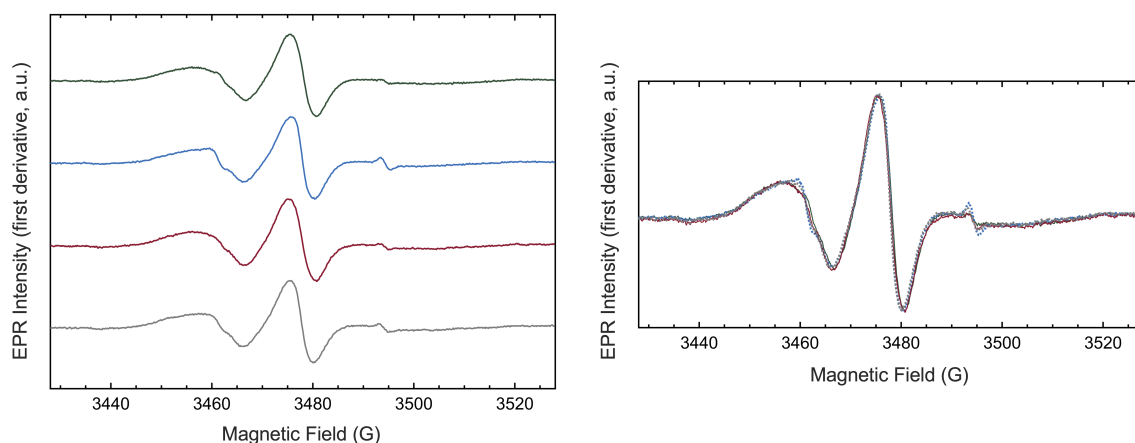


Figure 4.11: X-band RT cw-EPR spectra of apo GltPh_{SL} (green), and following incubation with excess Na⁺ (blue), Na⁺ + L-aspartate (maroon) and Na⁺ + DL-TBOA (grey). Spectra have been normalised to equivalent intensities and are shown in stacked (left) and overlaid (right) representations. Experimental parameters: microwave frequency = 9.759 GHz, centre field = 3478 G, sweep width = 100 G, modulation frequency = 100 KHz, modulation amplitude = 2 G, microwave power = 0.3965 mW, number averages = 64.

Initial tests were conducted at RT using cw-EPR measurements for the apo protein (as above), and following incubation with different mixtures of Na⁺ ions/substrate/inhibitor to probe any changes which might occur locally to the attached spin label. Figure 4.11 gives shows both stacked and overlaid RT cw-EPR X-band datasets for the apo GltPh_{SL} and following incubation with Na⁺ only, Na⁺ and L-aspartate, and Na⁺ and DL-TBOA. As can

be seen from the overlaid plot, changes in spectral line shapes are minimal between protein conditions, indicating that binding events do not induce local changes to the environment of the spin label, nor changes in the rotational motion of the protein itself or the attached spin label. The only notable differences are at approximately 3460 G and 3495 G, in which we can see a slightly different ratio of bound:unbound label between samples. The latter is evidenced by the sharper features centred around these magnetic field strengths.

4.5.2 DEER studies of GltPh_{SL}

In silico predictions of inter-spin distances (based upon MMM analyses) for GltPh_{SL} indicate some expected changes (between apo, ions-only-bound, substrate-bound or inhibitor-bound) depending on the resulting conformation of each protomer. Distances within this predicted range of approximately 2.2 - 3 nm can be effectively probed using DEER spectroscopy. As such, samples were prepared with GltPh_{SL} (a) when apo, and in the presence of excess (b) Na⁺, (c) Na⁺ and L-aspartate and (d) Na⁺ and DL-TBOA. As with the MamM experiments described in chapter 3, samples were mixed thoroughly with approximately 30 % d8-glycerol to act as a cryo-protectant, and to increase the T₂ relaxation time of perturbed spins. Initial FSE measurements afforded determination of pump and observe positions for use in the DEER experiment, which are highlighted in Figure 4.12.

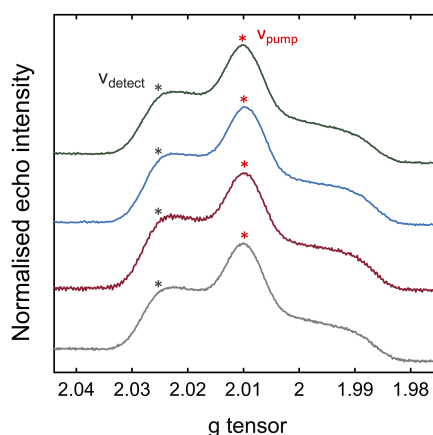


Figure 4.12: X-band field-swept echo spectra measured for apo GltPh_{SL} (T = 50 K) (green), and following incubation with excess Na⁺ (blue), Na⁺ + L-aspartate (maroon) and Na⁺ + DL-TBOA (grey). Pump and observe positions for DEER experiments are highlighted by red and grey asterisks, respectively. Experimental parameters: magnetic field sweep width = 120 G (Na⁺/DL-TBOA) and 150 G (all others), delay time (τ) = 120 ns, number of points = 400-600, shots per point = 50, shot repetition time = 2 ms, number of averages = 4.

The results of subsequent DEER measurements, using the pump and observe positions specified above, can be seen in Figure 4.13. Fitting of all the dipolar evolutions and subsequent distance distributions were calculated with DeerAnalysis¹⁹ using Tikhonov reg-

ularisation. Left panels give the background corrected, time domain dipolar evolutions whilst right panels give the resulting distance distributions for each sample. For the apo case, the distance distribution is very broad; this is indicative of multiple conformations of protomers within solution for which we see an overall envelope of multiple spin-spin distances. Addition of Na^+ alone, and then with L-aspartate, appears to induce a stepwise transition towards a favoured conformation with an average inter-spin distance of approximately 2.3 nm. With Na^+ alone, the protein is sampling fewer conformations than in the apo protein, indicating an effect of Na^+ binding on the overall conformational dynamics of the protomers. Two inter-spin distances are most prominent, centred at approximately 2.3 and 2.8 nm; these are comparable with MMM predictions for GltPh_{SL} in IF (predicted: 2.2 nm) and OF (predicted: 2.6-2.7 nm) conformations, respectively. Notably, the lower of these distances is more prominent and indeed becomes even more so upon addition of L-aspartate. The results for Na^+ and DL-TBOA are not as clear as with L-aspartate, however there is still evidence of a

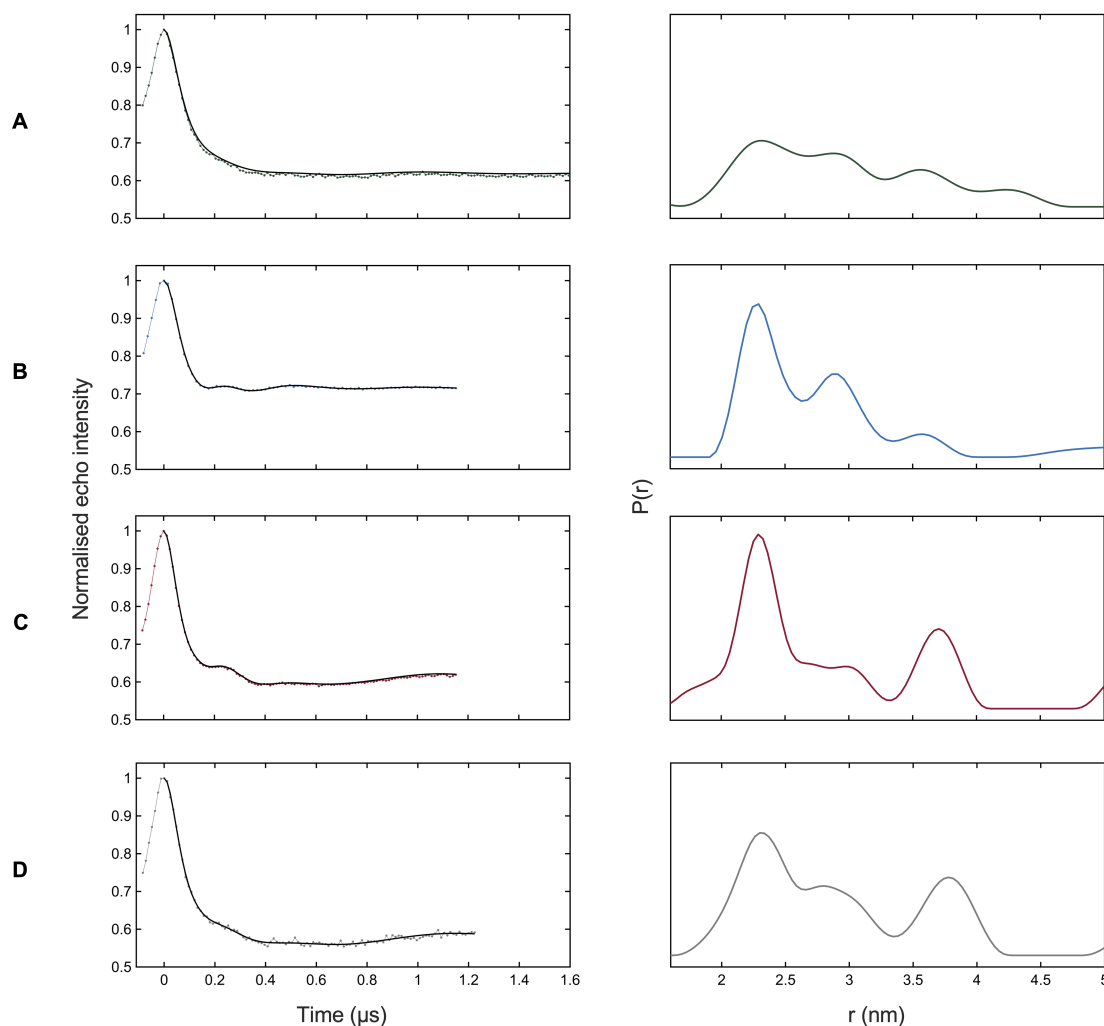


Figure 4.13: 4-pulse DEER experiments. (Left panel) Normalised background-corrected dipolar evolution for GltPh_{SL} under (A) apo, (B) Na^+ -only-, (C) Na^+ /L-aspartate- and (D) Na^+ /DL-TBOA-bound conditions. Fittings are given as black lines overlaid with these plots. (Right panel) Resulting distance distributions for (A), (B), (C) and (D) as described following fitting of the time domain traces by Tikhonov regularisation. Experimental parameters are given in chapter 5.2.6.

higher population of spins at approximately 2.3 nm from one another.

In all cases, and for (C) and (D) in particular, there appears to be an additional distribution of spins at approximately 3.8 nm from one another - for the length of the traces measured (all under 2 μ s), such distances are considered unreliable and are most likely artefacts introduced by the fitting process.

4.6 Conformational locking of GltPh protomers

As mentioned previously, the aim of these EPR investigations is to use GltPh as a model system on which to test approaches for the simplification of DEER results, which on a trimeric system such as this in which many conformational states of the protomers can be adopted, can be quite confusing. As such, we have utilised two additional variants of GltPh which were designed to prevent protomers adopting either OF or IF conformations. Results of EPR measurements conducted on these variants, which contain the same labelling site as GltPh_{SL}, can be used to solve which distances in the previous distance distributions presented correspond to which conformational states of the protomers.

4.6.1 Ideas behind GltPh 'conformation-locking' variants

The two additional GltPh variants produced contain either G79A or K290E point mutations, in addition to the background C321A/A71C mutations, affording spin labelling and direct comparison of EPR results with those described previously for GltPh_{SL}. The former of these mutations - G79A - is intended to cause protomers to favour the OF state. Residue G79 is found within the loop region between TM2 (of the trimerisation domain) and TM3 (of the transport domain); this short loop has been shown previously to undergo defined movements, allowing transition between OF and IF states through movements of the transport domain.²¹²

Specifically, to facilitate the downward movement of TM3 toward the cytoplasmic side it was shown that extension of the TM2-TM3 loop is required, by means of unwinding of the TM3 α -helix by approximately one turn. Three glycine residues within the TM2-3 loop - namely G69, G79 and G82 - were suggested to allow this α -helical unwinding, and to act as hinges. The overall movement of the transport domain and extension of the TM2-3 loop and descent of TM3 is shown in Figure 4.14. It is therefore thought that disruption of this conserved

set of glycines via substitution with the slightly larger, but still non-polar, alanine at position 79 might prevent such movements, resulting in preference by protomers for the OF state.

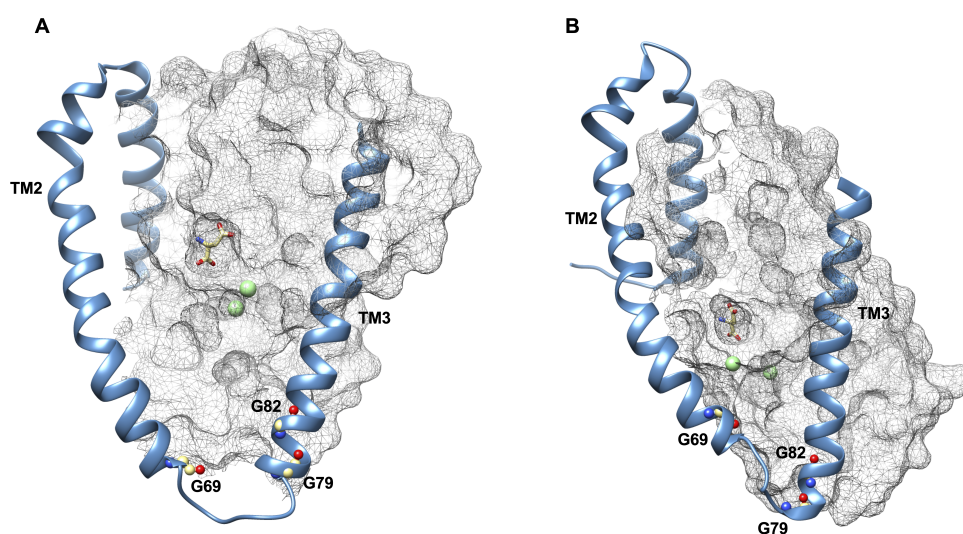


Figure 4.14: Structural representations of TM2-TM3 (cartoon) in (A) outward-facing and (B) inward-facing GltPh protomers, and overall movement of the TM3 and the rest of the transport domain (surface) following unwinding of the TM2-3 loop. Conserved glycines suggested to play a role in affording movement of TM3 are labelled. TM4 and TM5 have been omitted for clarity. Bound L-aspartate (sticks) and Na⁺ ions (spheres) are also shown. PDB codes used: 2NWX (OF) and 3KBC (IF).

Meanwhile the second of our additional mutations - K290E - was chosen with the opposite aim, to destabilise the OF state and thereby force the favouring of protomers for the IF state. K290 is found on the HP1 loop of the transport domain, and is known to participate in a salt bridge with residue E192, which is further stabilised by interactions with the aromatic ring of residue Y195. This network of interactions occurs at the interface between trimerisation and transport domains and stabilises the OF state of protomers.²²³ Previous studies have successfully utilised mutations at position 290 in GltPh for selective destabilisation of the OF state.^{233,240} Figure 4.15 gives the relative positions of residues K290 and E192 in OF and IF states of a GltPh protomer, demonstrating the vast separation required between these residues in the IF state. Our substitution of this basic lysine residue for acidic glutamate should abolish the aforementioned salt bridge, hence destabilising the OF state.

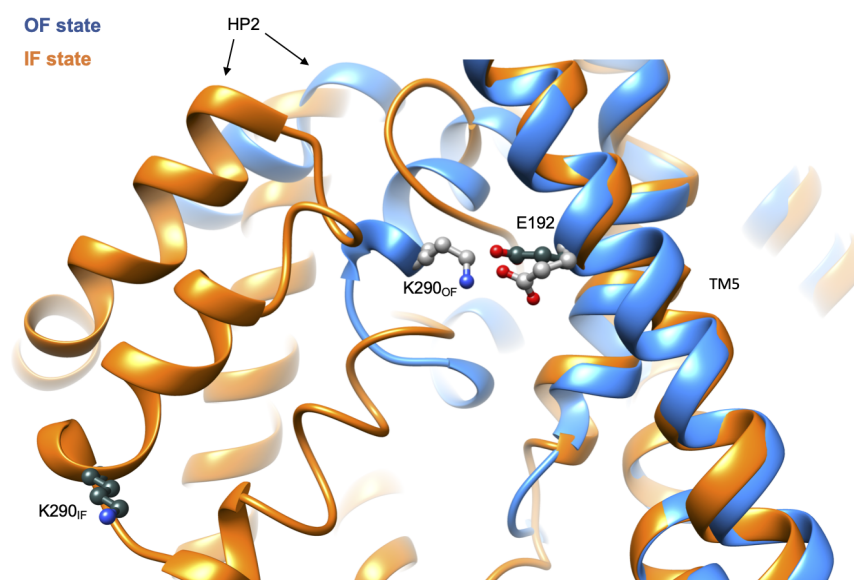


Figure 4.15: Cartoon representation of GltPh in OF (blue) and IF (orange) states, with residues E192 and K290 labelled (shown in stick and dot representation). TM5 and HP2 in each structure are also labelled, which house residues E192 and K290, respectively. PDB codes used: 2NWX (OF) and 3KBC (IF).

4.6.2 SDSL and cw-EPR measurements

SDSL was performed on both C321A/A71C/G79A and C321A/A71C/K290E GltPh variants, utilising the same method as was described for the non-lock variant previously. These will be referred to as GltPh_{SL_{OF}} and GltPh_{SL_{IF}} for variants housing the G79A and K290E mutations, respectively. Labelling efficiencies were equivalent to those accomplished in the absence of the additional conformation-locking mutations, at close to 100 %. RT cw-EPR X-band measurements following SDSL are presented in Figure 4.16, slightly offset from one another; this figure also includes the results for GltPh_{SL} for easy comparison of line shapes. It is clear that the EPR line shapes for each A71C GltPh variant when apo are essentially equivalent, indicating no major change in spin label environment or rotational motion. This is not surprising, since line shapes for GltPh_{SL} remained largely the same after addition of Na⁺/L-aspartate/DL-TBOA, despite evidence of conformational change seen in the DEER spectra. Furthermore, the results for GltPh_{SL_{OF}} and GltPh_{SL_{IF}} demonstrate that the introduction of these variants does not cause any dramatic structural effects on the protein in solution (for example unfolding) other than the conformation-locking effect that we hope to see via further DEER experiments.

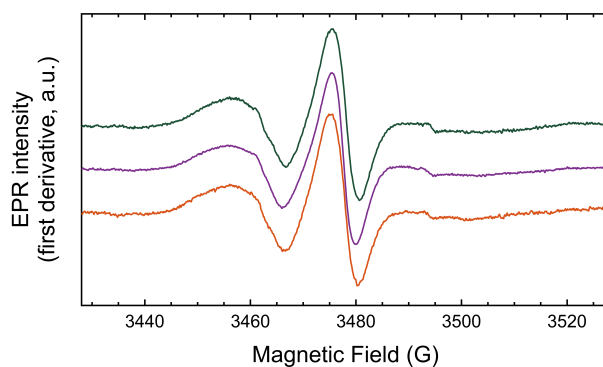


Figure 4.16: X-band RT cw-EPR spectra of GltPh_{SL} (green), GltPh_{SL_{OF}} (purple) and GltPh_{SL_{IF}} (orange) when apo, following SDSL. Experimental parameters: microwave frequency = 9.75899 GHz, centre field = 3478 G, sweep width = 100 G, modulation frequency = 100 KHz, modulation amplitude = 2 G, microwave power = 0.3965 mW, number averages = 64.

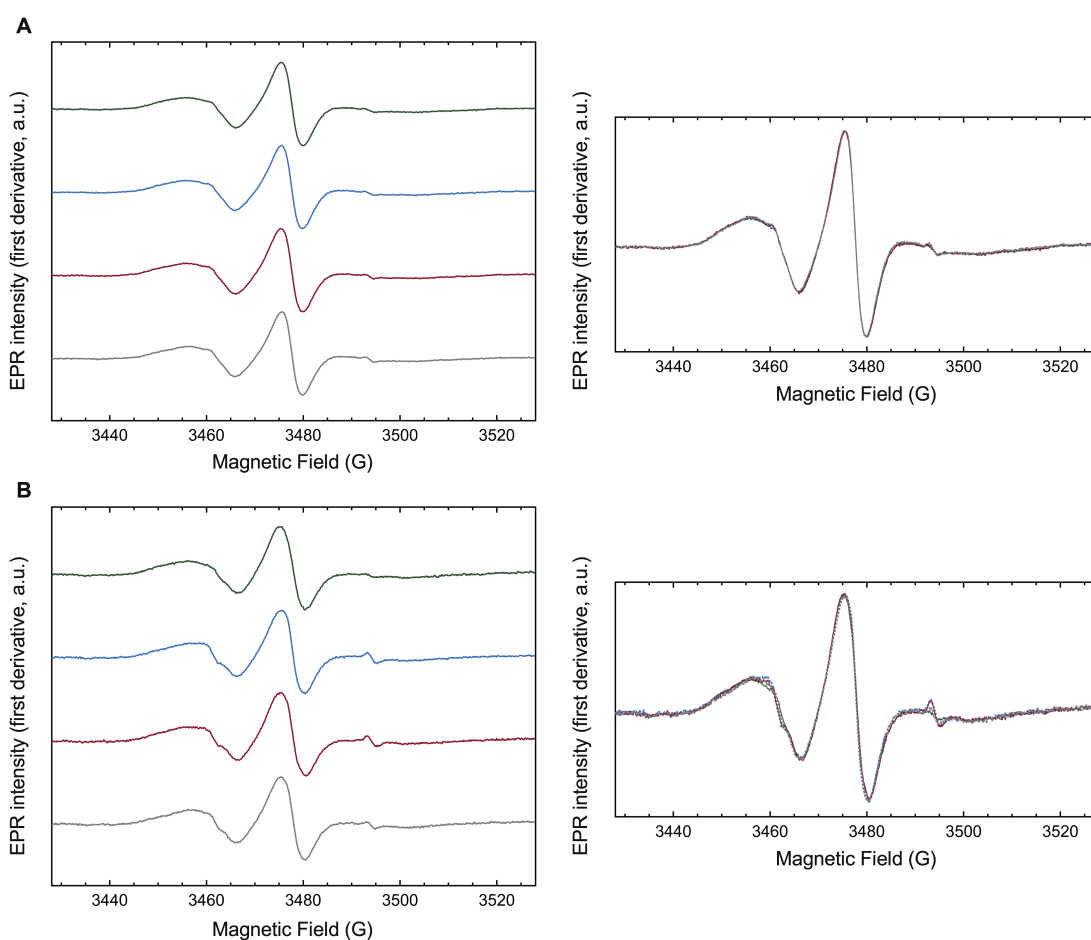


Figure 4.17: X-band RT cw-EPR spectra of (A) GltPh_{SL_{OF}} and (B) GltPh_{SL_{IF}} when apo (green), and following incubation with excess Na⁺ (blue), Na⁺ + L-aspartate (maroon) and Na⁺ + DL-TBOA (grey). Spectra have been normalised to equivalent intensities and are shown in stacked (left) and overlaid (right) representations. Experimental parameters: microwave frequency = 9.759 GHz, centre field = 3478 G, sweep width = 100 G, modulation frequency = 100 KHz, modulation amplitude = 2 G, microwave power = 0.3965 mW, number averages = 64.

Further RT X-band cw-EPR measurements were undertaken with both lock variants in the presence of Na⁺ alone, Na⁺ + L-aspartate, and Na⁺ + DL-TBOA under the same conditions as were used with GltPh_{SL}. The experimental results are presented in Figure 4.17A and 4.17B, for GltPh_{SL_{OF}} and GltPh_{SL_{IF}}, respectively. These spectra are once again

almost equivalent for each measurement of the series in both cases, indicating no effect of ion/substrate/inhibitor binding on the local environment or rotational motion of the attached spin label.

4.6.3 DEER studies of conformation-locked variants

The aim of the DEER experiments described in this section was to demonstrate the use of conformation-locking variants in clarification of conformations adopted in solution by GltPh which is not locked in either conformation. Thus DEER experiments have been undertaken with both GltPh_{SL_{OF}} and GltPh_{SL_{IF}} under all conditions as were measured for GltPh_{SL}. Equivalent pump and observe positions as are shown in Figure 4.12 were chosen in the experimental setup. Results of these experiments can then be directly compared to hopefully aid in the analysis of distance distributions seen in the absence of conformation locks. Firstly, the results of GltPh_{SL_{OF}} are presented in Figure 4.18, including background-corrected time domain spectra with their associated fittings, as well as the corresponding calculated distance distributions. Since this variant should cause protomers of GltPh to favour the OF state, we would expect that distances corresponding to inter-spin interaction between labelled protomers in IF conformations to be eradicated. In all cases (apo, with Na⁺, Na⁺/L-aspartate and Na⁺/DL-TBOA) the calculated distance distributions are very broad, with very little variance between different apo and bound protein states. The highest intensity point in each distribution does sit at around 2.8 nm (which is similar to MMM predictions based on all-OF GltPh structures), however these distributions are so broad that a single, stable conformation does not seem to be adopted by protomers in any case. Importantly, the affects of Na⁺ only, and Na⁺/L-aspartate addition are not the same as in the case of GltPh_{SL} - specifically their addition does not seem to induce conformations which result in an a more intense and narrow inter-spin distance distribution centred at 2.3 nm as was seen with GltPh_{SL}.

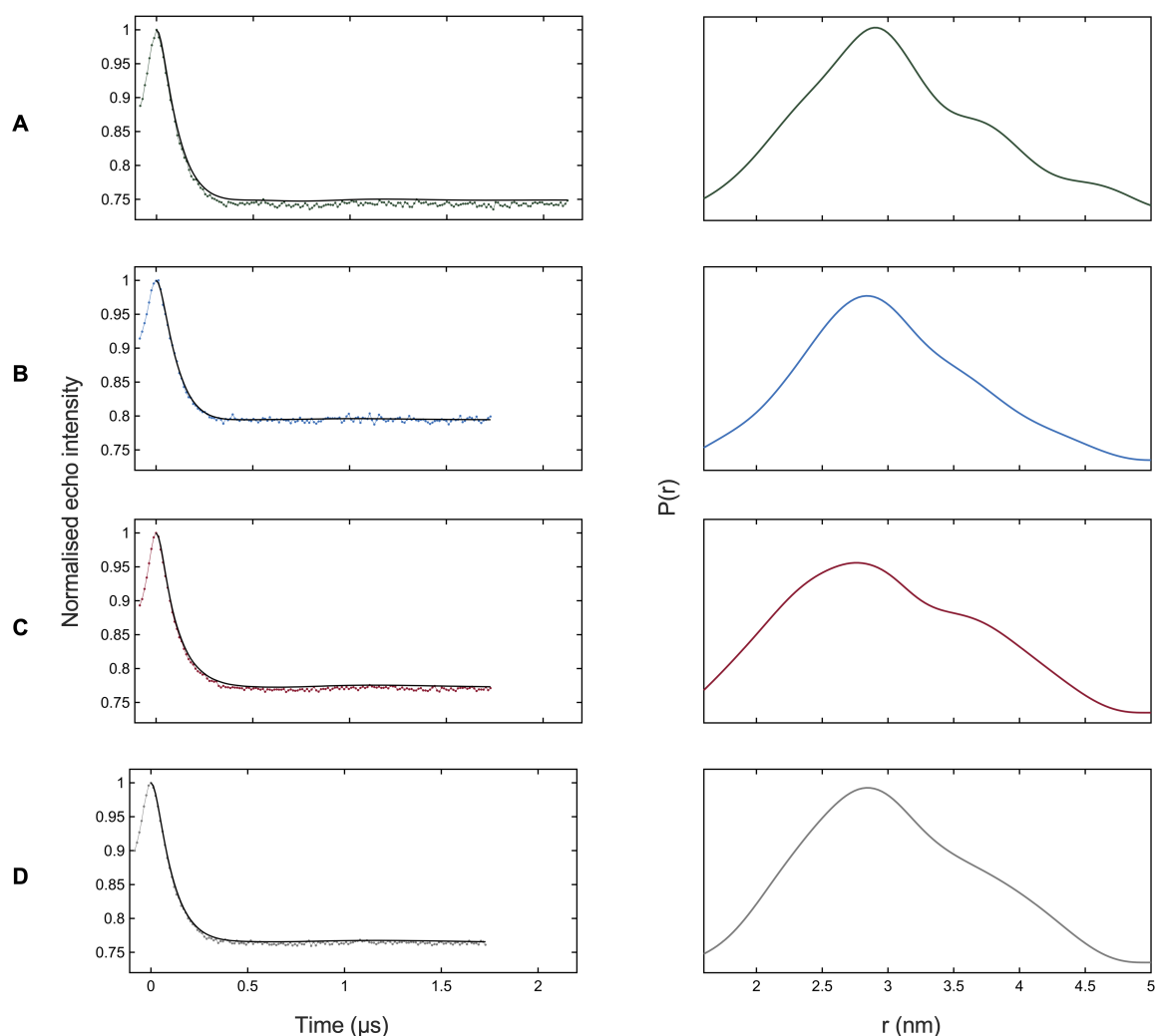


Figure 4.18: 4-pulse DEER experiments. (Left panel) Normalised background-corrected dipolar evolution for GltPh_{SL_{OF}} under (A) apo, (B) Na⁺-only-, (C) Na⁺/L-aspartate- and (D) Na⁺/DL-TBOA-bound conditions. Fittings are given as black lines overlaid with these plots. (Right panel) Resulting distance distributions for (A), (B), (C) and (D) as described following fitting of the time domain traces by Tikhonov regularisation. Experimental parameters are given in chapter 5.

Results for GltPh_{SL_{IF}}, which should in this case prevent OF conformations of GltPh protomers, are indeed different to those described above; these results are presented in Figure 4.19. Here we can see in all protein states that the inter-spin distances seen tend towards the short distance that was observed for GltPh_{SL} of 2.3 nm. The apo condition again gives a broad distance distribution, from which not a huge amount of information can be extracted. Addition of the Na⁺/L-aspartate pair results in a more highly populated state, resulting in a narrow distance distribution centred at approximately 2.3 nm, as was the case with GltPh_{SL}; similar conformational states appear to be adopted after the addition of Na⁺ with DL-TBOA as well. The 2.3 nm distance distribution observed agrees well with MMM predictions for all-IF protomers of trimeric GltPh.

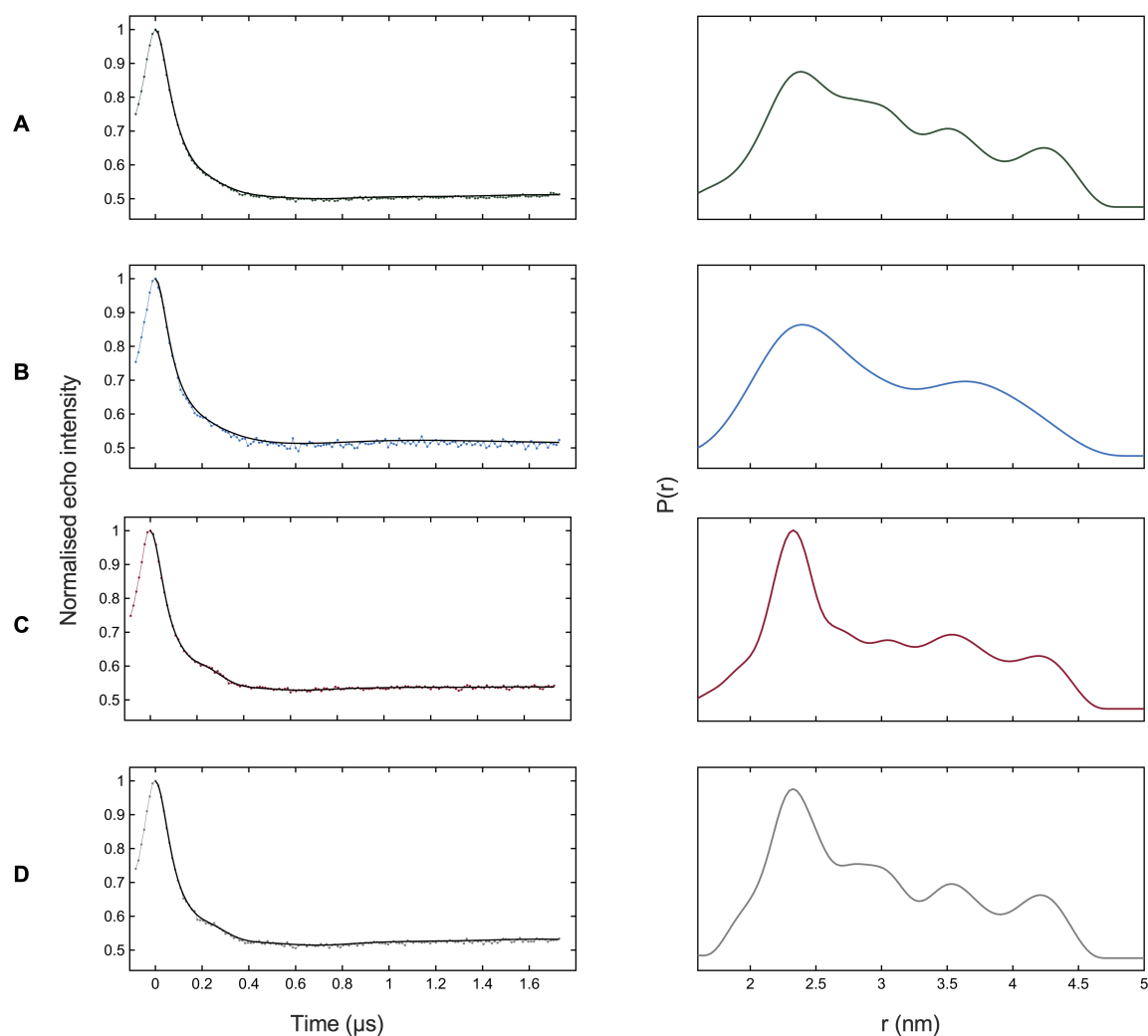


Figure 4.19: 4-pulse DEER experiments. (Left panel) Normalised background-corrected dipolar evolution for GlTPh_{SLIF} under (A) apo, (B) Na⁺-only-, (C) Na⁺/L-aspartate- and (D) Na⁺/DL-TBOA-bound conditions. Fittings are given as black lines overlaid with these plots. (Right panel) Resulting distance distributions for (A), (B), (C) and (D) as described following fitting of the time domain traces by Tikhonov regularisation. Experimental parameters are given in chapter 5.

4.7 Discussion and conclusions (EPR)

Starting with the basics, SDSL followed by sample preparation and measurement of GltPh variants under various conditions (including apo, Na^+ , Na^+ + L-aspartate and Na^+ + DL-TBOA) gives no evidence for any structural rearrangements or instabilities as a result of any mutation. To the eye, a slight increase in a more mobile SL species is observed in the RT X-band cw-EPR spectra for GltPh_{SL} and GltPh_{SL_{IF}} under certain conditions. However this effect is minimal, and is most likely attributable to the presence of a small amount of unbound, or non-specifically bound, MTSSL label. As such, these experiments suggest no local changes in spin label environment, nor changes in rotational motion of the protein itself at the SL position or indeed the attached probe, as a function of either the protein variant or protein state.

Following initial cw-EPR investigations, DEER spectroscopy was undertaken for all GltPh variants to test the use of conformation 'locking' variants in interpretation of complex and convoluted data. A key message to note from these studies is that our EPR investigations (specifically the DEER data) confirm that the introduction of the known and previously-used conformational 'lock' variant - K290E - does indeed stabilise the IF conformation in detergent-solubilised GltPh trimers. Meanwhile, although the G79A mutation has not been previously utilised for the purpose of an OF 'lock' mutation, DEER results clearly demonstrate the inhibition of IF conformations with the introduction of this mutation. Overall, the use of these two additional variants has afforded insight into the origin of inter-spin distances within GltPh_{SL} under specific conditions, from which we can conclude that an altogether very dynamic apoprotein becomes less-so upon binding of coupled ions/substrate/inhibitor. More specifically, GltPh tends to favour IF conformations in the presence of either Na^+ alone, Na^+ + L-aspartate, and Na^+ + DL-TBOA.

A direct comparison of these results with other published EPR studies is difficult, since the conclusions from most are quite conflicted. However one study (the most recent in which there are EPR investigations of GltPh) has tested the preparation of heterotrimers (in detergent micelles) as a means for more effective, refined, intra-protomer distance determination in DEER experiments.²³⁸ Interestingly, the results quite clearly show that the protein (at least one protomer) is preferentially adopting an IF conformation in both states for which experiments were measured - namely Na^+ only, and Na^+ + L-aspartate bound. The results of this particular study are extremely clear as a result of the heterotrimer produced, in which only one protomer contains cysteines for SDSL; they are also in good

agreement with those discussed in this chapter for GltPh_{SL} (also in detergent micelles). This IF conformation is in fact typically the opposite of what is observed throughout many X-ray crystal/cryo-EM structures of GltPh, at least for protein which hasn't been cross-linked in such a way to encourage the IF conformation. However, this does not mean to say that the observation (by DEER) of preferentially IF conformations for GltPh in solution should be surprising. Many recent cryo-EM structures of cross-linked IF GltPh give clear evidence that binding of Na⁺/L-aspartate/DL-TBOA is possible in protomers adopting an IF conformation.²¹⁷ Indeed MMM analysis of these IF crystal structures in bound/unbound states predict inter-spin distances between MTSSL-labelled A71C sites of approximately 2.3 - 2.4 nm. Many earlier structures were solved using X-ray crystallography, in which protein conditions were typically far from native. The crystal state represents a global minimum of free energy, however it might be the case that biologically relevant interactions are sacrificed in favour of unspecific crystal contacts. A crystal structure might therefore reflect a conformation that is most energetically favourable under the conditions used, but not necessarily one that is most abundant, biologically relevant or stable under native conditions, or in solution. Additionally, the solved crystal might only represent a very small proportion of the conformations actually present. This actually highlights the importance of biophysical, as well as structural, approaches for elucidation of protein structure and function.

As was highlighted during the introductory parts of this chapter, the majority of published GltPh EPR studies have included somewhat confusing, and often contradictory, conclusions. This is mainly due to complexities in inter-spin distance determination within a homotrimeric protein which is known to undergo substantial conformational movement during the transport cycle. To probe intra-protomer distances using current approaches has proven extremely difficult, which ultimately limits the application of DEER to inter-protomer distance determination; this is not especially useful when wanting to probe specific movements of helices/hairpins during transport. Thus this work provides a potential starting point for the strategic use of conformation-locking (but also structurally sound) variants for interpretation of complex results, by means of what is essentially the process of elimination. Such an approach could be useful not only for GltPh, but other multimeric proteins which undergo a vast amount of conformational change as a result of their biological function. Homologous mutations to the closely related human EAAT proteins, for example, might also be beneficial for their study via biophysical methods to complement recently solved 3D structures.^{241–244}

4.8 STD NMR for probing binding events in GltPh

The work presented in this section significantly extends the application of STD NMR to membrane proteins, for which results in literature are currently limited despite membrane proteins constituting more than half of all drug targets. This, most likely is due to the added complexity of the protein environment required for stabilisation. For example, with a membrane protein stabilised in detergent micelles, detergent molecules are in constant on-off interaction with the protein, resulting in potential additional resonances in the STD spectrum which might overlap with those of the ligand of interest. For this reason it is necessary to carefully assess the environment in which the protein is stabilised for the STD experiment. This work demonstrates the use of STD NMR as an alternative method for K_D determination not only for weakly binding ligands, but also for very tightly binding ligands, highlighting key differences which might relate to protein function in differing environments. Previous examples of STD NMR for reporting simply the observation of ligand binding to membrane proteins have been conducted in whole cells,^{245–248} cell membranes,²⁴⁹ and in proteins reconstituted into proteoliposomes²⁵⁰ or nanodiscs.²⁵¹ Only recently have examples of STD-based analysis of ligand binding to detergent solubilised proteins been reported.^{252–255} Meanwhile, estimation of ligand K_D to a membrane protein via STD NMR has been reported just once, in which an as-yet unspecified and unpublished fitting routine was employed.²⁵⁰

4.8.1 K_D determination by STD NMR and the effects of protein environment

For successful measurement of an STD NMR effect, particularly when using the method for quantitative determination of dissociation constants, several factors must be considered prior to running the experiment. As previously mentioned, the STD NMR technique is sensitive for weak-binding ligands (K_D values typically ranging from high nM to low mM), meaning that there is a fast exchange between the free and bound ligand states. This allows a build-up in concentration of saturated ligands in solution, giving greater STD signals. This means that the ligand k_{off} rate should be fast enough for the number of saturated molecules in solution to build-up, but not so fast that it binds for insufficient time to allow saturation transfer from the protein to the ligand to actually happen. When the ligand is bound to the protein, its relaxation behaviour takes on that of the macromolecule, meaning the rate at which its perturbed magnetisation relaxes back to equilibrium is now much faster than in the free (unbound) state. Thus it is necessary for the ligand to dissociate at a faster rate than the magnetisation relaxation rate, to avoid loss of magnetisation. For these reasons, STD

signals cannot be measured for tight-binding ligands. Additionally, build-up of saturation of ligand protons as a result of re-binding is possible, leading to artificially high STD intensities and thus higher estimations for K_D 's. Such inaccuracies can be accounted for by utilising an initial growth rates protocol,²⁵⁶ however this is time consuming and not appropriate for potentially fast-degrading membrane proteins. Hence a compromise must be reached between experimental sensitivity and accuracy of K_D determination when deciding on the length of the saturating pulse.

T_1 determination for L-glutamate

Saturation build-up is limited by relaxation ($R_1 = 1/T_1$) of the saturated ligand protons in the free state. This means that protons with a longer T_1 accumulate more saturation when bound to the receptor (through the NOE), and that magnetisation decays back to thermal equilibrium more slowly in solution, and hence the relative STD intensity is higher. When working with membrane proteins which have been extracted and purified from the cell, an environment is required which mimics the types of interactions that stabilise the protein in the native membrane lipid bilayer; this might mean protein is solubilised in detergent micelles, reconstituted into proteoliposomes or nanodiscs. The work presented here utilises stabilisation of WT GltPh in both of the former two environments. However, since proton relaxation rates are sensitive to the rate of molecular tumbling in solution, the presence of detergent micelles or proteoliposomes in solution might alter the longitudinal relaxation time, T_1 , of the ligand protons in the free state. In order to account for any differences in T_1 between detergent solubilised and proteoliposome-reconstituted states when setting up the STD experiment (specifically in choosing the saturation and delay times), proton longitudinal relaxation times were determined by means of the inversion-recovery experiment²⁵⁷. The principle being this experiment, including the pulse sequence and delay times, is much the same as that described for EPR measurements in chapter 3.8.1. Proton T_1 times were determined for L-glutamate under both conditions - specifically (a) in the appropriate buffer supplemented with DDM (0.075 %) and (b) in the appropriate buffer containing empty (not reconstituted) liposomes which had been subjected to extrusion through a 400 nm pore size filter.

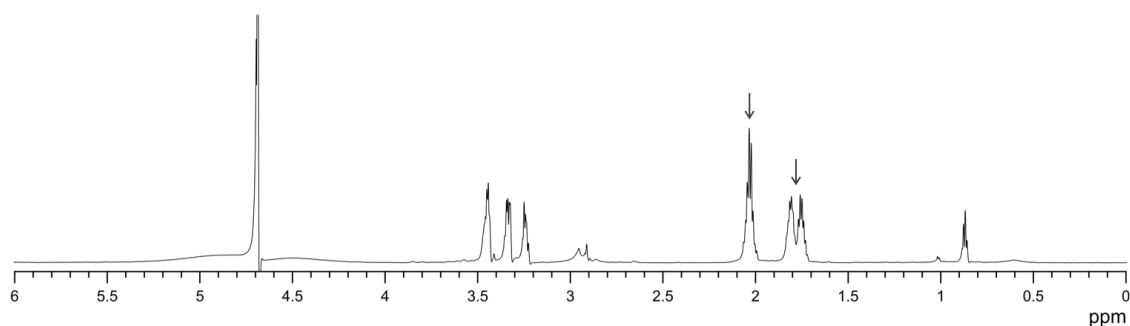


Figure 4.20: Full ^1H NMR spectra of L-glutamate in buffer with empty proteoliposomes. L-glutamate resonances are highlighted, and centred at 2.05 ppm and 1.8 ppm.

Figure 4.20 gives the 1D ^1H NMR spectra for L-glutamate (1.5 mM) in buffer with empty liposomes, in which the resonances corresponding to γ (2.05 ppm) and β (1.81 and 1.75 ppm) protons of L-glutamate are highlighted. It is the intensity of these peaks as a function of the time, τ , between the initial inversion (180°) pulse and the subsequent spin echo forming pulses, that is monitored for determination of proton T_1 . Equation 4.1 gives the fitting function for T_1 determination,

$$I = I_{max}(1 - 2e^{-\tau/T_1}) \quad (4.1)$$

in which I_{max} is the maximum signal intensity, and I is the signal intensity at a given delay time, τ . For T_1 analysis, signal intensities for each relevant L-glutamate proton peak are plotted as a function of τ , in both detergent and liposome environments (Figure 4.21). Curves were fitted according to equation 4.1, the results of which are plotted as black lines. Table 4.2 gives the T_1 values determined in each case as a result of this fitting for L-glutamate protons; results show that T_1 values remain largely the same despite the differences in surrounding environment. Thus it is suitable to use equivalent saturation and delay times between all STD NMR experiments, regardless of sample type.

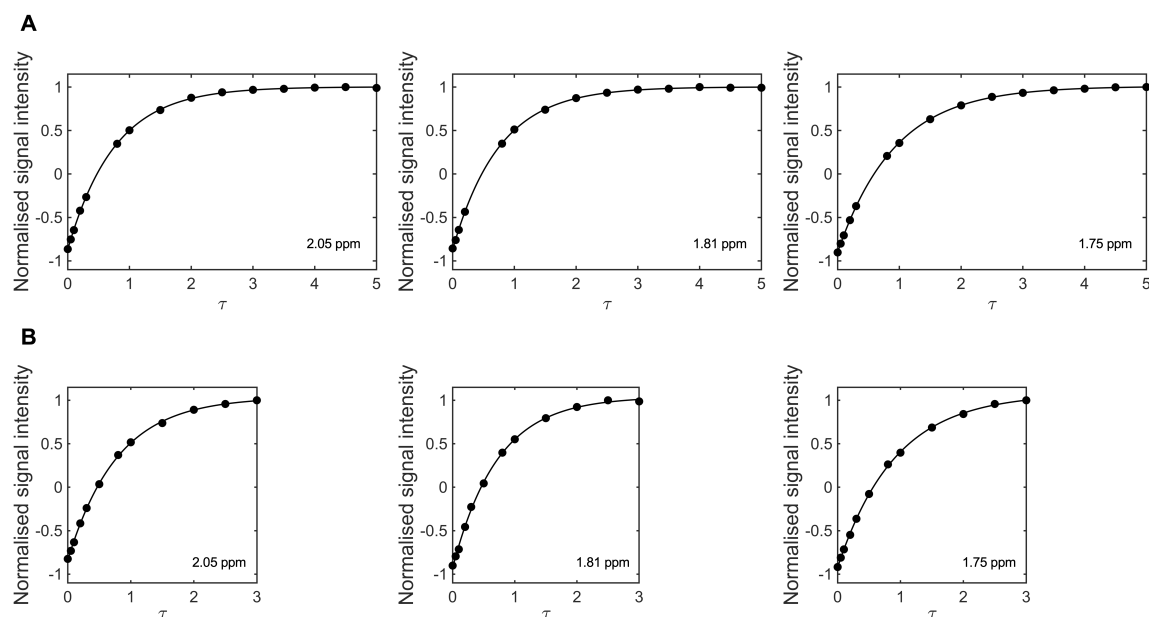


Figure 4.21: Plot of normalised ^1H NMR signal intensity as a function of delay time τ (dots) and the associated fittings according to the fitting function equation 4.1 (black lines), for (A) 1.5 mM L-glutamate in D_2O based buffer (20 mM potassium phosphate pH 7.5, 100 mM NaCl) supplemented with DDM (0.075 %), and (B) 1.5 mM L-glutamate added to empty proteoliposomes prepared in buffer (as above). Individual plots are labelled according to the resonance position analysed in ppm.

	T_1 (ms)		
	γ proton (2.05 ppm)	β proton (1.81 ppm)	β proton (1.75 ppm)
1.5 mM L-glu + DDM	923.4	755.0	758.8
1.5 mM L-glu + empty liposomes	914.2	718.3	786.3

Table 4.2: Table of the determined T_1 values for each observable L-glutamate proton resonance, calculated according to equation 4.1.

Initial observation of the STD effect and control experiments

Before attempting quantitative K_D determination for L-glutamate with WT GltPh by means of titration, it was first necessary to (a) ensure an STD signal was indeed measurable for L-glutamate with the protein, which would require (b) the measurement of specific control experiments to confirm the correct origin of saturation transfer to the ligand. Since it is well known that binding and transport of L-glutamate requires as a pre-requisite the binding of Na^+ ions, initial attempts were conducted for WT GltPh in the presence of a large excess of Na^+ ions. As mentioned, the STD spectrum is simply the difference spectrum between on-resonance (saturated) and off-resonance (reference) spectra. Any peaks in the STD spectrum indicate a transfer of saturation from the directly saturated protein protons, to the protons associated with those resonances; this of course requires direct interaction of molecules with the macromolecule for a sufficient period of time to receive the saturation.

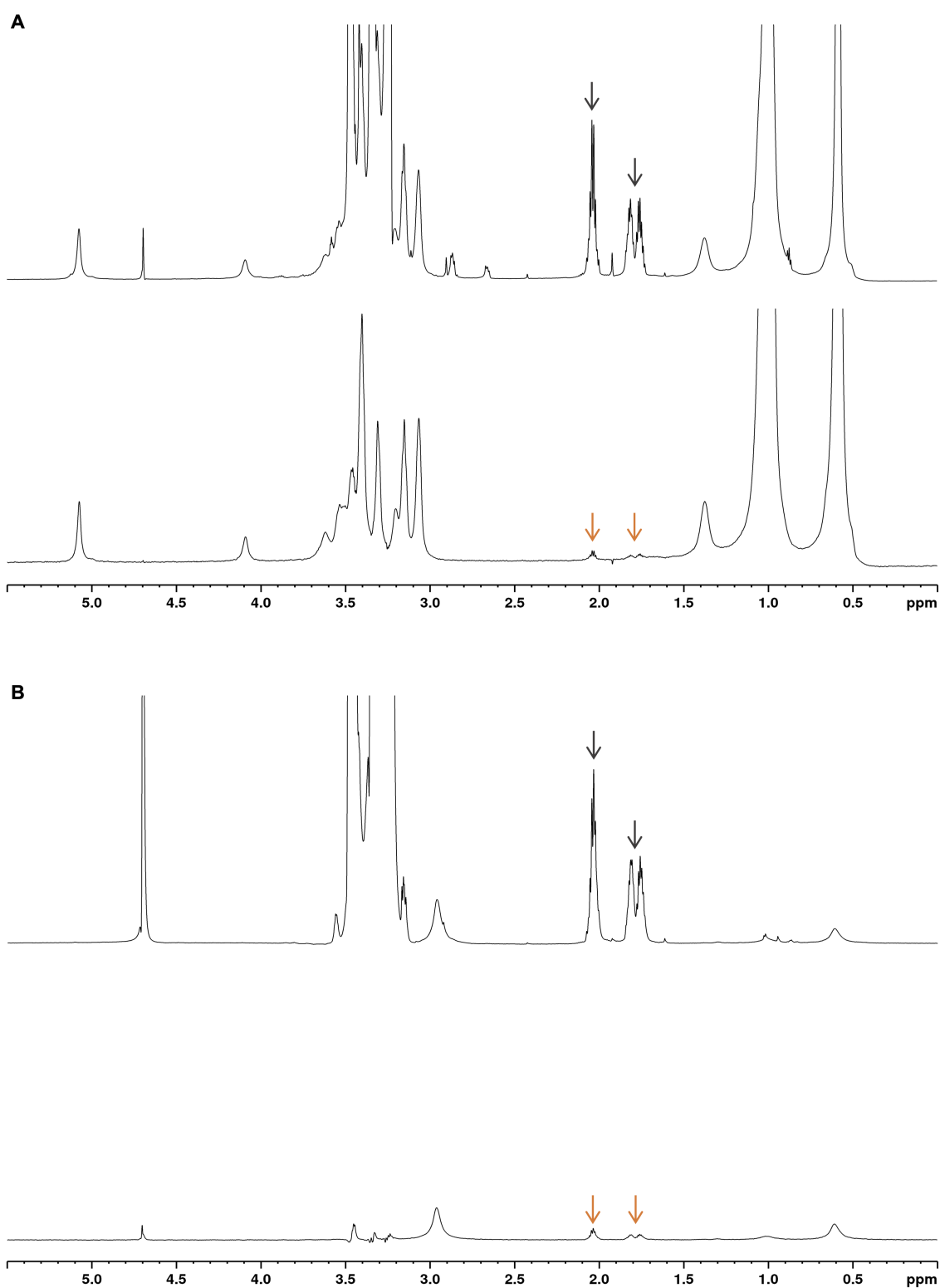


Figure 4.22: Full ^1H NMR reference spectra (top) and corresponding STD spectra (bottom) of (A) detergent solubilised WT GltPh ($19.4\ \mu\text{M}$), and (B) WT GltPh reconstituted into liposomes, both in the presence of NaCl ($100\ \text{mM}$) and following incubation with L-glutamate ($1.5\ \text{mM}$). NMR spectra were recorded at $800\ \text{MHz}$ and $278\ \text{K}$, with a saturation time of $2\ \text{s}$. Signals from L-glutamate centred at $2.05\ \text{ppm}$ and $1.8\ \text{ppm}$ have been highlighted in both reference and STD spectra, using grey and orange arrows respectively. Stacked spectra have been normalised according to signal:noise ratios.

Results of the STD experiment for WT GltPh, both stabilised in detergent micelles and reconstituted into proteoliposomes, following incubation with excess L-glutamate in the presence of Na^+ ions, are shown in Figures 4.22 A and B respectively. The full reference ^1H NMR spectrum measured in each case is given (top spectra in A and B); the spectrum of detergent-solubilised WT GltPh with L-glutamate is dominated by strong signals from detergent protons. However L-glutamate resonances for both γ (2.05 ppm) and β (centred at approximately 1.8 ppm) protons are in a region which is devoid of interferences from these strong detergent signals.

The presence, in both cases, of an L-glutamate STD effect in the presence Na^+ ions is a strong indication of ligand-protein interaction, however control experiments are required to ensure that the saturation transferred to the L-glutamate protons does indeed come from protein. This is especially important for the detergent solubilised case, in which the STD spectrum gives additional, strong peaks corresponding to detergent molecules as a result of two mechanisms: the constant on-off exchange within the detergent/protein complex in solutions; and direct saturation caused by the on-resonance irradiation pulse. The latter of these mechanisms means that interaction of L-glutamate with directly saturated detergent molecules might also occur, giving rise to an STD effect. In order to test this, STD NMR experiments were conducted under the same experimental conditions for WT GltPh (detergent solubilised and reconstituted into proteoliposomes) in the presence of L-glutamate, but without NaCl ions (which was substituted for KCl). Additionally, an STD NMR experiment was conducted on a sample containing empty liposomes, incubated with L-glutamate in the presence of NaCl . The results of these control experiments are shown in Figure 4.23, in which reference spectra and STD spectra are stacked on top of one another for each sample. In all cases, no STD effect is observed for L-glutamate; this acts as confirmation that the ligand only receives its saturation when bound to the protein. This also gives direct evidence for the Na^+ dependence of L-glutamate binding to GltPh.

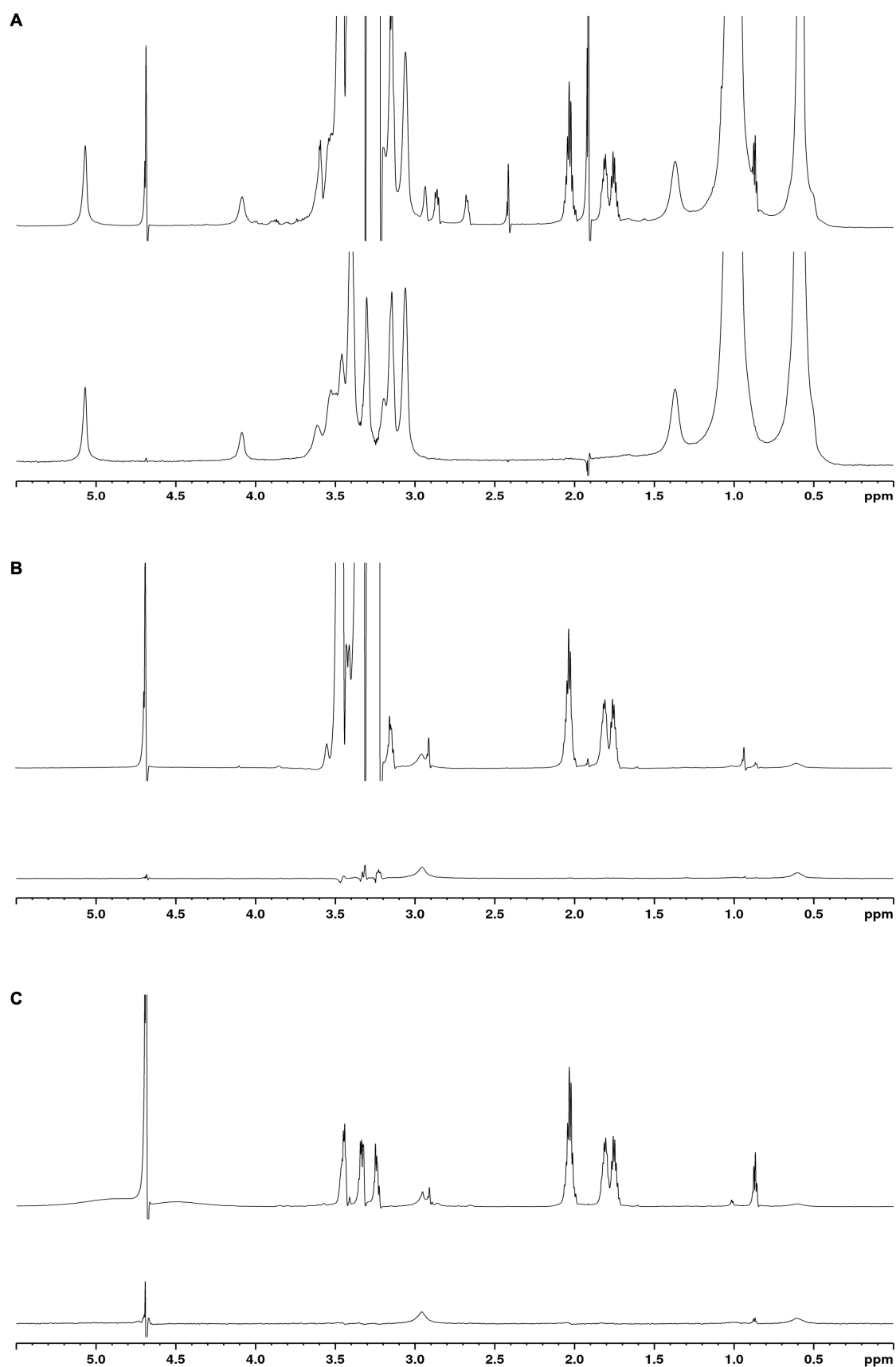


Figure 4.23: Full ^1H NMR reference spectra (top) and corresponding STD spectra (bottom) of (A) detergent solubilised WT GltPh (22.7 μM) with KCl (100 mM), (B) WT GltPh reconstituted into liposomes with KCl (100 mM), and (C) empty proteoliposomes with NaCl (100 mM), following incubation with L-glutamate (1.5 mM). NMR spectra were recorded at 800 MHz and 278 K, with a saturation time of 2 s. Stacked spectra have been normalised according to signal:noise ratios.

K_D determination for L-glutamate with WT GltPh

The method for K_D determination via STD NMR involves analysis of the proton STD signal intensity of interacting ligand as a function of its concentration. This is accomplished via titration of ligand into the protein sample, over a range of ligand concentrations which span either side of the ligand K_D , eventually reaching an end point of high ligand excess. For these studies, K_D determination was conducted for WT GltPh which was (a) solubilised in detergent micelles, and (b) reconstituted into proteoliposomes. Since initial tests with excess L-glutamate gave clear STD signals in both environments, titrations were carried out using the same NaCl concentration of 100 mM. Concentrations and titres required for the incremental additions of L-glutamate were decided based upon previously published K_D values, ranging from 120-250 μ M. STD spectra were recorded for each subsequent addition of L-glutamate, and the resulting experimentally observed L-glutamate STD signals were analysed. This is achieved via conversion of the measured STD effect to an STD amplification factor (STD-AF) using the following equation,

$$STD - AF = \frac{I_0 - I_{SAT}}{I_0} \times \frac{[L]_{Total}}{[P]_{Total}} \quad (4.2)$$

in which $(I_0 - I_{SAT})$ represents the ligand signal intensity in the STD spectrum, I_0 is the ligand signal intensity in the off-resonance (reference) 1H NMR spectrum, and $([L]_{Total}/[P]_{Total})$ is the ligand excess relative to the fixed protein concentration. Thus the amplification factor for a determined saturation time can be described as the average number of ligand molecules saturated per molecule of receptor. This was determined for each titre at the γ proton resonance position and plotted against the ligand concentration, resulting in a hyperbolic curve for both detergent solubilised and proteoliposome-reconstituted protein samples. The initial steep slope of the curve gradually decreases to an eventual plateau as a result of saturation of GltPh binding sites with ligand. For K_D determination from this data, a fitting function analogous to the Michaelis-Menten equation is used (Equation 4.3). Here, α_{STD} corresponds to the maximum STD-AF and $[L]$ is the total concentration of free ligand.

$$STD - AF = \frac{\alpha_{STD}[L]}{K_D + [L]} \quad (4.3)$$

Such treatment of the data is possible when $[L]$ tends to $[L]_{Total}$, which is true of the experiments undertaken here since $[L]_{Total} \gg [P]_{Total}$, and thus the STD-AF is expected to increase with increasing $[L]_{Total}$ until the GltPh binding sites are fully occupied. This occurs at a ligand concentration which is much greater than the ligand K_D . Fittings resulted

in determined K_D values for L-glutamate with WT GltPh of 414 μM and 1034 μM , for detergent solubilised and proteoliposome-reconstituted samples respectively. Results were very similar (yielding K_D s of 458 μM and 984 μM respectively) when the L-glutamate β proton resonance intensities were treated and analysed in the exact same fashion; since both appear to receive a similar amount of saturation from the protein, for the sake of this work γ proton resonances have been analysed due to their larger relative intensities (i.e. the signals are larger and thus more easy to analyse).

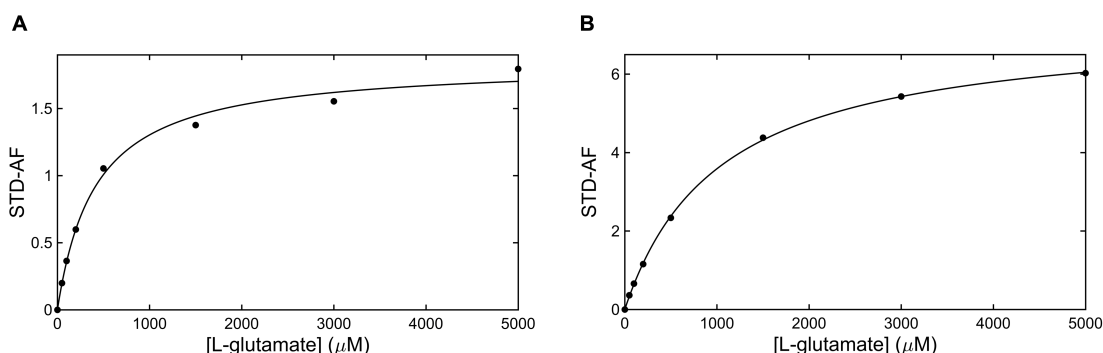


Figure 4.24: Plots of the STD-AF against total L-glutamate concentration for GltPh (A) solubilised in detergent micelles and (B) reconstituted into proteoliposomes (dots). Results of data fitting according to equation 4.3 are shown as black solid lines.

Literature values for L-glutamate K_D values with GltPh are typically of the same order of magnitude, but always lower than those determined here. However the concentration of Na^+ ions used here is lower than is used in most previous cases.^{211,222} Considering that binding and transport of ligands is dependent on binding of three Na^+ ions, these differences might be explained by the concentration of Na^+ ions present in solution. To test this hypothesis, a further experiment was carried out with WT GltPh solubilised in detergent micelles with a constant concentration of L-glutamate in which the intensity of the L-glutamate STD signal was monitored as a function of Na^+ concentration. This titration was performed from sub-stoichiometric concentrations, to a very large excess of NaCl as compared to the protein concentration. A plot of STD effect (%) of L-glutamate as a function of $[\text{Na}^+]$ is shown in Figure 4.25, in which it can be seen that the STD signal intensity does indeed vary considerably with NaCl concentration. It should be noted that concentrations of NaCl above 100 mM are not easily accessible by the NMR experiment due to difficulties in tuning and matching during experimental setup.

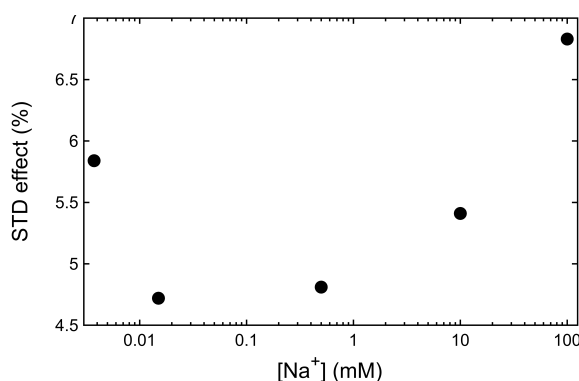


Figure 4.25: Plot of the L-glutamate STD effect (%) as a function of the concentration of Na⁺ for detergent solubilised WT GltPh (14 μ M) incubated with 1.5 mM L-glutamate.

K_D determination via competitive displacement

When considering a high affinity ligand (i.e. $K_D < \mu$ M), measurement and observation of an STD effect is not possible. The result of tight binding in the binding site is a much slower k_{off} rate; this means that saturation cannot be effectively transferred to solution for measurement. L-glutamate is a relatively weak-binding substrate of GltPh, and thus direct K_D determination via analysis of proton STD signals was possible. The more biologically relevant ligand of GltPh - L-aspartate - binds with much higher affinity, with published K_D 's typically in the nM range. As such, observation of an STD effect for L-aspartate would not be expected, meaning a direct approach for K_D determination would not be possible. To overcome this limitation, one can employ the use of a 'reporter' ligand, in combination with competition binding experiments.²⁵⁸ Thus for further demonstration of the STD NMR technique for analysis of ligand binding to membrane proteins, this type of approach was used for L-aspartate K_D determination.

The competition-based experiment involves the use of high and low affinity substrates which bind to the same site within the protein. The lower affinity ligand acts as a 'reporter', since its STD NMR signal is measurable (in this case, the reporter ligand is L-glutamate). This ligand is added to the protein sample and kept at a fixed concentration which affords a clearly measurable STD effect. The intensities of its STD NMR signals are then monitored following incremental titration of the high affinity ligand (in this case, L-aspartate). As competitive displacement of the reporter ligand (by the higher affinity ligand) occurs, the STD response of the reporter ligand should decrease. This method has been employed here to afford estimation of the upper limit for the K_D values of L-aspartate for GltPh, again in both detergent solubilised and proteoliposome-reconstituted environments. Experiments were conducted using an excess of L-glutamate (1.5 mM) for a clear STD NMR signal.

Data processing is slightly different in this case: the STD effect is determined as usual by comparison of the reference and STD spectra, however rather than calculation of an amplification factor, these values are simply normalised such that the STD effect in the absence of any competitor ligand is set to 100 %, and the rest calculated relative to this. These values are then plotted as a function of L-aspartate concentration, shown in Figure 4.26 (dots). The resulting hyperbolic curve can be fitted according to equation 4.4,^{259,260}

$$\%bound \propto -A \times \frac{[\frac{K_D}{L_0}][\frac{I_0}{K_I}]}{1 + [\frac{K_D}{L_0}][1 + (\frac{I_0}{K_I})]} + 100 \quad (4.4)$$

in which A is a proportionality constant, K_D is the dissociation constant of the reporter ligand (L-glutamate) and L_0 is its fixed concentration, I_0 is the concentration of L-aspartate and K_I is its dissociation constant. The resulting fits for experiments with WT GltPh in detergent solubilised and proteoliposome-reconstituted environments are also given in Figure 4.26 (solid black lines). These fittings afforded L-aspartate K_D values of 42 nM and 13 nM for detergent solubilised and proteoliposome-reconstituted GltPh, respectively.

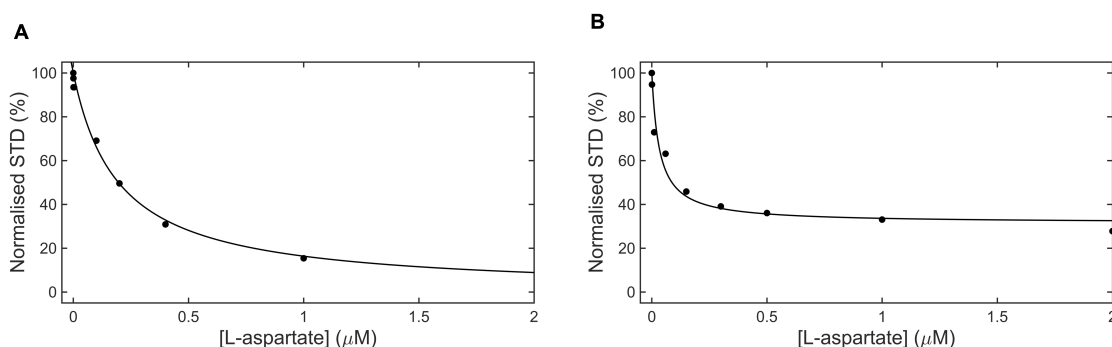


Figure 4.26: Plots of the normalised STD effect (or % bound) against total L-aspartate concentration for GltPh (A) solubilised in detergent micelles and (B) reconstituted into proteoliposomes (dots). Results of data fitting according to equation 4.4 are shown as black solid lines.

Interestingly, the extent of displacement of L-glutamate by L-aspartate appears to be different between detergent-solubilised and proteoliposome-reconstituted GltPh. Specifically, L-aspartate is only able to displace approximately two-thirds of the total bound L-glutamate when the protein is present within proteoliposomes. For further analysis of this behaviour, the data was plotted in a Hill plot, affording determination of Hill coefficients using equation 4.5,

$$y = \frac{1}{(\frac{K_D}{[L]})^n + 1} \quad (4.5)$$

where y is the normalised STD effect (or % bound) $^{-1}$, K_D is the determined dissociation constant for L-aspartate in each environment, and n is the Hill coefficient. Plots showing the resulting fittings are given in Figure 4.27, and determined Hill coefficients for detergent-solubilised and proteoliposome-reconstituted GltPh are equal to 1.06 (~ 1) and 0.68 ($\sim 2/3$), respectively.

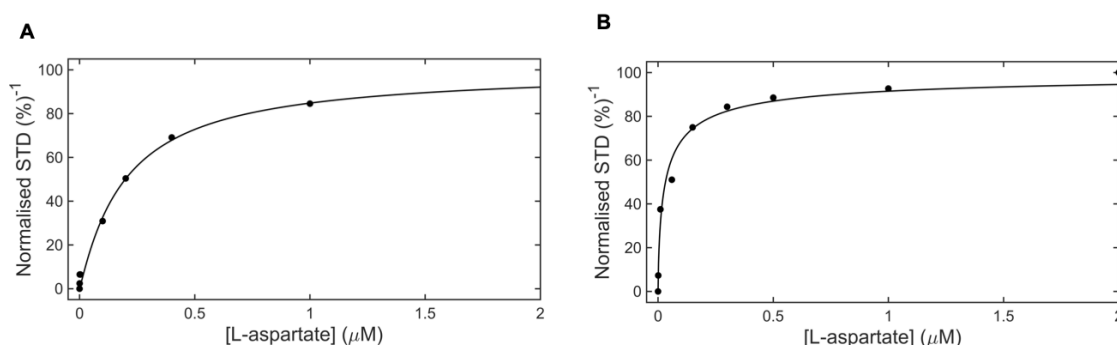


Figure 4.27: Plots of the normalised STD effect (or % bound) $^{-1}$, against total L-aspartate concentration for GltPh (A) solubilised in detergent micelles and (B) reconstituted into proteoliposomes (dots). Results of data fitting according to equation 4.5 are shown as black solid lines.

4.9 Discussion and conclusions (STD NMR)

Initial STD NMR experiments conducted for WT GltPh in (a) detergent micelles and (b) proteoliposomes demonstrated much clearer STD signals for the L-glutamate ligand in the latter case. For a detergent-solubilised sample, the detergent molecules will result in their own STD signals in addition to the ligand of interest, since they themselves are in constant on-off exchange with the saturated protein molecule. Nonetheless, STD signals were observed in both environments and were sufficient for further titration experiments, affording K_D determination. Control experiments confirmed that saturation transfer to the ligand was reliant on the protein being present, as oppose to saturation being transferred from other molecules present within the solution which might be receiving saturation directly from the RF pulse during the experiment. Additionally, no STD effect was observed for L-glutamate in the absence of (coupled) Na^+ ions, as expected.

K_D determination was successfully carried out by monitoring the L-glutamate STD NMR response as a function of total ligand concentration, via ligand titration. Values obtained for the K_D associated with L-glutamate binding to WT GltPh via this approach (414 and 1034 μM , for detergent-solubilised and proteoliposome-reconstituted protein respectively) are higher than typical published values, which range between 120 - 250 μM in detergent. This is likely (at least in part) due to the much lower Na^+ ion concentration used within the NMR

experiments (100 mM) as compared with that used in other studies (typically 200 mM). Binding of one substrate molecule by GltPh is well known to be dependent on the binding of at least two Na^+ ions (per protomer), whilst binding of three Na^+ ions is necessary for the transport event to take place. This principle was tested simply by the titration of NaCl into a sample of constant protein and L-glutamate concentrations, from which it is demonstrated that the concentration of Na^+ ions within the sample has a clear influence on the STD NMR response. It is however also well known that the STD NMR approach for K_D determination typically results in over-estimations, due to the build-up of the STD response from multiple binding events.

Since GltPh is actually a transporter, there are additional factors that need to be considered as part of the analysis process. The ligand is able to bind from either side of the transporter, and the STD effect is sensitive to the off rate. It is possible, therefore that the experiment might be reporting on transport events (i.e. a K_M). However, the isomerisation event which results in substrate and ion transport is an order of magnitude slower than the L-glutamate on/off rates; thus these results do indeed report an apparent K_D as oppose to rates associated with transport. This point is especially relevant to the results obtained in the context of proteoliposomes, in which (at least initially) there is a concentration gradient of NaCl and L-glutamate between the outside and inside of the proteoliposomes.

As previously mentioned, STD NMR is of limited use for the direct determination of K_D 's for 'high' affinity ligands - more specifically those with K_D s $< \mu\text{M}$. However, the competition experiment involving displacement of the reporter ligand, L-glutamate, by the higher affinity ligand L-aspartate has afforded successful (indirect) K_D determination for L-aspartate. The first point to note is that the reduction in STD response of L-glutamate as a function of increasing L-aspartate concentration gives direct evidence that both ligands bind within the same binding pocket, as is expected. K_D values determined for L-aspartate (equal to 42 and 13 nM for detergent-solubilised and proteoliposome-reconstituted GltPh, respectively) are comparable with literature values, which are also in the nM range. There appears to be little difference in observed K_D values for protein solubilised in the two environments, however the calculated Hill coefficients reveal insights into, and distinct differences in, cooperativity of ligand binding between these two environments. The Hill coefficient calculated for L-aspartate binding to GltPh in detergent is equal to 1.063 (~ 1) and indicative of typical non-cooperative, independent ligand binding by protomers of (trimeric) GltPh. However that calculated for proteoliposome-reconstituted protein, which is equal to 0.69, indicates negative cooperativity in which binding of substrate to one binding site makes binding at

another less likely. Thus this study has highlighted some potential interplay between GltPh protomers, which is unique and dependent upon the protein being stabilised within the proteoliposome environment. As has already been explained, ligand binding to protein within proteoliposomes might be associated with an actual transport event; this is much more representative of physiological conditions.

Altogether this body of work has demonstrated (for the first time) the application of STD NMR for determination of ligand binding constants, across at least four orders of magnitude, for membrane proteins. This is achievable in multiple protein environments, thus making the technique more universally applicable and lending further information on biologically relevant processes. There is the potential for further application of competition-based STD NMR techniques, including lead discovery involving the screening of large compound libraries against a wide range of drug targets in order to identify both high- and low-affinity ligands. This could afford optimisation of 'hit' compounds into drug leads, through derivation of structure-activity relationships. As such, this STD NMR approach could form an additional screening tool for structure-based and fragment-based drug design in key membrane systems. The STD NMR experiment might also be applicable to spin labelled systems, whereby knowledge of labelling sites and the phenomenon of paramagnetic relaxation enhancement could afford triangulation of unknown substrate binding sites, for example.

Chapter 5

Materials and methods

5.1 MamM-CTD

5.1.1 Biological material

Purified MamM-CTD protein, including the WT, C267S, C267S-D249A-H285A and C267S-H264A-E289A variants, was provided by Shiran Barber-Zucker and Raz Zarivach of the Ben-Gurion University of the Negev, Israel. All information regarding preparation of constructs, protein expression and purification can be found in our recent publication.¹¹⁵

5.1.2 Site-directed spin labelling

IAP

For site-directed spin labelling using IAP, protein solutions in buffer A (containing 10 mM Tris-HCl pH=8 and 150 mM NaCl) were labelled overnight at 4 °C with iodoacetamido proxyl (IAP, 5x excess per cysteine) in darkness and under gentle agitation. To remove excess unbound IAP, samples were washed using ZebaTM Spin desalting columns (7 kDa MWCO, 2 mL; Thermo Scientific, Rockford, IL, USA) and degassed buffer A. The final protein concentration post-labelling was determined using the optical absorbance at 280 nm, and an extinction coefficient for the dimer of 13993.3 M⁻¹.cm⁻¹.

MTSSL

For site-directed spin labelling using MTSSL, protein solutions in buffer A (containing 10 mM Tris-HCl pH=8 and 150 mM NaCl) were degassed under argon and treated with degassed 1,4- dithiothietol (1 mM) for 4 hours at 4 °C with gentle agitation. Subsequently, protein samples were washed once using ZebaTM Spin desalting columns and degassed buffer A. Protein was then labelled overnight at 4 °C with (1-oxyl-2,2,5,5-tetramethyl-D3-pyrroline-3-methyl) methane-thiosulfonate spin label (MTSSL, 10x excess per cysteine; Toronto Re-

search Chemicals Inc., North York, ON, Canada) in darkness and under gentle agitation. To remove excess unbound MTSSL, samples were washed using ZebaTM Spin desalting columns and degassed buffer A. The final protein concentration post-labelling was determined using the optical absorbance at 280 nm, and an extinction coefficient for the dimer of $13993.3 \text{ M}^{-1}\text{cm}^{-1}$.

5.1.3 Continuous wave X-band EPR spectroscopy

Sample preparation

For apo and metal-bound samples, freshly spin labelled MamM-CTD variants in buffer A (dimer concentrations ranging from 50-100 μM) were used. Stock solutions of metal salts (including $\text{ZnSO}_4 \cdot 5\text{H}_2\text{O}$, CdCl_2 , NiCl_2 , $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ and $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$) were prepared in H_2O to a final concentration of 100 mM. Protein aliquots of 15 μL were incubated with the desired excess of associated metal salts for 45 minutes on ice (details of excesses used are noted in the appropriate figure legends in chapter 3.8.1). For both apo and metal-bound samples, aliquots of $\sim 10 \mu\text{L}$ were transferred to capillary tubes (outer diameter = 0.8 mm) for measurement.

For metal-displacement experiments with C267S MamM-CTD, spin labelled C267S MamM-CTD at a (dimer) concentration of 160 μM was used. Metals were added for incubation and washed in sequence according to calculated K_D values. The same metal ion stocks as described for the regular metal-bound continuous wave X-band measurements above were used. The total sample (230 μM) was incubated with CdCl_2 (5x excess per dimer) for 45 minutes on ice. An aliquot of $\sim 10 \mu\text{L}$ was transferred to a capillary tube (outer diameter = 0.8 mm) for measurement. The remainder of the sample was then incubated with $\text{ZnSO}_4 \cdot 5\text{H}_2\text{O}$ (5x excess per dimer) for 45 minutes on ice. To remove any excess, unbound metal ions the sample was washed using a ZebaTM Spin desalting column. An aliquot of $\sim 10 \mu\text{L}$ was transferred to a capillary tube (outer diameter = 0.8 mm) for measurement. Protein concentration was determined again due to potential losses during the wash step. This process was repeated in sequence for $\text{ZnSO}_4 \cdot 5\text{H}_2\text{O}$, NiCl_2 , $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, and finally CdCl_2 once again.

Ambient temperature cw-EPR measurements

The ambient temperature (room temperature, RT) set-up for X-band cw-EPR consisted of a Bruker E500 eleXsys spectrometer fitted with an ER 4123D (dielectric RT cw-EPR) resonator. Specific measuring parameters for individual experiments are highlighted in the rele-

vant figure legends in chapter 3.8.1. Spin labelling efficiencies were calculated by comparing the double integral of the cw-EPR (first derivative) signal against that of known concentrations of 4-Maleimido-TEMPO (Sigma-Aldrich), forming a calibration curve.

Low temperature cw-EPR measurements

A liquid helium transfer setup was used in order to achieve low-temperatures for spectroscopy, with a turbomolecular vacuum pump (Oerlikon Leybold PT-50 Dry Divac) and a proportional-integral-derivative (PID) type temperature controller. The transfer setup involved a super-high Q (SHQ) resonator with the ER 4112HV low temperature (3.8-300 K) control system, in a Bruker E500 Eleksys spectrometer (Bruker, Rheinstetten, Germany). The ER 4112HV involves a dewar and quartz sample holder such that the 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 - the gas that passes back from the cryostat shields cold helium gas from the dewar. Helium flow is controlled manually using a needle valve on the transfer line.

EasySpin for analysis of cw-EPR data

All cw-EPR spectral simulated fits were carried out using EasySpin,¹⁵⁹ an open-source Matlab[®] toolbox. For solution spectra collected at RT, simulations were computed using the EasySpin simulation function *chili*, which is appropriate for slow/intermediate motion regimes. The fitting of simulations was computed using EasySpin's fitting function *esfit*. Figure 5.1 gives an example script used for this spectral fitting. The very bottom line of this script is the command to call the fitting function, containing multiple parameters which are explained below:

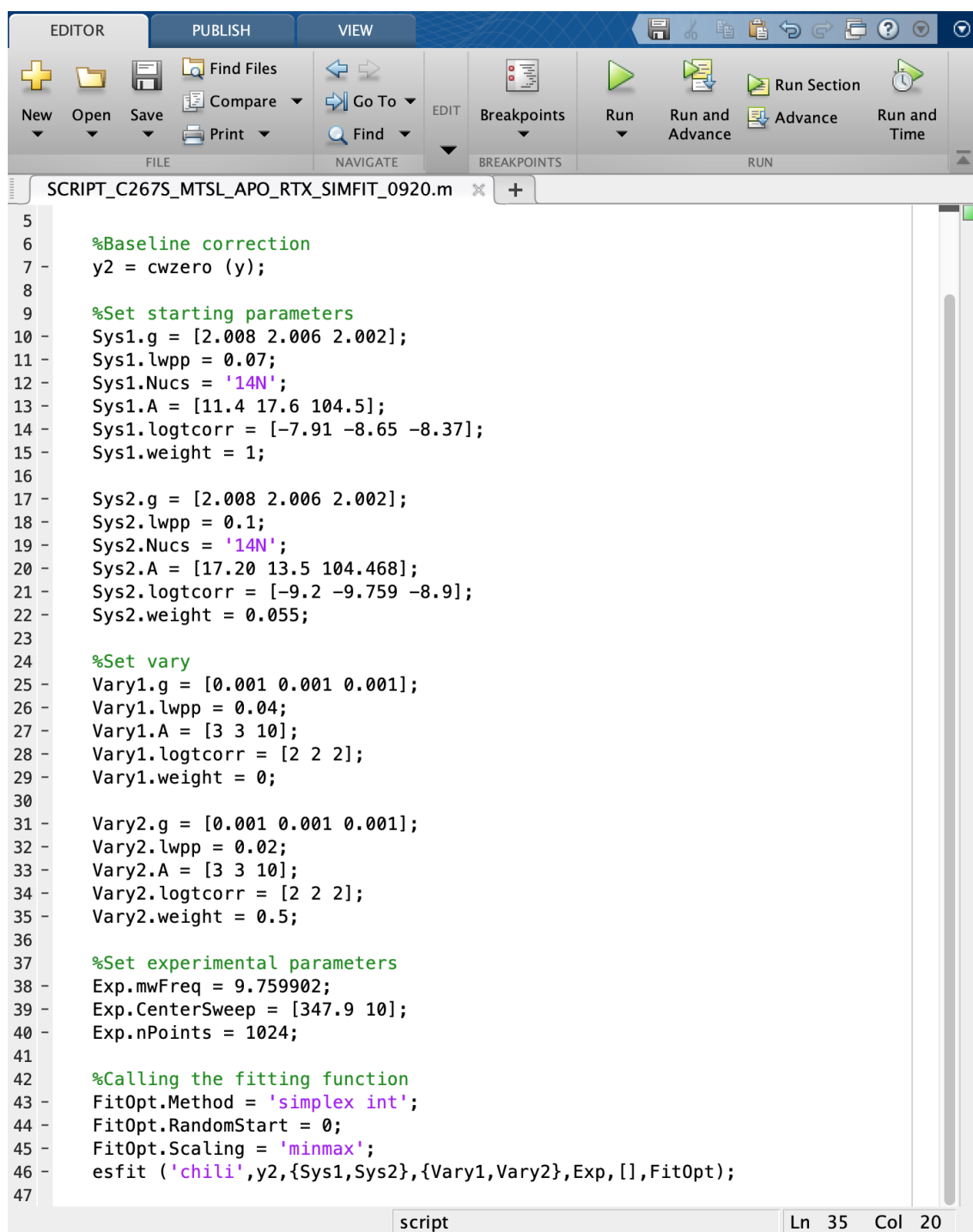
- '*chili*' as mentioned already, indicates that the simulation function *chili* should be used.
- *y2* specifies the experimental spectral y-axis data (i.e. intensities), to which the simulation will be fitted. This y-axis data has been zeroed using the command in the second line of the example script.
- *{Sys1,Sys2}* is a list of structures, each collecting the magnetic parameters of the two different spin systems defined (Sys1 and Sys2). The parameters given earlier on in the script (i.e. 'Sys1.g = [2.008 2.006 2.002]') are used as starting values for parameters in the fitting procedure.
- *{Vary1,Vary2}* is similar to *{Sys1,Sys2}*, but instead contains the search ranges of the

parameters that should be fitted. Here, '*Vary1*' defines the variance of *Sys1* parameters, whilst '*Vary2*' defines the variance of *Sys2* parameters.

- '*Exp*' is a structure containing the experimental parameters for the simulation function (i.e. the microwave frequency - '*Exp.mwFreq* = 9.759902').
- '*FitOpt*' contains settings for the least-squares fitting algorithm.

For each system, 6 parameters were defined as follows: (1) the three principle g values (g_{xx} , g_{yy} and g_{zz} , defined by '*Sys1/2.g*'); (2) the peak-to-peak line width (defined by '*Sys1/2.lwpp*'); (3) the magnetic nuclei present within the spin system (defined by '*Sys1/2.Nucs*'); (4) the three principle hyperfine couplings (A_{xx} , A_{yy} and A_{zz}) in MHz, for each of the magnetic nuclei (defined by '*Sys1/2.A*'); (5) the base-10 logarithm of rotational correlation time, again for each of the three principle axes (defined by '*Sys1/2.logtcrr*'); and (6) the weighting of each spin system (or component) relative to the other (given by '*Sys1/2.weight*'). Each of these parameters are varied during the fitting according to the variance set for each. As part of the fitting process, a number of optimization algorithms are possible, and defined under '*FitOpt*'. These include (1) the Nelder/Mead downhill simplex method, (2) the Levenberg/Marquardt algorithm, (3) Monte Carlo random search, (4) a genetic algorithm and, (5) a systematic grid search. The first two are quite fast, local search algorithms, whilst the latter three are global search methods. The Nelder/Mead downhill simplex method was chosen for all simulated fits discussed in this thesis.

Once the fitting routine has been set to run, local minima are eventually reached, at which point the fitting automatically stops and gives the results as 'best fit' values for each varied parameter. It is then a manual process to change the central values of the input parameters (under *Sys1/2.g* etc.) accordingly, before beginning the next fitting iteration. This is the basis by which all simulated fits were completed - best fit parameter values following one fitting were chosen as central values for the next iteration of fitting. Once accurate values for sets of parameters were reached, they were fixed in any further iterations. Once the final fitting was complete for each spectrum, the output was saved and exported into a format suitable for plotting. This exact protocol was used for simulated fits of LT spectra also, the only difference being the simulation function used (for solid samples the appropriate function is *pepper*).



```

5
6 %Baseline correction
7 y2 = cwzero (y);
8
9 %Set starting parameters
10 Sys1.g = [2.008 2.006 2.002];
11 Sys1.lwpp = 0.07;
12 Sys1.Nucs = '14N';
13 Sys1.A = [11.4 17.6 104.5];
14 Sys1.logtcorr = [-7.91 -8.65 -8.37];
15 Sys1.weight = 1;
16
17 Sys2.g = [2.008 2.006 2.002];
18 Sys2.lwpp = 0.1;
19 Sys2.Nucs = '14N';
20 Sys2.A = [17.20 13.5 104.468];
21 Sys2.logtcorr = [-9.2 -9.759 -8.9];
22 Sys2.weight = 0.055;
23
24 %Set vary
25 Vary1.g = [0.001 0.001 0.001];
26 Vary1.lwpp = 0.04;
27 Vary1.A = [3 3 10];
28 Vary1.logtcorr = [2 2 2];
29 Vary1.weight = 0;
30
31 Vary2.g = [0.001 0.001 0.001];
32 Vary2.lwpp = 0.02;
33 Vary2.A = [3 3 10];
34 Vary2.logtcorr = [2 2 2];
35 Vary2.weight = 0.5;
36
37 %Set experimental parameters
38 Exp.mwFreq = 9.759902;
39 Exp.CenterSweep = [347.9 10];
40 Exp.nPoints = 1024;
41
42 %Calling the fitting function
43 FitOpt.Method = 'simplex int';
44 FitOpt.RandomStart = 0;
45 FitOpt.Scaling = 'minmax';
46 esfit ('chili',y2,{Sys1,Sys2},{Vary1,Vary2},Exp,[],FitOpt);
47

```

script Ln 35 Col 20

Figure 5.1: An example script used for simulated fits of cw-EPR datasets within EasySpin,¹⁵⁹ an open-source Matlab® toolbox.

5.1.4 Pulsed X-band EPR spectroscopy

All X-band pulsed EPR spectra were recorded on a Bruker E580 spectrometer using a Bruker MD5-W1 EPR probehead equipped with a self-modified cryogen-free cryostat (Advanced Research Systems Inc., Macungie, PA, USA). The microwave pulses were amplified using a 1 kW-TWT (Applied Systems Engineering Inc., Fort Worth, TX, USA). The temperatures used for spectral acquisition are highlighted in the relevant figure legends.

Sample preparation

For apo samples, freshly spin labelled MamM-CTD variants in buffer A were concentrated using Corning Spin-X UF 500 concentrators (5 kDa MWCO, 500 μ L; Sigma-Aldrich) and dimer protein concentration was determined using the optical absorbance at 280 nm, and an extinction coefficient for the dimer of 13993.3 M⁻¹cm⁻¹. Concentrated samples were then lyophilised overnight and dissolved to the final desired concentration with D₂O (Sigma-Aldrich). Samples of $\sim$ 70 μ L were transferred to EPR Quartz tubes (WilmaD SQ-707; WilmaD-LabGlass, Vinland, NJ, USA), and mixed well with Glycerol-d₈ (final concentration $\sim$ 30 %; Sigma-Aldrich). Samples were then carefully flash-frozen in liquid nitrogen.

For metal-bound samples, freshly spin labelled MamM-CTD variants were concentrated, lyophilised and dissolved in D₂O as described above. Stock solutions of metal salts (including ZnSO₄·5H₂O, CdCl₂, NiCl₂, CuSO₄·5H₂O, CoCl₂·6H₂O and Fe(NH₄)₂(SO₄)₂·6H₂O) were prepared in D₂O to a final concentration of 100 mM. Protein aliquots of 100 μ L were incubated with the desired excess of associated metal salts for 45 minutes on ice (details of excesses are noted in associated figure legends). Aliquots of metal-bound samples were transferred to EPR Quartz tubes, and mixed well with Glycerol-d₈ to a final concentration of $\sim$ 30 %. Samples were then carefully flash-frozen in liquid nitrogen. For the Fe²⁺-bound sample, each step was carried out under a steady stream of argon to eliminate oxygen during sample preparation.

For Zn²⁺ titration experiments of C267S and CSM-C267S variants, freshly spin labelled samples were concentrated, lyophilised and dissolved in D₂O as described above. ZnSO₄·5H₂O was added (1x excess) and allowed to incubate for 45 minutes on ice, followed by Glycerol-d₈ addition as previously described, sample transfer to EPR Quartz tubes, flash-freezing and measurement. Each first titre was carefully thawed over liquid nitrogen and removed from its EPR tube to an eppendorf. Further ZnSO₄·5H₂O for the second titre was then added, allowed to incubate for 2 hours on ice, transferred back to the EPR tube and measured. This was repeated for all other Zn²⁺ excesses; excesses used for titrations of C267S were 1x, 2x, 3x, 4x, 5x, 6x, 8x and 10x, whilst those used for titration of CSM-C267S were 1x, 2x, 3x and 5x.

Field-swept echo (FSE) measurements

Field-swept echo measurements were undertaken for each sample prior to further pulsed experiments. The field-swept spectrum was obtained by integrating the Hahn echo signal as

a function of the magnetic field after the following two-pulse sequence: $\pi/2(10 \text{ ns}) - \tau(120 \text{ ns}) - \pi(20 \text{ ns}) - \tau$ -. Experiments were typically conducted with a shot repetition time of 2 ms with 50 shots per point (unless otherwise stated).

T₂ measurements

T₂ measurements were conducted by integrating the Hahn echo signal as a function of time (T) after the following two-pulse sequence: $\pi/2(10 \text{ ns}) - \tau(120 \text{ ns}) - \pi(20 \text{ ns}) - \tau$ -. The π pulse is displaced incrementally through time for a set number of points, leading to an overall decrease in echo intensity over T. Experiments were undertaken at a single magnetic field value, corresponding to the detect position of associated DEER experiments. Experiments were typically conducted with a shot repetition time of 2 ms with 50 shots per point (unless otherwise stated).

Inversion recovery versus field measurements

Inversion recovery measurements were conducted over a specific range of magnetic field values by integrating the Hahn echo signal as a function of time after the following three-pulse sequence: $\pi(24 \text{ ns}) - T(1000 \text{ ns}) - \pi/2(12 \text{ ns}) - \tau(120 \text{ ns}) - \pi(20 \text{ ns}) - \tau$ -. The Hahn echo sequence (including the second and third pulses, and the delay between them) is displaced by 1000 ns increments for 1000 points in time, affording recovery of the inverted Hahn echo to be measured. further acquisition parameters are given in the associated figure legends. In some cases, field-swept spectra were measured at a single delay time (T); these are termed inversion recovery-filtered field swept echos (IFfFSE).

Double electron-electron resonance (DEER)

For DEER experiments a dead time free four-pulse sequence was applied. Observer pulses ($\pi/2$ and π) were set to 20 ns in length, with a pump pulse of 16 ns. The frequency offset $\Delta\nu$ ($\nu_{obs} - \nu_{pump}$) was set to between 60 - 80 MHz for SL-SL measurements, and between 210 - 230 MHz for SL-Cu measurements. The observer pulse separation τ_1 was determined according to the maximum echo intensity in the T₂ measurement, and thus differed between experiments; values for τ_1 were typically 120 ns, and in a few cases between 320 - 390 ns. The chosen length of τ_2 was dependent upon the lifetime of the echo following excitation, and varied between 1.3 - 4 μs . Echo signals were integrated using a video amplifier bandwidth of 20 MHz. The pump pulse was stepped out in 12 ns intervals for a total number of points in T that depended upon the τ_1 and τ_2 used. Nuclear modulation effects were

suppressed using a systematic variation of the inter-pulse delay time τ_1 and an appropriate phase cycling routine.

Inversion recovery-filtered DEER (*i*DEER)

For inversion recovery-filtered DEER experiments, an initial π inversion pulse was applied before the regular four-pulse DEER sequence. The delay time, T_F , between this inversion pulse and the four-pulse DEER sequence is known as the filter time, and is dependent on the species you are aiming to filter out detection of. All other aspects of the measurement were carried out in much the same way as that described for the regular DEER experiment.

5.2 GltPh

5.2.1 Plasmids

GltPh (including wild-type, C321A-A71C, C321A-A71C-G79A and C321A-A71C-K290E) was previously cloned into pBAD24 vector by Eric Gouaux (Oregon Health and Science University, USA), and was supplied by Harald H. Sitte and Thomas Stockner (Medical University of Vienna, Vienna, Austria).

5.2.2 Expression, purification and site-directed spin labelling of GltPh in detergent

Plasmids containing the required GltPh variant were introduced into chemically competent TOP10 *Escherichia coli* bacterial cells by heat-shock transformation ($1 \text{ ng} \cdot \mu\text{L}^{-1}$, $100 \mu\text{L}$). Cells harbouring the GltPh plasmid were grown in LB medium containing ampicillin ($100 \mu\text{g} \cdot \text{mL}^{-1}$) at 37°C , 190 rpm shaking until they reached an optical density of 0.7-0.8 at 600 nm. The temperature was lowered to 30°C and expression of the GltPh gene for protein production was induced by addition of 0.2 % L-(+)-arabinose (Sigma-Aldrich) for a further 3 hours. Harvesting of cells was carried out by centrifugation (Thermo Scientific Sorvall RC6 Plus) at 6000 rcf and 4°C for 15 minutes. Cell pellets (when not being used immediately) were stored at -80°C .

Cells were resuspended in lysis buffer (20 mM HEPES pH 7.5, 200 mM KCl, 5 mM MgCl_2) with DNase, PMSF, lysozyme, cCompleteTM mini EDTA-free protease inhibitor cocktail tabs and glycerol (10 %; Sigma-Aldrich) for 30 minutes at room temperature. Cells were then disrupted by two cycles of sonication and cellular debris was pelleted by centrifugation at 6000 rcf and 4°C for 10 minutes. Membrane fragments were pelleted from the resulting supernatant by ultracentrifugation at 176000 rcf and 4°C for 2 hours. Membrane pellets were resuspended in binding buffer (20 mM HEPES pH 7.5, 200 mM KCl) supplemented with n-Dodecyl- β -D-Maltoside (DDM) detergent (1 %). Protein was solubilised for 90 minutes at 4°C with gentle agitation, after which the solution was diluted using binding buffer to a final DDM concentration of 0.5 %. Insoluble protein/protein aggregates were pelleted via ultracentrifugation at 176000 rcf and 4°C for 15 minutes.

TALON Metal Affinity Resin ($500 \mu\text{L}$ resuspended resin per 1L of culture medium) was pre-equilibrated with binding buffer. For protein binding and purification, the supernatant from ultracentrifugation was supplemented with imidazole (pH 8.0) to a final concentration

of 10 mM, and incubated with TALON resin for 1 hour at 4 °C with gentle mixing. Washes 1 and 2 were carried out using purification buffer (20 mM HEPES, 200 mM KCl, 0.075 % DDM) supplemented with 10 mM (wash buffer 1) and 20 mM imidazole (wash buffer 2), in 40 and 5 column volumes, respectively. Protein was eluted in 10 x 1 mL elution fractions of purification buffer supplemented with 350 mM imidazole. The purity of the eluted protein was analysed via SDS-PAGE gel electrophoresis. The eluted protein was exchanged into storage buffer (20 mM HEPES pH 7.5, 100 mM KCl, 0.075 % DDM) using disposable PD-10 Desalting Columns by GE Healthcare. Protein was concentrated using Amicon Ultra-2 Centrifugal Filter Units (50 kDa MWCO), and protein concentration was determined using the optical absorbance at 280 nm and a molar extinction coefficient of $26820 \text{ M}^{-1}.\text{cm}^{-1}$. Purified protein not for immediate use was kept in 20 % glycerol, aliquoted, flash frozen and stored at -80 °C.

For site-directed spin labelling, some slight modifications were made to the above binding and purification protocol. Following binding to TALON resin, the flow through was discarded and the resin was washed in 20 column volumes of wash buffer 1. Binding buffer, supplemented with MTSSL (1 mM), was added to the top of the bed (2 column volumes). The mixture was covered with foil and left at 4°C overnight with gentle mixing. The resin was then washed in a further 20 column volumes of wash buffer 1. Steps following wash 1 were then carried out in the same manner as described above.

5.2.3 Reconstitution into liposomes

For NMR experiments of reconstituted protein, WT GltPh was first exchanged into buffer B (20 mM potassium phosphate pH 7.5, 100 mM KCl, 0.075 % DDM), and concentrated to approximately 50 μM using Amicon Ultra-2 Centrifugal Filter Units (50 kDa MWCO). *Escherichia coli* polar lipid extract (Avanti Polar Lipids Inc.) and 1,2-dioleoyl-*sn*-glycerophosphocholine (DOPC) (Avanti Polar Lipids Inc.) were mixed at a 1:3 ratio, and then dried under a constant, steady stream of argon, forming a thin lipid film. This was resuspended in degassed buffer C (20 mM potassium phosphate pH 7.5, 100 mM KCl) to a final concentration of 20 mg.mL^{-1} and subjected to 5 freeze-thaw cycles using liquid nitrogen and thawing at room temperature, forming large multilamellar vesicles (LMVs). The mixture was extruded through polycarbonate membranes with a pore size of 400 nm (Avanti Polar Lipids Inc.) to form large unilamellar vesicles (LUVs). Liposomes were diluted to 4 mg.mL^{-1} using buffer D (20 mM potassium phosphate pH 7.5, 100 mM KCl, 25 % glycerol), and then destabilised by dropwise addition of Triton X-100 (10 % w/v), adding in total 150 μL per 5 mL of liposomes. WT GltPh in buffer B was added at 20 μg protein per mg

of lipid and left to mix gently for 30 minutes at 4 °C. Detergent removal was afforded by stepwise addition of Bio-Beads SM-2, whereby the proteoliposome mixture was incubated and gently mixed with four successive additions 200 mg Bio-Beads SM2 per 5 mL of mixture (1 addition per hour), with the final addition left overnight. These were removed by filtration. Proteoliposomes were concentrated by ultracentrifugation at 140000 rcf for 2 hours at 4 °C. Buffers E (20 mM potassium phosphate pH 7.5, 100 mM NaCl) and F (20 mM potassium phosphate pH 7.5, 100 mM KCl) were prepared in D₂O, and were used for resuspension of proteoliposome pellets. Pellets were resuspended in either buffer for a final lipid concentration of approximately 40 mg.mL⁻¹. Freshly prepared proteoliposomes (PL-WT GltPh) were used for NMR experiments immediately, and not flash frozen between times.

The reconstitution density can be calculated using equation 5.1 below,

$$\chi_{reconstituted} = \frac{[P]/M_r^P}{[L]/M_r^L} \quad (5.1)$$

where $\chi_{reconstituted}$ is the reconstitution density (subunits/lipid), [P] and [L] are the protein and lipid concentrations, respectively, M_r^P and M_r^L are the molecular weights of the protein and lipid, respectively.²⁶¹ *E. coli* polar lipid extract is composed (wt/wt%) of 67 % phosphatidylethanolamine (PE), 23.2 % phosphatidylglycerol (PG) and 9.8 % cardiolipin (CA), and as such has an average molecular weight of 786 g.mol⁻¹ which is equal to that of DOPC.²⁶² Thus the average molecular mass of the 1:3 lipid mixture is simply 786 g.mol⁻¹. Using a molecular mass for trimeric GltPh of 133779 g.mol⁻¹, and considering that 20 µg of protein was added per mg of lipid, the reconstitution density is equal to 1.18 x 10⁻⁴ subunits/lipid.

Calculation of the number of lipid molecules in a liposome of known diameter can then be used, alongside the reconstitution density, to approximate the number of protein molecules per liposome. Equation 5.2 is used to calculate the number of lipid molecules in a unilamellar liposome (N_{tot}), where d is the diameter of the liposome (400 nm in this case), h is the thickness of the bilayer (typically around 5 nm) and a is the lipid head group area (approximately 0.7 nm for DOPC).

$$N_{tot} = \frac{[4\pi \cdot (\frac{d}{2})^2 + 4\pi \cdot (\frac{d}{2} - h)^2]}{a} \quad (5.2)$$

Substitution of each value into the above equation gives a value for N_{tot} equal to 1365437.5 lipids per (400 nm) liposome. Multiplication of subunits/lipid ($\chi_{reconstituted}$) by lipids/liposome

(N_{tot}) gives an estimate for the number of proteins per liposome of 161.²⁶³

Empty liposomes for control experiments were prepared in the same manner as above, and at the same lipid concentration, however steps including liposome destabilisation, protein integration and detergent removal were removed in this protocol. The external buffer E was used for resuspension of the liposome pellets in the final step.

5.2.4 Continuous wave X-band EPR spectroscopy

Sample preparation

For apo/ion/ion+substrate/ion+inhibitor bound samples, purified, MTSSL labelled protein in storage buffer was used. Aliquots of 15 μ L were taken for additions of ions/substrate/inhibitor. Stock solutions of (a) NaCl (1 M), (b) L-aspartate (20 mM) and (c) DL-TBOA (20 mM) were prepared in H₂O. For ions-only bound samples, NaCl was added to a final concentration of 50 mM. For ions and substrate bound samples, in addition to 50 mM NaCl, L-aspartate was added to a final concentration of 1 mM. For ions and inhibitor bound samples, in addition to 50 mM NaCl, DL-TBOA was added to a final concentration of 1 mM. All samples were incubated on ice for approximately 1 hour. For both apo and bound samples, aliquots of $\sim$ 10 μ L (apo/bound samples) were transferred to capillary tubes (outer diameter = 0.8 mm) for measurement.

Ambient temperature cw-EPR measurements

The ambient temperature (room temperature, RT) set-up for X-band cwEPR consisted of a Bruker E500 eleXsys spectrometer fitted with an ER 4123D (dielectric RT cw-EPR) resonator. The magnetic field was swept over a width of 100-150 G, with a centre field of 3478 G. The incident microwave frequency during measurements was $\sim$ 9.759 GHz, with a constant power of 0.3965 mW. For each spectrum, background noise was averaged over 100 sweeps, each with 1024 points across the magnetic field axis. Spin labelling efficiencies were calculated by comparing the double integral of the cw-EPR (first derivative) signal against that of known concentrations of 4-Maleimido-TEMPO (Sigma-Aldrich), forming a calibration curve.

5.2.5 Pulsed X-band EPR spectroscopy

All X-band pulsed EPR spectra were recorded on a Bruker E580 spectrometer (Bruker, Rheinstetten, Germany) using a Bruker MD5-W1 EPR probehead equipped with a self-

modified cryogen-free cryostat (Advanced Research Systems Inc., Macungie, PA, USA). The microwave pulses were amplified using a 1 kW-TWT (Applied Systems Engineering Inc., Fort Worth, TX, USA). The temperatures used for spectral acquisition are highlighted in the relevant figure legends.

Sample preparation

For apo/ion/ion+substrate/ion+inhibitor bound samples, purified, MTSSL labelled protein in storage buffer was used. Aliquots of 100 μ L were taken for additions of ions/substrate/inhibitor. Stock solutions of (a) NaCl (1 M), (b) L-aspartate (20 mM) and (c) DL-TBOA (20 mM) were prepared in D₂O. For ions-only bound samples, NaCl was added to a final concentration of 50 mM. For ion and substrate bound samples, in addition to 50 mM NaCl, L-aspartate was added to a final concentration of 1 mM. For ion and inhibitor bound samples, in addition to 50 mM NaCl, DL-TBOA was added to a final concentration of 1 mM. All samples were incubated on ice for approximately 1 hour. For both apo and bound samples, aliquots of $\sim$ 70 μ L were transferred to EPR Quartz tubes (Wilmad SQ-707; Wilmad-LabGlass, Vinland, NJ, USA), and mixed well with Glycerol-d₈ (final concentration $\sim$ 30 %; Sigma-Aldrich).

Field swept echo (FSE) measurements

Field-swept echo measurements were undertaken for each sample prior to further pulsed experiments. The field-swept spectrum was obtained by integrating the Hahn echo signal as a function of the magnetic field after the following two-pulse sequence: $\pi/2$ (10 ns) - τ (120 ns) - π (20 ns) - τ -. Experiments were typically conducted with a shot repetition time of 2 ms with 50 shots per point (unless otherwise stated).

T₂ measurements

T₂ measurements were conducted by integrating the Hahn echo signal as a function of time (T) after the following two-pulse sequence: $\pi/2$ (10 ns) - τ (120 ns) - π (20 ns) - τ -. The π pulse is displaced incrementally through time for a set number of points, leading to an overall decrease in echo intensity over T. Experiments were undertaken at a single magnetic field value, corresponding to the detect position of associated DEER experiments. Experiments were typically conducted with a shot repetition time of 2 ms with 50 shots per point (unless otherwise stated).

Double electron-electron resonance (DEER)

For DEER experiments a dead time free four-pulse sequence was applied. Observer pulses ($\pi/2$ and π) were set to 20 ns in length, with a pump pulse of 16 ns. The frequency offset $\Delta\nu$ ($\nu_{obs} - \nu_{pump}$) was set to 80 MHz for SL-SL measurements. The observer pulse separation τ_1 was determined according to the maximum echo intensity in the T_2 measurement, and thus differed between experiments; values for τ_1 were typically 120 ns. The chosen length of τ_2 was dependent upon the lifetime of the echo following excitation, and varied between 1.3 - 2.2 μ s. Echo signals were integrated using a video amplifier bandwidth of 20 MHz. The pump pulse was stepped out in 12 ns intervals for a total number of points in T that depended upon the τ_1 and τ_2 used. Nuclear modulation effects were suppressed using a systematic variation of the inter-pulse delay time τ_1 and an appropriate phase cycling routine. Experiments were typically conducted with a shot repetition time of 2 ms with 50 shots per point (unless otherwise stated).

5.2.6 NMR spectroscopy

NMR spectra were recorded at 278 K with D₂O as solvent, on an Ultra-Compact 800 MHz Bruker Avance III NMR spectrometer equipped with an inverse triple resonance (H/C/N) z-gradient probe head. STD NMR experiments were carried out using a pseudo 2D pulse sequence including spoil pulses to destroy residual magnetisation during the relaxation decay delay (two trim pulses of 2.5 and 5 ms, followed by a 3 ms gradient pulse on the z-axis). For selective saturation of GltPh, cascades of 49 ms Gaussian-shaped pulses (field strength of 90 Hz) were used with a 1 ms delay between successive pulses. Total saturation time for STD measurements was 2 s with a recycling delay of 2 s in experiments consisting of 512 scans. The saturation time of 2 s was selected as a compromise between sensitivity and accuracy in K_D determination, as a long full initial-slope STD NMR approach is not appropriate for relatively fast-degrading membrane proteins. Selective saturation of the protein was achieved by setting the on-resonance frequency at 0.6 ppm, to produce efficient saturation of the aliphatic side chains of the protein. This saturating radio frequency indeed also directly hit on the most up-field proton signal of the detergent molecules, contributing to a very efficient saturation of the whole system. Potential saturation transfer effects arising from non-specific interactions of the ligands with the detergent could be ruled out by negative results of STD NMR experiments of the protein–ligand system in the absence of Na⁺. For the reference (off-resonance) spectrum, the irradiation frequency was shifted to 40 ppm.

For detergent solubilised WT GltPh, protein was washed, exchanged and concentrated into either buffer G (20 mM potassium phosphate pH 7.5, 100 mM NaCl, 0.075 % DDM) or H (20 mM potassium phosphate pH 7.5, 100 mM KCl, 0.075 % DDM) prepared in D₂O, as described previously, for removal of glycerol and other small molecule impurities. Protein concentration was determined using the optical absorbance at 280 nm and a molar extinction coefficient of 26820 M⁻¹.cm⁻¹. An initial series of STD NMR experiments was conducted using WT GltPh in buffer G and increasing L-glutamate concentrations ranging from 0 to 5 mM via titration. A control experiment was also conducted using WT GltPh in buffer H (without Na⁺ ions) supplemented with 0.5 mM L-glutamate. Another series of STD NMR experiments were measured following a competition-based titration of L-aspartate (ranging from 0 to 1 μ M) into WT GltPh in buffer G supplemented with 1.5 mM L-glutamate. Finally, STD NMR experiments were conducted following titration of NaCl (from 3.75 μ M to 100 mM) into WT GltPh in buffer H, supplemented with 0.5 mM L-glutamate.

For PL-WT GltPh, the first set of STD NMR experiments were carried out using proteoliposomes in external buffer E and increasing L-glutamate concentrations from 0 to 5 mM. Control experiments were also conducted, using (1) empty liposomes in external buffer E and (2) PL-WT GltPh in buffer F, each supplemented with 1.5 mM L-glutamate. Finally, a series of STD NMR experiments were undertaken following a competition-based titration of L-aspartate (ranging from 0 to 2 μ M) into PL-WT GltPh in buffer E supplemented with 1.5 mM L-glutamate. Bruker Topspin (version 3.5) was employed for processing the NMR spectra and determination of peak integrals for STD-AF calculations. The K_D determination was performed using the curve fitting toolbox in Matlab.

Spin-lattice (longitudinal) relaxation times (T_1) were determined using the inversion recovery method²⁵⁷ for (a) 1.5 mM L-glutamate in buffer G, and (b) 1.5 mM L-glutamate added to empty liposomes in external buffer E. For these experiments, spectra of 16 K data points were obtained for 16 and 12 delay time (τ) values in cases (a) and (b) respectively, ranging from 1 ms to 5 s. Initial quick T_1 estimation was performed for all samples in order to set the appropriate relaxation delay, which was set to 5 s in all samples.

Bibliography

- (1) Schweiger, A.; Jeschke, G., *Principles of Pulse Electron Paramagnetic Resonance*; Oxford University Press: 2001.
- (2) Odom, B.; Hanneke, D.; D'Urso, B.; Gabrielse, G. *Physical Review Letters* **2006**, *97*, 030801.
- (3) Weil, J. A.; Bolton, J. R., *Electron Paramagnetic Resonance*, 2nd ed.; John Wiley & Sons, Inc.: 2006.
- (4) Chechik, V.; Carter, E.; Murphy, D., *Electron Paramagnetic Resonance*; Oxford Chemistry Primers: 2016.
- (5) Roessler, M. M.; Salvadori, E. *Chemical Society Reviews* **2018**, *47*, 2534–2553.
- (6) Grinberg, O. Y.; Dadali, A. A.; Dubinskii, A. A.; Vasserman, A. M.; Buchachenko, A. L.; Lebedev, Y. S. *Theoretical and Experimental Chemistry* **1980**, *15*, 457–461.
- (7) Eaton, G. R.; Eaton, S. S.; Barr, D. P.; Weber, R. T., *Quantitative EPR*; Springer Vienna: Vienna, 2010, pp 1–30.
- (8) Eaton, G. R.; Eaton, S. S.; Salikhov, K. M., *Foundations of Modern EPR*; World Scientific: 1998.
- (9) Hubbell, W. L.; Mchaourab, H. S.; Altenbach, C.; Lietzow, M. A. *Structure* **1996**, *4*, 779–783.
- (10) Klare, J. P.; Steinhoff, H.-J. *Photosynthesis Research* **2009**, *102*, 377–390.
- (11) Klug, C. S.; Feix, J. B. *Methods in Cell Biology* **2008**, *84*, 617–658.
- (12) Marsh, D., *Spin-Label Electron Paramagnetic Resonance Spectroscopy*; CRC Press: 2019.
- (13) Goldfarb, D.; Stoll, S., *EPR Spectroscopy: Fundamentals and Methods*; John Wiley & Sons: 2018.
- (14) Schmidt, T.; Wälti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M. *Angewandte Chemie - International Edition* **2016**, *55*, 15905–15909.

- (15) Hemminga, M. A.; Berliner, L., *ESR Spectroscopy in Membrane Biophysics*; Biological Magnetic Resonance, Vol. 27; Springer US: 2007.
- (16) Schiemann, O.; Prisner, T. F. *Quarterly Reviews of Biophysics* **2007**, *40*, 1–53.
- (17) Reginsson, G. W.; Schiemann, O. *Biochemical Journal* **2011**, *434*, 353–363.
- (18) Tkach, I.; Pornsuwan, S.; Höbartner, C.; Wachowius, F.; Sigurdsson, S. T.; Baranova, T. Y.; Diederichsen, U.; Sicoli, G.; Bennati, M. *Physical Chemistry Chemical Physics* **2013**, *15*, 3433.
- (19) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. *Applied Magnetic Resonance* **2006**, *30*, 473–498.
- (20) Worswick, S. G.; Spencer, J. A.; Jeschke, G.; Kuprov, I. *Science Advances* **2018**, *4*, eaat5218.
- (21) Jeschke, G. *Protein Science* **2018**, *27*, 76–85.
- (22) Yu, H. *Proceedings of the National Academy of Sciences* **1999**, *96*, 332–334.
- (23) Liang, B.; Tamm, L. K. *Nature Structural & Molecular Biology* **2016**, *23*, 468–474.
- (24) Levitt, M. H., *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, 2nd ed.; Wiley: 2008.
- (25) Cavanagh, J.; Fairbrother, W. J.; Palmer III, A. G.; Rance, M.; Skelton, N. J., *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed.; Elsevier Academic Press: 2007, pp 2–27.
- (26) Kleckner, I. R.; Foster, M. P. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2011**, *1814*, 942–968.
- (27) Sugiki, T.; Furuita, K.; Fujiwara, T.; Kojima, C. *Molecules* **2018**, *23*, 148.
- (28) Williamson, M. In *Encyclopedia of Analytical Science*; Elsevier: 2005, pp 342–349.
- (29) Ur Rahman, A.; Choudhary, M. I. In *Solving Problems with NMR Spectroscopy*; Elsevier: 1996; Chapter 4, pp 187–211.
- (30) Mayer, M.; Meyer, B. *Journal of the American Chemical Society* **2001**, *123*, 6108–6117.
- (31) Monaco, S.; Tailford, L. E.; Juge, N.; Angulo, J. *Angewandte Chemie - International Edition* **2017**, *56*, 15289–15293.
- (32) Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. *Journal of Chemical Education* **2011**, *88*, 990–994.
- (33) Hediger, M. A.; Romero, M. F.; Peng, J.-B.; Rolfs, A.; Takanaga, H.; Bruford, E. A. *Pflügers Archiv European Journal of Physiology* **2004**, *447*, 465–468.

- (34) Lin, L.; Yee, S. W.; Kim, R. B.; Giacomini, K. M. *Nature Reviews Drug Discovery* **2015**, *14*, 543–560.
- (35) Henderson, R. K.; Fendler, K.; Poolman, B. *Current Opinion in Biotechnology* **2019**, *58*, 62–71.
- (36) Desuzinges Mandon, E.; Agez, M.; Pellegrin, R.; Igonet, S.; Jawhari, A. *Analytical Biochemistry* **2017**, *517*, 40–49.
- (37) Kwan, T. O. C.; Reis, R.; Siligardi, G.; Hussain, R.; Cheruvara, H.; Moraes, I. *International Journal of Molecular Sciences* **2019**, *20*, 2605.
- (38) Quistgaard, E. M.; Löw, C.; Guettou, F.; Nordlund, P. *Nature Reviews Molecular Cell Biology* **2016**, *17*, 123–132.
- (39) Sahu, I. D.; Lorigan, G. A. *BioMed Research International* **2018**, *2018*, 1–13.
- (40) Husada, F.; Gouridis, G.; Vietrov, R.; Schuurman-Wolters, G. K.; Ploetz, E.; de Boer, M.; Poolman, B.; Cordes, T. *Biochemical Society Transactions* **2015**, *43*, 1041–1047.
- (41) Stone, T. J.; Buckman, T.; Nordio, P. L.; McConnell, H. M. *Proceedings of the National Academy of Sciences* **1965**, *54*, 1010–1017.
- (42) Hubbell, W. L.; Altenbach, C. In *Membrane Protein Structure*; Springer New York: 1994, pp 224–248.
- (43) Jeschke, G. *Emerging Topics in Life Sciences* **2018**, *2*, 9–18.
- (44) Boudker, O.; Verdon, G. *Trends in Pharmacological Sciences* **2010**, *31*, 418–426.
- (45) Jung, H. *FEBS Letters* **2002**, *529*, 73–77.
- (46) Wegener, C.; Tebbe, S.; Steinhoff, H. J.; Jung, H. *Biochemistry* **2000**, *39*, 4831–4837.
- (47) Jung, H.; Rübenhagen, R.; Tebbe, S.; Leifker, K.; Tholema, N.; Quick, M.; Schmid, R. *Journal of Biological Chemistry* **1998**, *273*, 26400–26407.
- (48) Jeschke, G.; Wegener, C.; Nietschke, M.; Jung, H.; Steinhoff, H. J. *Biophysical Journal* **2004**, *86*, 2551–2557.
- (49) Hilger, D.; Polyhach, Y.; Jung, H.; Jeschke, G. *Biophysical Journal* **2009**, *96*, 217–225.
- (50) Hilger, D.; Raba, M.; Lipiszko, K.; Polyhach, Y.; Jeschke, G.; Dunkel, S.; Klare, J.; Steinhoff, H.-J.; Bracher, S.; Quick, M.; Jung, H. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2012**, *1817*, S34–S35.
- (51) Bracher, S.; Guérin, K.; Polyhach, Y.; Jeschke, G.; Dittmer, S.; Frey, S.; Böhm, M.; Jung, H. *Journal of Biological Chemistry* **2016**, *291*, 4998–5008.

- (52) Raba, M.; Dunkel, S.; Hilger, D.; Lipiszko, K.; Polyhach, Y.; Jeschke, G.; Bracher, S.; Klare, J. P.; Quick, M.; Jung, H.; Steinhoff, H. J. *Structure* **2014**, *22*, 769–780.
- (53) Faham, S.; Watanabe, A.; Besserer, G. M.; Cascio, D.; Specht, A.; Hirayama, B. A.; Wright, E. M.; Abramson, J. *Science* **2008**, *321*, 810–814.
- (54) Krishnamurthy, H.; Gouaux, E. *Nature* **2012**, *481*, 469–474.
- (55) Claxton, D. P.; Quick, M.; Shi, L.; De Carvalho, F. D.; Weinstein, H.; Javitch, J. A.; McHaourab, H. S. *Nature Structural and Molecular Biology* **2010**, *17*, 822–829.
- (56) Kazmier, K.; Sharma, S.; Quick, M.; Islam, S. M.; Roux, B.; Weinstein, H.; Javitch, J. A.; McHaourab, H. S. *Nature Structural and Molecular Biology* **2014**, *21*, 472–479.
- (57) Weyand, S. et al. *Science* **2008**, *322*, 709–713.
- (58) Shimamura, T.; Weyand, S.; Beckstein, O.; Rutherford, N. G.; Hadden, J. M.; Sharpies, D.; Sansom, M. S.; Iwata, S.; Henderson, P. J.; Cameron, A. D. *Science* **2010**, *328*, 470–473.
- (59) Kazmier, K.; Sharma, S.; Islam, S. M.; Roux, B.; Mchaourab, H. S. *Proceedings of the National Academy of Sciences* **2014**, *111*, 14752–14757.
- (60) Ressler, S.; Terwisscha Van Scheltinga, A. C.; Vonnrhein, C.; Ott, V.; Ziegler, C. *Nature* **2009**, *458*, 47–52.
- (61) Tsai, C. J.; Khafizov, K.; Hakulinen, J.; Forrest, L. R.; Krämer, R.; Kühlbrandt, W.; Ziegler, C. *Journal of Molecular Biology* **2011**, *407*, 368–381.
- (62) Nicklisch, S. C.; Wunnicke, D.; Borovykh, I. V.; Morbach, S.; Klare, J. P.; Steinhoff, H. J.; Krämer, R. *Biochimica et Biophysica Acta - Biomembranes* **2012**, *1818*, 359–366.
- (63) Leone, V.; Wacławska, I.; Kossmann, K.; Koshy, C.; Sharma, M.; Prisner, T. F.; Ziegler, C.; Endeward, B.; Forrest, L. R. *Journal of General Physiology* **2019**, *151*, 381–394.
- (64) Saier, M. H.; Beatty, J. T.; Goffeau, A.; Harley, K. T.; Heijne, W. H.; Huang, S. C.; Jack, D. L.; Jähn, P. S.; Lew, K.; Liu, J.; Pao, S. S.; Paulsen, I. T.; Tseng, T. T.; Virk, P. S. *Journal of Molecular Microbiology and Biotechnology* **1999**, *1*, 257–279.
- (65) Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H. R.; Iwata, S. *Science* **2003**, *301*, 610–615.
- (66) Smirnova, I.; Kasho, V.; Kaback, H. R. *Biochemistry* **2011**, *50*, 9684–9693.
- (67) Marsh, D.; Henderson, P. J. *Biochimica et Biophysica Acta - Biomembranes* **2001**, *1510*, 464–473.

- (68) Hubert, A.; Henderson, P. J.; Marsh, D. *Biochimica et Biophysica Acta - Biomembranes* **2003**, *1611*, 243–248.
- (69) Smirnova, I.; Kasho, V.; Choe, J. Y.; Altenbach, C.; Hubbell, W. L.; Kaback, H. R. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 16504–16509.
- (70) Padan, E.; Venturi, M.; Gerchman, Y.; Dover, N. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2001**, *1505*, 144–157.
- (71) Hunte, C.; Screpanti, E.; Venturi, M.; Rimon, A.; Padan, E.; Michel, H. *Nature* **2005**, *435*, 1197–1202.
- (72) Hilger, D.; Jung, H.; Padan, E.; Wegener, C.; Vogel, K. P.; Steinhoff, H. J.; Jeschke, G. *Biophysical Journal* **2005**, *89*, 1328–1338.
- (73) Hilger, D.; Polyhach, Y.; Padan, E.; Jung, H.; Jeschke, G. *Biophysical Journal* **2007**, *93*, 3675–3683.
- (74) 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. *Journal of Biological Chemistry* **2011**, *286*, 634–648.
- (75) Anantharam, V.; Allison, M. J.; Maloney, P. C. *The Journal of Biological Chemistry* **1989**, *264*, 7244–50.
- (76) Putman, M.; Van Veen, H. W.; Degener, J. E.; Konings, W. N. *Microbiology* **2001**, *147*, 2873–2880.
- (77) Yan, N. *Annual Review of Biophysics* **2015**, *44*, 257–283.
- (78) Steed, P. R.; Zou, P.; Trone, K. E.; McHaourab, H. S. *Biochemistry* **2013**, *52*, 7964–7974.
- (79) Iyalomhe, O.; Herrick, D. Z.; Cafiso, D. S.; Maloney, P. C. *Biochemistry* **2014**, *53*, 7735–7744.
- (80) Masureel, M.; Smriti, S.; Martens, C.; Zou, P.; Ruyschaert, J.-M.; Mchaourab, H. S.; Govaerts, C. *Biophysical Journal* **2012**, *102*, 660a.
- (81) Masureel, M.; Martens, C.; Stein, R. A.; Mishra, S.; Ruyschaert, J.-M.; Mchaourab, H. S.; Govaerts, C. *Nature Chemical Biology* **2014**, *10*, 149–155.
- (82) Yin, Y.; He, X.; Szewczyk, P.; Nguyen, T.; Chang, G. *Science* **2006**, *312*, 741–744.
- (83) Otsuka, M.; Yasuda, M.; Morita, Y.; Otsuka, C.; Tsuchiya, T.; Omote, H.; Moriyama, Y. *Journal of Bacteriology* **2005**, *187*, 1552–1558.
- (84) Van Veen, H. W. *Nature* **2010**, *467*, 926–927.

- (85) Kuroda, T.; Tsuchiya, T. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2009**, *1794*, 763–768.
- (86) He, X.; Szewczyk, P.; Karyakin, A.; Evin, M.; Hong, W. X.; Zhang, Q.; Chang, G. *Nature* **2010**, *467*, 991–994.
- (87) Steed, P. R.; Stein, R. A.; Mishra, S.; Goodman, M. C.; Mchaourab, H. S. *Biochemistry* **2013**, *52*, 5790–5799.
- (88) Claxton, D. P.; Jagessar, K. L.; Steed, P. R.; Stein, R. A.; Mchaourab, H. S. *Proceedings of the National Academy of Sciences* **2018**, *115*, E6182–E6190.
- (89) Paulsen, I. T.; Saier, M. H. *Journal of Membrane Biology* **1997**, *156*, 99–103.
- (90) Lasry, I.; Golan, Y.; Berman, B.; Amram, N.; Glaser, F.; Assaraf, Y. G. *Journal of Biological Chemistry* **2014**, *289*, 7275–7292.
- (91) Silver, S.; Nucifora, G.; Chu, L.; Misra, T. K. *Trends in Biochemical Sciences* **1989**, *14*, 76–80.
- (92) Haney, C. J.; Grass, G.; Franke, S.; Rensing, C. *Journal of Industrial Microbiology and Biotechnology* **2005**, *32*, 215–226.
- (93) Anton, A.; Grosse, C.; Reissmann, J.; Pribyl, T.; Nies, D. H. *Journal of bacteriology* **1999**, *181*, 6876–81.
- (94) Delhaize, E.; Kataoka, T.; Hebb, D. M.; White, R. G.; Ryan, P. R. *The Plant Cell* **2003**, *15*, 1131–1142.
- (95) Montanini, B.; Blaudez, D.; Jeandroz, S.; Sanders, D.; Chalot, M. *BMC Genomics* **2007**, *8*, 107.
- (96) Cubillas, C.; Vinuesa, P.; Tabche, M. L.; García-De Los Santos, A. *Metallomics* **2013**, *5*, 1634–1643.
- (97) Fang, C. T.; Chen, H. C.; Chuang, Y. P.; Chang, S. C.; Wang, J. T. *Antimicrobial Agents and Chemotherapy* **2002**, *46*, 2024–2028.
- (98) Uebe, R.; Junge, K.; Henn, V.; Poxleitner, G.; Katzmann, E.; Plitzko, J. M.; Zarivach, R.; Kasama, T.; Wanner, G.; Pósfai, M.; Böttger, L.; Matzanke, B.; Schüller, D. *Molecular Microbiology* **2011**, *82*, 818–835.
- (99) Huang, L.; Tepaamorndech, S. *Molecular Aspects of Medicine* **2013**, *34*, 548–560.
- (100) Maret, W. *Advances in Nutrition* **2013**, *4*, 82–91.
- (101) Qin, Y.; Thomas, D.; Fontaine, C. P.; Colvin, R. A. *Neuroscience Letters* **2009**, *450*, 206–210.

- (102) Ohana, E.; Hoch, E.; Keasar, C.; Kambe, T.; Yifrach, O.; Hershfinkel, M.; Sekler, I. *Journal of Biological Chemistry* **2009**, *284*, 17677–17686.
- (103) Shusterman, E.; Beharier, O.; Shiri, L.; Zarivach, R.; Etzion, Y.; Campbell, C. R.; Lee, I. H.; Okabayashi, K.; Dinudom, A.; Cook, D. I.; Katz, A.; Moran, A. *Metallomics* **2014**, *6*, 1656–1663.
- (104) Chao, Y.; Fu, D. *Journal of Biological Chemistry* **2004**, *279*, 12043–12050.
- (105) Guffanti, A. A.; Wei, Y.; Rood, S. V.; Krulwich, T. A. *Molecular Microbiology* **2002**, *45*, 145–153.
- (106) Lu, M.; Fu, D. *Science* **2007**, *317*, 1746–1748.
- (107) Lu, M.; Chai, J.; Fu, D. *Nature Structural and Molecular Biology* **2009**, *16*, 1063–1067.
- (108) Coudray, N.; Valvo, S.; Hu, M.; Lasala, R.; Kim, C.; Vink, M.; Zhou, M.; Provasi, D.; Filizola, M.; Tao, J.; Fang, J.; Penczek, P. A.; Ubarretxena-Belandia, I.; Stokes, D. L. *PNAS* **2013**, *110*, 2140–2145.
- (109) Lopez-Redondo, M. L.; Coudray, N.; Zhang, Z.; Alexopoulos, J.; Stokes, D. L. *Proceedings of the National Academy of Sciences* **2018**, *115*, 3042–3047.
- (110) Cherezov, V.; Höfer, N.; Szebenyi, D. M.; Kolaj, O.; Wall, J. G.; Gillilan, R.; Srinivasan, V.; Jaroniec, C. P.; Caffrey, M. *Structure* **2008**, *16*, 1378–1388.
- (111) Higuchi, T.; Hattori, M.; Tanaka, Y.; Ishitani, R.; Nureki, O. *Proteins: Structure, Function, and Bioinformatics* **2009**, *76*, 768–771.
- (112) Zeytuni, N.; Offer, T.; Davidov, G.; Zarivach, R. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **2012**, *68*, 927–930.
- (113) Zeytuni, N.; Uebe, R.; Maes, M.; Davidov, G.; Baram, M.; Raschdorf, O.; Nadav-Tsubery, M.; Kolusheva, S.; Bitton, R.; Goobes, G.; Friedler, A.; Miller, Y.; Schüler, D.; Zarivach, R. *PLoS ONE* **2014**, *9*, e92141.
- (114) Zeytuni, N.; Uebe, R.; Maes, M.; Davidov, G.; Baram, M.; Raschdorf, O.; Friedler, A.; Miller, Y.; Schüler, D.; Zarivach, R. *PLoS ONE* **2014**, *9*, e97154.
- (115) Barber-Zucker, S.; Hall, J.; Mangapuram, S. V.; Kass, I.; Kolusheva, S.; MacMillan, F.; Zarivach, R.; Henn, A. *FEBS Journal* **2019**, *286*, 2193–2215.
- (116) Uebe, R.; Keren-Khadmy, N.; Zeytuni, N.; Katzmann, E.; Navon, Y.; Davidov, G.; Bitton, R.; Plitzko, J. M.; Schüler, D.; Zarivach, R. *Molecular Microbiology* **2018**, *107*, 542–557.
- (117) Finney, L. A. *Science* **2003**, *300*, 931–936.

- (118) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *Journal of Computational Chemistry* **2004**, *25*, 1605–1612.
- (119) Bellini, S. *Chinese Journal of Oceanology and Limnology* **2009**, *27*, 3–5.
- (120) Bellini, S. *Chinese Journal of Oceanology and Limnology* **2009**, *27*, 6–12.
- (121) Blakemore, R. *Science* **1975**, *190*, 377–379.
- (122) Gorby, Y. A.; Beveridge, T. J.; Blakemore, R. P. *Journal of bacteriology* **1988**, *170*, 834–841.
- (123) Balkwill, D. L.; Maratea, D.; Blakemore, R. P. *Journal of Bacteriology* **1980**, *141*, 1399–1408.
- (124) Barber-Zucker, S.; Keren-Khadmy, N.; Zarivach, R. *Protein Science* **2016**, *25*, 338–351.
- (125) Barber-Zucker, S.; Zarivach, R. *ACS Chemical Biology* **2017**, *12*, 13–22.
- (126) Komeili, A. *FEMS Microbiology Reviews* **2012**, *36*, 232–255.
- (127) Frankel, R. B.; Bazyliniski, D. A. *Hyperfine Interactions* **1994**, *90*, 135–142.
- (128) Komeili, A.; Vali, H.; Beveridge, T. J.; Newman, D. K. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 3839–3844.
- (129) Komeili, A.; Li, Z.; Newman, D. K.; Jensen, G. J. *Science* **2006**, *311*, 242–245.
- (130) Katzmann, E.; Scheffel, A.; Gruska, M.; Plitzko, J. M.; Schüler, D. *Molecular Microbiology* **2010**, *77*, 208–224.
- (131) Arakaki, A.; Shimizu, K.; Oda, M.; Sakamoto, T.; Nishimura, T.; Kato, T. *Organic & Biomolecular Chemistry* **2015**, *13*, 974–989.
- (132) Raschdorf, O.; Forstner, Y.; Kolinko, I.; Uebe, R.; Plitzko, J. M.; Schüler, D. *PLOS Genetics* **2016**, *12*, e1006101.
- (133) Uebe, R.; Schüler, D. *Nature Reviews Microbiology* **2016**, *14*, 621–637.
- (134) Raschdorf, O.; Bonn, F.; Zeytuni, N.; Zarivach, R.; Becher, D.; Schüler, D. *Journal of Proteomics* **2018**, *172*, 89–99.
- (135) Quinlan, A.; Murat, D.; Vali, H.; Komeili, A. *Molecular Microbiology* **2011**, *80*, 1075–1087.
- (136) Nguyen, H. V.; Suzuki, E.; Oestreicher, Z.; Minamide, H.; Endoh, H.; Fukumori, Y.; Taoka, A. *Biochemistry and Biophysics Reports* **2016**, *7*, 39–44.
- (137) Lohße, A.; Borg, S.; Raschdorf, O.; Kolinko, I.; Tompa, É.; Pósfai, M.; Faivre, D.; Baumgartner, J.; Schülere, D. *Journal of Bacteriology* **2014**, *196*, 2658–2669.

- (138) Draper, O.; Byrne, M. E.; Li, Z.; Keyhani, S.; Barrozo, J. C.; Jensen, G.; Komeili, A. *Molecular Microbiology* **2011**, *82*, 342–354.
- (139) Sonkaria, S.; Fuentes, G.; Verma, C.; Narang, R.; Khare, V.; Fischer, A.; Faivre, D. *PLoS ONE* **2012**, *7*, e34189.
- (140) Cornejo, E.; Subramanian, P.; Li, Z.; Jensen, G. J.; Komeili, A. *mBio* **2016**, *7*, DOI: 10.1128/mBio.01898-15.
- (141) Raschdorf, O.; Müller, F. D.; Pósfaï, M.; Plitzko, J. M.; Schüler, D. *Molecular Microbiology* **2013**, *89*, 872–886.
- (142) Hershey, D. M.; Ren, X.; Melnyk, R. A.; Browne, P. J.; Ozyamak, E.; Jones, S. R.; Chang, M. C.; Hurley, J. H.; Komeili, A. *PLoS Biology* **2016**, *14*, 1002402.
- (143) Siponen, M. I.; Adryanczyk, G.; Ginet, N.; Arnoux, P.; Pignol, D. In *Biochemical Society Transactions*, Portland Press: 2012; Vol. 40, pp 1319–1323.
- (144) Siponen, M. I.; Legrand, P.; Widdrat, M.; Jones, S. R.; Zhang, W. J.; Chang, M. C.; Faivre, D.; Arnoux, P.; Pignol, D. *Nature* **2013**, *502*, 681–684.
- (145) Jones, S. R.; Wilson, T. D.; Brown, M. E.; Rahn-Lee, L.; Yu, Y.; Fredriksen, L. L.; Ozyamak, E.; Komeili, A.; Chang, M. C. Y. *Proceedings of the National Academy of Sciences* **2015**, *112*, 3904–3909.
- (146) Scheffel, A.; Gärdes, A.; Grünberg, K.; Wanner, G.; Schüler, D. *Journal of Bacteriology* **2008**, *190*, 377–386.
- (147) Arakaki, A.; Yamagishi, A.; Fukuyo, A.; Tanaka, M.; Matsunaga, T. *Molecular Microbiology* **2014**, *93*, 554–567.
- (148) Yoshino, T.; Matsunaga, T. *Applied and Environmental Microbiology* **2006**, *72*, 465–471.
- (149) Valverde-Tercedor, C.; Montalbán-López, M.; Perez-Gonzalez, T.; Sanchez-Quesada, M. S.; Prozorov, T.; Pineda-Molina, E.; Fernandez-Vivas, M. A.; Rodriguez-Navarro, A. B.; Trubitsyn, D.; Bazylnski, D. A.; Jimenez-Lopez, C. *Applied Microbiology and Biotechnology* **2015**, *99*, 5109–5121.
- (150) Nudelman, H.; Zarivach, R. *Frontiers in Microbiology* **2014**, *5*, 9.
- (151) Rawlings, A. E.; Bramble, J. P.; Hounslow, A. M.; Williamson, M. P.; Monnington, A. E.; Cooke, D. J.; Staniland, S. S. *Chemistry - A European Journal* **2016**, *22*, 7885–7894.
- (152) Rawlings, A. E.; Bramble, J. P.; Walker, R.; Bain, J.; Galloway, J. M.; Staniland, S. S. *Proceedings of the National Academy of Sciences* **2014**, *111*, 16094–16099.

- (153) Barber-Zucker, S.; Uebe, R.; Davidov, G.; Navon, Y.; Sherf, D.; Chill, J. H.; Kass, I.; Bitton, R.; Schöler, D.; Zarivach, R. *Scientific Reports* **2016**, *6*, 31933.
- (154) Barber-Zucker, S.; Hall, J.; Froes, A.; Kolusheva, S.; MacMillan, F.; Zarivach, R. *Journal of Biological Chemistry* **2020**, *295*, 16614–16629.
- (155) Barber-Zucker, S.; Shahar, A.; Kolusheva, S.; Zarivach, R. *Scientific Reports* **2020**, *10*, 14022.
- (156) Lopez-Redondo, M. L.; Coudray, N.; Zhang, Z.; Alexopoulos, J.; Stokes, D. L. *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115*, 3042–3047.
- (157) Mchaourab, H. S.; Lietzow, M. A.; Hideg, K.; Hubbell, W. L. *Biochemistry* **1996**, *35*, 7692–7704.
- (158) Bordignon, E. In *Encyclopedia of Biophysics*; Springer Berlin Heidelberg: 2013, pp 2439–2447.
- (159) Stoll, S.; Schweiger, A. *Journal of Magnetic Resonance* **2006**, *178*, 42–55.
- (160) El Mkami, H.; Norman, D. G. In *Methods in Enzymology*, 2015; Vol. 564, pp 125–152.
- (161) Vera, I. M. S.; Blackburn, M. E.; Galiano, L.; Fanucci, G. E. *Current Protocols in Protein Science* **2013**, *74*, DOI: 10.1002/0471140864.ps1717s74.
- (162) Gayler, K. H.; Woontner, L. *The Journal of Physical Chemistry* **1956**, *60*, 1569–1571.
- (163) Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. *Physical Chemistry Chemical Physics* **2012**, *14*, 10762–10773.
- (164) Richert, S.; Cremers, J.; Kuprov, I.; Peeks, M. D.; Anderson, H. L.; Timmel, C. R. *Nature Communications* **2017**, *8*, 14842.
- (165) Gatteschi, D.; Barra, A. L.; Caneschi, A.; Cornia, A.; Sessoli, R.; Sorace, L. *Coordination Chemistry Reviews* **2006**, *250*, 1514–1529.
- (166) Caneschi, A.; Gatteschi, D.; Sessoli, R.; Rey, P. *Accounts of Chemical Research* **1989**, *22*, 392–398.
- (167) Benelli, C.; Gatteschi, D. *Chemical Reviews* **2002**, *102*, 2369–2388.
- (168) Veber, S. L.; Fedin, M. V.; Potapov, A. I.; Maryunina, K. Y.; Romanenko, G. V.; Sagdeev, R. Z.; Ovcharenko, V. I.; Goldfarb, D.; Bagryanskaya, E. G. *Journal of the American Chemical Society* **2008**, *130*, 2444–2445.
- (169) Fedin, M.; Veber, S.; Gromov, I.; Maryunina, K.; Fokin, S.; Romanenko, G.; Sagdeev, R.; Ovcharenko, V.; Bagryanskaya, E. *Inorganic Chemistry* **2007**, *46*, 11405–11415.

- (170) Van Amsterdam, I. M. C.; Ubbink, M.; Canters, G. W.; Huber, M. *Angewandte Chemie International Edition* **2003**, *42*, 62–64.
- (171) Ji, M.; Ruthstein, S.; Saxena, S. *Accounts of Chemical Research* **2014**, *47*, 688–695.
- (172) Meng, R.; Becker, J.; Lin, F.-T.; Saxena, S.; Weber, S. G. *Inorganica Chimica Acta* **2005**, *358*, 2933–2942.
- (173) Yang, Z.; Ji, M.; Saxena, S. *Applied Magnetic Resonance* **2010**, *39*, 487–500.
- (174) Lawless, M. J.; Sarver, J. L.; Saxena, S. *Angewandte Chemie* **2017**, *129*, 2147–2149.
- (175) Lewin, A.; Hill, J. P.; Boetzel, R.; Georgiou, T.; James, R.; Kleanthous, C.; Moore, G. R. *Inorganica Chimica Acta* **2002**, *331*, 123–130.
- (176) Cunningham, T. F.; Shannon, M. D.; Putterman, M. R.; Arachchige, R. J.; Sengupta, I.; Gao, M.; Jaroniec, C. P.; Saxena, S. *The Journal of Physical Chemistry B* **2015**, *119*, 2839–2843.
- (177) Peisach, J.; Blumberg, W. E. *Archives of Biochemistry and Biophysics* **1974**, *165*, 691–708.
- (178) Kozlevčar, B.; Šegedin, P. *Croatica Chemica Acta* **2008**, *81*, 369–379.
- (179) Maly, T.; Prisner, T. F. *Journal of Magnetic Resonance* **2004**, *170*, 88–96.
- (180) Maly, T.; MacMillan, F.; Zwicker, K.; Kashani-Poor, N.; Brandt, U.; Prisner, T. F. *Biochemistry* **2004**, *43*, 3969–3978.
- (181) Maly, T.; Zwicker, K.; Cernescu, A.; Brandt, U.; Prisner, T. F. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2009**, *1787*, 584–592.
- (182) Van Wonderen, J. H.; Kostrz, D. N.; Dennison, C.; MacMillan, F. *Angewandte Chemie - International Edition* **2013**, *52*, 1990–1993.
- (183) Amor, M.; Busigny, V.; Durand-Dubief, M.; Tharaud, M.; Ona-Nguema, G.; Gélabert, A.; Alphandéry, E.; Menguy, N.; Benedetti, M. F.; Chebbi, I.; Guyot, F. *Proceedings of the National Academy of Sciences* **2015**, *112*, 1699–1703.
- (184) Keim, C. N.; Lins, U.; Farina, M. *FEMS Microbiology Letters* **2009**, *292*, 250–253.
- (185) Tanaka, M.; Brown, R.; Hondow, N.; Arakaki, A.; Matsunaga, T.; Staniland, S. *Journal of Materials Chemistry* **2012**, *22*, 11919–11921.
- (186) Focke, P. J.; Wang, X.; Larsson, H. P. *Structure* **2013**, *21*, 694–705.
- (187) Clements, J. D. *Trends in Neurosciences* **1996**, *19*, 163–171.
- (188) Danbolt, N. C. *Progress in Neurobiology* **2001**, *65*, 1–105.
- (189) Otis, T. S.; Kavanaugh, M. P.; Jahr, C. E. *Science* **1997**, *277*, 1515–1518.

- (190) Tanaka, K.; Watase, K.; Manabe, T.; Yamada, K.; Watanabe, M.; Takahashi, K.; Iwama, H.; Nishikawa, T.; Ichihara, N.; Kikuchi, T.; Okuyama, S.; Kawashima, N.; Hori, S.; Takimoto, M.; Wada, K. *Science* **1997**, *276*, 1699–1702.
- (191) Vandenberg, R. J.; Ryan, R. M. *Physiological Reviews* **2013**, *93*, 1621–1657.
- (192) Storck, T.; Schulte, S.; Hofmann, K.; Stoffel, W. *Proceedings of the National Academy of Sciences* **2006**, *89*, 10955–10959.
- (193) Pines, G.; Danbolt, N. C.; Bjørås, M.; Zhang, Y.; Bendahan, A.; Eide, L.; Koepsell, H.; Storm-Mathisen, J.; Seeberg, E.; Kanner, B. I. *Nature* **1992**, *360*, 464–467.
- (194) Kanai, Y.; Hediger, M. A. *Nature* **1992**, *360*, 467–471.
- (195) Arriza, J.; Fairman, W.; Wadiche, J.; Murdoch, G.; Kavanaugh, M.; Amara, S. *The Journal of Neuroscience* **1994**, *14*, 5559–5569.
- (196) Fairman, W. A.; Vandenberg, R. J.; Arriza, J. L.; Kavanaugh, M. P.; Amara, S. G. *Nature* **1995**, *375*, 599–603.
- (197) Arriza, J. L.; Eliasof, S.; Kavanaugh, M. P.; Amara, S. G. *Proceedings of the National Academy of Sciences* **2002**, *94*, 4155–4160.
- (198) Kanai, Y.; Nussberger, S.; Romero, M. F.; Boron, W. F.; Hebert, S. C.; Hediger, M. A. *Journal of Biological Chemistry* **1995**, *270*, 16561–16568.
- (199) Owe, S. G.; Marcaggi, P.; Attwell, D. *Journal of Physiology* **2006**, *577*, 591–599.
- (200) Zerangue, N.; Kavanaugh, M. P. *Nature* **1996**, *383*, 634–637.
- (201) Mitchell, P. *Nature* **1957**, *180*, 134–136.
- (202) Jardetzky, O. *Nature* **1966**, *211*, 969–970.
- (203) Edwards, P. W. *BBA - Biomembranes* **1974**, *345*, 373–386.
- (204) Patlak, C. S. *The Bulletin of Mathematical Biophysics* **1957**, *19*, 209–235.
- (205) Slotboom, D. J. *Nature Reviews Microbiology* **2014**, *12*, 79–87.
- (206) Yernool, D.; Boudker, O.; Jin, Y.; Gouaux, E. *Nature* **2004**, *431*, 811–818.
- (207) Slotboom, D. J.; Konings, W. N.; Lolkema, J. S. *Microbiology and Molecular Biology Reviews* **1999**, *63*, 293–307.
- (208) Slotboom, D. J.; Konings, W. N.; Lolkema, J. S. *Trends in Biochemical Sciences* **2001**, *26*, 534–539.
- (209) Ryan, R. M.; Mitrovic, A. D.; Vandenberg, R. J. *Journal of Biological Chemistry* **2004**, *279*, 20742–20751.

- (210) Ryan, R. M.; Mindell, J. A. *Nature Structural and Molecular Biology* **2007**, *14*, 365–371.
- (211) Boudker, O.; Ryan, R. M.; Yernool, D.; Shimamoto, K.; Gouaux, E. *Nature* **2007**, *445*, 387–393.
- (212) Reyes, N.; Ginter, C.; Boudker, O. *Nature* **2009**, *462*, 880–885.
- (213) Verdon, G.; Boudker, O. *Nature Structural and Molecular Biology* **2012**, *19*, 355–357.
- (214) Reyes, N.; Oh, S.; Boudker, O. *Nature Structural and Molecular Biology* **2013**, *20*, 634–640.
- (215) Verdon, G.; Oh, S. C.; Serio, R.; Boudker, O. *eLife* **2014**, *2014*, 2283.
- (216) Oh, S.; Boudker, O. *eLife* **2018**, *7*, DOI: 10.7554/eLife.37291.
- (217) Wang, X.; Boudker, O. *eLife* **2020**, *9*, DOI: 10.7554/eLife.58417.
- (218) Ryan, R. M.; Compton, E. L.; Mindell, J. A. *Journal of Biological Chemistry* **2009**, *284*, 17540–17548.
- (219) Shiwei, H.; Ryan, R. M.; Vandenberg, R. J. *Journal of Biological Chemistry* **2009**, *284*, 4510–4515.
- (220) Akyuz, N.; Georgieva, E. R.; Zhou, Z.; Stolzenberg, S.; Cuendet, M. A.; Khelashvili, G.; Altman, R. B.; Terry, D. S.; Freed, J. H.; Weinstein, H.; Boudker, O.; Blanchard, S. C. *Nature* **2015**, *518*, 68–73.
- (221) Hänelt, I.; Jensen, S.; Wunnicke, D.; Slotboom, D. J. *Journal of Biological Chemistry* **2015**, *290*, 15962–15972.
- (222) Silverstein, N.; Ewers, D.; Forrest, L. R.; Fahlke, C.; Kanner, B. I. *Journal of Biological Chemistry* **2015**, *290*, 28988–28996.
- (223) Venkatesan, S.; Saha, K.; Sohail, A.; Sandtner, W.; Freissmuth, M.; Ecker, G. F.; Sitte, H. H.; Stockner, T. *PLOS Computational Biology* **2015**, *11*, ed. by Punta, M., e1004551.
- (224) Erkens, G. B.; Hänelt, I.; Goudsmits, J. M.; Slotboom, D. J.; Van Oijen, A. M. *Nature* **2013**, *502*, 119–123.
- (225) Grewer, C.; Watzke, N.; Wiessner, M.; Rauen, T. *Proceedings of the National Academy of Sciences* **2000**, *97*, 9706–9711.
- (226) Holley, D. C.; Kavanaugh, M. P. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2009**, *364*, 155–161.

- (227) Larsson, H. P.; Wang, X.; Lev, B.; Bacongus, I.; Caplan, D. A.; Vyleta, N. P.; Koch, H. P.; Diez-Sampedro, A.; Noskov, S. Y. *Proceedings of the National Academy of Sciences* **2010**, *107*, 13912–13917.
- (228) Tao, Z.; Rosental, N.; Kanner, B. I.; Gameiro, A.; Mwaura, J.; Grever, C. *Journal of Biological Chemistry* **2010**, *285*, 17725–17733.
- (229) Huang, Z.; Tajkhorshid, E. *Biophysical Journal* **2010**, *99*, 1416–1425.
- (230) Bastug, T.; Heinzelmann, G.; Kuyucak, S.; Salim, M.; Vandenberg, R. J.; Ryan, R. M. *PLoS ONE* **2012**, *7*, 33058.
- (231) Guskov, A.; Jensen, S.; Faustino, I.; Marrink, S. J.; Slotboom, D. J. *Nature Communications* **2016**, *7*, 13420.
- (232) Jensen, S.; Guskov, A.; Rempel, S.; Hänelt, I.; Slotboom, D. J. *Nature Structural and Molecular Biology* **2013**, *20*, 1224–1227.
- (233) Akyuz, N.; Altman, R. B.; Blanchard, S. C.; Boudker, O. *Nature* **2013**, *502*, 114–118.
- (234) Ruan, Y.; Miyagi, A.; Wang, X.; Chami, M.; Boudker, O.; Scheuring, S. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114*, 1584–1588.
- (235) Focke, P. J.; Moenne-Loccoz, P.; Larsson, H. P. *Journal of Neuroscience* **2011**, *31*, 6255–6262.
- (236) Georgieva, E. R.; Borbat, P. P.; Ginter, C.; Freed, J. H.; Boudker, O. *Nature Structural and Molecular Biology* **2013**, *20*, 215–221.
- (237) Hänelt, I.; Wunnicke, D.; Bordignon, E.; Steinhoff, H. J.; Slotboom, D. J. *Nature Structural and Molecular Biology* **2013**, *20*, 210–214.
- (238) Riederer, E. A.; Focke, P. J.; Georgieva, E. R.; Akyuz, N.; Matulef, K.; Borbat, P. P.; Freed, J. H.; Blanchard, S. C.; Boudker, O.; Valiyaveetil, F. I. *eLife* **2018**, *7*, DOI: 10.7554/eLife.36478.
- (239) Heinzelmann, G.; Bastug, T.; Kuyucak, S. *Journal of Physical Chemistry B* **2013**, *117*, 5486–5496.
- (240) Huang, Y.; Wang, X.; Lv, G.; Razavi, A. M.; Huysmans, G. H.; Weinstein, H.; Bracken, C.; Eliezer, D.; Boudker, O. *Nature Chemical Biology* **2020**, *16*, 1006–1012.
- (241) Canul-Tec, J. C.; Assal, R.; Cirri, E.; Legrand, P.; Brier, S.; Chamot-Rooke, J.; Reyes, N. *Nature* **2017**, *544*, 446–451.
- (242) Qiu, B.; Matthies, D.; Fortea, E.; Yu, Z.; Boudker, O. *Science Advances* **2021**, *7*, eabf5814.

- (243) Garaeva, A. A.; Guskov, A.; Slotboom, D. J.; Paulino, C. *Nature Communications* **2019**, *10*, 3427.
- (244) Garaeva, A. A.; Oostergetel, G. T.; Gati, C.; Guskov, A.; Paulino, C.; Slotboom, D. J. *Nature Structural and Molecular Biology* **2018**, *25*, 515–521.
- (245) Mari, S.; Serrano-Gómez, D.; Cañada, F. J.; Corbí, A. L.; Jiménez-Barbero, J. *Angewandte Chemie - International Edition* **2004**, *44*, 296–298.
- (246) Claasen, B.; Axmann, M.; Meinecke, R.; Meyer, B. *Journal of the American Chemical Society* **2005**, *127*, 916–919.
- (247) Airoldi, C.; Giovannardi, S.; Laferla, B.; Jiménez-Barbero, J.; Nicotra, F. *Chemistry - A European Journal* **2011**, *17*, 13395–13399.
- (248) Cox, B. D.; Mehta, A. K.; Diraddo, J. O.; Liotta, D. C.; Wilson, L. J.; Snyder, J. P. *Biochemical and Biophysical Research Communications* **2015**, *466*, 28–32.
- (249) Venkitakrishnan, R. P.; Benard, O.; Max, M.; Markley, J. L.; Assadi-Porter, F. M. In *Membrane Protein Structure and Dynamics*; Humana Press: 2012; Vol. 914, pp 47–63.
- (250) Meinecke, R.; Meyer, B. *Journal of Medicinal Chemistry* **2001**, *44*, 3059–3065.
- (251) Fredriksson, K.; Lottmann, P.; Hinz, S.; Onila, I.; Shymanets, A.; Harteneck, C.; Müller, C. E.; Griesinger, C.; Exner, T. E. *Angewandte Chemie - International Edition* **2017**, *56*, 5750–5754.
- (252) Igonet, S.; Raingeval, C.; Cecon, E.; Pučić-Baković, M.; Lauc, G.; Cala, O.; Baranowski, M.; Perez, J.; Jockers, R.; Krimm, I.; Jawhari, A. *Scientific Reports* **2018**, *8*, 1–14.
- (253) Yong, K. J.; Vaid, T. M.; Shilling, P. J.; Wu, F. J.; Williams, L. M.; Deluigi, M.; Plückthun, A.; Bathgate, R. A.; Gooley, P. R.; Scott, D. J. *ACS Chemical Biology* **2018**, *13*, 1090–1102.
- (254) Vaid, T. M.; Chalmers, D. K.; Scott, D. J.; Gooley, P. R. *Chemistry - A European Journal* **2020**, *26*, 11796–11805.
- (255) Bumbak, F.; Thomas, T.; Noonan-Williams, B. J.; Vaid, T. M.; Yan, F.; Whitehead, A. R.; Bruell, S.; Kocan, M.; Tan, X.; Johnson, M. A.; Bathgate, R. A.; Chalmers, D. K.; Gooley, P. R.; Scott, D. J. *ACS Pharmacology and Translational Science* **2020**, *3*, 690–705.
- (256) Angulo, J.; Enríquez-Navas, P. M.; Nieto, P. M. *Chemistry - A European Journal* **2010**, *16*, 7803–7812.

- (257) Vold, R. L.; Waugh, J. S.; Klein, M. P.; Phelps, D. E. *The Journal of Chemical Physics* **1968**, *48*, 3831–3832.
- (258) Wang, Y. S.; Liu, D.; Wyss, D. F. *Magnetic Resonance in Chemistry* **2004**, *42*, 485–489.
- (259) Cantor, C. R.; Schimmel, P. R. In *Chapter 15 - Ligand interactions at equilibrium*; WH Freeman and Co., Oxford: 1980; Chapter 15, pp 849–1371.
- (260) Cheng, Y.; Prusoff, W. H. *Biochemical Pharmacology* **1973**, *22*, 3099–3108.
- (261) Cliff, L.; Chadda, R.; Robertson, J. L. *Biochimica et Biophysica Acta - Biomembranes* **2020**, *1862*, 183033.
- (262) Wu, C. H.; Bialecka-Fornal, M.; Newman, D. K. *eLife* **2015**, *2015*, DOI: 10 . 7554 / eLife.05663.
- (263) Hutchinson, F. J.; Francis, S. E.; Lyle, I. G.; Jones, M. N. *BBA - Biomembranes* **1989**, *978*, 17–24.
- (264) Hall, J. L.; Sohail, A.; Cabrita, E. J.; Macdonald, C.; Stockner, T.; Sitte, H. H.; Angulo, J.; MacMillan, F. *Scientific Reports* **2020**, *10*, 16483.

Appendices

I Published research - Metal binding to MamM-CTD

The following publication (of which I am a co-author) includes the study of Zn^{2+} binding in MamM-CTD variants via a range of biophysical and biochemical methods.¹¹⁵ Results of this research give insights into the conformations and flexibility of MamM-CTD dimers in solution in both apoprotein and Zn^{2+} -bound conditions, as well as the nature of metal binding sites and a suggested mechanism of Zn^{2+} binding. My contributions include the acquisition and analysis of EPR data, and composition of the text explaining these data. These results have been edited and reformatted for inclusion in chapter 3.8.1. Any additional results within this publication, which have been discussed to aid in the understanding of results presented in chapter 3.8.1 of this thesis, have been referenced at the relevant points.



Metal binding to the dynamic cytoplasmic domain of the cation diffusion facilitator (CDF) protein MamM induces a 'locked-in' configuration

Shiran Barber-Zucker^{1,2,3}, Jenny Hall⁴, Sivasubramanyan Venkata Mangapuram⁵, Itamar Kass², Sofiya Kolusheva³, Fraser MacMillan⁴ , Raz Zarivach^{1,2,3} and Arnon Henn⁵

1 Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel

2 The National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer Sheva, Israel

3 Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer Sheva, Israel

4 School of Chemistry, University of East Anglia, Norwich Research Park, UK

5 Faculty of Biology, Technion- Israel Institute of Technology, Haifa, Israel

Keywords

cation diffusion facilitator; PELDOR; protein kinetics; protein structure–function relationship; protein–metal interaction

Correspondence

A. Henn, Faculty of Biology, Technion- Israel Institute of Technology, Haifa 3200003, Israel

Tel: +972 (4) 8294839

E-mail: arnon.henn@technion.ac.il

R. Zarivach, Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel

Tel: +972 (8) 6461999

E-mail: zarivach@bgu.ac.il

and

F. MacMillan, School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK

Tel: +44 (0) 1603-592766

E-mail: fraser.macmillan@uea.ac.uk

(Received 7 August 2018, revised 2 February 2019, accepted 25 February 2019)

doi:10.1111/febs.14795

Cation diffusion facilitator (CDF) proteins are a conserved family of transmembrane transporters that ensure cellular homeostasis of divalent transition metal cations. Metal cations bind to CDF protein's cytoplasmic C-terminal domain (CTD), leading to closure from its apo open V-shaped dimer to a tighter packed structure, followed by a conformational change of the transmembrane domain, thus enabling transport of the metal cation. By implementing a comprehensive range of biochemical and biophysical methods, we studied the molecular mechanism of metal binding to the magnetotactic bacterial CDF protein MamM CTD. Our results reveal that the CTD is rather dynamic in its apo form, and that two dependent metal-binding sites, a single central binding site and two symmetrical, peripheral sites, are available for metal binding. However, only cation binding to the peripheral sites leads to conformational changes that lock the protein in a compact state. Thus, this work reveals how metal binding is regulating the sequential uptakes of metal cations by MamM, and extends our understanding of the complex regulation mechanism of CDF proteins.

Database

Structural data are available in RCSB Protein Data Bank under the accession numbers: 6G64, 6G55, 6G5E and 6G6I (for CS, C267S, CS-C267S and W247A, respectively).

Introduction

Protein–metal cation interactions are key players in the function of many biological processes, whether they serve as structural anchors (as in zinc fingers [1]),

participate directly in enzymatic reactions (as in the calcium/manganese cluster in photosystem II [2] or copper in tyrosinases' active site [3]) or stabilize

Abbreviations

CDF, cation diffusion facilitator; CS, central site; CTD, C-terminal domain; EPR, electron paramagnetic resonance; ITC, isothermal titration calorimetry; MD, molecular dynamics; MMM, multiscale modelling of macromolecules; MTB, magnetotactic bacteria; PCA, principal component analysis; PDB, Protein Data Bank; PELDOR, pulsed electron–electron double resonance; PS, peripheral site; WT, wild-type.

enzymes' active sites (as in magnesium in metallonucleases' active sites [4]). Since metal cations are rather small ligands, they can easily diffuse into empty but charged binding sites, thus metal-ligand binding processes usually occur quite quickly [5]. These cations must bind selectively in their specific binding sites to ensure the correct conformational changes and function [6]. Understanding the exact mechanism of this type of binding, especially in terms of how these interactions occur and lead to conformational changes on short time scales, is of huge value for the study of the overall mechanism of a specific protein. However, the molecular mechanism of protein–metal binding is often not well characterized, and such knowledge is lacking in many protein families.

The cation diffusion facilitator (CDF) family is an example of a protein class where such protein–metal binding is not well understood. CDFs are a divalent transition metal cation transporter family that maintain such cation homeostasis, and are conserved throughout all domains of life [7]. CDF proteins export a range of divalent metal cations, such as Zn^{2+} , Mn^{2+} , Fe^{2+} , Cd^{2+} and Co^{2+} out of the cytoplasm, typically by exploiting the proton motive force while undergoing conformational change [8–10]. CDF proteins form hetero- or homodimers and contain two major domains: the transmembrane domain and the cytoplasmic C-terminal domain (CTD). The transmembrane domain is found in all known CDF proteins and includes six helices per monomer as well as a well-characterized metal-binding site through which the metal cations are transported [9–11]. The CTD domain is abundant in CDF proteins, adopting a metallocharopone-like fold and undertaking metal-binding-related conformational changes [12,13]. The CTD has been better characterized both structurally and biochemically than the transmembrane domain, and it has been proposed to act as a regulatory domain. The CTD contains metal-binding sites which are found on the protein surface, thus exposed to cytoplasmic metal ions. With quite low affinity to metals, it has been speculated that in high cytoplasmic metal concentrations, the metal cations are bound to these CTD-binding sites. This causes a conformational change where the two monomers, that were found in an apo V-shaped conformation, approach each other to form a tighter V-shaped conformation. This, in turn, induces a conformational change of the transmembrane domain allowing the efficient, fast transport of cations through this domain (which has greater affinity) [8,13]. Namely, only at high levels of metals (concentrations that approach cytotoxicity) will they be bound to the CTD to activate the actual transport. In contrast to

the transmembrane binding site, the metal-binding sites in the CTD are not fully conserved in terms of properties, location and number (although in most proteins more than one binding site has been proposed) [14]. The exact binding mechanism and its energetics, the kinetic pathway by which metal binding induces these proposed structural changes, and how the ions bind and interact with the CTD, remain poorly understood. A number of specific questions are yet to be addressed, including how the metal-binding sites are populated during activation, and if there is an ordered binding sequence. Furthermore, whether allostery and cooperativity exist and play a functional role in CTD activation, and finally if all metal-binding sites are required to fully induce the conformational change and activation of the CTD.

To address these questions, the metal-binding mechanism in MamM CTD, a CDF protein from magnetotactic bacteria (MTB), has been investigated. MTB are a group of Gram-negative bacteria that biosynthesize iron-based magnetic particles in designated organelles named magnetosomes. These magnetosomes are arranged in a chain-like fashion, thus enabling them to align themselves to the geomagnetic field in order to reach their preferred habitat, usually the oxic-anoxic zones in aquatic environments [15,16]. MamM transports iron from the bacterium cytoplasm into such magnetosomes, which enables synthesis of iron-based magnetic particles inside the vesicles. In the absence of MamM or in case of mutations that cause its dysfunction, these particles are either not formed or formed defectively [13,17–19]. Since such differences in magnetic particle formation in MTB are easy to study, MamM serves as an effective model to study CDF proteins. The MamM CTD structure has been solved in its apo form, and a comprehensive biophysical study has revealed that it undergoes conformational change when it binds Fe^{2+} and Zn^{2+} in three proposed binding sites: two symmetrical peripheral binding sites and one central binding site. Mutations in these postulated binding site residues lead to a decreased function both *in vivo* and *in vitro* [13]. Here, a combination of different biophysical methods is used to study MamM CTD dynamics and kinetics, in order to decipher the protein–metal binding mechanism. Overall, the unique combination of X-ray crystallography, CD spectroscopy, pulsed electron paramagnetic resonance [EPR and pulsed electron–electron double resonance (PELDOR)] spectroscopy, molecular dynamics (MD), isothermal titration calorimetry (ITC), fluorescence spectrometry and stopped-flow results reveal a dependence between the central and peripheral binding sites, and that rapid

binding occurs at the central binding site while only binding to the peripheral sites (PSs) induces significant conformational change. These results reveal the complexity of metal binding to MamM specifically and to CDF proteins' cytoplasmic domain in general, and constitute a significant step towards the understanding of the diverse regulation mechanisms in CDF proteins and their contribution to CDF protein function.

Results

Previously it was shown that metal binding to MamM CTD *in vitro* occurs only at basic pH due to the involvement of histidine residues in the coordination of the metal ions [13]. Fe^{2+} tends to precipitate at basic pH; however, its effective concentration is highly important for stoichiometry determination and interpretation of the kinetic mechanism of ligand binding. Hence, for this reason and also to enable a comparison with previous studies utilizing Zn^{2+} as a model for metal binding [13], we performed a structural, thermodynamic and kinetic investigation of Zn^{2+} to study the metal-binding mechanism in MamM CTD.

Structural studies of MamM CTDs indicate binding-dependent conformational change

MamM CTD is arranged as a stable dimer forming two putative symmetrical PSs (histidine 264 and glutamate 289 of monomer one and two, respectively) and one central binding site formed with a quartet of residues (aspartate 249 and histidine 285 from each monomer; Fig. 1A) [13]. To understand the MamM CTD- Zn^{2+} -binding mechanism, constructs termed wild-type (WT; MamM CTD residues: 215–318), PS (WT with D249A and H285A mutations which abolishes the central binding site, construct contains only the PSs) and central site (CS; WT with H264A and E289A mutations which abolishes the peripheral binding sites, construct contains only the CS) were used.

The crystal structure of the apo WT protein, which has been solved previously, displays the characteristic fold of CDF proteins [Protein Data Bank (PDB) ID: 3W5X, 3W5Y] [13]. The crystal structures of both the apo PS construct (PDB ID: 3W8P, 3W66) [13] as well as the apo CS construct (PDB ID: 6G64, Tables 1–3, Fig. 1B) which was solved for this study reveal no differences as compared to the WT structure (see Table 4 for RMSD data between the different structures),

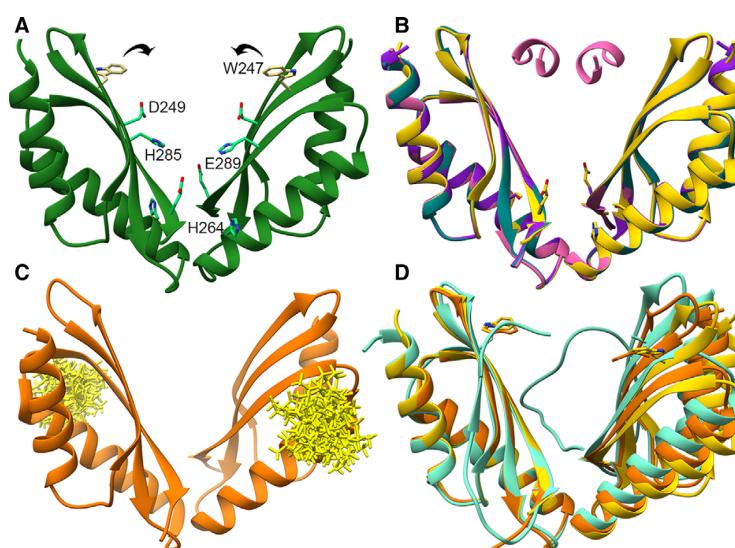


Fig. 1. MamM CTD crystal structures and models. (A) The dimeric apo-MamM CTD WT structure (PDB ID: 3W5X). The binding site residues (D249, H264, H285, E289, light green) and the fluorescence-active tryptophan residue (W247, yellow) are indicated. Black arrows correspond to the proposed movement of the monomers towards each other. (B) The new crystal structures of C267S (pink, PDB ID: 6G55), CS-C267S (purple, PDB ID: 6G5E) and CS (cyan, PDB ID: 6G64) overlapped onto the WT structure (yellow, PDB ID: 3W5X [13]). All mutation site residues (both in mutants and WT proteins) are indicated as sticks. (C) Using the PDB ID: 3W5Y [13] an MMM-modelled structure of WT MamM CTD is presented including the computed rotamer library of MTSSL spin labels attached to position C275 [22]. (D) The new crystal structure of W247A (light blue, PDB ID: 6G6I) overlapped onto the WT structures (PDB ID: 3W5X in yellow and PDB ID: 3W5Y in orange). All mutation site residues (both in mutants and WT proteins) are indicated as sticks.

Table 1. Crystallization of MamM CTD mutants.

Protein name	MamM CTD C267S	MamM CTD W247A	MamM CTD H264A-E289A-C267S	MamM CTD H264A-E289A
PDB ID	6G55	6G6I	6G5E	6G64
Crystallization conditions	0.2 M Li ₂ SO ₄ , 0.1 M BIS-TRIS pH = 5.5, 25% PEG 3350	0.1 M BIS-TRIS pH = 5.5, 20% PEG 3350	0.2 M Li ₂ SO ₄ , 0.1 M BIS-TRIS pH = 5.5, 25% PEG 3350	0.2 M (NH ₄) ₂ SO ₄ , 0.1 M BIS-TRIS pH = 5.5, 22% PEG 3350
Cryo protectant	—	50% PEG 3350	—	—
Protein concentration (mg·mL ⁻¹)	7.5	6.08	5	10
Crystallization type	Vapour diffusion sitting drop			
Data collection	P13 – DESY	ID29 – ESRF	ID29 – ESRF	Home source
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M	MAR345

Table 2. Data collection and refinement statistics of MamM CTD mutants. Values in parentheses are for the highest resolution shell. One crystal was used per dataset. Data were collected at 100K. A, allowed region; O, outliers; P, preferred region.

Protein name	MamM CTD C267S	MamM CTD W247A	MamM CTD H264A-E289A-C267S	MamM CTD H264A-E289A
PDB ID	6G55	6G6I	6G5E	6G64
Data collection	P13 – DESY	ID29 – ESRF	ID29 – ESRF	Home source
Space group	C 2 2 2 ₁	P 2 ₁ 2 ₁ 2 ₁	C 2 2 2 ₁	C 2 2 2 ₁
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	37.31, 94.37, 53.64	38.65, 65.75, 68.65	37.03, 94.79, 53.50	37.11, 95.09, 53.73
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	1.65–35.43 (1.65–1.68)	2.40–47.49 (2.40–2.49)	1.60–47.40 (1.60–1.62)	1.90–50.00 (1.90–1.93)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.044 (1.635)	0.193 (1.500)	0.029 (0.572)	0.042 (0.273)
<i>I</i> / σ <i>I</i>	16.8 (1.2)	9.7 (1.5)	26.9 (2.2)	48.5 (5.1)
CC _{1/2}	0.999 (0.636)	0.996 (0.665)	0.999 (0.931)	0.986 (0.892)
Completeness (%)	99.5 (97.9)	99.7 (97.2)	99.2 (86.5)	97.2 (88.9)
Redundancy	4.6 (4.5)	8.7 (8.7)	6.4 (5.8)	8.5 (4.0)
Wavelength (Å)	0.97631	0.975	0.975	1.5406
No. unique reflections	11 722	7305	12 814	7559
Refinement				
Resolution (Å)	1.65–35.43	2.40–47.49	1.60–35.48	1.90–47.54
Rwork/Rfree	19.04/21.51	22.10/26.18	20.61/25.68	20.51/23.90
No. atoms				
Protein	A- 658, D-57	A-728, B-815	639	622
Ligand/ion	14	4	10	14
Water	41	23	52	53
<i>B</i> -factors				
Protein	A- 39.861, D-76.573	A- 44.222, B-47.739	34.768	25.620
Ligand/ion	SO ₄ -43.610, BCT-57.580	β ME-79.395	SO ₄ -31.602	SO ₄ -32.880, β ME-52.630
Water	43.445	34.486	40.492	30.642
RMSD				
Bond lengths (Å)	0.014	0.006	0.006	0.013
Bond angles (°)	1.602	1.09	0.805	1.598
Ramachandran statistics	P: 75 (98.68%), A: 1 (1.32%), O: 0 (0%)	P: 187 (98.42%), A: 3 (1.58%), O: 0 (0%)	P: 76 (98.70%), A: 1 (1.30%), O: 0 (0%)	P: 77 (98.72%), A: 1 (1.28%), O: 0 (0%)
Missing residues	A- 211, 293–318, D- 211–306, 314–318	A: 307–318. B: 211, 316–318	293–318	211, 293–318

suggesting that the apo forms do not differ in their three dimensional crystal structures. Since, it has not yet proven possible to crystallize the different MamM CTD

constructs in the presence of Zn²⁺, other structural methods have been applied to better understand the conformational changes occurring due to metal binding.

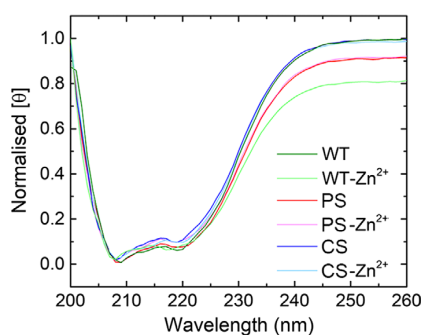
Table 3. Crystallographic software used for solving the structure of MamM CTD mutants. Manual refinement was performed using coot version 0.8.9 [44]. Aimless, Phaser MR and Refmac5 were used through the ccp4i package [45].

Protein name	MamM CTD C267S	MamM CTD W247A	MamM CTD H264A-E289A-C267S	MamM CTD H264A-E289A
PDB ID	6G55	6G6I	6G5E	6G64
Data reduction	XDS [46]	XDS	XDS	HKL2000 [47]
Data scaling	Aimless [48]	Aimless	Aimless	HKL2000
Structure solution method	Molecular replacement – using MamM CTD WT structure (PDB ID: 3W5X)			
Phasing	Phaser MR [49]			
Refinement	Refmac5 [50], PDB-REDO [51]	Phenix [52], PDB-REDO	Refmac5, Phenix	Refmac5, PDB-REDO

Table 4. RMSDs between crystal structures of the different MamM CTD constructs. RMSDs between structures (carbon alpha) were calculated using the iterative magic fit tool in SWISS-PDBVIEWER [36].

Protein name	WT – 3W5X	WT – 3W5Y	PS – 3W66	PS – 3W8P	CS – 6G64
WT – 3W5X	–	0.82	0.88	1.04	0.19
WT – 3W5Y	–	–	1.00	0.74	0.92
PS – 3W66	–	–	–	0.97	0.90
PS – 3W8P	–	–	–	–	1.11
CS – 6G64	–	–	–	–	–

CD spectra of all apo protein constructs (WT, PS and CS) confirmed that the secondary structures of all three proteins are also conserved in solution (Fig. 2). The CD spectral trends after addition of Zn^{2+} (protein : Zn^{2+} at 1 : 15 ratio) are consistent with the apo forms, suggesting that the conformational changes occurring due to metal binding do not involve substantial changes in their secondary structures, but most

**Fig. 2.** CD spectra of MamM CTD in the apo and metal-bound forms. To ensure that no metals are bound to MamM CTD proteins, proteins were pre-incubated with EDTA for 5 h, which was subsequently removed by dialysis overnight. The proteins were diluted to a final concentration of 0.3 mg mL^{-1} , with or without Zn^{2+} (15 : 1 Zn^{2+} to protein ratio), and spectral profiles were measured in a wavelength range of 200–260 nm. Darker colours refer to the apo forms and lighter colours to the metal-bound forms; WT in green, PS in red and CS in blue.

likely at the dimerization interface, similar to the CDF protein CzbB [12].

To further understand the MamM CTD metal-dependent conformational changes, site-directed spin labelling was applied in combination with PELDOR spectroscopy [20,21], which affords the detection of both distances and distance distributions between labelled sites within the dimer. During site-directed spin labelling, the paramagnetic (1-oxyl-2,2,5,5-tetra-methyl- Δ^3 -pyrroline-3-methyl) methanethiosulfonate spin label (MTSSL) reacts specifically with cysteine residues of the protein. MamM CTD contains two intrinsic cysteine residues, C267 and C275, the second of which is found on the periphery of the protein in a solvent-exposed region. Both *in silico* modelling and spin labelling efficiencies indicate that C267 does not label well, hence PELDOR experiments were performed using the C267S variant protein which provides only distances between the two labelled C275 sites on the protein (Fig. 1C). *In silico* prediction of interspin label distances was carried out using multi-scale modelling of macromolecules (MMM) [22], a Matlab[®]-based open-source modelling toolbox.

This C267S mutation was introduced by site-directed mutagenesis to all three MamM CTD constructs: C267S (WT with C267S mutation), PS-C267S (PS with C267S mutation) and CS-C267S (CS with C267S mutation). Crystallization trials were performed on these constructs to confirm that the C267S mutation also does not alter the protein structure. Only the C267S and CS-C267S constructs formed crystals

allowing a structure to be determined (PDB IDs: 6G55 and 6G5E, respectively, see Tables 1–3 and Fig. 1B). In both cases, the structure revealed no difference to the wild-type structure. All variant proteins labelled efficiently, indicating good accessibility of the C275 site. Further, by using the structure of the apo C267S form, the MMM study predicted an interspin distance of 4.0 nm between the labelled cysteines at this position (C275, Fig. 1C).

In Fig. 3 PELDOR experiments of all three apo forms (as well as Zn^{2+} -bound) are presented. The upper panels of Fig. 3 present the background-subtracted time-domain traces for both apo and metal-bound protein, while the lower panels represent the distance distribution probability of the metal-bound state after data analysis using DeerAnalysis 2016 [23]. Both the apo forms of C267S and CS-C267S revealed no defined distance between the two labelled cysteines, suggesting that these proteins do not adopt a single stable conformation in this state, but are rather dynamic (Fig. 3A,B). In the PS-C267S apo protein, however, a strong initial decay of the PELDOR signal in the time-domain was observed which is quite different from the C267S and CS-C267S time-domain spectra (Fig. 3C). Such initial steep decays can be

indicative of very short distances, typically < 2 nm between the labels. In the apo WT protein, the negatively charged aspartate and bulky histidine residues that form the putative CS protrude towards one another in the space between monomers, resulting in clashes and/or charge repulsion. This could increase the distance between monomers. Removal of these residues from the dimer interface could in principle allow for more stable and tighter packing as compared to the apo WT protein; however, the direct closure of the V-shape dimer would not be in agreement with a short interspin distance inferred from the PELDOR time-domain results. It was previously suggested that the two monomers could also swivel along a perpendicular axis [13] rather than simply moving laterally towards one another. Hence, removal of these charged, bulky residues in PS-C267S could allow for a larger degree of swivelling movement and thus explain the proposed short interspin distance in PS-C267S.

Addition of zinc chloride ($5\times$ excess) to each MamM CTD form (C267S- Zn^{2+} , PS-C267S- Zn^{2+} and CS-C267S- Zn^{2+}) results in large changes in the PELDOR time-domain traces (Fig. 3 upper panels). For C267S- Zn^{2+} , a narrow distance distribution was determined with a mean distance of 3.3 nm (Fig. 3A, lower

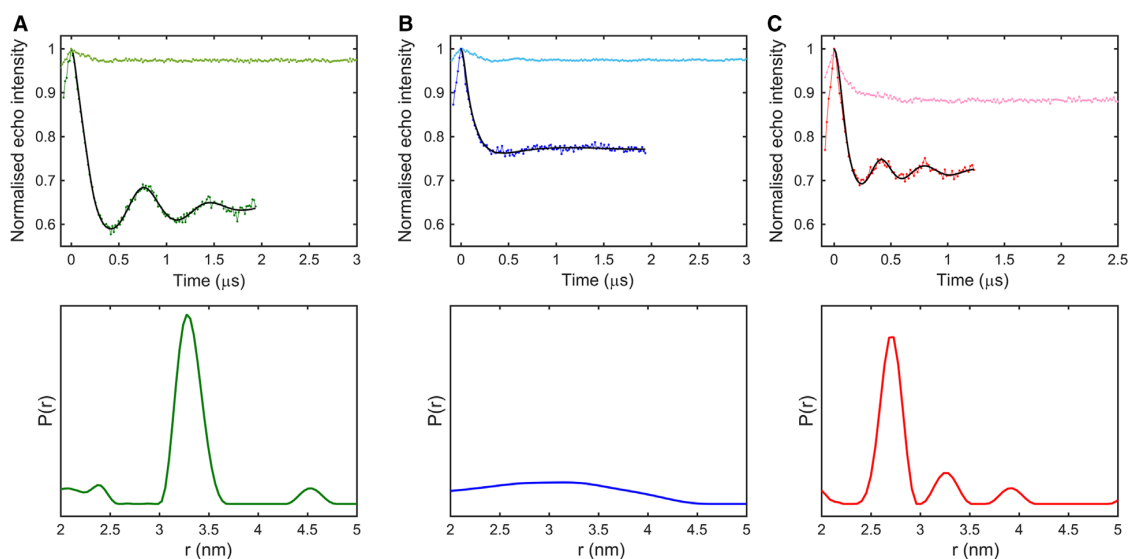


Fig. 3. Four-pulse PELDOR experiments of doubly spin-labelled MamM CTD constructs containing C275, in both apo and Zn^{2+} -bound forms, measured at 60 K. (Top panel) Green, blue and red traces are the normalized background-corrected dipolar evolution of (A) C267S, (B) CS-C267S and (C) PS-C267S MamM CTD respectively. The traces for both apo and Zn^{2+} -bound forms are presented in light and dark colours, respectively. All raw time traces were subjected to a first-order polynomial background correction. Fittings of Zn^{2+} -bound corrected traces are shown as black lines. (Bottom panel) Distance distributions of Zn^{2+} -bound MamM CTD constructs (A) C267S, (B) CS-C267S and (C) PS-C267S. Fitting of all the dipolar evolutions and subsequent distance distributions were calculated with DeerAnalysis2016 [23] using a Tikhonov regularization.

panel), indicative of a predominantly closed conformation upon metal binding which is rigid compared to the apo form and slightly shorter than the MMM prediction. A systematic Zn^{2+} titration (from 1 : 1 to 1 : 10 protein : metal ratios) revealed that the observed modulation depth increases with the successive addition of Zn^{2+} up until a ratio of 1 : 3–1 : 4 where it reaches a saturation point, which can be indicative of 3–4 Zn^{2+} ions bound per dimer (Fig. 4 left panels and Table 5). Beyond this ratio, no further increase in modulation depth is observed and indeed there is only a slight decrease which is due to the high metal ion and protein concentrations, which is known to cause aggregation in solution. An increase in modulation depth was also observed for CS-C267S- Zn^{2+} compared to its apo state; however, the distance distribution obtained was extremely broad, indicating that a single stable conformation was not adopted in the presence of the Zn^{2+} (Fig. 3B). This suggests that whilst the addition of metal has some effect on the structure and dynamics of the protein, an entirely closed and rigid Zn^{2+} -bound state is unlikely to be present. Hence, mutation of the peripheral binding sites does affect metal binding and does not allow meaningful conformational changes to be observed. For PS-C267S- Zn^{2+} , again a predominant, but distinct narrow distance distribution is obtained, with a mean distance of 2.7 nm (Fig. 3C).

This result would suggest that the addition of Zn^{2+} causes PS-C267S to adopt a slightly different rigid conformation, where the spin labels are now further apart as compared to the apo state. Such a 2.7-nm distance would fit with a tight V-shape packing that is not possible in the WT protein due to the presence of the histidine and aspartate residues in the metal-binding site, which are proposed to increase the gap between the monomers as compared to this alanine mutant. Repeating the Zn^{2+} titration experiment with PS-C267S MamM CTD with Zn^{2+} (from 1 : 1 to 1 : 5 protein : metal ratios) reveals an increase in modulation depth with successive addition of Zn^{2+} up until a ratio of 1 : 2, at which point saturation is already reached and no further increase is observed (Fig. 4 right panels and Table 5). This clearly indicates that less Zn^{2+} ions are bound per dimer of PS and indeed fits well with the proposal that only two Zn^{2+} ions are bound, one in each PS. In summary, it is proposed that Zn^{2+} addition to PS results in its binding to the PSs only, which restricts the swivelling movement and stabilizes the protein in a 'native-like' closed conformation, but with tighter packing due to removal of the bulky residues at the CS.

Overall, the PELDOR results obtained, both when comparing apo and Zn^{2+} -bound states of each variant forms as well as between the Zn^{2+} -bound forms of the

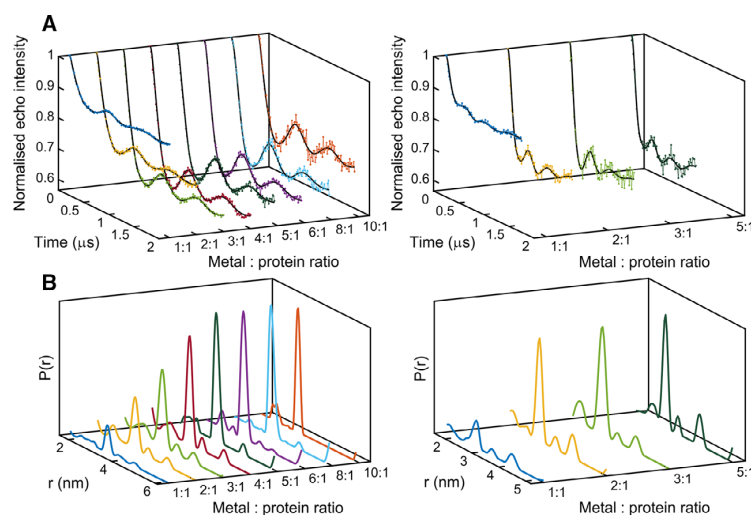


Fig. 4. Four-pulse PELDOR titrations of doubly spin-labelled MamM CTD C267S and PS-C267S constructs containing C275, measured at 60 K. (A) The normalized and background-corrected dipolar evolution of C267S (left) and PS-C267S (right) MamM CTD following Zn^{2+} titration, with metal : protein ratios ranging from 1 : 1 to 10 : 1 excess and 1 : 1 to 5 : 1 excess, respectively (coloured lines). All raw time traces were subjected to a first-order polynomial background correction. Spectral fittings are shown as black lines. (B) Distance distributions corresponding to each titre of Zn^{2+} -bound C267S (left) and PS-C267S (right). Fitting of the dipolar evolutions and subsequent distance distributions were calculated with DeerAnalysis2016 [23] using a Tikhonov regularization.

different MamM CTDs, suggest that the protein adopts a closer packing in solution upon addition of metal ions, and that the residues in both peripheral and CSs have an influence on, and are important for metal binding and conformational change.

Apo-MamM CTDs demonstrate different dimerization dynamics

In order to study the effect of mutations on apo-MamM dynamics, and to assess whether the proposed swivelling movement is possible, MD simulations were also used. As CS was shown to behave similarly to WT in their apo forms, MD was only applied to WT and PS. Later, the MD trajectories were studied using principal component analysis (PCA). PCA, also known as essential dynamics, is an analysis method that reduces the dimensionality of a complex system while retaining the most important features of the distribution of conformations [24–26].

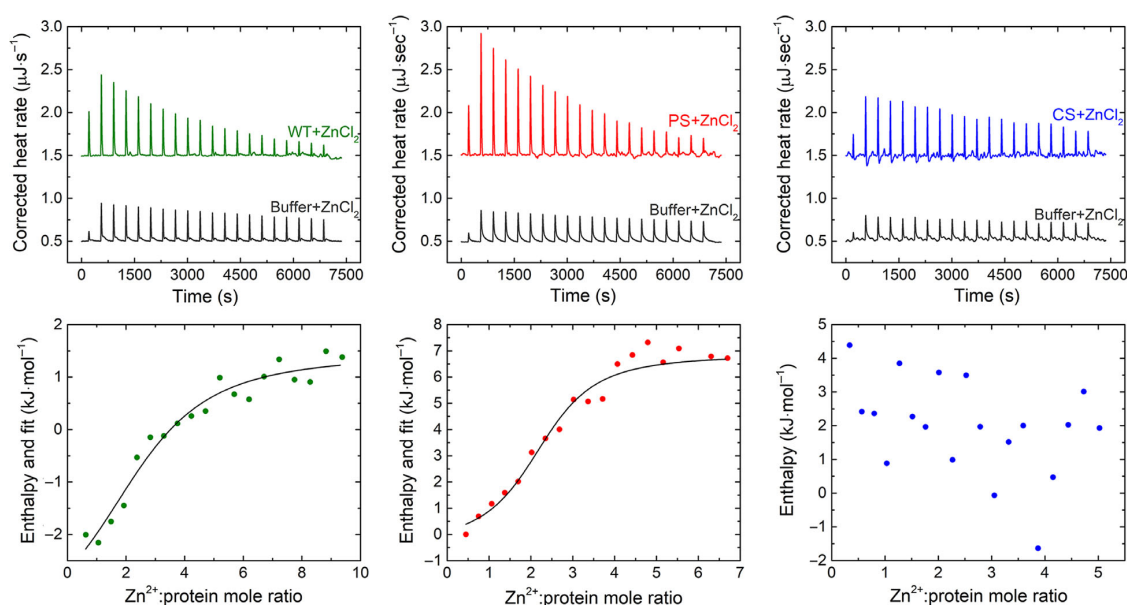
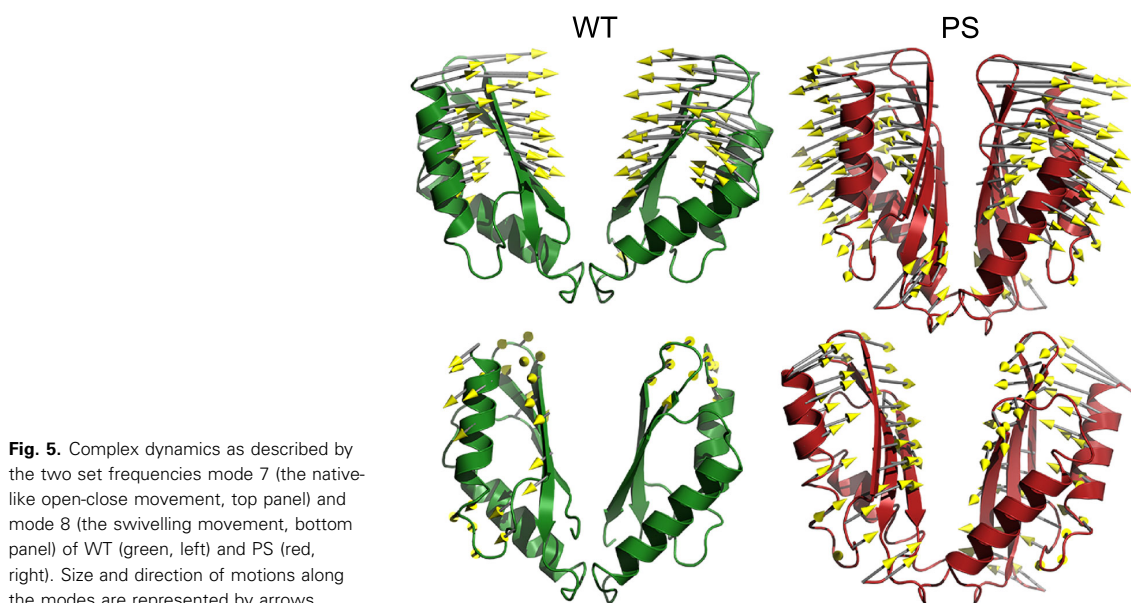
Initially, the stability of each simulated protein was validated, and the RMSD values of the backbone were calculated. As MamM CTD is a homodimer, the RMSDs for each monomer and of the dimer were calculated separately. Calculated values of backbone RMSD for the production stage (after the initial 20 ns) of each simulation are 0.13 ± 0.02 nm (monomer 1, WT), 0.11 ± 0.01 nm (monomer 2, WT), 0.31 ± 0.10 nm (dimer, WT), 0.11 ± 0.01 nm (monomer 1, PS), 0.13 ± 0.02 nm (monomer 2, PS) and 0.31 ± 0.13 nm (dimer, PS). The outcome indicates that following an initial reengagement, the 3D structure of all monomers is highly stable. On the other hand, while both WT and PS maintain their dimers, their averaged RMSDs are higher. In addition, their RMSDs show a higher rate of fluctuation (data not shown). This RMSD profile is likely to be due to the movement of one monomer towards the other.

To further study the dynamics of MamM dimers, the first two modes of motions (modes 7 and 8) calculated using PCA were used. Modes 7 and 8 are the largest collective motions of groups of atoms following the removal of translation (modes: 1–3) and rotation (mode: 4–6) movements of the WT and PS complexes. The translation and rotations are removed as they are not contributing to the molecule flexibility. First, the calculated fluctuations of modes 7 and 8 were compared with those calculated from the MD trajectories. The fluctuation profiles were found to be very similar (Pearson correlation coefficient are $R = 0.97$ and 0.96 for WT and PS, respectively), indicating that modes 7 (the native-like open-close movement) and 8 (the swivelling movement) account for most of the overall

protein motion. Furthermore, this confirms that low-frequency motions govern MamM-complex's overall dynamics. A comparison of WT and PS revealed that PS is more flexible than WT (Fig. 5). Whereas most fluctuations in the WT are located at the top of the beta-sheet, in PS a larger portion of the protein fluctuates. Moreover, the magnitude of fluctuations in PS is larger compared to the WT's. In addition, fluctuations at mode 8 of PS reveal a large amplitude of swivelling movements, unlike WT (Fig. 5, bottom panel). Taken together, the simulations suggest that the lack of the binding residues in the CS effects the apo-MamM dynamics as was also proposed by the EPR results, with the apo-PS undergoing a more pronounced dynamic swivelling movement.

Stoichiometry of Zn^{2+} binding to MamM CTD constructs

The occupancy of the three MamM CTD-binding sites during its activation in solution has not been determined previously. PELDOR results would suggest that three or four ions are bound to the WT construct while only two ions bind to the PS construct. In order to formulate a kinetic model for metal binding during activation of MamM CTD, the binding stoichiometry of divalent ions must be precisely determined for all protein constructs. For that, Zn^{2+} and the three MamM CTD constructs WT, PS and CS were utilized. Equilibrium ligand binding was performed by ITC to determine MamM CTD- Zn^{2+} binding thermodynamic parameters and stoichiometry, that is, how many metal ions are bound per MamM CTD homodimer. In Table 5, the summary of the ITC experiments with WT and the PS mutant is presented. The CS revealed no heat change on Zn^{2+} concentration-dependent binding (Fig. 6). The PELDOR experiments and fluorescence measurements (see below) did, however, indicate Zn^{2+} concentration-dependent binding. It can be speculated that Zn^{2+} binding to CS does not generate a large enough heat capacity change in the system to be detected. When considering that Zn^{2+} binding to the central binding site leads to minor conformational changes, as is evident from results obtained by the other methods employed, it is more probable that the entropy and enthalpy changes are balanced such that overall nearly no change in the Gibbs free energy of the system is observed. Both WT and PS show similar parameters (Table 5), but PS shows a slightly preferred binding to Zn^{2+} in terms of thermodynamics parameters ($\sim 57\%$ lower K_D , as well as slightly lower ΔG , which is characterized by both smaller ΔH and larger ΔS). However, WT shows a greater N value



(2.80 ± 0.24) as compared to PS (2.41 ± 0.24). This would mean that while WT can bind $\sim 3 \text{ Zn}^{2+}$ ions per dimer, in clear agreement with the PELDOR titration data, PS can only bind $\sim 2\text{--}2.5$ ions but with

slightly higher affinity, again corroborated by the PELDOR data. These ITC results suggest that each PS can bind one ion, and the CS is occupied by a single ion that has almost no impact on the heat of the

Table 5. Thermodynamic parameters of Zn^{2+} binding to WT and PS determined by ITC. Data were fitted to independent + blank (constant) models in the TA NanoAnalyze Data Analysis software. Each parameter represents three independent measurements after control reduction (Zn^{2+} at the same concentration, titrated into buffer solution). Errors are reported as the standard deviation. In all measurements, the sample in cell was in 10 mM Tris-HCl pH 8.0 and 150 mM NaCl, while the titrant Zn^{2+} was in 1 mM Tris-HCl pH = 8.0 and 150 mM NaCl.

Parameter	WT ^a	PS ^b
K_D (μM)	31.5 ± 3.4	17.9 ± 2.6
N	2.80 ± 0.24 (3–4) ^c	2.41 ± 0.24 (2) ^c
ΔH ($\text{kJ}\cdot\text{mol}^{-1}$)	-4.20 ± 0.10	-5.42 ± 2.36
K_{ass} (mM^{-1})	32.07 ± 3.22	56.78 ± 7.45
$-T\Delta S$ ($\text{kJ}\cdot\text{mol}^{-1}$)	-21.51 ± 0.29	-21.69 ± 2.34
ΔG ($\text{kJ}\cdot\text{mol}^{-1}$)	-25.71 ± 0.26	-27.11 ± 0.34
ΔS ($\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$)	72.15 ± 0.98	72.76 ± 7.84

^a $[\text{WT}]_i = 50 \mu\text{M}$, $[\text{WT}]_f = 37.45 \mu\text{M}$, $[\text{Zn}^{2+}]_i = 1.4 \text{ mM}$ & $[\text{Zn}^{2+}]_f = 351 \mu\text{M}$, ratio: 1 : 9.4

^b $[\text{PS}]_i = 50 \mu\text{M}$, $[\text{PS}]_f = 37.45 \mu\text{M}$, $[\text{Zn}^{2+}]_i = 1.0 \text{ mM}$ & $[\text{Zn}^{2+}]_f = 251 \mu\text{M}$, ratio: 1 : 6.7

^c N values from PELDOR titrations.

system. In the case of the PS construct, the mutated CS that lacks the histidine and aspartate residues cannot bind the Zn^{2+} ion effectively, but the binding to this PS construct is thermodynamically preferred, most likely due to a more stable conformational change caused by Zn^{2+} binding to the PSs, as compared to the WT protein.

Interpretation of the fluorescence signal arising from metal binding to MamM CTD

Upon analysis of the MamM CTD sequence, one tryptophan (W247) residue per monomer was identified, juxtaposed to the central binding site (Fig. 1A), and one tyrosine residue, Y241, that is remote from both the central and peripheral binding sites, but which may still contribute to any intrinsic fluorescence change upon binding of metals after excitation at λ_{ex} 280 nm. Indeed, metal binding to MamM CTD has been shown to induce a change in the intrinsic fluorescence signal upon excitation at λ_{ex} 280 nm [13]. To detect whether similar differences in the fluorescence signal occur due to Zn^{2+} binding in the different MamM CTD constructs, fluorescence measurements were conducted upon Zn^{2+} binding to MamM CTD as a function of Zn^{2+} concentration (Zn^{2+} :MamM CTD ratios were varied from 0 : 1 to 100 : 0; Fig. 7A–C). In order to be assured that the fluorescence signal observed is only contributed by the W247 residue and not by the Y241 residue, a more specific

excitation wavelength of λ_{ex} 297 nm was used which is towards the tryptophan. Indeed, a point mutation to remove this particular tryptophan residue, W247A (WT with W247A mutation), abolishes completely the fluorescence signal (Fig. 7D). The crystal structure of this mutant (Fig. 1D), which also confirms the substitution, reveals the same monomeric fold as that of the WT, and a very similar dimerization interface with only a minor shift of the monomers towards one another, probably due to the crystal packing. The tryptophan fluorescence emission spectra of the MamM CTD proteins reveal a Zn^{2+} concentration-dependent blue shift of the spectrum for all the different MamM CTD constructs, as well as an increase in fluorescence intensity. These results suggest a shift of W247 towards a more hydrophobic environment in the presence of Zn^{2+} , hinting at a Zn^{2+} -dependent conformational change. When comparing the MamM CTD WT data with those of PS (Fig. 7A,B,E,F), WT demonstrates a greater blue shift and an increase in fluorescence intensity in the presence of Zn^{2+} . This is most likely due to: (a) The alanine substitution at position 249 in PS which is juxtaposed to the W247 residue, which will influence the initial tryptophan environment hydrophobicity and lead to a lower initial λ_{max} ; and (b) a closer, or different, apo V-shaped conformer in solution, which is possible due to the lack of D249–D249 repulsion and the removal of the bulky H285 residues from the centre of the V-shape dimer, as is also proposed from both the PELDOR and MD data. In contrast, upon addition of Zn^{2+} to the CS construct, there is a large increase in intensity as compared to the other MamM CTD constructs while only a minor blue shift is observed. This would suggest that the CS construct is somehow affected differently by the presence of Zn^{2+} which is reinforced by the lack of clear detectable binding as determined by both the ITC and PELDOR spectroscopy. The Zn^{2+} concentration-dependent maximum wavelength plots were fitted to a hyperbolic equation with K_D values determined using the point of half saturation for protein- Zn^{2+} binding (WT and PS are $18.66 \pm 3.54 \mu\text{M}$ and $82.72 \pm 9.23 \mu\text{M}$, respectively). These values differ from values calculated from the ITC, but with similar order of magnitude. These differences are most likely due to the fluorescence curves reflecting only binding that is affected by the fluorescence signal and therefore should only provide a first estimate of metal affinities to MamM CTD. Importantly, the ITC values should be considered as the hard parameters for the model. Nonetheless, both methods give K_D values of the same order of magnitude although the concentrations used in the fluorescence measurements were 10 times lower

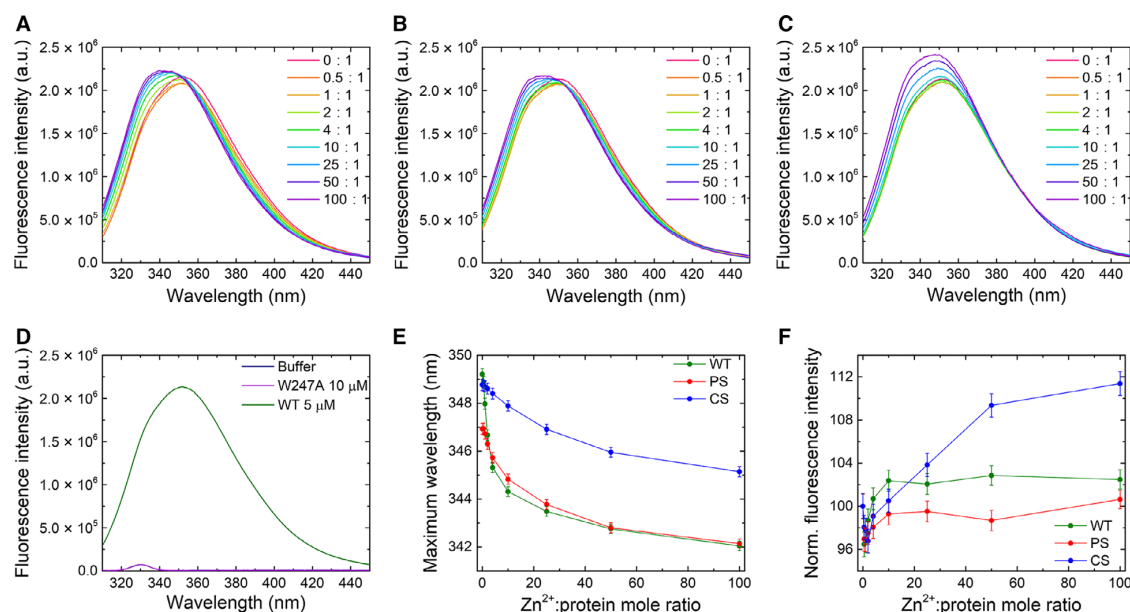


Fig. 7. Fluorescence scans of MamM CTD constructs with varying Zn^{2+} concentrations. MamM CTD proteins at 5 μM concentration were titrated using a ZnCl_2 solution to reach different MamM CTD : Zn^{2+} ratios (intensity was normalized due to the change in MamM CTD concentration) and emission spectra were recorded. Samples were measured at an excitation of λ_{ex} 297 nm and the emission spectrum for each ZnCl_2 concentration was recorded between 310 and 450 nm. For each protein, the value presented is the average of three independent measurements. (A) Emission spectra of WT with different Zn^{2+} concentrations, (B) Emission spectra of PS with different Zn^{2+} concentrations, (C) Emission spectra of CS with different Zn^{2+} concentrations, (D) Emission spectra of 5 μM WT, 10 μM D247A and buffer as a control, (E) λ_{max} as function of MamM CTD : Zn^{2+} ratio (WT in green, PS red and CS in blue), (F) Normalized fluorescence intensity compared to no Zn^{2+} as function of MamM CTD : Zn^{2+} ratio (WT in green, PS red and CS in blue). Errors are reported as the standard deviation.

than that used in the ITC, suggesting that indeed the affinity is in the micromolar range, as can be expected from protein surface-bound metals in general, and from the CDFs' CTD in particular [13,27].

Kinetics of Zn^{2+} binding to MamM CTD constructs

The kinetics of Zn^{2+} binding to MamM CTD has been measured to investigate the dynamics of metal binding and to identify any biochemical intermediates during the time course of its binding reaction. Upon Zn^{2+} binding, an increase in the intrinsic tryptophan fluorescence of both WT and PS can be nicely observed (Fig. 8A,B). However, for CS no time-dependent changes in tryptophan fluorescence are detected (Fig. 8C). Nevertheless, careful inspection of CS time traces suggests that a rapid binding event, within the dead time of the stopped-flow experiment (1.5 ms), might be occurring, due to the higher fluorescence already observed at time zero as a function of the

increasing Zn^{2+} concentration. In addition, the fluorescence changes upon Zn^{2+} binding to the CS construct are nicely shown in Fig. 7C,E,F.

The time courses of the fluorescence increase upon rapid mixing of Zn^{2+} with WT, follows a single exponential up to 500 ms. Beyond this time regime, the fluorescence intensity decreases, most likely due to relatively high photobleaching of the W247 residue in each monomer. The ITC stoichiometric model fitting predicts three Zn^{2+} ions binding per WT dimer (Table 5). Structurally, this can account for the two identical PSs and the one CS which is formed at the interface of the two monomers. Multiple binding sites which are not identical in terms of their chemical composition are less likely to yield an identical spectroscopic signature. Therefore, it was not expected that rapid binding kinetics exhibits only a single kinetic phase as shown by the mono-exponential fluorescence intensity. The observed single phase may be due to binding occurring during the stopped-flow dead time (~1.5 ms) to either sites, to a fluorescence silent

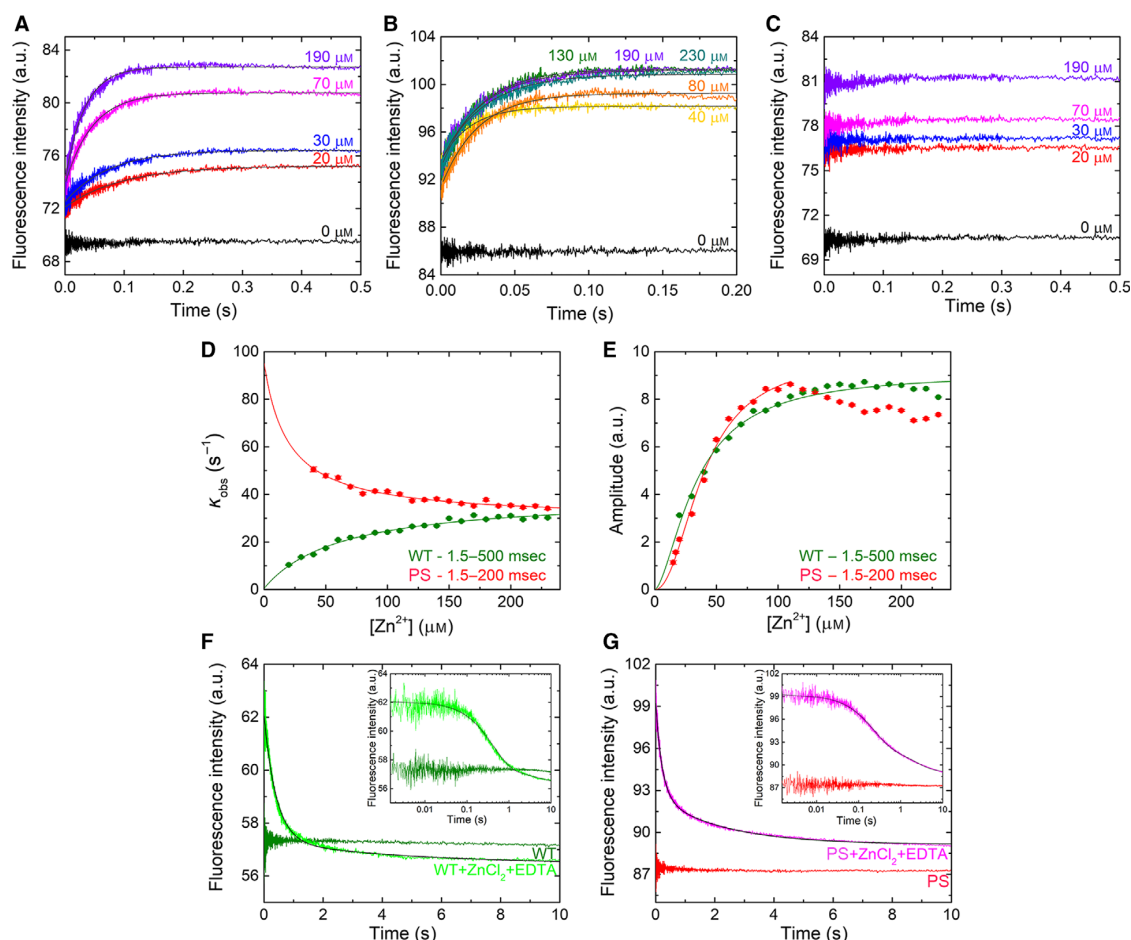


Fig. 8. Stopped-flow measurements of MamM CTD binding to Zn^{2+} . (A) Representative time courses of WT tryptophan fluorescence enhancement after a rapid mixing of WT (2.5 μM , final protein concentration) with different Zn^{2+} concentrations. Datasets are averaged transients ($n = 2$ and for each n , at least 3 repeats were also averaged). Datasets were fitted to Eqn (4) using $n = 1$. (B) Representative time courses of PS tryptophan fluorescence enhancement after a rapid mixing of WT (2.5 μM , final concentration) with different Zn^{2+} concentrations. Datasets are averaged transients ($n = 2$ and for each n , at least three repeats were averaged). Datasets were fitted to Eqn (4) using $n = 1$. (C) Representative time courses of CS tryptophan fluorescence enhancement after a rapid mixing of CS (2.5 μM , final concentration) with different Zn^{2+} concentrations. Datasets are averaged transients (using $n = 2$ while for each n , at least three repeats were averaged). (D) k_{obs} of WT (green) and PS (red) plotted as function of Zn^{2+} concentration. Datasets were fitted to a hyperbolic function (Eqns (1) and (3)). For PS, only the second phase (40–230 μM Zn^{2+}) could be fitted. Error bars are shown within the symbols. (E) Amplitude of WT and PS plotted as function of Zn^{2+} concentration. Datasets were fitted to the Hill equation (Eqn (2)) for WT and PS's first phase (15–110 μM Zn^{2+}). Error bars are shown within the symbols. (F + G) Zn^{2+} dissociation experiments for MamM CTD: Representative time courses of tryptophan fluorescence enhancement after a rapid mixing of: (F) dark green – 2.5 μM WT with buffer, light green – 2.5 μM WT + 150 μM Zn^{2+} with 2.5 mM EDTA, (G) red – 2.5 μM PS with buffer, magenta – 2.5 μM PS + 150 μM Zn^{2+} with 2.5 mM EDTA (final concentrations). Datasets are averaged transients (with $n = 1$, at least 4 repeats were averaged). Datasets were fitted using Eqn (4) with $n = 2$. Insets: the same data presented on a logarithmic scale.

binding site or to a combination of both. Based on the WT and the mutants' PELDOR, fluorescence and ITC data, we favour that the measured fluorescence increase in the transient time course reflects the binding of Zn^{2+} to the PSs according to the general

Scheme 1. The initial binding of Zn^{2+} to the CS is rapid, occurring within the dead time of the stopped-flow with a binding rate $> 1000 \text{ s}^{-1}$ (shown in squared parenthesis: CTD^{O} and CTD^{C} are open and the slightly more closed conformers of CTD, respectively;

complexes without a dot represent a strongly bound state). This is followed by a rapid collision complex formation of two Zn^{2+} ions binding to the PSs followed by an additional isomerization reflected by formation of a high fluorescence state termed 'compact closed' (CTD^{CC} , Fig. 9). The microscopic rate constants are global binding parameters that reflect binding to both PSs since we cannot distinguish between the two (Scheme 1). The observed rate constant (k_{obs}) of Zn^{2+} binding to WT exhibits a hyperbolic dependence on Zn^{2+} concentration (Fig. 8D). Based on the kinetic Scheme 1, and under the assumption that $k_{+p}[\text{Zn}^{2+}], k_{-ps} \gg k_{+ps}^{\text{CC}} \sim (K_{\text{PS}} = k_{-ps}/k_{+ps})$, the observed rate constant for Zn^{2+} binding is predicted to follow a rectangular hyperbola and can be fitted accordingly to Eqn (1):

$$k_{\text{obs}} = \left(\frac{K_{\text{PS}} \cdot k_{+ps}^{\text{CC}} [\text{Zn}^{2+}]}{1 + K_{\text{PS}} [\text{Zn}^{2+}]} \right) + k_{-ps}^{\text{CC}} \quad (1)$$

where K_{PS} represents the equilibrium constant for Zn^{2+} binding to the PSs and k_{+ps}^{CC} reflects the isomerization rate constant suggesting further closure of the CTD central cleft. The best fit of Zn^{2+} concentration dependence binding to the WT according to Eqn (1) yields an equilibrium constant $1/K_{\text{PS}} = 62.6 \pm 13.0 \mu\text{M}$ and an isomerization rate constant of $k_{+ps}^{\text{CC}} =$

$39.2 \pm 1.4 \text{ s}^{-1}$ (Table 6). The intercept of the fit is indistinguishable from the origin; thus, Zn^{2+} binding is essentially irreversible ($k_{+ps}^{\text{CC}} \gg k_{-ps}^{\text{CC}} \sim 0$), consistent with the results of slow Zn^{2+} dissociation kinetics (see below, Fig. 8F). The apparent second order binding constant $1/K_{\text{PS}} \cdot k_{+ps}^{\text{CC}} = 0.63 \pm 0.14 \mu\text{M}^{-1} \cdot \text{s}^{-1}$, is surprisingly slow in comparison to a small ligand such as a charged ion binding to a protein [5], suggesting that the reaction is not diffusion-limited. Interestingly, dissociation kinetics of Zn^{2+} measured by mixing a pre-equilibrated $\text{WT} \cdot \text{Zn}^{2+}$ with excess of EDTA exhibited double exponentials (Fig. 8F). The dissociation rate constants, $k_{\text{diss,fast}} = 2.4 \pm 0.1 \text{ s}^{-1}$ and $k_{\text{diss,slow}} = 0.20 \pm 0.07 \text{ s}^{-1}$ may reflect dissociation events from the two different Zn^{2+} binding sites, the PSs and the CS. However, the fact that similar dissociation kinetics are seen with PS (Fig. 8G) and with similar magnitudes to the WT, $k_{\text{diss,fast}} = 4.6 \pm 0.1 \text{ s}^{-1}$ and $k_{\text{diss,slow}} = 0.46 \pm 0.02 \text{ s}^{-1}$, that was determined to not bind Zn^{2+} in the CS by the PELDOR and ITC measurements, would suggest otherwise. Furthermore, the faster dissociation rates suggest that PS binds Zn^{2+} with a different conformation at the PSs. The intercept in Fig. 8D is similar to the slower dissociation rate measured directly ($k_{-ps}^{\text{CC}} \sim 0$ & $k_{\text{diss,slow}} = 0.20 \pm 0.07 \text{ s}^{-1}$, see Table 6) favouring a slower isomerization occurring before complete dissociation.

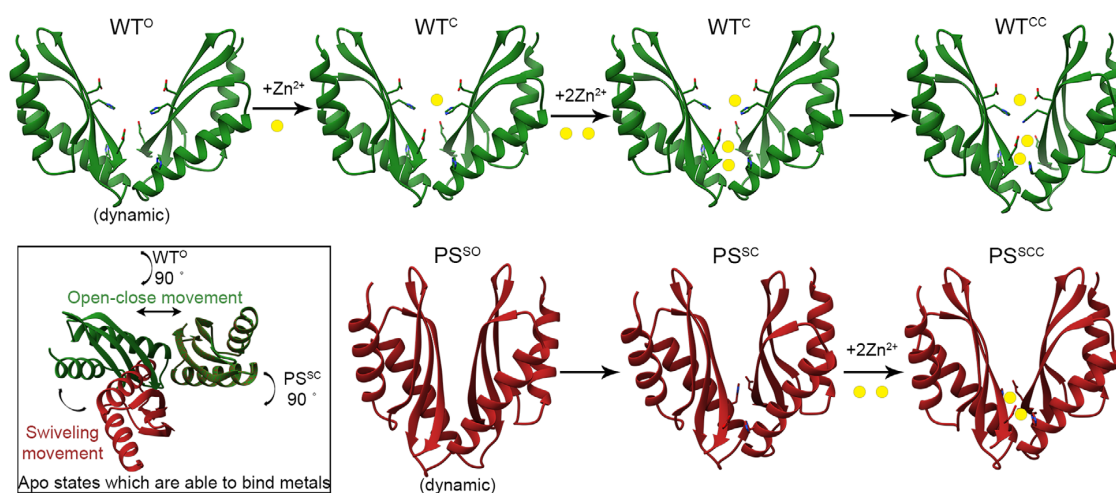
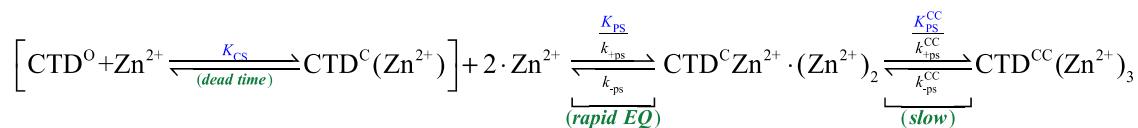


Fig. 9. A binding mechanism model for WT and PS. Top row, green- A three-step binding mechanism for WT as described in Scheme 1: first, a rapid step in which one Zn^{2+} binds to the CS, followed by fast Zn^{2+} binding to the PSs and a slower isomerization step to achieve the final, three-ion bound-closed state. Bottom row, red- A two-step binding mechanism for PS as described in Scheme 2: a slow step in which an isomerization between different apo swivelling forms occurs, followed by fast Zn^{2+} binding to the PSs that is only allowed by one apo form, and packing of the dimer to a similar, but yet tighter conformation than that of the WT. Lower left box; An overlapped representation of the WT^{O} and PS^{SC} forms to demonstrate the difference between the swivelling movement and the standard CTD's closure movement.

**Scheme 1.** Kinetic reaction mechanism of Zn^{2+} binding to WT.

Both WT and PS dissociation kinetics demonstrate that structurally the PSs exchange Zn^{2+} with EDTA which can compete over the coordination of Zn^{2+} at these sites. Along with these arguments, we favour a two-step dissociation kinetic from the PSs.

The amplitudes of the PSs' binding kinetics are related to the fractional bound Zn^{2+} , and therefore to their overall affinity. Since these amplitudes report binding of two ligands to two identical sites, we utilized a simple Hill model. The fitting of the amplitudes for the WT according to Eqn (2) is shown in Fig. 8E with a $K_{\text{H,WT}} = 33.9 \pm 1.3 \mu\text{M}$ and a moderate Hill coefficient $n_{\text{H,WT}} = 1.61 \pm 0.13$ (Table 6).

$$A = \frac{V_{\text{max}} \cdot [\text{Zn}^{2+}]^n}{K_{\text{H}} + [\text{Zn}^{2+}]^n} \quad (2)$$

Comparison of the equilibrium constant measured by ITC for WT ($K_{\text{d}} = 31.5 \pm 3.4 \mu\text{M}$, Table 5) with the equilibrium constant determined by the Hill model suggests that the overall affinities of both the CS and PSs in the WT protein are only marginally different, though their kinetics seems to be on a very different time scale. This further agrees with the rapid dissociation from the CS as the 'on rate' is also very fast. This is not to classify them as similar in terms of their binding kinetic pathways and the impact of their binding on the allostery they may induce on each other.

Although PS displays a similar single-phase behaviour (Fig. 8B), with fluorescence increase upon Zn^{2+} binding, the k_{obs} dependence on Zn^{2+} concentration decreases rather than increasing (Fig. 8D, Table 6). Therefore, it may be that PS exhibits a different mechanism for Zn^{2+} PS binding than WT. In PS, the fast binding step to the CS does not occur in the dead time as it does for WT. This is also evidenced by the PELDOR and ITC results which reveal a lower stoichiometry of Zn^{2+} binding to PS (Table 5, $N = 2$ and $N = 2.41 \pm 0.24$ respectively), most likely only to the PSs. This behaviour is well-matched with a slower step of 'close' to 'open' conformational change preceding the metal-ligand binding step to a 'selected conformation' (open state) [28]. The presence of an additional isomerization following the fast ligand binding step $\text{CTD}_{\text{CSM}}^{\text{SC}} \cdot (\text{Zn}^{2+})_2$ is not justified based on this hyperbolic Zn^{2+} concentration dependence of the k_{obs} (see

below), as this may occur too fast to be detected. When we consider the dissociation kinetics of PS, the presence of an isomerization step $\text{CTD}_{\text{CSM}}^{\text{SC}}(\text{Zn}^{2+})_2$ as observed for the WT before the peripheral Zn^{2+} completely dissociated may be justified. This two-step dissociation most likely originates from a tighter complex, similar to the WT but albeit different structurally. Therefore, we propose the reaction mechanism in Scheme 2 for Zn^{2+} binding to PS stringently based on the observable 'on' rate. $\text{CTD}_{\text{CSM}}^{\text{SO}}$ and $\text{CTD}_{\text{CSM}}^{\text{SC}}$ are the 'swivel open' and 'swivel closed' conformers of the PS CTD, respectively where $\text{CSM}_{\text{CSM}}^{\text{SCC}}$ is the 'swivel compact closed' state (CSM for CS mutant represents the PS construct, so as to avoid confusion with the PS used to represent the PSs in all the kinetic schemes; Fig. 9). At lower Zn^{2+} concentration, the binding of Zn^{2+} is limited by the presence of $\text{CTD}_{\text{CSM}}^{\text{SC}}$ and hence the transition from $\text{CTD}_{\text{CSM}}^{\text{SO}}$ to $\text{CTD}_{\text{CSM}}^{\text{SC}}$ contributes to k_{obs} . At very high Zn^{2+} concentration, the formation of $\text{CTD}_{\text{CSM}}^{\text{SC}} \cdot (\text{Zn}^{2+})_2$ is fast and k_{obs} is dominated by $k_{\text{+csm}}^{\text{S}}$ ($K_{\text{CSM}}^{\text{S}} = k_{\text{+csm}}^{\text{S}}/k_{\text{-csm}}^{\text{S}}$) and will approach the net rate constant of this isomerization step. We further reason that the binding steps are faster than isomerization based on the relatively slow dissociation kinetics and the high affinity binding of Zn^{2+} to the mutant (ITC results, Table 5), hence at high Zn^{2+} concentration, $k_{\text{+csm}}^{\text{S}} \ll k_{\text{+csm}}^{\text{PS}}[\text{Zn}^{2+}]$ ($K_{\text{PS}} = k_{\text{-csm}}^{\text{PS}}/k_{\text{+csm}}^{\text{PS}}$). Therefore, in accordance with such a mechanism, Eqn (3)

Table 6. Kinetic parameters for Zn^{2+} binding to MamM CTD WT and PS. Left parameters relate to WT, right parameters relate to PS. Conditions were as follows: 10 mM Tris (pH 8.0) and 150 mM NaCl at 25 °C.

Parameter	WT	PS
$k_{\text{+ps}}^{\text{CC}}/k_{\text{-ps}}^{\text{S}}$ (s^{-1})	39.2 ± 1.4	29.4 ± 1.6
$K_{\text{PS}}(1/K_{\text{PS}})$ (μM)	62.6 ± 13.0	19.4 ± 17.2
$K_{\text{PS}}k_{\text{+ps}}^{\text{CC}}$ ($\mu\text{M}\cdot\text{s}^{-1}$) ^a	0.63 ± 0.14	–
$k_{\text{-ps}}^{\text{CC}}/k_{\text{-csm}}^{\text{S}} + k_{\text{+csm}}^{\text{S}}$ (s^{-1})	0.52 ± 2.32	95.1 ± 40.4
$k_{\text{-csm}}^{\text{S}}$ (s^{-1})	–	65.7 ± 38.8
$K_{\text{CSM}}^{\text{S}}$	–	0.45 ± 0.27
$k_{\text{diss,fast}}$ (s^{-1})	2.4 ± 0.1	4.6 ± 0.1
$k_{\text{diss,slow}}$ (s^{-1})	0.20 ± 0.07	0.46 ± 0.02
K_{H} (μM) ^b	33.9 ± 1.3	39.5 ± 2.2
n_{H} ^b	1.61 ± 0.13	2.1 ± 0.2

^a Relates to WT.^b Determined from the total amplitudes.

that although exhibiting a different distance range as compared to the apo form, the CS does not adopt a closed, distinct conformation when Zn^{2+} is added. In CS, it is more probable that entropy decreases due to the more ordered zinc ions bound at the CS. Furthermore, enthalpy is also slightly decreased, in a way that leads to nearly no net overall change in the Gibbs free energy of the system. This is further supported by the WT and PS ITC results which indicate that in equilibrium their energetic parameters are very similar. As their final states are expected to be similar, binding to the CS does not contribute much to the overall heat change of the system. In terms of structural changes, one Zn^{2+} binds initially to the CS residues of only one monomer (nonspecifically), which occurs very fast (< 1.5 ms, as resolved from the WT kinetic results). The binding of one such ion could potentially change the local electrostatic surface, driving the second monomer to close the binding site, thus neutralizing the positive charge. Our results, however, indicate that binding at the CS only leads to minor conformational changes and the protein remains in a dynamic, open state. This is only altered when the PSs are occupied, leading to the tighter, locked state (Fig. 9), as evidenced from the PS and WT fluorescence and PELDOR results.

PS exhibits a different initial conformation and $[\text{Zn}^{2+}]$ -dependent movement as compared to WT

The N value derived from ITC experiments of the WT and the PS mutants reflects the number of bound Zn^{2+} ions. These N values that were calculated from the ITC data indicate that the PSs bind two ions ($N_{\text{PS}} = 2.4$), hence the CS populates only one Zn^{2+} ion ($N_{\text{WT}} = 2.8$, N_{CS} could not be calculated), which is lower than previous published results ($N_{\text{WT}} = 4$) [13]. PS exhibits a slightly lower N than WT; however, it has a twofold tighter K_{D} than WT (lower ΔG , suggesting that the binding occupies less sites but with a tighter affinity. This can also be inferred from the PELDOR results where both a less dynamic conformation and/or a shorter smaller distance distribution for the apo PS is observed as compared to the apo WT, as well as a reduction in the number of bound Zn^{2+} ions. Structurally, this can be reasoned by a swivelling movement of the PS dimer (relative to the dimerization interface at the bottom of the V-shape; Fig. 9). Hence, and as was supported by the MD results, PS adopts a scissor-like conformation instead WT's V-shape apo conformation. This would suggest that the initial states of the two apo forms are different and hence their binding mechanism is different. In the final state, both

WT and PS involve the binding of Zn^{2+} in the two PSs and adopt a more similar conformation as compared to their initial states, which is supported by both the fluorescence and PELDOR results. These different initial states are not reflected in crystal structure models of the apo forms which are likely restrained by crystal contacts. Further, the only apo crystal structure which could not be detected is that of PS-C267S, further supporting that it might adopt a different conformation.

Comparison of MamM CTD metal-binding sites and conformational changes to other CDF proteins' CTDs suggest a complex regulation mechanism in CDF proteins

A comparison of our results in the wider context of the whole CDF protein family in terms of the binding sites and conformational changes observed in this study reveal further significant structural insights within the known CDFs' CTDs-bound structures (Fig. 10A). The evolutionary-closest CDF CTD structure that was solved in a Zn^{2+} -bound state is of MamB, another iron transporter CDF protein from MTB [14] (Fig. 10B). Zn^{2+} in the MamB structure is chelated by D247 (homologous to D249 in MamM), H283 (homologous to H285 in MamM) and H245 (found in the same location as W247 in MamM; Fig. 10C). Since MamM and MamB CTD sequences are highly conserved [14,17], MamB's crystal structure strengthens the evidence for the ability of MamM's CS to bind metals. However, in MamB, no conformational changes have yet been detected in the metal-bound state compared to the apo structure (Fig. 10B). This could be due to two reasons: (a) Zn^{2+} was only soaked into the crystals after their formation [14], a method that may not invoke significant further conformational changes due to pre-existing strong crystal contacts. (b) The MamB structure does not involve Zn^{2+} binding at a PS (which contains aspartate and threonine residues rather than the histidine and glutamate residues found in MamM), which has been shown here to induce such major conformational changes.

All other CDF structures that were solved are of bacterial Zn^{2+} transporters (Fig. 10D). CzcB is the only CDF protein whose CTD structure has been solved in both the apo and bound-closed forms. CzcB structures exhibit much larger conformational change between the apo and the bound-closed states [12] when compared to the degree-of-closure indicated by the PELDOR results for MamM. Moreover, the metal-binding sites are not completely conserved between MamM and CzcB. The PS of MamM involves H264

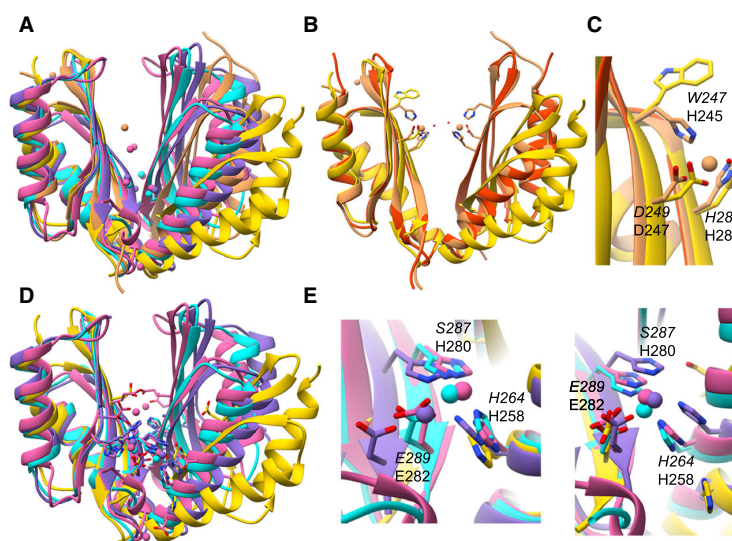


Fig. 10. CDF protein CTDs conserved binding sites. (A) The dimeric CTD crystal structures of Zn^{2+} -bound MamB (light orange, PDB ID: 5HO1 [14]), Zn^{2+} -bound CzrB (pink, PDB ID: 3BYR [12]), Zn^{2+} -bound *Escherichia coli* YiiP (blue, PDB ID: 3H90 [9]) and Zn^{2+} -bound *Shewanella oneidensis* (strain MR-1) YiiP (purple, PDB ID: 5VRF [10]) overlapped onto the dimeric apo-MamM CTD WT structure (yellow, PDB ID: 3W5X [13]). For all the bound structures, the Zn^{2+} ions are indicated in the structure colour. (B) The dimeric crystal structures of Zn^{2+} -bound MamB (light orange) and apo-MamB CTD (dark orange, PDB ID: 5HO5 [14]) overlapped onto the dimeric apo-MamM CTD WT structure (yellow). MamB's Zn^{2+} ions, the Zn^{2+} ions chelating water molecules, MamB's central binding site residues and their homologous residues from MamM are indicated as sticks. (C) Magnification of the monomeric central binding site of MamB and MamM. The residues are labelled, so that the upper labels in italics refer to MamM's residues and the lower labels to MamB's residues. The D249 and H285 residues in MamM CTD central binding site are conserved in MamB, where they were shown to participate directly in the chelation of the metal ion. (D) The dimeric CTD crystal structures of Zn^{2+} -bound CzrB (pink), Zn^{2+} -bound *Escherichia coli* YiiP (blue) and Zn^{2+} -bound *S. oneidensis* YiiP (purple) overlapped onto the dimeric apo-MamM CTD WT structure (yellow). All proteins' bound Zn^{2+} ions, all the Zn^{2+} ions binding site residues and their homologous residues from MamM are indicated as sticks. (E) Magnifications of the peripheral binding site of CzrB, YiiP proteins and MamM (only this site is shown for clarity). The residues are labelled, so that the upper labels in italics refer to MamM's residues and the lower labels to CzrB's residues. The H264 (left panel) and E289 (right panel) residues in MamM CTD peripheral binding site are conserved in YiiP and CzrB proteins, where they were shown to participate directly in the chelation of the Zn^{2+} ion.

from one monomer and E289 from the other. In CzrB, one of the Zn^{2+} ions is chelated by homologous histidine and glutamate residues that are found in the same positions (H258 and E282), but also by an additional histidine (H280, homologous to S287 in MamM; Fig. 10E). Additionally, there are two Zn^{2+} ions bound per monomer in CzrB, that do not involve homologous residues to those of MamM-binding sites and which involve residues only from one monomer (E243, homologous to D249 in MamM, is found at a distance of 3.71 Å from one of these ions, a long distance for an expected strong electrostatic interaction, therefore it is not expected to contribute much to the binding of the ion in this site). However, one of these Zn^{2+} ions (namely the second Zn^{2+} ion) is also bound by E282 and by other residues that are involved in the binding of the third Zn^{2+} [12] (Fig. 10D), suggesting a net of Zn^{2+} -protein interactions that can lead to the

stable binding of Zn^{2+} in the CzrB PS. The first two Zn^{2+} ions discussed in the context of CzrB are also found in the same locations and are ligated by homologous residues in the bacterial CDF model protein YiiP, for which only its closed structure has been solved [9,10,30,31] (Fig. 10D,E). Such a multiple- Zn^{2+} -binding state is not possible in MamM as some of the homologous residues in MamM are not suited to metal binding, yet despite this both the CzrB and YiiP closed structures support the importance of the PS residues for metal binding and conformational changes, as revealed in this study.

The comparison between all CDFs CTDs' structures and binding sites reveals a complex regulatory mechanism for CDF proteins. The Zn^{2+} transporters show a network of Zn^{2+} -binding sites that includes the PSs and extends towards the centre of the proteins, while the iron transporter MamB binds ion only in the

centre of the V-shaped dimer. MamM contains both the central and PSs, suggesting a double-safe mechanism for this specific protein compared to MamB, and a different regulation mechanism compared to the Zn^{2+} transporters: one site allows rapid binding and a change of the charge distribution, which facilitates the binding to the second site. This allows the conformational lock-down that will lead to further transmembrane domain conformational changes and active transport. Previous *in vivo* studies of MamM showed that mutations in both the central and peripheral CTD-binding sites lead to small variations in magnetosome size, number and magnetic response [13]. These rather small changes might suggest that although the regulation is important for normal function, the transport can occur even when this regulation path is defected, meaning that metal can diffuse to the transmembrane domain independently, yet less efficiently. However, even mutations in the conserved transmembrane transport site of MamM, which were shown previously in many CDF proteins to have a drastic influence on the proteins' function, showed relatively small impact on magnetite formation [8,17]. Hence, we speculate that MamB somehow compensates MamM function as opposed to mutations in the CTD-binding sites having only a minor impact on the protein function.

All the known CDF proteins' CTD structures support our key finding of the capability of the CS to bind metal readily without any further large conformational changes. In addition, binding to the PSs affords the induction of conformational changes in the dimer. While the transmembrane transport site is conserved between the different CDF proteins (in terms of location and importance for function) which implies a similar transport mechanism [8,11], the apparent versatility in the binding sites and conformations between the proteins CTDs suggest diverse regulation mechanisms for different proteins. The fine tuning in the regulation might be evolution-related, since more highly evolved organisms emerge more sophisticated and complex regulation mechanisms. For example, recent studies of eukaryotic CDF proteins (plant MTP proteins and human ZnT proteins) suggest that other elements such as a His-rich loop – a cytoplasmatic loop between two transmembrane helices that is not found in the structurally characterized CDF proteins – might play a role in protein regulation [32,33]. Further biochemical and structural studies of CDF proteins that are not Zn^{2+} or Fe^{2+} transporters proteins, and of eukaryotic CDF proteins, would be required to decipher the evolutionary pathway of the regulation mechanism in CDF proteins.

To conclude, the kinetics of the transport rate of the full length CDF protein has been previously studied by the detection of change in Zn^{2+} concentration inside a vesicle. Here for the first time, the conformational dynamics and kinetics of the CTD upon Zn^{2+} binding have been directly studied, revealing a molecular mechanism for binding, including stoichiometry, allostery, energetics and the order of binding. The structure–function relationships revealed here demonstrate how metal binding to a CTD of a CDF protein leads to distinct conformational changes within this domain in solution, thus significantly strengthening previous evidence highlighting the importance of CTD for overall protein regulation and function.

Materials and methods

Site-directed mutagenesis and protein expression

mamM CTD gene (UniProt Q6NE57 residues: 215–318) was previously cloned into pET28a(+) vector (Novagen, Merck Biosciences, Darmstadt, Germany) [34]; in this construct, pET28a-MamM-CTD-MSR1, the *mamM* gene was fused in-frame to express a six-His tag at the N terminus of the protein followed by a thrombin proteolysis site. All MamM CTD mutations were applied to the pET28a-MamM-CTD-MSR1 vector using the QuickChange site-directed mutagenesis method (Stratagene, La Jolla, CA, USA). Primers containing single mutation sites (Hylabs, Rehovot, Israel) were designed and used for PCR amplifications. All MamM CTD forms were expressed similarly to previously described for MamM CTD WT [34]: *Escherichia coli* Rosetta(DE3) pLysS strain cells (Novagen) harbouring the pET28a-MamM-CTD-MSR1 plasmid were grown in auto-induction medium [35] containing kanamycin ($100 \text{ mg}\cdot\text{mL}^{-1}$) and chloramphenicol ($30 \text{ mg}\cdot\text{mL}^{-1}$) at 310 K for 6 h. The cultivation temperature was then reduced from 310 to 300 K for a further 18 h. The cells were harvested by centrifugation at $7438 g$ for 8 min at 277 K.

Protein purification

All proteins used in this study were purified similarly to previously described for MamM CTD M250L mutant [19,34]. All MamM CTD forms expressing cells were suspended in binding buffer containing 50 mM Tris-HCl pH = 8, 300 mM NaCl, 20 mM imidazole, 5 mM β -mercaptoethanol and 0.01% Triton X-100 at a weight ratio of 1 : 2, with DNaseI ($10 \text{ }\mu\text{g}\cdot\text{mL}^{-1}$) and a protease inhibitor cocktail (containing 100 μM PMSF, $1.2 \text{ }\mu\text{g}\cdot\text{mL}^{-1}$ leupeptin and 1 μM pepstatin A) for 30 min at 277 K. The cells were then disrupted by three cycles of French press pressure cell (Thermo Scientific, Waltham, MA, USA) at

207 MPa. Cell debris was separated by centrifugation at 203 000 *g* (ECF) (60 Ti fixed angle rotor; Beckman Coulter, Brea, CA, USA) for 45 min at 277 K and the soluble fraction was applied onto a home-made gravity Ni-NTA column (2.5 cm diameter Econo-Column by Bio-Rad, Hercules, CA, USA; 5 mL bed volume of HisPur Ni-NTA resin by Thermo Scientific) pre-equilibrated with the binding buffer. Protein were washed with 50 mL of two washing buffers for further purification, containing: (a) 20 mM Tris-HCl pH = 8, 1 M NaCl, 40 mM imidazole and 5 mM β -mercaptoethanol; (b) 20 mM Tris-HCl pH = 8, 150 mM NaCl, 40 mM imidazole and 5 mM β -mercaptoethanol, and eluted with buffer containing 20 mM Tris-HCl pH = 8, 150 mM NaCl, 500 mM imidazole and 5 mM β -mercaptoethanol. To remove the six-His tag, bovine thrombin (10 U·mL⁻¹; t4648-10KU; Sigma-Aldrich, St. Louis, MO, USA) was added to the eluted protein and the mixture dialysed against the final buffer, containing 10 mM Tris-HCl pH = 8, 150 mM NaCl and 5 mM β -mercaptoethanol for 16 h at 277 K. Protein was then concentrated to a volume of 4 mL using a Vivaspin-4 (3000 Da MW cutoff; Sartorius Stedim Biotech GmbH, Goettingen, Germany) and applied onto a size-exclusion column (HiLoad 26/60 Superdex 75; GE Healthcare Biosciences, Danyel Biotech Ltd., Rehovot, Israel) pre-equilibrated with the final buffer. Selected peaks of pure protein were collected and concentrated using a Vivaspin-4 to a final concentration of 10–60 mg·mL⁻¹, flash-frozen in liquid nitrogen and stored at 193 K for further use. Throughout the purification process, protein concentration was determined by measuring absorption at 280 nm, and protein purity was analysed by SDS/PAGE. For all C267S constructs, β -mercaptoethanol was not used during purification and in the final storage solutions. For all further experiments, protein concentration was also determined by measuring absorption at 280 nm.

Crystallization and structure determination

Purified MamM CTD proteins: W247A in buffer containing 10 mM Tris pH = 8.0, 150 mM NaCl, and 5 mM β -mercaptoethanol; CS in buffer containing 10 mM Tris pH = 8.0, 150 mM NaCl, 3.375 mM MnCl₂ and 5 mM β -mercaptoethanol; CS-C267S and C267S in buffer containing 10 mM Tris pH = 8.0, 150 mM NaCl (buffer A); were crystallized using the vapour diffusion method at 293 K (0.3 μ L protein with a 0.3 μ L reservoir solution for all proteins but CS, where 1 μ L protein with 1 μ L reservoir solution were used). Crystals were harvested with or without treatment of cryo agent and flash-frozen in liquid nitrogen. Data collection was performed on a single-crystal at a temperature of 100 K. All cryo conditions, data collection details, data reduction and scaling, phasing and refinement details are given in Tables 1–3.

Least-squares overlaps

MamM CTD structures were overlaid and RMSDs calculated using the iterative magic fit tool of SWISS-PDBVIEWER 4.1.0 [36]. All MamM CTD structures and all overlapped structure figures were prepared using UCSF CHIMERA package, version 1.12 [37].

Circular dichroism measurements

CD measurements were performed using a J750 spectropolarimeter (Jasco Inc, Easton, MD, USA) at ambient temperature. MamM CTD proteins were incubated with EDTA for 5 h, which was subsequently removed by dialysis. The proteins were diluted to a final concentration of 0.3 mg·mL⁻¹, with or without Zn²⁺ (15 : 1 Zn²⁺ to protein ratio), in a buffer containing 10 mM Tris-HCl pH = 8, 50 mM NaCl and 0.188 mM β -mercaptoethanol, and measured with a 0.1 cm optical path Suprasil quartz cuvette (Hellma GmbH & Co., Müllheim, Germany). Spectra profiles of the samples were measured in a wavelength range of 200–260 nm, with a bandwidth of 1 nm, a scan speed of 5 nm·min⁻¹ and a time constant of 8 s.

Site-directed spin labelling

For site-directed spin labelling, protein solutions in buffer A (containing 10 mM Tris-HCl pH = 8 and 150 mM NaCl) were degassed under argon and treated with degassed 1,4-dithiothreitol (1 mM) for 4 h at 4 °C with gentle agitation. Subsequently, protein samples were washed once using ZebaTM Spin desalting columns (7 kDa MWCO, 2 mL; Thermo Scientific, Rockford, IL, USA) and degassed buffer A. Protein was then labelled overnight at 4 °C with (1-oxyl-2,2,5,5-tetramethyl- Δ 3-pyrroline-3-methyl) methanethiosulfonate spin label (MTSSL, 20 $\times$ excess; Toronto Research Chemicals Inc., North York, ON, Canada) under gentle agitation. To remove excess unbound MTSSL, samples were washed twice using ZebaTM Spin desalting columns and degassed buffer A.

Pulsed electron–electron double resonance measurements

Spin-labelled variants were lyophilized before being taken up in D₂O to a final concentration of 150–200 μ M, and EPR samples (both apo and Zn²⁺-containing) were prepared with 30% Glycerol-d₈ (Sigma-Aldrich). For Zn²⁺-containing samples, ZnCl₂ stock solutions were prepared in buffer A (in D₂O). Each spin-labelled variant C267S, PS-C267S and CS-C267S was incubated with a five-fold excess of Zn²⁺ for 1 h at room temperature, then Glycerol-d₈ was added. A 100 μ L sample was loaded into an EPR quartz tube (Wilmad SQ-707; Wilmad-LabGlass, Vinland, NJ, USA) and flash-frozen in liquid nitrogen. For the titration

experiment of the C267S variant, the previously measured apo sample was first thawed slowly over liquid-N₂ and removed from its EPR tube to an eppendorf. Zn²⁺ was then added (1× excess) and the sample was allowed to incubate at room temperature for ~ 2 h before transferring back into the tube, flash freezing and re-measuring. This was repeated for Zn²⁺ excesses of 2×, 3×, 4×, 5×, 6×, 8×, and 10× using the same sample. The same experimental approach was taken for the titration experiment of the PS-C267S variant, this time with Zn²⁺ excesses of 1×, 2×, 3× and 5×. X-band pulsed EPR spectra were recorded on a Bruker E580 spectrometer (Bruker, Rheinstetten, Germany) using a Bruker MD5-W1 EPR probehead equipped with a self-modified cryogen-free cryostat (Advanced Research Systems Inc, Macungie, PA, USA). The microwave pulses were amplified using a 1 kW-TWT (Applied Systems Engineering Inc., Fort Worth, TX, USA). All EPR experiments were carried out at 60 K. The field-swept spectrum was obtained by integrating the Hahn echo signal as a function of the magnetic field after a two-pulse sequence.

For the four-pulse PELDOR experiments recorded at 60 K, pulse lengths were 20 ns for $\pi/2$ and 20 ns for π . The pump pulse length was 16 ns and $\Delta\nu$ ($\nu_{\text{obs}} - \nu_{\text{pump}}$) was 70 MHz. The pulse separation τ_1 was 352 ns (C267S apo), 386 ns (CS-C267S apo) or 120 ns (all others), while τ_2 varied between 2000 and 4000 ns, and the echo signals were integrated using a video amplifier bandwidth of 20 MHz. The pump pulse was stepped out by 16 ns for a total number of points in T that depended upon the τ_2 used. Nuclear modulation artefacts were suppressed using a systematic variation of the interpulse delay time τ_1 and an appropriate phase cycling routine.

Molecular dynamics and principal component analysis

All MD simulations were performed using the NAMD MD software version 2.13b2 [38]. The CHARMM 36/CMAP force field [39] with TIP3P water model [40] was used to describe the system.

During simulations, a time step of 2 fs was used. In all simulations, periodic boundary conditions with the particle mesh Ewald algorithm for electrostatics were employed [41]. All simulations were performed under the same conditions: temperature was held constant using Langevin dynamics at 300 K with a damping coefficient, 1 ps⁻¹; pressure was held constant using Langevin piston method at 1 atm also with a damping coefficient, 1 ps⁻¹, piston's oscillation period to 100 fs and pistons decay time of 50 fs; switching and cutoff distances, 1 nm and 1.2 nm, respectively; pairlist distance, 1.4 nm.

System preparation was done using the VMD software package version 1.93 [42]. Initial models of WT and PS were based on the PDB IDs: 3W5X and 3W66, respectively [13]. All simulated proteins were solvated into a

water box with a minimum solvation shell of 14 Å thickness. Simulated systems were then neutralized by adding counter ions at random positions. Before initiating free MD simulations, all systems were minimized and equilibrated. Initially, each system was submitted to 10 K steps of conjugate gradient minimization where protein's heavy atoms are positionally restrained. This minimization was followed by two steps of equilibration. First, each system was heated to 300 K while simulated for 200 ps under NVT conditions with the solute's heavy atoms restrained. Second, each system was simulated using the NPT ensemble for 200 ns while restraints are gradually removed. Finally, each system was submitted to free NPT 200 ns length MD simulations.

Analyses were performed using GROMACS version 2018.1 [43]. Analyses were performed on the productive stage of the simulations (i.e. after the first 20 ns). For the RMSD analysis, the positional deviations of backbone heavy atoms (N, C α , C), with respect to their initial structure, were calculated every 10 ps (after performing a least square fit to their initial structure). For fluctuation profile analysis, fluctuations of backbone heavy atoms (N, C α , C and O) with respect to their initial structure were calculated every 10 ps (after performing a least square fit to their initial structure).

All visualization was done using the PYMOL MOLECULAR GRAPHICS SYSTEM, Version 2.3 (Schrödinger, LLC, New York, NY, USA). Principle movements were drawn using the *modevectors.py* script by Sean M. Law.

Isothermal titration calorimetry measurements and analysis

Isothermal titration calorimetry measurements were performed in low-volume Nano ITC calorimeter (TA Instruments, New Castle, DE, USA) at 298 K. Proteins were prepared in buffer A and Zn²⁺ solutions in buffer containing 1 mM Tris pH = 8 and 150 mM NaCl. All proteins were diluted to 50 μ M and titrated with: Zn²⁺ 1.4 mM (WT), Zn²⁺ 1 mM (PS) or Zn²⁺ 750 μ M (CS). The protein samples were injected into the instrument cell (170 μ L) and 20 aliquots of 2.5 μ L of a suitable ZnCl₂ solution were titrated into the cell every 350 s. For each protein, three independent titrations were recorded. As a control, each ZnCl₂ solution was also titrated into buffer A under identical experimental conditions, and all measurements were compared to a Double-distilled water-containing reference cell. Data were analysed using TA NANOANALYZE DATA ANALYSIS SOFTWARE, version 3.7.5 (TA Instruments). For WT and PS, data for each measurement were fitted after reduction of the control to an independent model combined with a blank constant model, and the thermodynamics values determined are the average of three different titrations for each protein. Errors are reported as the standard deviation.

Fluorescence spectrometry measurements and analysis

Changes in tryptophan intrinsic fluorescence were monitored using a Fluorolog®-3 (HORIBA Scientific, Edison, NJ, USA) equipped with 1 cm optical path length quartz cell at ambient temperature. One microlitre sample MamM CTD proteins at 5 μM in buffer A was titrated using a 1-mM ZnCl_2 solution in the same buffer to reach different Zn^{2+} concentrations. Samples were measured at an excitation of λ_{ex} 297 nm, and the emission spectrum for each Zn^{2+} concentration recorded between 310 and 450 nm. Each protein–metal titration was replicated three times, and each spectrum was fitted to an extreme function using ORIGINPRO [R^2 (COD) > 0.98; OriginLab Corporation, Northampton, MA, USA]. The maximum wavelength (wavelength of maximum intensity) and the intensity at that wavelength (aka maximum intensity) were averaged for each Zn^{2+} concentration. Errors are reported as the standard deviation.

Stopped-flow measurements and analysis

Stopped-flow experiments were performed using a temperature-controlled Tech Scientific SF-61DX2 stopped-flow instrument (TgK Scientific Ltd., Bradford-on-Avon, UK) at 298 K. All proteins, Zn^{2+} and EDTA solutions were prepared in buffer A. Proteins at 5 μM initial concentration were injected at a 1 : 1 ratio with Zn^{2+} solutions (40–460 μM initial concentrations for WT, 30–460 μM initial concentrations for PS) into the 20 μL reaction cell. For the dissociation experiment, protein at 5 μM pre-equilibrated with 300 μM Zn^{2+} was injected in a 1 : 1 ratio with 2 mM/5 mM EDTA solutions (initial concentrations) into the 20 μL reaction cell. Samples were excited at λ_{ex} 297 nm and the tryptophan fluorescence signal was recorded through a long-pass 320 nm filter. For each injection, 2000 data points were collected for 10 s at logarithmic time scale containing 10 cycles of 200 points/cycle.

For the association measurements of MamM CTD with Zn^{2+} $n = 2$ was used, while for the dissociation measurements of the MamM CTD – Zn^{2+} with EDTA $n = 1$ was used, where n is the number of independent measurements (different protein and Zn^{2+} stocks and initialized instrument; in each measurement of protein- Zn^{2+} or protein- Zn^{2+} -EDTA ratio, at least three injections were averaged). Data were averaged and analysed using the Kinetic Studio software provided with the instrument or with ORIGINPRO. Time-dependent changes of the fluorescence signal were fitted to the exponential function given in Eqn (4):

$$F(t) = F_{\infty} + \sum_{i=1}^n A_i \cdot e^{-k_i t} \quad (4)$$

where $F(t)$ is the fluorescence signal at time t , F_{∞} is the signal at t_{10} s, A_i is the amplitude, k_i is the observed rate constant, i represents the i th relaxation

processes, and n is the total number of relaxation processes, which is either 1 (for a single exponential, e.g. for PS and WT) or 2 (for a double exponential, dissociation experiments). Data were fitted in the time range 1.5–200 ms for PS or 1.5–500 ms for WT for a single exponential, or in time range of 0.0015–10 s for the double exponential EDTA experiments, to exclude the continuous-flow phase of mixing and the tryptophan photobleaching phase.

Acknowledgements

We thank Dr. Anat Shahar for her help with the crystallographic data collection and analysis, to Dr. Tamar Kurzion-Zilbermann for her assistance in CD measurements, to Geula Davidov and to Dr. Nitzan Kutnowski for their help with the ITC measurements and to Prof. Ofer Yifrach for fruitful discussion. SBZ and RZ are supported by the Israel Ministry of Science, Technology and Space; the Israel Science Foundation (grant no. 167/16); the European Molecular Biology Organization and the EU (CMST COST Action CM1306 Understanding Movement and Mechanism in Molecular Machines). This work was supported in part by the Royal Society (FM was a Wolfson Research Merit Award Holder); the EU (COST Action CM1306 Understanding Movement and Mechanism in Molecular Machines) to FM. JH was supported by the BBSRC Norwich Research Park Biosciences Doctoral Training Partnership grant number BB/M011216/1. AH and MVS are supported by the Israel Science Foundation (grants 296/13) and Marie Curie Career Integration Award to AH (grants 1403705/11). Finally, we thank Geffen Zucker for being an integral part of this work and influencing every part of it.

Conflicts of interest

The authors declare no conflicts of interest.

Authors contributions

All authors designed the study, analysed the results and wrote the paper. SBZ and JH conducted the experiments, IK performed the MD simulations and MVS assisted with kinetic modelling. All authors approved the final version of the manuscript.

References

- 1 Laity JH, Lee BM & Wright PE (2001) Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11**, 39–46.

- 2 Cox N, Retegan M, Neese F, Pantazis DA, Boussac A & Lubitz W (2014) Photosynthesis. Electronic structure of the oxygen-evolving complex in photosystem II prior to O-O bond formation. *Science* **345**, 804–808.
- 3 Kanteev M, Goldfeder M & Fishman A (2015) Structure-function correlations in tyrosinases. *Protein Sci* **24**, 1360–1369.
- 4 Dupureur CM (2008) Roles of metal ions in nucleases. *Curr Opin Chem Biol* **12**, 250–255.
- 5 Berg OG & von Hippel PH (1985) Diffusion-controlled macromolecular interactions. *Annu Rev Biophys Chem* **14**, 131–160.
- 6 Dudev T & Lim C (2014) Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. *Chem Rev* **114**, 538–556.
- 7 Montanini B, Blaudez D, Jeandroz S, Sanders D & Chalot M (2007) Phylogenetic and functional analysis of the cation diffusion facilitator (CDF) family: improved signature and prediction of substrate specificity. *BMC Genom* **8**, 107.
- 8 Kolaj-Robin O, Russell D, Hayes KA, Pembroke JT & Soulimane T (2015) Cation diffusion facilitator family: structure and function. *FEBS Lett* **589**, 1283–1295.
- 9 Lu M, Chai J & Fu D (2009) Structural basis for autoregulation of the zinc transporter YiiP. *Nat Struct Mol Biol* **16**, 1063–1067.
- 10 Lopez-Redondo ML, Coudray N, Zhang Z, Alexopoulos J & Stokes DL (2018) Structural basis for the alternating access mechanism of the cation diffusion facilitator YiiP. *Proc Natl Acad Sci* **115**, 3042–3047.
- 11 Barber-Zucker S, Shaanan B & Zarivach R (2017) Transition metal binding selectivity in proteins and its correlation with the phylogenomic classification of the cation diffusion facilitator protein family. *Sci Rep* **7**, 16381.
- 12 Cherezov V, Höfer N, Szebenyi DME, Kolaj O, Wall JG, Gillilan R, Srinivasan V, Jaroniec CP & Caffrey M (2008) Insights into the mode of action of a putative zinc transporter CzcB in *Thermus thermophilus*. *Structure* **16**, 1378–1388.
- 13 Zeytuni N, Uebe R, Maes M, Davidov G, Baram M, Raschdorf O, Nadav-Tsubery M, Kolusheva S, Bitton R, Goobes G *et al.* (2014) Cation diffusion facilitators transport initiation and regulation is mediated by cation induced conformational changes of the cytoplasmic domain. *PLoS ONE* **9**, e92141.
- 14 Uebe R, Keren-Khadmy N, Zeytuni N, Katzmam E, Navon Y, Davidov G, Bitton R, Plitzko JM, Schüler D & Zarivach R (2018) The dual role of MamB in magnetosome membrane assembly and magnetite biomineralization. *Mol Microbiol* **107**, 542–557.
- 15 Barber-Zucker S & Zarivach R (2017) A look into the biochemistry of magnetosome biosynthesis in magnetotactic bacteria. *ACS Chem Biol* **12**, 13–22.
- 16 Uebe R & Schüler D (2016) Magnetosome biogenesis in magnetotactic bacteria. *Nat Rev Microbiol* **14**, 621–637.
- 17 Uebe R, Junge K, Henn V, Poxleitner G, Katzmam E, Plitzko JM, Zarivach R, Kasama T, Wanner G, Pósfai M *et al.* (2011) The cation diffusion facilitator proteins MamB and MamM of *Magnetospirillum gryphiswaldense* have distinct and complex functions, and are involved in magnetite biomineralization and magnetosome membrane assembly. *Mol Microbiol* **82**, 818–835.
- 18 Zeytuni N, Uebe R, Maes M, Davidov G, Baram M, Raschdorf O, Friedler A, Miller Y, Schüler D & Zarivach R (2014) Bacterial magnetosome biomineralization—a novel platform to study molecular mechanisms of human CDF-related Type-II diabetes. *PLoS ONE* **9**, e97154.
- 19 Barber-Zucker S, Uebe R, Davidov G, Navon Y, Sherf D, Chill JH, Kass I, Bitton R, Schüler D & Zarivach R (2016) Disease-homologous mutation in the cation diffusion facilitator protein mamM causes single-domain structural loss and signifies its importance. *Sci Rep* **6**, 31933.
- 20 Jeschke G (2016) Dipolar spectroscopy - double-resonance methods. eMagRes (Ashbrook S, ed), pp. 1459–1476. John Wiley & Sons Ltd, Chichester, UK.
- 21 Jeschke G (2012) DEER distance measurements on proteins. *Annu Rev Phys Chem* **63**, 419–446.
- 22 Jeschke G (2018) MMM: a toolbox for integrative structure modeling. *Protein Sci* **27**, 76–85.
- 23 Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, Timmel CR, Hilger D & Jung H (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Appl Magn Reson* **30**, 473–498.
- 24 David CC & Jacobs DJ (2014) Principal component analysis: a method for determining the essential dynamics of proteins. In *Methods in Molecular Biology* (Livesay D, ed), pp. 193–226. Humana Press, Totowa, NJ.
- 25 Ichiye T & Karplus M (1991) Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins Struct Funct Genet* **11**, 205–217.
- 26 Amadei A, Linssen ABM & Berendsen HJC (1993) Essential dynamics of proteins. *Proteins Struct Funct Genet* **17**, 412–425.
- 27 Xiao Z & Wedd AG (2010) The challenges of determining metal–protein affinities. *Nat Prod Rep* **27**, 768–789.
- 28 Fersht A (1999) Measurement and magnitude of individual rate constants. In *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Hadler GL, ed), pp. 148–153. W.H. Freeman and Company, NY.
- 29 Hannemann DE, Cao W, Olivares AO, Robblee JP & De La Cruz EM (2005) Magnesium, ADP, and actin binding linkage of myosin V: evidence for multiple

- myosin V–ADP and actomyosin V–ADP states. *Biochemistry* **44**, 8826–8840.
- 30 Coudray N, Valvo S, Hu M, Lasala R, Kim C, Vink M, Zhou M, Provati D, Filizola M, Tao J *et al.* (2013) Inward-facing conformation of the zinc transporter YiiP revealed by cryoelectron microscopy. *Proc Natl Acad Sci USA* **110**, 2140–2145.
 - 31 Lu M & Fu D (2007) Structure of the zinc transporter YiiP. *Science* **317**, 1746–1748.
 - 32 Fukue K, Isumura N, Tsuji N, Nishino K, Nagao M, Narita H & Kambe T (2018) Evaluation of the roles of the cytosolic N-terminus and His-rich loop of ZnT proteins using ZnT-2 and ZnT-3 chimeric mutants. *Sci Rep* **8**, 14084.
 - 33 Tanaka N, Fujiwara T, Tomioka R, Krämer U, Kawachi M & Maeshima M (2015) Characterization of the histidine-rich loop of *arabidopsis* vacuolar membrane zinc transporter AtMTP1 as a sensor of zinc level in the cytosol. *Plant Cell Physiol* **56**, 510–519.
 - 34 Zeytuni N, Offer T, Davidov G & Zarivach R (2012) Crystallization and preliminary crystallographic analysis of the C-terminal domain of MamM, a magnetosome-associated protein from *Magnetospirillum gryphiswaldense* MSR-1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **68**, 927–930.
 - 35 Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* **41**, 207–234.
 - 36 Guex N & Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.
 - 37 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera-A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
 - 38 Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L & Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781–1802.
 - 39 Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M & MacKerell AD (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *J Chem Theory Comput* **8**, 3257–3273.
 - 40 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926.
 - 41 Essmann U, Perera L, Berkowitz ML, Darden T, Lee H & Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* **103**, 8577–8593.
 - 42 Humphrey W, Dalke A & Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–38, 27–28.
 - 43 Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25.
 - 44 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
 - 45 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.
 - 46 Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
 - 47 Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
 - 48 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr* **69**, 1204–1214.
 - 49 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
 - 50 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.
 - 51 Joosten RP, Long F, Murshudov GN & Perrakis A (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220.
 - 52 Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Derivation of Eqn (3) for the PS kinetic model.

II Published research - Metal-type dependent binding modes and discriminatory behaviour of MamM-CTD

The following publication (of which I am the joint primary author) includes the study of metal selectivity exhibited by MamM-CTD, via a range of structural, and biophysical techniques.¹⁵⁴ Results include the characterisation of binding of different metals to MamM-CTD, both in crystal form and in solution. This research also provides the first evidence that the CTDs of CDF proteins play a role in metal selectivity, shown by the discriminatory behaviour exhibited by MamM-CTD against Mn^{2+} . My contributions include the acquisition and analysis of EPR data (the results of which have been edited and reformatted for inclusion in chapter 3.8.1), as well as interpretation of results as a whole in the context of their biological relevance. Any text relating to these data was composed by me. Any additional results within this publication, which have been discussed to aid in the understanding of results presented in chapter 3.8.1 of this thesis, have been referenced at the relevant points.



The cation diffusion facilitator protein MamM's cytoplasmic domain exhibits metal-type dependent binding modes and discriminates against Mn^{2+}

Received for publication, April 30, 2020, and in revised form, September 11, 2020. Published, Papers in Press, September 23, 2020. DOI 10.1074/jbc.RA120.014145

Shiran Barber-Zucker^{1,2,3,*}, Jenny Hall^{4,†}, Afonso Froes⁴, Sofiya Kolusheva³, Fraser MacMillan^{4,*}, and Raz Zarivach^{1,2,3,*}

From the ¹Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel, the ²National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer Sheva, Israel, the ³Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer Sheva, Israel, and the ⁴Henry Wellcome Unit for Biological EPR, School of Chemistry, University of East Anglia, Norwich, United Kingdom

Edited by Joseph M. Jez

Cation diffusion facilitator (CDF) proteins are a conserved family of divalent transition metal cation transporters. CDF proteins are usually composed of two domains: the transmembrane domain, in which the metal cations are transported through, and a regulatory cytoplasmic C-terminal domain (CTD). Each CDF protein transports either one specific metal or multiple metals from the cytoplasm, and it is not known whether the CTD takes an active regulatory role in metal recognition and discrimination during cation transport. Here, the model CDF protein MamM, an iron transporter from magnetotactic bacteria, was used to probe the role of the CTD in metal recognition and selectivity. Using a combination of biophysical and structural approaches, the binding of different metals to MamM CTD was characterized. Results reveal that different metals bind distinctively to MamM CTD in terms of their binding sites, thermodynamics, and binding-dependent conformations, both in crystal form and in solution, which suggests a varying level of functional discrimination between CDF domains. Furthermore, these results provide the first direct evidence that CDF CTDs play a role in metal selectivity. We demonstrate that MamM's CTD can discriminate against Mn^{2+} , supporting its postulated role in preventing magnetite formation poisoning in magnetotactic bacteria via Mn^{2+} incorporation.

Cation diffusion facilitator (CDF) proteins are a conserved family of divalent transition metal cation transporters. CDF proteins form dimers, commonly homodimers, usually containing two domains: the N-terminal transmembrane domain (TMD) through which cations are transported and the frequently found regulatory cytoplasmic C-terminal domain (CTD) (1–5). Each CDF protein transports specific metal cations (Zn^{2+} , Mn^{2+} , Cu^{2+} , Co^{2+} , Ni^{2+} , Cd^{2+} , or Fe^{2+}) from the cytoplasm to either the extracellular environment or inner-cellular compartments depending on its membranal location, typically by exploiting the proton-motive-force or by utilizing other cation gradients (6–10). Because they control metal cat-

ion homeostasis at the cellular level, CDF proteins are crucial for proper cell function. As an example, malfunction of human CDF proteins leads to severe diseases including cardiovascular diseases, type-II diabetes, parkinsonism, and Alzheimer's disease (11, 12). Although a greater understanding of CDF protein function at the molecular level has been achieved in recent years, the CDF transport mechanism is not yet fully understood. Furthermore, factors governing metal selectivity of these proteins are still under investigation (7); hence the exact structure-function relationship of CDF proteins remains unresolved, specifically how such relationships affect metal selectivity.

Magnetotactic bacteria (MTB) are a group of gram-negative bacteria that align with and move along magnetic fields. The bacterium contains subcellular organelles named magnetosomes, each of which is composed of ~30–120-nm iron-based magnetic particles (magnetite or greigite) enclosed in a protein-rich lipid membrane. The magnetosome is able to discriminate against different cations, resulting in pure magnetite or greigite particles (13–15). The magnetosomes are arranged on an actin-like polymer in a chain arrangement, creating a magnetic dipole moment allowing these bacteria to make use of external magnetic fields. In nature, this property enables the bacteria to orient themselves to the geomagnetic field, thus navigating to their preferred habitat, usually oxic-anoxic zones in aquatic environments (16, 17). MamM, a CDF protein found uniquely on the magnetosome membrane, transports iron from the cytoplasm into the magnetosomes, enabling the synthesis of pure magnetic particles. It has been demonstrated that the deletion of the *mamM* gene, or even of its CTD, abolishes magnetic particle formation, emphasizing its importance for the biomineralization process (18). In previous studies, WT MamM CTD was shown to bind iron, and the crystal structure of the apo form adopts the characteristic fold of CDF protein CTDs, the presence of which is crucial for the overall protein function. Additionally, three putative MamM CTD metal-binding sites were proposed that included two symmetrical, allosteric peripheral sites (PSs, composed of His-264 from one monomer and Glu-289 from the second monomer) and one central site (CS, composed of Asp-249 and His-285 from both monomers). Biophysical evidence revealed that zinc binding to the PSs leads to a change in protein conformation from an open, dynamic V-shaped dimer to a much

This article contains supporting information.

[†]These authors contributed equally to this work.

*For correspondence: Raz Zarivach, zarivach@bgu.ac.il; Fraser MacMillan, Fraser.MacMillan@uea.ac.uk.

MamM CTD exhibits metal-type dependent binding modes

more rigid, tighter dimeric packing (5, 19–21). These studies and those of other bacterial CDF proteins suggest the following common mechanism for CDF proteins: at a certain cytoplasmic metal cation concentration, cations bind to the CTD, locking it in this tighter packed state as compared with its apo open and dynamic V-shape. This, in turn, facilitates a conformational change of the TMD, thus enabling cations to be transported through this domain to the other side of the membrane (4–6, 21–23). Although previous MamM studies characterized some of the roles and the mechanism of the CTD in CDF proteins, one question has remained unanswered: whether the CTD has an active regulatory role in metal recognition and discrimination during cation transport.

In this study, the well-characterized MamM CTD is used as a model to investigate the role of CTDs in metal recognition and selectivity *in vitro*, and this work presents, for the first time, clear structural and spectroscopic evidence of metal binding of any CDF CTD to multiple metals. The crystal structure of MamM CTD was solved in the presence of Cu^{2+} , Cd^{2+} , and Ni^{2+} , revealing the propensity of different metals to bind at different metal-specific sites, leading to different conformations in the crystalline state. To further characterize the binding abilities of additional metal cations to MamM CTD in solution, tryptophan-fluorescence spectrometry, isothermal titration calorimetry (ITC), and pulsed electron-electron double resonance (PELDOR) spectroscopy were employed. Results indicate that all these metal cations have similar binding abilities to MamM CTD, with the exception of Mn^{2+} . Their binding typically leads to the same or a similar tighter, more rigid conformation as that detected for the Cu^{2+} -bound crystallographic conformation. Each bound metal, however, does exhibit slightly different thermodynamic binding parameters, which could imply a degree of selectivity for the effectivity of metal binding to the CTD. Overall, the results presented here demonstrate that MamM CTD can recognize different metals and discriminates against Mn^{2+} , suggesting an inherent mechanism to prevent its transport by hindering the CDF transport activation, presumably to avoid impure iron-based magnetic particles synthesis. Thus, we see here the first direct evidence for the role of the CTD in metal selectivity in CDF proteins.

Results and Discussion

Crystal structures of MamM CTD with different metals exhibit different conformations and binding sites

The crystal structure of apo MamM CTD has been previously solved in two space groups, both revealing the same monomeric fold and dimerization interface (PDB IDs: 3W5X and 3W5Y) (Fig. 1, A–C, see RMSDs in Table 1 and distances and dihedral angles in Table S4) (5). To better achieve an understanding of the conformational changes occurring during metal binding in MamM (5, 21), co-crystallization trials of MamM CTD (residues 215–318) with various divalent transition metal cations (Fe^{2+} , Mn^{2+} , Zn^{2+} , Ni^{2+} , Co^{2+} , Cu^{2+} , and Cd^{2+}) under a range of different conditions were performed. Crystal structure models of MamM CTD were only obtained in the presence of three different metal cations: Cd^{2+} , Ni^{2+} , and Cu^{2+} (although crystals containing Zn^{2+} diffracted well, Zn^{2+}

could not be detected in the electron density map). The structures with bound Cd^{2+} and Ni^{2+} indicate no conformational change both in their monomeric fold and their dimeric interface, as compared with the apo form (Fig. 1, A and C, Table 1, and Table S4). In contrast, the Cu^{2+} -bound form reveals a clear change in the dimeric conformation to a tighter dimer interface, as has already been demonstrated in other CDF proteins and proposed for zinc-bound MamM CTD in solution using PELDOR spectroscopy (Fig. 1, A and B, Table 1, and Table S4) (21, 23). Furthermore, in this specific Cu^{2+} -bound structure, most of the C-terminal tail of one of the monomers, which was not previously identified in the apo forms' electron density maps, is now resolved. The C-terminal residues (292–314) form a short β -strand that adds to the previously solved three-stranded β -sheet, resulting in a 4-stranded β -sheet, followed by an α -helix. The C-terminal helix is found in the space between the two monomers and forms a stabilizing bonding network with the adjacent monomer's residues. This nonsymmetric interaction may be due to crystal packing or some as yet unknown biological function. The α -helical tail stabilizes the CTD closed state and may interact with the TMD loops to also stabilize the transporter's transmembrane conformation in its activated state as could be seen in the recently determined structure of ZnT-8 (24). However, because CDFs homodimers have not previously revealed nonsymmetric characteristics, we propose that these C-terminal interactions may be locking the dimer in its closed state, resulting in the stabilized crystal packing.

Closer inspection of the different structures' metal-binding sites confirms the identity of the previously proposed residues involved in metal binding (5, 21). The Cu^{2+} -bound structure contains three binding sites: a central binding site (CS) in which one copper ion is bound by His-285 from both monomers and an additional two water molecules, and two "symmetric" peripheral binding sites (PSs), each of which binds one copper ion (Fig. 1, A, B, and D, see metal-ligand distances in Table 2). In both PSs, His-264 chelates the copper ion. In one of the PSs, three water molecules appear to stabilize the ion at this location, and one of them bridges between Glu-289 in the adjacent monomer and the metal ion. Glu-268 is also close to the copper ion at this site; however it does not appear to ligate directly or through a water molecule, whereas in the second PS, Glu-268 is much closer to the ion and participates directly in its chelation (Table 2). In this second PS (much like in the first), a water molecule also bridges between Glu-289 in the adjacent monomer and the metal ion. Because Glu-289 has previously been shown to affect MamM function (5), and because binding to the PSs appears to facilitate tighter dimeric packing (21), we propose that this water molecule plays an important role in the stabilization of the closed state.

In the Cd^{2+} -bound structure, which contains two bound metals per monomer (*i.e.* four metal sites in total), Asp-249 and His-285 each bind one cadmium ion in each monomer. Both cadmium ions are also chelated by Glu-215 from a nonbiological-assembly monomer (Fig. 1, A, C, and E and Table 2) and possess half-occupancy in the electron density map, suggesting that they may be bound alternately. Mutations of both Asp-249 and His-285 to alanine, either separately or together, were

MamM CTD exhibits metal-type dependent binding modes

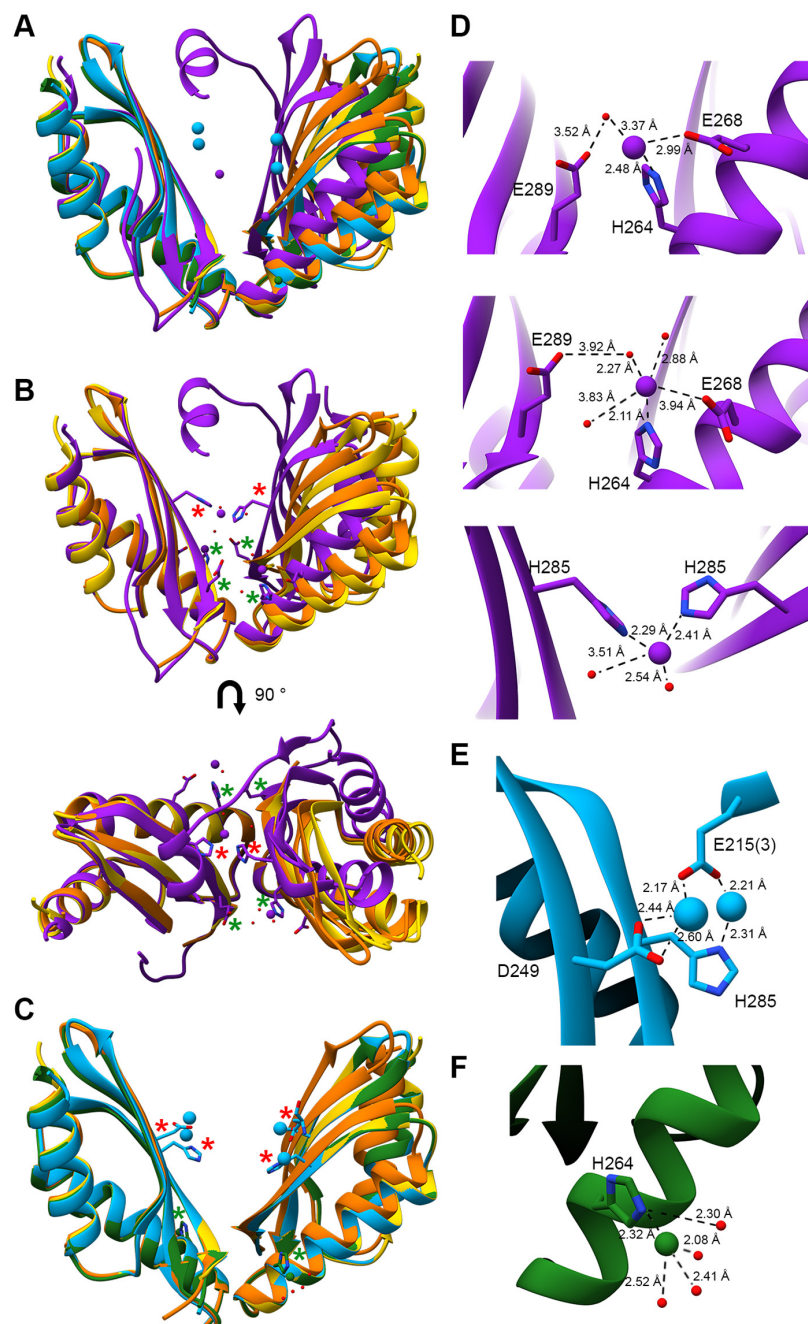


Figure 1. Crystal structures of MamM CTD with different metal cations. A, the three MamM CTD bound structures (Cd²⁺ (PDB ID:6GMT), blue; Ni²⁺ (PDB ID:6GMV), green; Cu²⁺ (PDB ID: 6GP6), purple) and the two apo structures (PDB ID: 3W5X in yellow, PDB ID: 3W5Y in orange), overlapped onto each other. Metal cations are presented for each bound structure in the same colors. B, apo forms and Cu²⁺-bound structures overlapped onto each other from two angles, showing the conformational change exhibited by the Cu²⁺-bound structure. The metal-binding residues are presented as sticks and pointed out with asterisks (CS residues in red and PS residues in green). C, apo forms, Ni²⁺-bound and Cd²⁺-bound structures overlapped onto each other, indicating no conformational changes between all structures. The metal-binding residues are presented as sticks and pointed out with asterisks (CS residues in red and PS residues in green). D, close-up of the Cu²⁺-bound structure binding sites (PSs, upper and middle panels; CS, lower panel). E, close-up of the Cd²⁺-bound structure binding site (of one monomer). (3) refers to a nonbiological-assembly monomer. F, close-up of the Ni²⁺-bound structure binding site (of one monomer). In (D), (E), and (F), the metals, metal-binding residues, metal-bound water molecules, and relevant bond distances are presented.

previously shown to affect the function of MamM CTD both *in vitro* and *in vivo* (5, 21). This Cd²⁺-bound structural model indicates that both residues are able to bind metals, although

this binding alone does not appear sufficient to induce the same conformational change as observed with Cu²⁺. Our previous observation of conformational change toward a closed

MamM CTD exhibits metal-type dependent binding modes

Table 1
RMSDs between different CDFs' CTD dimeric structures

Protein form	MamM apo (3W5X)	MamM apo (3W5Y)	MamM Cu ²⁺ -bound (6GP6)	MamM Cd ²⁺ -bound (6GMT)
MamM apo (3W5X)	-	0.82	1.56	0.35
MamM apo (3W5Y)	-	-	1.60	0.89
MamM Cu ²⁺ -bound (6GP6)	-	-	-	1.62
MamM Cd ²⁺ -bound (6GMT)	-	-	-	-
MamM Ni ²⁺ -bound (6GMV)	0.18	0.91	1.58	0.44
CzrB apo (3BYP)	1.49	1.59	-	-
CzrB bound (3BYR)	-	-	1.43	1.64
EcYiiP bound (3H90)	-	-	1.33	1.19
SoYiiP bound (5VRF) ^a	-	-	-	-
MamB apo (5HO5)	1.21	1.20	1.32	1.26 ^b
MamB bound (5HO1)	1.76	1.62	1.37	1.13

All RMSDs (given in Ångströms) between the dimeric structures were calculated using the iterative magic fit tool of Swiss-PdbViewer 4.1.0.

^aSatisfying fit directly between the dimeric MamM and SoYiiP structures could not be obtained.

^bRMSD was calculated using the domain fragment alternate fit tool of Swiss-PdbViewer 4.1.0.

Table 2
Coordination distances to ions in MamM CTD metal-bound crystal structures

Protein form	Glu-215 3 rd -mer	Asp-249	His-285	His-264	Glu-289 to water	Glu-268	Water
Cu ²⁺ -bound (6GP6)	-	-	2.41 ³ 2.29 ³	2.48 ¹ 2.11 ²	3.52 ^{a,1} 3.92 ^{b,2}	2.99 ¹ 3.94 ²	3.37 ^{a,1} 2.27 ^{b,2} 2.88 ² 3.83 ² 3.51 ³ 2.54 ³
Cd ²⁺ -bound (6GMT)	2.17 ¹ 2.21 ²	2.44 ¹ 2.60 ¹	2.31 ²	-	-	-	-
Ni ²⁺ -bound (6GMV)	-	-	-	2.32	-	-	2.30 2.08 2.41 2.52

All distances were measured using UCSF Chimera (34).

For the Cu²⁺-bound structure, Glu-289 and the water molecule that coordinates it to the Cu²⁺ ion are marked as ^a for one peripheral site (Fig. 1D, upper panel) or ^b for the second peripheral site (Fig. 1D, middle panel).

For the Cu²⁺-bound structure, ¹ relates to coordination of the Cu²⁺ ion in one peripheral site (Fig. 1D, upper panel), ² relates to coordination of the Cu²⁺ ion in the second peripheral site (Fig. 1D, middle panel), and ³ relates to coordination of the Cu²⁺ ion in the central site (Fig. 1D, lower panel).

For the Cd²⁺-bound structure, ¹ relates to coordination of one Cd²⁺ ion in the central site (Fig. 1E, left ion) and ² relates to coordination of the second Cd²⁺ ion in the central site (Fig. 1E, right ion).

stable state upon Zn²⁺ binding was only induced upon metal binding to both the PSs. Hence, binding of Cd²⁺ to the CS alone (in agreement with Zn²⁺ binding to the CS alone (21)) is not expected to cause a distinct conformational change.

In the Ni²⁺-bound structure, which only contains one bound metal per monomer in the peripheral site, this nickel ion is bound to His-264 and to water molecules, possibly implying a nonspecific interaction of the Ni²⁺ ion with the histidine (Fig. 1, A, C, and F and Table 2). However, specific chelation to this histidine rather than to other metal-binding residues might suggest a higher affinity for nickel, whereas its location in the protein periphery could also suggest it plays an important role in attracting the metal ions. This proposal is supported by previous studies that demonstrated the impact of His-264 on MamM function (5) and the new Cu²⁺-bound structure.

Interestingly, all the metal-dependent crystallographic binding sites can be well-explained by the metals' propensities to be bound by specific amino acids: Cu²⁺ has the highest affinity for histidine, followed by Ni²⁺ with Cd²⁺ having the lowest affinity, whereas the aspartate and glutamate binding tendencies show the opposite trend (7). Overall, the MamM CTD metal-bound crystal structures confirm the previously proposed binding sites and reveal diversity among the different metals' chela-

tion. Because each metal is bound by different residues and only the Cu²⁺-bound structure reveals the apparent conformational change, these crystallographic results would indicate that each metal exhibits a different binding mode. These binding modes may well underpin the prevention of the correct transport activation via such altered conformational changes and thus may distinctly regulate the overall protein.

Dimer packing of the MamM CTD Cu²⁺-bound structure is less tight compared with other CDF CTD bound structures

Previously, a number of CDF protein CTD structures were solved in their apo state, their Zn²⁺-bound state (YiiP from *Escherichia coli* and YiiP homolog from *Shewanella oneidensis*, EcYiiP and SoYiiP respectively, and human ZnT-8), or in both apo and Zn²⁺-bound states (MamB and CzrB) (1–4, 23–25) (Fig. 2A). According to these solved structures, the only protein to exhibit a distinct closure of the V-shape upon metal binding compared with its apo state is CzrB (PDB IDs: 3BYP and 3BYR) (23). The bound state of CzrB exhibits a larger degree-of-closure than that of the MamM CTD Cu²⁺-bound state, demonstrated by the larger differences in the distances and the dihedral angles between equivalent residues in the apo and closed

MamM CTD exhibits metal-type dependent binding modes

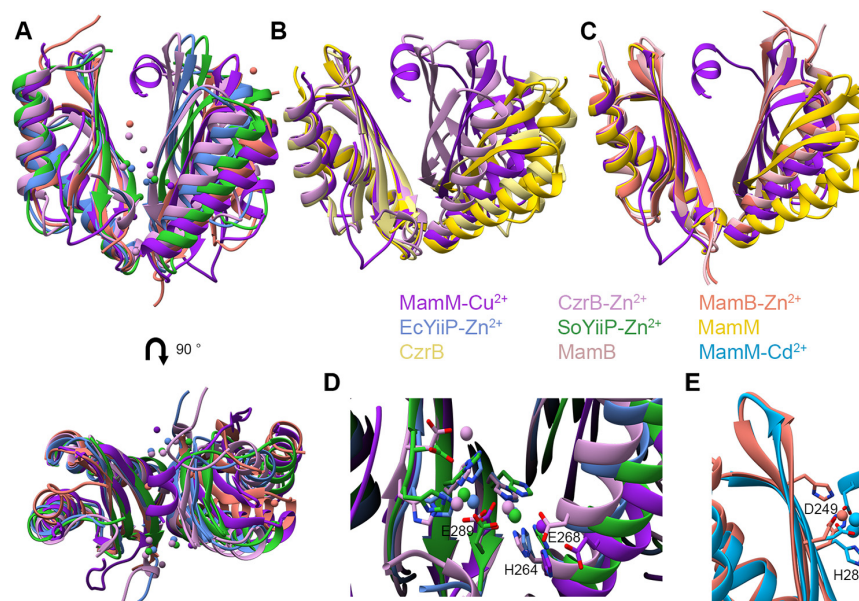


Figure 2. Comparison of the conformational changes and binding sites of different bacterial CDFs' CTDs. Colors legend: MamM Cu^{2+} -bound (PDB ID: 6GP6, dark purple), CzfB Zn^{2+} -bound (PDB ID: 3BYR, light purple), MamB Zn^{2+} -bound (PDB ID: 5HO1, salmon), EcYiiP Zn^{2+} -bound (PDB ID: 3H90, cornflower blue), SoYiiP Zn^{2+} -bound (PDB ID: 5VRF, green), MamM apo (PDB ID: 3W5X, dark yellow), CzfB apo (PDB ID: 3BYR, light yellow), MamB apo (PDB ID: 5HO5, light pink), MamM Cd^{2+} -bound (PDB ID: 6GMT, sky blue). **A**, comparison between all bound CDF CTDs structures at two angles: MamM Cu^{2+} -bound, MamB, CzfB, EcYiiP, and SoYiiP (all Zn^{2+} -bound). Metal cations are presented for each bound structure in the same colors. **B**, CzfB exhibits much larger conformational change between apo and bound states compared with MamM; MamM apo, MamM Cu^{2+} -bound, CzfB apo, and CzfB Zn^{2+} -bound structures are overlapped. **C**, MamB exhibits no conformational changes between the bound and apo states and shows tighter packing compared with MamM apo form; MamM apo, MamM Cu^{2+} -bound, MamB apo, and MamB Zn^{2+} -bound structures are overlapped. **D**, peripheral site binding residues are conserved between MamM, CzfB, and YiiP proteins. Metal cations are presented for each bound structure in the same colors; metal-binding residues are presented as sticks. The presented residue labels refer to MamM. **E**, central site binding residues are conserved between MamM and MamB proteins. Metal cations are presented for each bound structure in the same colors; metal-binding residues are presented as sticks. The presented residue labels refer to MamM.

states (Fig. 2B, Table 1, and Table S4). This is due not only to a tighter closed state but also to a slightly more open apo form compared with those of MamM CTD. MamB, another CDF protein from MTB, exhibits very little conformational change between the apo and Zn^{2+} -bound states (25). However, the degree-of-closure in both MamB CTD structures is larger than that of MamM CTD apo structures and is similar to that of the MamM Cu^{2+} -bound structure (PDB IDs: 5HO1 and 5HO5; Fig. 2C, Table 1, and Table S4). The structures of the entire CDF proteins EcYiiP, SoYiiP, and human ZnT-8 (including both the CTD and TMD) have also been solved with the CTD in the Zn^{2+} -bound state (PDB IDs: 3H90, 5VRF, and 6XPD) (2, 4, 24). The CTD structures are highly conserved in the high-resolution YiiP structures, and the degree-of-closure is also greater than that of the MamM Cu^{2+} -bound state (Fig. 2A, Table 1, and Table S4). However, this cannot be directly compared with the overall conformational change exhibited by MamM between apo and bound states, because the apo structures of EcYiiP and SoYiiP are not currently available and because ZnT-8 is stabilized in a permanent closed conformation via its C and N termini. In conclusion, the MamM CTD Cu^{2+} -bound structure shows moderate dimerization closure compared with other CDF bound proteins, as was also proposed in our previous study of Zn^{2+} -binding and associated conformational change in solution (21). This may be due to the presence of an additional CTD terminal helix within MamM, which is not

found in other CDF protein structures and which potentially extends the interdimer gap. However, there is variation in the degree-of-closure between the CDF CTDs of other proteins, suggesting that some degree of flexibility in the CTD conformational changes between the different CDF family members exists. Noticeably, this also depends on the CTD-TMD interface that occurs mainly between the CTD and TMD loops. This interface can vary between the different proteins and thus affect the degree-of-closure exhibited by each individual protein's CTD.

The binding sites present in all MamM CTD metal-bound structures are consistent with those of other CDF protein binding sites. MamM PSs are composed of His-264 (Cu^{2+} -bound and Ni^{2+} -bound structures) and Glu-289 via a bridging water molecule (Cu^{2+} -bound structure). In CzfB, six zinc ions are bound to the CTD through three sets of symmetric binding sites (Fig. 2D) (23). Two Zn^{2+} ions are bound in symmetric PSs each by His-258 (from one monomer), Glu-282, and His-280 (from the second monomer), where residues His-258 and Glu-282 are homologous to His-264 and Glu-289 in MamM PSs (His-280 is homologous to Ser-287 in MamM). Two additional Zn^{2+} ions are symmetrically bound by Glu-282, and additional residues of the same monomer that are not homologous to any of the MamM CTD binding residues but to residues that do not usually participate in metal binding. These residues also participate in the binding of the last two symmetric Zn^{2+} ions.

MamM CTD exhibits metal-type dependent binding modes

Overall, this protein- Zn^{2+} interaction network stabilizes Zn^{2+} binding in the peripheral sites, which are the only sites to exhibit binding by residues in both monomers, indicating that the PSs are important for conformational change in CrzB. In the YiiP bound structures, two symmetrical pairs of Zn^{2+} ions are bound in the same locations and by homologous residues to the first two Zn^{2+} pairs in CsrB (Fig. 2D) (2, 4), further highlighting the importance of the peripheral sites. Although the Glu-289 residue in MamM does not chelate the Cu^{2+} directly in the crystal structure, Zn^{2+} is bound by the homologous residues to Glu-289 in the crystal structures of the other Zn^{2+} -bound CDF proteins. Furthermore, Glu-289 is crucial for proper function *in vitro* and *in vivo* (5). Thus, it is clear that Glu-289 in MamM plays an important role in metal binding for biological function.

The MamM Cu^{2+} -bound structure also involves one Cu^{2+} ion bound centrally by His-285 from each monomer and an additional two water molecules. Homologous to His-285 of MamM is residue His-283 in MamB, involved in the chelation of Zn^{2+} . In MamB, one Zn^{2+} ion is coordinated in each monomer by residues His-283, Asp-247 (homologous to Asp-249 in MamM), and His-245 (found in the same location as Trp-247 in MamM) (Fig. 2E) (25). Whereas the MamM Cu^{2+} -bound structure involves only His-285 (but from both monomers, together binding a single copper ion), the Cd^{2+} -bound structure exhibits binding more similar to that of MamB. In the Cd^{2+} -bound structure, no cadmium ions are bound to the PSs, and residues of MamB that are homologous to the MamM PSs would not afford effective metal cation binding. Additionally, both metal-bound structures exhibit no conformational changes compared with their apo forms. However, unlike the MamB Zn^{2+} -bound structure, Cd^{2+} -bound MamM includes two Cd^{2+} ions per monomer (one bound by His-285 and the other by Asp-249). Nonetheless, because these Cd^{2+} ions possess only half-occupancy in the electron density map, on average just one ion is bound in each monomer, as is the case with MamB. When considering all CDF metal-bound structures, a central site exists only in MamM and MamB, and mutations of these CS residues in both proteins were shown to lead to lower functionality both *in vivo* and *in vitro* (5, 25). Noticeably, copper binding in a unique coordination in the CS facilitates subsequent binding to the PS (as was shown for Zn^{2+} (21)), or vice versa. Overall, the analysis of all known structures suggests that the CS and PSs bind different metals distinctively in MamM, strengthening the previous observation that only when metals are bound to the PSs can the conformational changes effectively occur.

MamM CTD shows metal-type-dependent change in tryptophan fluorescence signal

Crystal structures often only provide limited and sometimes biased information on the conformation of proteins. Hence, to further characterize MamM CTD binding to different metal cations, we have investigated these interactions in solution. Each MamM CTD monomer contains one tryptophan residue, Trp-247, juxtaposed to the CS. When metal cations bind to MamM CTD, it is proposed that the monomers approach each other, and hence the tryptophan environment will be changed,

which should lead to a shift in the emission spectrum of the tryptophan residue when excited at the tryptophan-specific wavelength ($\lambda = 297$ nm). In the case of paramagnetic metals, their binding near this tryptophan will cause the tryptophan fluorescence intensity to be quenched. We have shown previously that the addition of Zn^{2+} causes a blue shift in the emission spectrum of MamM CTD and that the addition of the paramagnetic Fe^{2+} causes signal quenching (5, 21), indicating that MamM CTD binds both cations. In this study, we have titrated further metal cations (Cd^{2+} , Ni^{2+} , and Cu^{2+} , in which their bound crystal structures were detected, and Mn^{2+} , which was shown to be transported by numerous CDF proteins (7, 8)) into MamM CTD solution samples and measured the concentration-dependent changes in the tryptophan emission spectra (ratios of 0:1 to 100:1 metal:protein, Fig. 3). As shown previously for Zn^{2+} (Fig. 3A, adopted from (21)), in the case of the diamagnetic Cd^{2+} , a shift in maximum wavelength could be observed (Fig. 3B), suggesting that the binding of Cd^{2+} leads to a change in the tryptophan environment. However, although the spectral shift is accompanied by an increase in intensity for Zn^{2+} , the intensity decreases when Cd^{2+} is added and the blue shift is smaller (Fig. 3, F and G). The Cd^{2+} -bound crystal structure suggests no conformational change compared with the apo form, whereas it was previously shown that the addition of Zn^{2+} causes a tight closure of the CTD (21). Because cadmium is much bigger than zinc, Cd^{2+} binding might restrict the ability of the dimer to pack tighter, which could result in a smaller shift and the difference in the fluorescence intensity observed for Cd^{2+} . Alternatively, it is also possible that the Cd^{2+} ions, which are bound in close proximity to Trp-247, simply influence the tryptophan fluorescence signal differently. Mn^{2+} , Ni^{2+} , and Cu^{2+} are all paramagnetic; hence their binding to the CTD would be expected to cause observable fluorescence quenching leading to a stronger signal than their binding-dependent spectral shift. However, the addition of Mn^{2+} only reveals minor changes in the fluorescence intensity (Fig. 3, C and G) and in the maximum wavelength (Fig. 3F), suggesting that Mn^{2+} cannot bind to the CTD in a way that would cause a change in the fluorescence spectra. In contrast, Ni^{2+} and Cu^{2+} demonstrate clear metal concentration-dependent quenching (Fig. 3, D, E, and G) and a slight spectral shift (Fig. 3F), with Cu^{2+} exhibiting stronger quenching and a larger shift. Overall, these results indicate that MamM CTD can bind all the examined metals with the exception of Mn^{2+} .

MamM CTD exhibits metal-type-dependent binding thermodynamics

To accurately calculate the affinity of each metal to MamM CTD, other thermodynamic parameters, and the number of metal cations that can bind to the CTD, ITC measurements were performed (Fig. 4). Here Cd^{2+} , Ni^{2+} , Cu^{2+} , and Mn^{2+} were titrated into MamM CTD solution samples. All ITC parameters are summarized in Table 3, including those for previously titrated Zn^{2+} (21) (note: metal binding to MamM CTD requires a slightly basic pH because of the metal chelation by histidine residues; therefore titration of Fe^{2+} leads to iron precipitation and hence its binding parameters cannot be

MamM CTD exhibits metal-type dependent binding modes

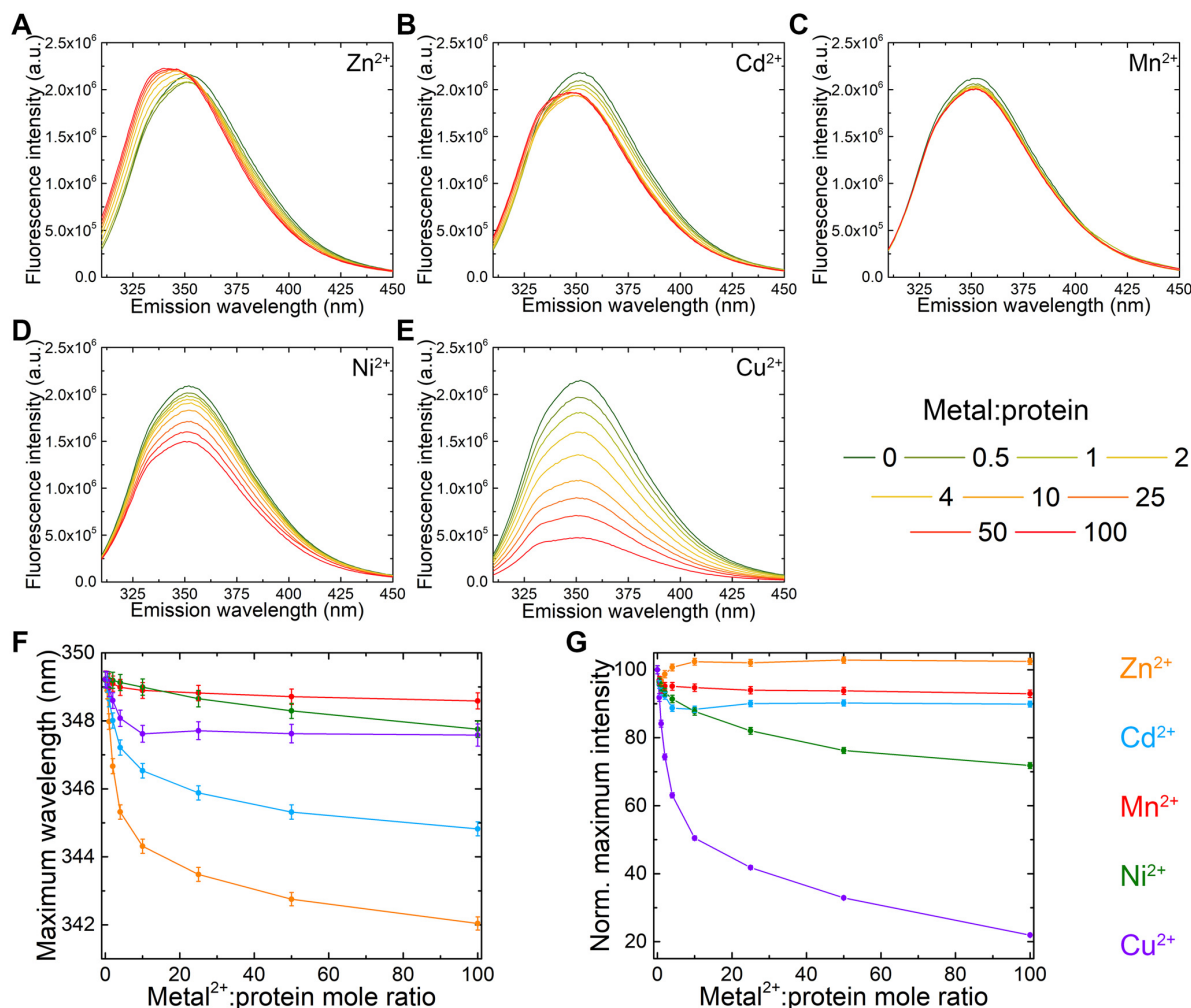


Figure 3. Fluorescence spectral scans of MamM CTD with different metals. All emission spectra of 5 μM MamM CTD were titrated using different metal solutions to reach different MamM CTD:Metal²⁺ ratios (intensities were normalized because of the change in MamM CTD concentration). Samples were measured at λ_{ex} 297 nm, and the emission spectrum for each metal concentration was recorded at 310–450 nm and presented as arbitrary units (a. u.). Each presented spectrum is the average of three independent measurements. A, emission spectra of MamM CTD with different Zn²⁺ concentrations. B, emission spectra of MamM CTD with different Cd²⁺ concentrations. C, emission spectra of MamM CTD with different Mn²⁺ concentrations. D, emission spectra of MamM CTD with different Ni²⁺ concentrations. E, emission spectra of MamM CTD with different Cu²⁺ concentrations. F, λ_{max} as function of Metal²⁺:MamM CTD mole ratio. G, normalized fluorescence intensity compared with no Metal²⁺ as function of Metal²⁺:MamM CTD mole ratio. In (F) and (G), Zn²⁺ is presented in orange, Cd²⁺ in blue, Mn²⁺ in red, Ni²⁺ in green, and Cu²⁺ in purple, and error bars are shown within the symbols.

accurately determined). ITC results indicate that all metals except Mn²⁺ induce a heat change when titrated into a MamM CTD solution. Hence taken together with the fact that Mn²⁺ revealed no fluorescence shift, we propose that MamM CTD is unable to bind Mn²⁺.

The K_d values determined vary from 6–48 μM , with Cu²⁺ having the greatest affinity followed by Ni²⁺, Zn²⁺, and Cd²⁺ (Table 3). As revealed in all metal-bound crystal structures and previous *in vivo* studies (5), histidine residues are the major components for metal binding. Again, the larger size of the Cd²⁺ ion may explain its low affinity because its coordination may be less preferred; however, a lower propensity to bind histidine residues compared with the other metals (7) may also result in a lower affinity to proteins with multiple histidine-containing binding sites such as MamM

CTD. Cu²⁺ has the highest propensity to bind histidine, which clearly explains its high affinity to MamM. Ni²⁺ also exhibits a high propensity for histidine binding compared with Zn²⁺ and Mn²⁺ (7). Although all binding reactions are exothermic (Table 3), Cu²⁺-binding displays much lower enthalpy, which means it is favored in terms of the heat release. Moreover, whereas Ni²⁺ and Zn²⁺, but also Cd²⁺, display high entropy, the entropy of Cu²⁺ binding is negative, suggesting that its binding results in a more ordered state of the cations compared with its free state in solution. Summarizing all of the above, the negative Gibbs free energy of the binding reactions implies that all metal-binding reactions are preferred to the apo state (excluding Mn²⁺), yet MamM CTD shows clear preferences for the binding of specific metals over others.

MamM CTD exhibits metal-type dependent binding modes

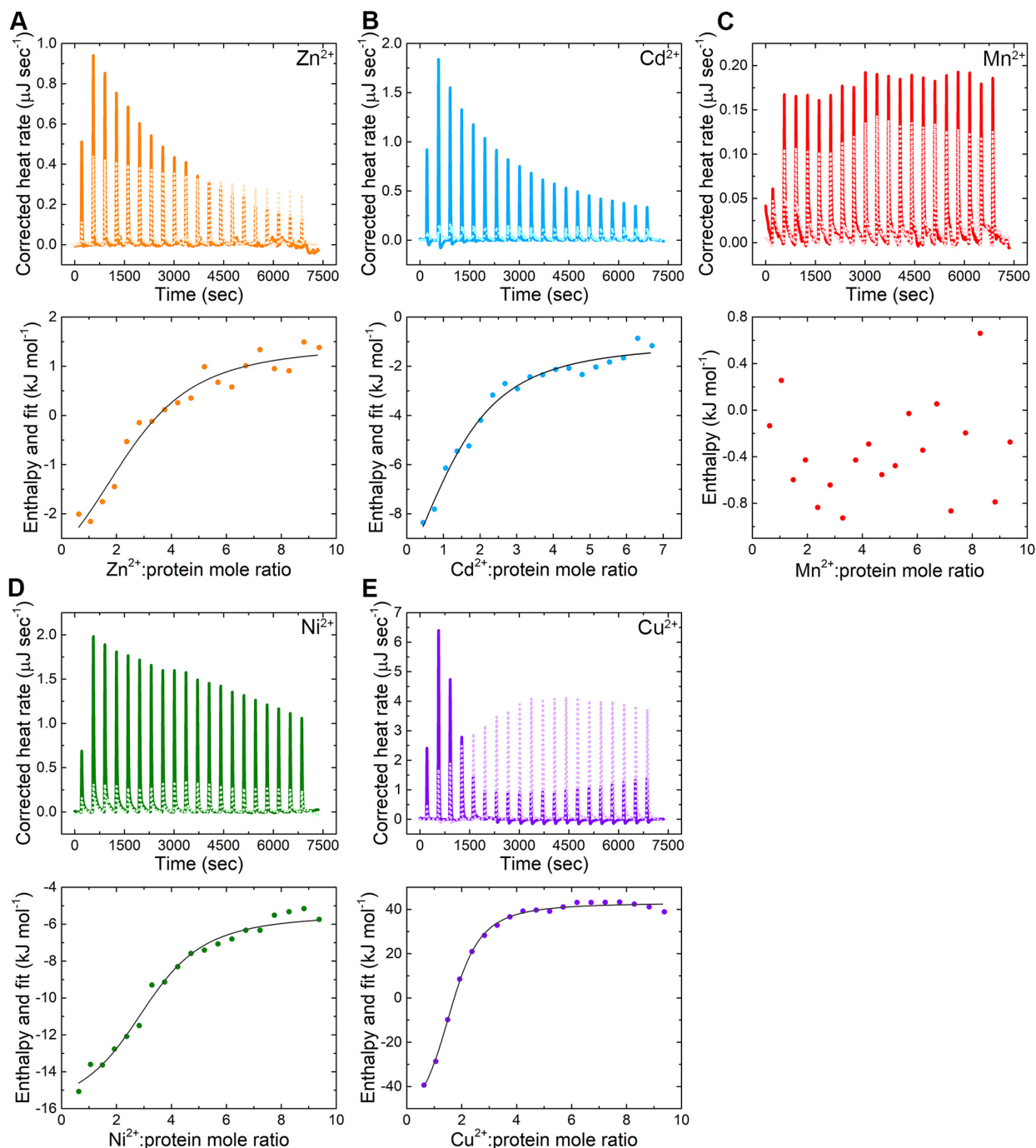


Figure 4. ITC titration measurements of MamM CTD-metals interactions, with (A) Zn^{2+} , (B) Cd^{2+} , (C) Mn^{2+} , (D) Ni^{2+} , and (E) Cu^{2+} . Top panels: Corrected heat rate curves for a representative titration of MamM CTD with each metal as a function of time. For each metal, the heat flow curve of the control is given as a reference (metal titration into buffer, dotted line in brighter color). Bottom panels: Titration data after peak integration as a function of Metal^{2+} :MamM CTD mole ratio, and data fitting for all metals except Mn^{2+} . See Table 3 for binding parameters.

The number of metal cations bound per dimer is also given by the ITC measurements and varies significantly between metals (Table 3). For Zn^{2+} and Ni^{2+} , ~ 3 cations are bound per dimer, in agreement with the Cu^{2+} -bound structure. The Ni^{2+} -bound crystal structure exhibits two cations per

dimer in the PSs. However, it shows no conformational changes when compared with the apo state, and binding appears to be nonspecific; disagreements between the metal-bound structure and ITC results can therefore be expected.

MamM CTD exhibits metal-type dependent binding modes

Table 3
Thermodynamic parameters of MamM CTD binding to different metals

Parameter	Zn ²⁺ ^a	Cd ²⁺	Ni ²⁺	Cu ²⁺
K _d (μM)	31.50 ± 3.40	48.39 ± 2.61	17.71 ± 3.35	6.00 ± 3.83
N	2.80 ± 0.24	1.23 ± 0.08	3.26 ± 0.19	1.56 ± 0.15
ΔH (kJ mol ⁻¹)	-4.20 ± 0.10	-13.08 ± 1.56	-7.78 ± 2.38	-71.80 ± 23.59
ΔG (kJ mol ⁻¹)	-25.71 ± 0.26	-24.64 ± 0.13	-27.17 ± 0.48	-30.13 ± 1.59
ΔS (J mol ⁻¹ K ⁻¹)	72.15 ± 0.98	38.76 ± 4.81	65.03 ± 9.59	-136.72 ± 83.76

ITC data were fitted to independent + blank (constant) models in the TA NanoAnalyze Data Analysis software.

Each parameter represents three independent measurements after control reduction (metal at the same concentration titrated into buffer solution).

The buffer used for all titrations: 10 mM Tris-HCl, pH 8.0, and 150 mM NaCl.

^aThe data are taken from (21).

For Cd²⁺ and Cu²⁺, lower numbers of metal cations bound per dimer are observed: 1.23 ± 0.08 for Cd²⁺ and 1.56 ± 0.15 for Cu²⁺. The numbers of metals observed in metal-bound structures are clearly not in agreement with these ITC observations, which are much lower. The Cd²⁺-bound structure reveals four metal-binding sites per dimer; however, the low occupancy observed in the electron density map would suggest that actually only two cations are bound at any one time. Moreover, binding in the crystal structure is mediated by a glutamate residue from a third monomer (from an adjacent unit cell), which may simply be a crystallographic artifact not necessarily expected to be observed in solution. Nevertheless, although the tryptophan-fluorescence scans suggest otherwise, this metal-bound structure does not exhibit any conformational change as also observed for the Ni²⁺-bound structure. Hence, the ~1 Cd²⁺ ion bound per dimer determined from the ITC results suggests that binding at the CS is of a different nature from that displayed in the crystal structure.

The unexpectedly low number of Cu²⁺ ions bound in solution, as compared with the crystal structure, could have several explanations. It is quite possible that Cu²⁺ binding to the CS is silent in terms of the heat change (as was also shown previously for Zn²⁺-binding to the CS (21)); hence the thermodynamic parameters may relate only to the binding of two ions to the PSs (1.56 ~ 2, and thus three Cu ions in total). Alternatively, in solution only one ion might bind at the CS, and none at the PSs (then 1.56 ~ 1, in clear disagreement with the structure). However, the opposite may also be true, in which two ions are bound to the PSs and none to the CS (then again, 1.56 ~ 2, and thus two overall, again in disagreement with the structural data). Because binding to the PSs has previously been shown to be the only cause of conformational change in MamM and other CDF CTDs, one would expect that if Cu²⁺ only binds to the CS, then no conformational changes should occur in solution, and that if Cu²⁺ only binds to the PSs, then a tighter conformation compared with the apo form would be detected in solution. The large tryptophan-fluorescence signal quenching demonstrated by titration of Cu²⁺ into MamM CTD indicates that the copper ions are bound in close proximity to Trp-247, hinting at an occupied CS. However, the tryptophan-fluorescence resolution is not sufficient to distinguish between these proposed options.

EPR and PELDOR spectroscopy observe metal-bound MamM CTD conformational changes

To probe divalent metal binding and the associated conformational changes reflected by intra-molecular distance, site-

directed spin labeling was employed in combination with both continuous-wave EPR (cw-EPR) and pulsed electron-electron double resonance (PELLDOR) spectroscopic studies (21). These methods provide a wealth of dynamic and structural information via observations of both changes in the local environment of an attached exogenous spin label elucidated by cw-EPR and the determination of long-range (typically between 2–8 nm) distance measurements and distance distributions afforded by PELDOR spectroscopy (26). We have previously reported (21) that using only one of the two intrinsic MamM CTD cysteine residues simplifies EPR measurements; hence the studies reported here use the same C267S variant of MamM CTD, in which only the Cys-275 site is available for labeling. *In silico* analysis using Multiscale Modelling of Macromolecules (MMM), a Matlab©-based open-source modeling toolbox (27), with both the apo WT and C267S CTD crystal structures (PDB IDs: 3W5Y and 6G55, respectively) predicts an average inter-spin distance of 4.0 nm between the labeled cysteines (Cys-275) in each protomer (Fig. 5A) (21).

cw-EPR studies at X-band (9 GHz) of complex dynamic proteins afford the observation of changes in the local environment of an attached exogenous spin label and the overall mobility of the protein in different states. Studies were undertaken to ascertain what limiting effects, if any, the addition of zinc and copper metals (first-row transition metals to which the protein has a lower and higher affinity, respectively) has on the local spin label environment of C267S_{SL}. Indeed, cw-EPR spectra of these metal-bound C267S_{SL} samples clearly differ substantially from those of the apo protein (Fig. 6A). The dynamics exhibited by spin labels attached to a protein have three main contributions: 1) the rotational correlation time of the protein itself; 2) the backbone motions of the protein at the point of spin label attachment; and 3) the localized intrinsic motion of the spin label. The latter two modes of motion are dominating factors in the line shapes of cw-EPR spectra at X-band, in particular their spectral broadness. Binding of metal ions to the protein resulting in conformational change affects these modes of motion, hence resulting in different spectral line shapes observed in Fig. 6A.

Our previous PELDOR studies of spin-labeled C267S MamM CTD (C267S_{SL}), undertaken on apo and Zn²⁺-bound forms of the protein, afforded a direct comparison of the inter-spin distances giving insight into the conformational changes associated with Zn²⁺ binding (21). Although results for the apo protein clearly indicated that no single conformation is adopted, a very narrow distance distribution upon addition of Zn²⁺ (forming C267S_{SL}-Zn²⁺) with a mean distance of 3.3 nm was interpreted as a change from a broad dynamically labile

MamM CTD exhibits metal-type dependent binding modes

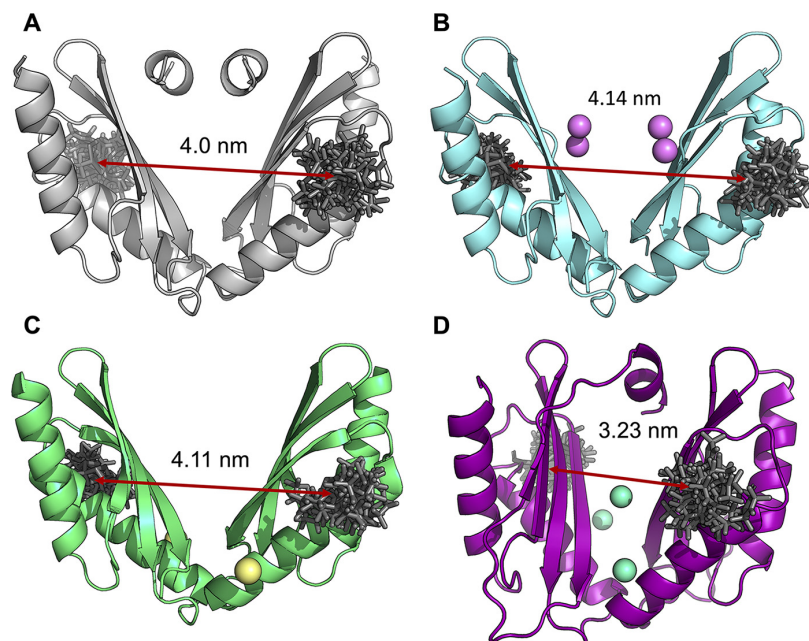


Figure 5. MamM CTD structural models built using the published crystal structures. Cartoon representation of (A) dimeric apo MamM CTD C267S (PDB ID: 6G55), (B) Cd^{2+} -bound MamM CTD (PDB ID: 6GMT), (C) Ni^{2+} -bound MamM CTD (PDB ID: 6GMV), and (D) Cu^{2+} -bound MamM CTD (PDB ID: 6GP6) with the calculated rotamer libraries of MTSL attached to position Cys-275 shown as sticks, and with metal ions shown as spheres. *In silico* modeling of MTSL rotamer libraries and calculation of the average inter-spin distance was performed using MMM, a Matlab[®]-based open-source modeling toolbox.

state of apo C267S (Fig. 6C) to a single closed conformation upon Zn^{2+} -binding. Further, stoichiometric Zn^{2+} titrations, followed by PELDOR, suggested a maximum binding of 3–4 Zn^{2+} ions per dimer, in good agreement with ITC experiments (21). Here we have undertaken further cw-EPR and PELDOR studies on C267S_{SL} incubated with a range of divalent metal cations, including Cd^{2+} , Ni^{2+} , Cu^{2+} , Mn^{2+} , and Fe^{2+} (the natural substrate of MamM). Overall, our results indicate that although most PELDOR traces and associated distance distributions in the presence of divalent metal ions do show distinct differences from the apo protein, not all metal ion effects are equivalent.

The addition of both CdCl_2 and NiCl_2 to C267S_{SL} ($\times 5$ excess, forming C267S_{SL}- Cd^{2+} and C267S_{SL}- Ni^{2+} , respectively) results in a distinct change in the PELDOR time trace obtained as compared with the apo case. A single and quite narrow distance distribution is predominant in both cases with an average distance of 3.3 nm (Fig. 6, D and E) and is very similar to the distribution obtained previously with Zn^{2+} (Fig. 6C, adopted from (21)). Whereas the nature of this binding appears to be similar to that of Zn^{2+} -bound MamM CTD, neither Cd^{2+} - nor Ni^{2+} -bound crystal structure models would suggest such a metal-dependent conformational change. Consequently, MMM inter-spin distance predictions are 4.14 and 4.11 nm for Cd^{2+} - and Ni^{2+} -bound structures, respectively (Fig. 5, B and C). Because the occupancy for each metal-binding site in the Cd^{2+} -bound solved crystal structure is rather low (0.5), and considering the ITC data which indicate that only one Cd^{2+} ion is bound per dimer, and the observed change in tryptophan fluorescence, this would suggest that in solution only one Cd^{2+}

binds at the CS. This is in contrast to Zn^{2+} binding because it appears that only metal binding at the CS is required to induce the conformational change. This is possible because of the larger size of the Cd^{2+} ion which, once bound to the CS residues of just one monomer, is in close enough proximity to attract the CS binding residues of the second monomer, forcing the observed conformational change. The observed conformational change with Ni^{2+} might be rationalized by nonspecific Ni^{2+} -binding events, as suggested by the metal-ligand variations within the solved structure. Nonetheless, the shallower modulation depth (Fig. 6E) may indicate that a smaller fraction of nitroxide spins are actually dipolar coupled and thus interacting. ITC results indicate binding of three Ni^{2+} ions per dimeric protein with mild affinity compared with the other metals, and the observation of tryptophan fluorescence quenching in the presence of Ni^{2+} would also imply Ni^{2+} binding either at or close to the CS. Taken together, and considering that the nickel ion's radius is comparable with that of a zinc ion, we propose that Ni^{2+} binds both to the CS (one ion) and the PSs (two ions), with binding to the latter being associated with a conformational change similar to that observed for Zn^{2+} .

The room-temperature cw-EPR experiments of C267S_{SL} incubated with an excess of Cu^{2+} , which is itself paramagnetic (forming C267S_{SL}- Cu^{2+}), suggest a partial quenching of the EPR signal as compared with both the Zn^{2+} bound and apo states (Fig. 6A). The Cu^{2+} ions are EPR active ($S = 1/2$), giving EPR signals that spectrally overlap with that of the spin label; thus the excess chosen for PELDOR measurements was equal to the number of proposed binding sites to avoid any exogenous, spurious Cu^{2+} EPR signals arising from nonspecifically

MamM CTD exhibits metal-type dependent binding modes

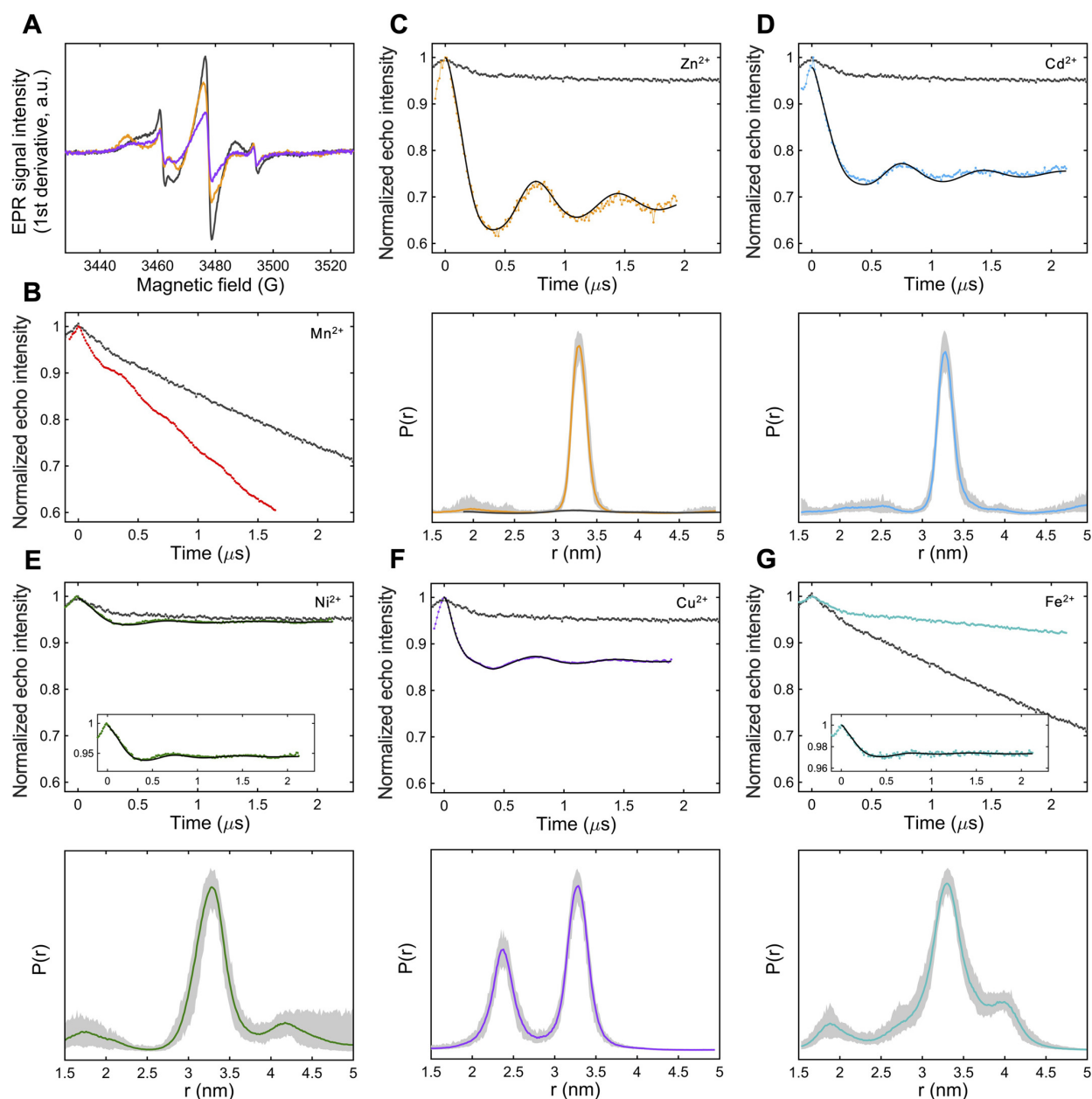


Figure 6. Room temperature continuous-wave EPR measurements (A) performed on apo MamM C267_{SL} (gray), C267_{SL}-Zn²⁺ (orange), and C267_{SL}-Cu²⁺ (purple) revealing quenching of the EPR signal intensity upon addition of Cu²⁺ as compared with both the apo and Zn²⁺-bound states. B–G, four-pulse PELDOR experiments of doubly spin-labeled MamM CTD C267_{SL}, containing Cys-275 as the labeling position, for the measurement of inter-spin label distances. Results were obtained for samples incubated with a range of divalent transition metal ions, including Mn²⁺ (B), Zn²⁺ (C), Cd²⁺ (D), Ni²⁺ (E), Cu²⁺ (F), and Fe²⁺ (G). *Top panel:* The normalized background-corrected dipolar evolution of metal-bound samples (bright colors) presented together with the apo spin-labeled protein (gray). Exceptions: For both Mn²⁺ and Fe²⁺, the nonbaseline-corrected time traces are shown to clearly demonstrate the change in decay caused by specific versus nonspecific metal binding. *Bottom panel:* normalized distance distributions (where appropriate) of metal-bound C267_{SL}. The distance distribution of apo C267_{SL} is also shown for reference, plotted against the Zn²⁺-bound results and shown in dark gray. Standard deviations in the form of 95% confidence intervals are also given in each case and represented by light gray areas around the lines. Fitting of all dipolar evolutions and subsequent distance distributions were calculated with DEERNet.

bound copper. MMM analysis of the Cu²⁺-bound structure (Fig. 5D) predicts an inter-spin label distance of 3.23 nm, whereas the resultant distance distribution of the experimental PELDOR trace reveals a bimodal distribution with two mean

distances of 2.4 and 3.24 nm (Fig. 6F). At first glance this could imply two different inter-spin label (SL–SL) distances; however, because the Cu²⁺ species overlaps with the spin label in the EPR spectrum, this PELDOR trace may not only comprise

MamM CTD exhibits metal-type dependent binding modes

SL-SL distances but in addition contributions from SL-Cu²⁺ and Cu²⁺-Cu²⁺ distances. The 3.24-nm distance agrees well with both the predicted MMM SL-SL distance and other metal-bound PELDOR results, and hence we assign this to the SL-SL distance, again indicative of a single closed V-shaped conformation. The second observed distance at 2.4 nm could correspond to either a Cu²⁺-SL or Cu²⁺-Cu²⁺ distance. Considering the fluorescence data, which indicates binding of Cu²⁺ in the CS, and previous PELDOR measurements in which conformational change was shown to be associated with metal binding to the PSs (21), we propose that binding to the CS is silent in terms of heat change and that the Cu²⁺-bound crystal structure represents well both the conformation in solution and the copper binding sites.

Current experimental evidence for Fe²⁺ binding to MamM CTD, both via biophysical and structural methods, is very limited (5). This is due largely to the difficulty associated with preparation of bound Fe²⁺ samples; when working at basic pH, Fe²⁺ is readily oxidized to Fe³⁺, which notoriously forms insoluble complexes. Nevertheless, it proved possible to prepare an Fe²⁺-incubated sample using argon gas to largely eliminate any oxygen present with C267S_{SL} and an excess of Fe²⁺ (×5 excess, forming C267S_{SL}-Fe²⁺). Although there was slight indication of Fe³⁺ precipitation, PELDOR yielded results clearly different from the apo state: nonbaseline-corrected, time-domain data are shown to indicate these differences more clearly, and a single distance distribution with a mean distance of 3.3 nm is obtained (Fig. 6G). Despite the low modulation depth, probably due to very low levels of Fe²⁺ binding, the difference to the apo state is quite apparent, providing a good indication of a more rigid structure although the distance distribution is slightly broader than for Zn²⁺/Ni²⁺/Cu²⁺/Cd²⁺. Interestingly, the sample became slightly yellow upon incubation with the metal salt, which could not be reproduced upon incubation of the buffer alone with the same concentration of salt. It is therefore likely that the protein also catalyzes the iron oxidation process, resulting in the formation of unbound (or nonspecifically bound) yellow Fe³⁺ complexes. More importantly, this is the first high-resolution structural observation of Fe²⁺-binding to MamM CTD which corroborates our earlier findings of iron binding and the associated conformational change of MamM CTD upon binding (5). These results indicate that studies with other metals (including Zn²⁺, for which substantial work has been done previously) do reflect biologically relevant states of the protein, because all the metal-bound proteins appear to adopt a similar conformation in solution, which is most likely similar to the crystal Cu²⁺-bound conformation.

In contrast, the addition of paramagnetic Mn²⁺ (×5 excess) to C267S_{SL} does not result in a change in the dipolar evolution time trace seen when other divalent metals are added (Fig. 6B). There are no periodic oscillations arising from dipolar coupling between the two nitroxide moieties but simply only a slightly modified decay function modulated by the presence of deuterated glycerol (which are not present when protonated glycerol is used, data not shown). Thus, PELDOR spectroscopy, taken together with the tryptophan fluorescence and ITC data, provide clear evidence of the inability of MamM CTD to bind Mn²⁺.

Previous analysis of metal-binding sites of all PDB structures revealed that the typical amino acid composition in Mn²⁺ binding sites is similar to those that bind Fe²⁺, Cd²⁺, Zn²⁺, and Ni²⁺ (7), and MamM's cation binding sites are composed of histidine residues and aspartic or glutamic acids; however, MamM CTD can bind these cations but still discriminates against Mn²⁺. This phenomenon cannot be rationalized by comparison of other properties important for metal binding such as dominant binding geometry (tetrahedral, octahedral, or trigonal plane), metal size, solvation sphere, or electrostatic charges. Attempts to actively modify the MamM CTD metal-binding sites to allow Mn²⁺ binding *in vitro* failed to alter the CTD discrimination abilities (28). Overall, the common binding capabilities of Mn²⁺ throughout biology indicate that the observed discrimination by MamM CTD must be more complex and may include quantum effects that cannot be calculated for this system. The inability of MamM CTD to bind Mn²⁺, in contrast to the other cations, is further indication of an evolutionary pressure that discriminates against Mn²⁺.

Our results ultimately raise the question of what benefits there are for MamM CTD as a result of its discrimination against Mn²⁺. Mn²⁺ has a greater affinity to magnetite, both when abiotically synthesized and in magnetotactic bacteria, as compared with other transition metal cations (29). It was shown that for iron oxides formation *in vitro*, the presence of manganese leads to oxidation of Fe²⁺ to Fe³⁺ by manganese oxides (30); thus it may change the proper magnetite formation in the magnetosome by changing the Fe²⁺:Fe³⁺ balance. Furthermore, uncultivated magnetotactic bacteria, which are exposed to very high concentrations of manganese, can incorporate it into the iron-based magnetic particles, yet to a limited amount (31). Based on our results and these observations, we propose that to avoid overflow of manganese in the magnetosomes, leading to sequential incorporation of it into magnetite or altering magnetite formation by affecting the oxidation/reduction of iron, evolutionary demands have driven MTB to develop a CDF CTD-related mechanism for discrimination against manganese. Such discrimination will prevent both Mn²⁺ ion transport and any initial activation of the transporter by not facilitating any binding of Mn²⁺ to the CTD (Fig. 7). Whereas all the other tested metals were shown to bind to the CTD, leading to a tighter conformation similar to that of the iron-bound CTD, only rarely are contaminated magnetite particles formed in MTB *in vivo*. Our results would thus indicate that any discrimination against other metals should occur at the transport site level, which is unique compared with other CDF proteins (contains a rare tyrosine residue) (7, 18) (Fig. 7).

Concluding remarks

The participation of CDF protein CTDs in metal selectivity was previously studied at the cellular level. Here, for the first time, structural characterization of a CDF protein CTD in the presence of various metals reveals the direct effect of different metals on the CTD, with each metal exhibiting distinct binding parameters, sites, and conformation. This study uniquely includes high-resolution structural data both in the crystalline form and in solution, which are complementary to one another

MamM CTD exhibits metal-type dependent binding modes

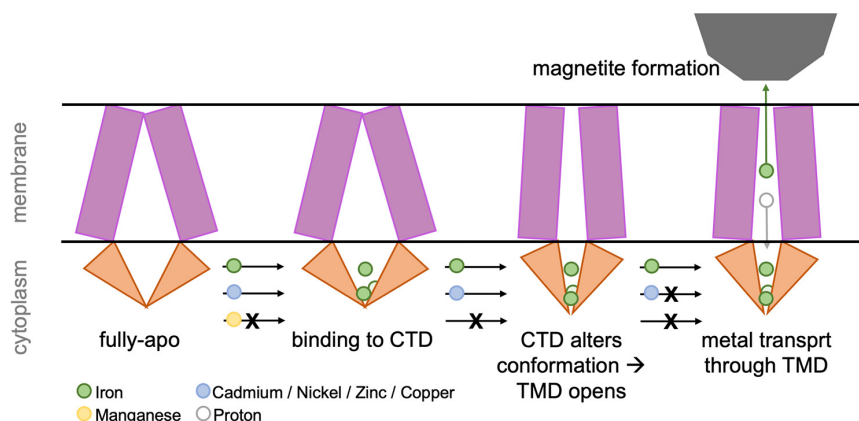


Figure 7. Metal-binding mechanism to MamM CTD. MamM CTD can bind divalent iron (green sphere), cadmium, zinc, nickel, and copper (blue sphere), but not manganese (yellow sphere). Metal binding to the CTD leads to its tighter packing, which in turn facilitates the conformational change of the transmembrane domain. However, only iron can be transported through this domain, in exchange for protons, which allows the synthesis of pure magnetic particles.

and in combination exhibit the differences between the binding modes to the native metal and other metals. Furthermore, these results present subtle differences in the degree-of-closure between the apo and bound CTDs among different CDF proteins, suggesting that it not only depends on the protein studied but, importantly, also on the nature of the metal. Based on these distinct binding modes, we propose that the CDF CTD is the first recognition site able to alter or stop the transport of the unwanted metals prior to their reaching the transport site, which provides a second recognition site. This could be because of the slightly different CTD closed conformations induced by differences in the metal size, metal coordination geometry, or binding of the flexible C-terminal tail. Such an additional ability of the CTD to discriminate would also mean that the CDF is not simply limited by a single site, making it hard to discriminate against all the different cations that exist in the cytoplasm because of their relative similarity. One surprising piece of evidence for such varying levels of discrimination is found in the MamM system where manganese is fully discriminated against by the CTD alone, which may reflect the specific evolutionary pressure to specifically transport iron.

Experimental procedures

Expression and purification

mamM CTD gene (UniProt Q6NE57 residues 215–318) was previously cloned into pET28a(+) vector (Novagen, Merck Biosciences, Germany) (32). MamM CTD was expressed and purified as previously described for MamM CTD M250L mutant (20, 32). The C267S mutation for PELDOR studies was applied to the pET28a-MamM CTD vector, expressed, and purified as previously described (21). For all experiments, protein concentration was determined by measuring protein absorption at 280 nm.

Crystallization and structure determination

20 mg ml⁻¹ purified MamM CTD in buffer containing 10 mM Tris, pH 8.0, 150 mM NaCl, 5 mM β -mercaptoethanol, and either 3.375 mM CuSO₄ (Cu²⁺-bound structure), 3.375 mM

CdCl₂ (Cd²⁺-bound structure), or 3.375 mM NiCl₂ (Ni²⁺-bound structure) was crystallized using the vapor diffusion method at 293 K (0.3 μ l of protein with 0.3 μ l of reservoir solution for all protein-metal pairs). Crystals were harvested with or without treatment of cryo-agent and flash-frozen in liquid nitrogen. Data collection was performed on a single crystal at 100 K. All cryo conditions, data collection details, data reduction, and scaling, phasing, and refinement details are given in Tables S1–S3 (37–43).

Least-squares overlaps

All CDFs' CTD structures were overlapped and RMSDs were calculated using the iterative magic fit tool or the domain fragment alternate fit tool of Swiss-PdbViewer 4.1.0 (33). MamM CTD structures and overlapped structures' figures were prepared using UCSF Chimera package, version 1.12 (34) or PyMOL, version 2.0.6 (PyMOL Molecular Graphics System, Schrödinger).

Fluorescence spectrometry

Changes in tryptophan intrinsic fluorescence were recorded using Fluorolog®-3 (HORIBA, Edison, NJ, USA) equipped with a quartz cell with a 1-cm optical path at ambient temperature. Samples of 1 ml of MamM CTD at 5 μ M concentration in buffer A (containing 10 mM Tris, pH 8.0, and 150 mM NaCl) were titrated using 2.5 mM metal solution (CdCl₂, NiCl₂, MnCl₂, and CuSO₄) in the same buffer to reach different concentrations. Samples were measured at λ_{ex} 297 nm, and the emission spectrum for each metal concentration was recorded at 310–450 nm. For each metal, the titration was replicated three times, and each spectrum was fitted to Extreme function by OriginPro (R-Square > 0.98) (OriginLab Corporation). The maximum wavelength (wavelength at maximum intensity) and the intensity at that wavelength (maximum intensity) were averaged for each metal concentration. Errors are reported as the standard deviation.

MamM CTD exhibits metal-type dependent binding modes

Isothermal titration calorimetry

ITC measurements were performed in a low-volume Nano ITC calorimeter (TA Instruments) at 298 K. Proteins and metals (CdCl_2 , NiCl_2 , MnCl_2 , and CuSO_4) were prepared in buffer A. MamM CTD was diluted to 50 μM concentration and was titrated with 1.4 mM (NiCl_2 , MnCl_2 , and CuSO_4) or 1 mM (CdCl_2) metal solutions. The protein samples were injected into the instrument cell (170 μl), and 20 aliquots of 2.5 μl of the suitable metal solution were titrated into the cell every 350 s. For each protein, three independent titrations were measured. As a control, each metal solution was titrated into buffer A under the same experimental conditions, and all measurements were compared with a DDW-containing reference cell. Data were analyzed using TA NanoAnalyze Data Analysis software, version 3.7.5 (TA Instruments). The data of each measurement were fitted to an independent model combined with a blank constant model, and the given thermodynamics values are the average of the three different titrations for each protein. Errors are reported as standard deviation.

Site-directed spin labeling

For site-directed spin labeling, protein solutions in buffer A were degassed under argon and treated with degassed DTT (1 mM) for 4 h at 4 °C with gentle agitation. Subsequently, protein samples were washed once using ZebaTM Spin desalting columns (7 kDa molecular weight cutoff, 2 ml) (Thermo Fisher Scientific) equilibrated with degassed buffer A. Protein was then labeled overnight at 4 °C with (1-oxyl-2,2,5,5-tetramethylpyrrolidin-3-yl)methylthiosulfonate spin label (MTSL, 20 $\times$ excess) (Toronto Research Chemicals) with gentle agitation. To remove excess unbound MTSL, the samples were washed twice using ZebaTM Spin desalting columns equilibrated with degassed buffer A.

Continuous-wave X-band EPR spectroscopy

For metal-bound samples, freshly spin-labeled C267S MamM CTD in buffer A at $\sim 70 \mu\text{M}$ was used. Stock solutions of metal salts ($\text{ZnSO}_4 \cdot 5\text{H}_2\text{O}$ and $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$) were prepared in H_2O to a final concentration of 100 mM. Aliquots (15 μl) of spin-labeled protein were incubated with a 3-fold excess of each metal salt for 40 min on ice. Aliquots of $\sim 10 \mu\text{l}$ were transferred to 0.8-mm (outer diameter) capillary tubes for measurement.

The ambient temperature setup for X-band cw-EPR consisted of a Bruker E500 eleXsys spectrometer fitted with an ER 4123D (dielectric RT cw-EPR) resonator. The following measuring parameters were used for data acquisition: a microwave frequency of 9.758 GHz, a modulation frequency of 100 KHz, a modulation amplitude of 1 gauss, and a microwave power of 0.2 milliwatt.

Pulsed EPR spectroscopy

Spin-labeled C267S MamM CTD samples were lyophilized and subsequently dissolved in D_2O to a final concentration of 250 μM (Cd^{2+} - and Fe^{2+} -bound samples) or 500 μM (Mn^{2+} -, Ni^{2+} -, and Cu^{2+} -bound samples). All EPR samples were pre-

pared with 30% Glycerol- d_8 (Sigma-Aldrich). Stock solutions of metal salts (CdCl_2 , $\text{MnCl}_2 \cdot 7\text{H}_2\text{O}$, NiCl_2 , $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, and $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$) were prepared to a final concentration of 100 mM in D_2O . Spin-labeled samples were incubated with a 5-fold (Cd^{2+} -, Mn^{2+} -, Ni^{2+} -, and Fe^{2+} -bound samples) or 3-fold (Cu^{2+} -bound) excess according to the dimeric protein concentration. Incubation was carried out at room temperature for 1 h, before the addition of glycerol- d_8 to a final concentration of $\sim 30\%$. 80–100- μl samples were transferred to EPR quartz tubes (Wilma SQ-707; Wilma-LabGlass) and flash-frozen in liquid nitrogen. For the Fe^{2+} -bound sample, each step was carried out under a steady stream of argon to eliminate oxygen during sample preparation. X-band pulsed EPR spectra were recorded on a Bruker E580 spectrometer (Bruker, Rheinstetten, Germany) using a Bruker MD5-W1 EPR probehead equipped with a self-modified cryogen-free cryostat (Advanced Research Systems, Macungie, PA, USA). The microwave pulses were amplified using a 1-kW TWT amplifier (Applied Systems Engineering, Fort Worth, TX, USA). All EPR experiments were carried out at 50 K. The field-swept spectrum was obtained by integrating the Hahn echo signal as a function of the magnetic field after a two-pulse sequence (10 ns, τ (120 ns), 20 ns). For PELDOR experiments a four-pulse sequence was applied. Observer pulses ($\pi/2$ and π) were set to 20 ns, with a pump pulse of 16 ns. The frequency offset $\Delta\nu$ ($\nu_{\text{obs}} - \nu_{\text{pump}}$) was set to between 70–80 MHz for SL-SL measurements. The observer pulse separation τ_1 was 120 ns in all cases apart from with Ni^{2+} , where τ_1 was 350 ns, whereas τ_2 varied between 1300–2200 ns, and echo signals were integrated using a video amplifier bandwidth of 20 MHz. The pump pulse was stepped out in 12-ns intervals to a total number of points that depended upon the τ_1 and τ_2 used. Nuclear modulation effects were suppressed using a systematic variation of the inter-pulse delay time τ_1 and an appropriate phase cycling routine.

PELDOR data processing was performed using both Tikhonov regularization and DEERNet (35) within the Matlab-based DEER Analysis (36) toolbox. The recently developed DEERNet method uses artificial neural networks for prediction of an inter-spin label distance distribution and gives a measure of uncertainty of the resulting distance distributions represented by 95% confidence bounds, which are indicated as the *gray bands* in Fig. 6.

Data availability

All structural data are available in RCSB Protein Data Bank under the accession numbers 6GP6, 6GMT, and 6GMV. All other data are contained within this article and in the supporting information.

Acknowledgments—We thank Dr. Anat Shahar for help with the crystallographic data collection and analysis.

Author contributions—S. B.-Z., J. H., F. M., and R. Z. conceptualization; S. B.-Z., J. H., A. F., and F. M. data curation; S. B.-Z., J. H., A. F., S. K., F. M., and R. Z. formal analysis; S. B.-Z., J. H., A. F., S. K., F. M., and R. Z. validation; S. B.-Z., J. H., F. M., and R. Z. investigation;

MamM CTD exhibits metal-type dependent binding modes

S. B.-Z., J. H., F. M., and R. Z. visualization; S. B.-Z., J. H., A. F., S. K., F. M., and R. Z. methodology; S. B.-Z., J. H., F. M., and R. Z. writing-original draft; S. B.-Z. project administration; S. B.-Z., J. H., A. F., S. K., F. M., and R. Z. writing-review and editing; F. M. and R. Z. resources; F. M. and R. Z. supervision; F. M. and R. Z. funding acquisition.

Funding and additional information—S. B. Z. and R. Z. are supported by the Israel Ministry of Science, Technology and Space, the Israel Science Foundation (Grant 167/16), the European Molecular Biology Organization, and the EU (CMST COST Action CM1306 Understanding Movement and Mechanism in Molecular Machines). J. H. was supported by the UKRI Biotechnology and Biological Sciences Research Council Norwich Research Park Biosciences Doctoral Training Partnership (Grant BB/M011216/1). F. M. was supported in part by the Royal Society (F. M. was a Wolfson Research Merit Award Holder). J. H. and F. M. are also supported by the EU (CMST COST Action CM1306 Understanding Movement and Mechanism in Molecular Machines).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CDF, cation diffusion facilitator; CS, central site; CTD, C-terminal domain; EPR, electron paramagnetic resonance; ITC, isothermal titration calorimetry; MMM, Multiscale Modelling of Macromolecules; MTB, magnetotactic bacteria; MTSL, (1-oxyl-2,2,5,5-tetramethylpyrrolidin-3-yl) methylthiosulfonate spin label; PDB, Protein Data Bank; PELDOR, pulsed electron-electron double resonance; PS, peripheral site; SL, spin label; TMD, transmembrane domain; RMSD, root mean square deviation; cw-EPR, continuous-wave EPR.

References

- Lu, M., and Fu, D. (2007) Structure of the zinc transporter YiiP. *Science* **317**, 1746–1748 [CrossRef Medline](#)
- Lu, M., Chai, J., and Fu, D. (2009) Structural basis for autoregulation of the zinc transporter YiiP. *Nat. Struct. Mol. Biol.* **16**, 1063–1067 [CrossRef Medline](#)
- Coudray, N., Valvo, S., Hu, M., Lasala, R., Kim, C., Vink, M., Zhou, M., Provasi, D., Filizola, M., Tao, J., Fang, J., Penczek, P. A., Ubarretxena-Belandia, I., and Stokes, D. L. (2013) Inward-facing conformation of the zinc transporter YiiP revealed by cryoelectron microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2140–2145 [CrossRef Medline](#)
- Lopez-Redondo, M. L., Coudray, N., Zhang, Z., Alexopoulos, J., and Stokes, D. L. (2018) Structural basis for the alternating access mechanism of the cation diffusion facilitator YiiP. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3042–3047 [CrossRef Medline](#)
- Zeytuni, N., Uebe, R., Maes, M., Davidov, G., Baram, M., Raschdorf, O., Nadav-Tsubery, M., Kolusheva, S., Bitton, R., Goobes, G., Friedler, A., Miller, Y., Schüler, D., and Zarivach, R. (2014) Cation diffusion facilitators transport initiation and regulation is mediated by cation induced conformational changes of the cytoplasmic domain. *PLoS ONE* **9**, e92141 [CrossRef Medline](#)
- Kolaj-Robin, O., Russell, D., Hayes, K. A., Pembroke, J. T., and Soulimane, T. (2015) Cation diffusion facilitator family: structure and function. *FEBS Lett.* **589**, 1283–1295 [CrossRef Medline](#)
- Barber-Zucker, S., Shaan, B., and Zarivach, R. (2017) Transition metal binding selectivity in proteins and its correlation with the phylogenomic classification of the cation diffusion facilitator protein family. *Sci. Rep.* **7**, 16381 [CrossRef Medline](#)
- Cubillas, C., Vinuesa, P., Tabche, M. L., and García-de los Santos, A. (2013) Phylogenomic analysis of cation diffusion facilitator proteins uncovers $\text{Ni}^{2+}/\text{Co}^{2+}$ transporters. *Metalomics* **5**, 1634–1643 [CrossRef Medline](#)
- Montanini, B., Blaudez, D., Jeandroz, S., Sanders, D., and Chalot, M. (2007) Phylogenetic and functional analysis of the cation diffusion facilitator (CDF) family: improved signature and prediction of substrate specificity. *BMC Genomics* **8**, 107 [CrossRef Medline](#)
- Levy, M., Elkoshi, N., Barber-Zucker, S., Hoch, E., Zarivach, R., Hershfinkel, M., and Sekler, I. (2019) Zinc transporter 10 (ZnT10)-dependent extrusion of cellular Mn^{2+} is driven by an active Ca^{2+} -coupled exchange. *J. Biol. Chem.* **294**, 5879–5889 [CrossRef Medline](#)
- Kambe, T., Yamaguchi-Iwai, Y., Sasaki, R., and Nagao, M. (2004) Overview of mammalian zinc transporters. *Cell. Mol. Life Sci.* **61**, 49–68 [CrossRef Medline](#)
- Huang, L., and Tepasorndech, S. (2013) The SLC30 family of zinc transporters - a review of current understanding of their biological and pathophysiological roles. *Mol. Aspects Med.* **34**, 548–560 [CrossRef Medline](#)
- Uebe, R., and Schüler, D. (2016) Magnetosome biogenesis in magnetotactic bacteria. *Nat. Rev. Microbiol.* **14**, 621–637 [CrossRef Medline](#)
- Barber-Zucker, S., Keren-Khadmy, N., and Zarivach, R. (2016) From invagination to navigation: The story of magnetosome-associated proteins in magnetotactic bacteria. *Protein Sci.* **25**, 338–351 [CrossRef Medline](#)
- Komeili, A. (2012) Molecular mechanisms of compartmentalization and biomineralization in magnetotactic bacteria. *FEMS Microbiol. Rev.* **36**, 232–255 [CrossRef Medline](#)
- Komeili, A., Li, Z., Newman, D. K., and Jensen, G. J. (2006) Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK. *Science* **311**, 242–245 [CrossRef Medline](#)
- Barber-Zucker, S., and Zarivach, R. (2017) A look into the biochemistry of magnetosome biosynthesis in magnetotactic bacteria. *ACS Chem. Biol.* **12**, 13–22 [CrossRef Medline](#)
- Uebe, R., Junge, K., Henn, V., Poxleitner, G., Katzmman, E., Plitzko, J. M., Zarivach, R., Kasama, T., Wanner, G., Pósfai, M., Böttger, L., Matzanke, B., and Schüler, D. (2011) The cation diffusion facilitator proteins MamB and MamM of *Magnetospirillum gryphiswaldense* have distinct and complex functions, and are involved in magnetite biomineralization and magnetosome membrane assembly. *Mol. Microbiol.* **82**, 818–835 [CrossRef Medline](#)
- Zeytuni, N., Uebe, R., Maes, M., Davidov, G., Baram, M., Raschdorf, O., Friedler, A., Miller, Y., Schüler, D., and Zarivach, R. (2014) Bacterial magnetosome biomineralization—a novel platform to study molecular mechanisms of human CDF-related type-II diabetes. *PLoS ONE* **9**, e97154 [CrossRef Medline](#)
- Barber-Zucker, S., Uebe, R., Davidov, G., Navon, Y., Sherf, D., Chill, J. H., Kass, I., Bitton, R., Schüler, D., and Zarivach, R. (2016) Disease-homologous mutation in the cation diffusion facilitator protein MamM causes single-domain structural loss and signifies its importance. *Sci. Rep.* **6**, 31933 [Medline CrossRef](#)
- Barber-Zucker, S., Hall, J., Mangapuram, S. V., Kass, I., Kolusheva, S., MacMillan, F., Zarivach, R., and Henn, A. (2019) Metal binding to the dynamic cytoplasmic domain of the cation diffusion facilitator (CDF) protein MamM induces a 'locked-in' configuration. *FEBS J.* **286**, 2193–2215 [CrossRef Medline](#)
- Cotrim, C. A., Jarrott, R. J., Martin, J. L., and Drew, D. (2019) A structural overview of the zinc transporters in the cation diffusion facilitator family. *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 357–367 [CrossRef Medline](#)
- Cherezov, V., Höfer, N., Szebenyi, D. M. E., Kolaj, O., Wall, J. G., Gillilan, R., Srinivasan, V., Jaroniec, C. P., and Caffrey, M. (2008) Insights into the mode of action of a putative zinc transporter CzcB in *Thermus thermophilus*. *Structure* **16**, 1378–1388 [CrossRef Medline](#)
- Xue, J., Xie, T., Zeng, W., Jiang, Y., and Bai, X. (2020) Cryo-EM structures of human ZnT8 in both outward- and inward-facing conformations. *Elife* **9**, e58823 [CrossRef Medline](#)
- Uebe, R., Keren-Khadmy, N., Zeytuni, N., Katzmman, E., Navon, Y., Davidov, G., Bitton, R., Plitzko, J. M., Schüler, D., and Zarivach, R. (2018) The dual role of MamB in magnetosome membrane assembly and magnetite biomineralization. *Mol. Microbiol.* **107**, 542–557 [CrossRef Medline](#)

MamM CTD exhibits metal-type dependent binding modes

26. Mullen, A., Hall, J., Diegel, J., Hassan, I., Fey, A., and MacMillan, F. (2016) Membrane transporters studied by EPR spectroscopy: structure determination and elucidation of functional dynamics. *Biochem. Soc. Trans.* **44**, 905–915 [CrossRef Medline](#)
27. Jeschke, G. (2018) MMM: A toolbox for integrative structure modeling. *Protein Sci.* **27**, 76–85 [CrossRef Medline](#)
28. Barber-Zucker, S., Shahar, A., Kolusheva, S., and Zarivach, R. (2020) The metal binding site composition of the cation diffusion facilitator protein MamM cytoplasmic domain impacts its metal responsivity. *Sci. Rep.* **10**, 14022 [CrossRef Medline](#)
29. Amor, M., Busigny, V., Durand-Dubief, M., Tharaud, M., Ona-Nguema, G., Gélalbert, A., Alphandéry, E., Menguy, N., Benedetti, M. F., Chebbi, I., and Guyot, F. (2015) Chemical signature of magnetotactic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1699–1703 [CrossRef Medline](#)
30. Krishnamurti, G. S. R., and Huang, P. M. (1988) Influence of manganese oxide minerals on the formation of iron oxides. *Clays Clay Miner.* **36**, 467–475 [CrossRef](#)
31. Keim, C. N., Lins, U., and Farina, M. (2009) Manganese in biogenic magnetite crystals from magnetotactic bacteria. *FEMS Microbiol. Lett.* **292**, 250–253 [CrossRef Medline](#)
32. Zeytuni, N., Offer, T., Davidov, G., and Zarivach, R. (2012) Crystallization and preliminary crystallographic analysis of the C-terminal domain of MamM, a magnetosome-associated protein from *Magnetospirillum gryphiswaldense* MSR-1. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **68**, 927–930 [CrossRef Medline](#)
33. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 [CrossRef Medline](#)
34. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 [CrossRef Medline](#)
35. Worswick, S. G., Spencer, J. A., Jeschke, G., and Kuprov, I. (2018) Deep neural network processing of DEER data. *Sci. Adv.* **4**, eaat5218 [CrossRef Medline](#)
36. Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* **30**, 473–498 [CrossRef](#)
37. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 [CrossRef Medline](#)
38. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 [CrossRef Medline](#)
39. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 [CrossRef Medline](#)
40. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 [CrossRef Medline](#)
41. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 [CrossRef Medline](#)
42. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 [CrossRef Medline](#)
43. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 [CrossRef Medline](#)

III Published research - STD NMR investigation of substrate binding in GltPh

The following publication (of which I am the primary author) includes the study of substrate binding by GltPh, via STD NMR.²⁶⁴ This research demonstrates for the first time, the application of STD NMR for determination of dissociation constants for two competing substrates of a membrane protein, in multiple environments. Differences in determined K_D 's, as well as insights into cooperativity exhibited by protomers of the trimeric GltPh, are discussed. My contributions include all experimental data acquisition, analysis of these data with the input of other authors, composition of the text and all figures included within the publication. The results included within this manuscript have been edited and reformatted for inclusion in chapter 4.9 of this thesis.



OPEN

Saturation transfer difference NMR on the integral trimeric membrane transport protein GltPh determines cooperative substrate binding

Jenny L. Hall¹, Azmat Sohail², Eurico J. Cabrita³, Colin Macdonald¹, Thomas Stockner², Harald H. Sitte², Jesus Angulo⁴ & Fraser MacMillan^{1✉}

Saturation-transfer difference (STD) NMR spectroscopy is a fast and versatile method which can be applied for drug-screening purposes, allowing the determination of essential ligand binding affinities (K_D). Although widely employed to study soluble proteins, its use remains negligible for membrane proteins. Here the use of STD NMR for K_D determination is demonstrated for two competing substrates with very different binding affinities (low nanomolar to millimolar) for an integral membrane transport protein in both detergent-solubilised micelles and reconstituted proteoliposomes. GltPh, a homotrimeric aspartate transporter from *Pyrococcus horikoshii*, is an archaeal homolog of mammalian membrane transport proteins—known as excitatory amino acid transporters (EAATs). They are found within the central nervous system and are responsible for fast uptake of the neurotransmitter glutamate, essential for neuronal function. Differences in both K_D 's and cooperativity are observed between detergent micelles and proteoliposomes, the physiological implications of which are discussed.

L-glutamate is the primary excitatory neurotransmitter in the mammalian central nervous system. Excitatory amino acid transporters (EAATs) are membrane proteins that play an important role in regulation of the glutamate concentration within the synaptic cleft of glutamatergic neurons, thereby terminating synaptic transmission. Efficient uptake prevents a build-up of high extracellular glutamate levels leading to excitotoxicity (one of the hallmark features of Amyotrophic lateral sclerosis (ALS)), and also allows glutamate to be recycled¹. Substrate transport by EAATs is energised by the co-transport of up to three Na^+ ions and one proton. The counter-transport of one K^+ ion completes the putative transport cycle^{2,3}. In addition, EAATs exhibit uncoupled chloride-conductance, making them anion-selective ion channels⁴. Various neurological diseases are linked to their dysfunction in glutamatergic synapses, including ALS, stroke, epilepsy and Huntington's disease¹.

GltPh, a bacterial homolog of mammalian glutamate transporters from *Pyrococcus horikoshii* was the first member of the Solute Carrier 1 (SLC1) family of membrane transporters to have been crystallized while very recently a first crystal structure of human EAAT1 was determined facilitating extrapolation of results obtained for GltPh to its human counterparts^{5–8}. GltPh is often used to investigate structure–function relationships of the SLC1 transporter family via spectroscopic techniques including electron paramagnetic resonance (EPR) spectroscopy and single molecule fluorescence resonance energy transfer (smFRET)⁹. It shares approximately 35% amino acid sequence identity with human EAATs; and importantly, many of the amino acids involved in ion and substrate binding are conserved, making it an excellent model system¹⁰. The outward-occluded state, captured as the first crystal structure of GltPh, revealed a homotrimer assembly with a bowl-shaped extracellular basin whose surface is hydrophobic and reaches halfway down the trimer's height (Fig. 1)^{6–8}.

Each protomer consists of a trimerisation domain involved in stabilising subunit interactions, and a transport domain, containing the substrate and sodium ion binding sites. Further crystal conformations have been

¹Henry Wellcome Unit for Biological EPR, School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. ²Institute of Pharmacology, Medical University of Vienna, Währingerstrasse 13A, 1090 Vienna, Austria. ³UCIBIO, Chemistry Department, Faculty of Sciences and Technology, NOVA University of Lisbon, 2829-516 Caparica, Portugal. ⁴School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. ✉email: fraser.macmillan@uea.ac.uk

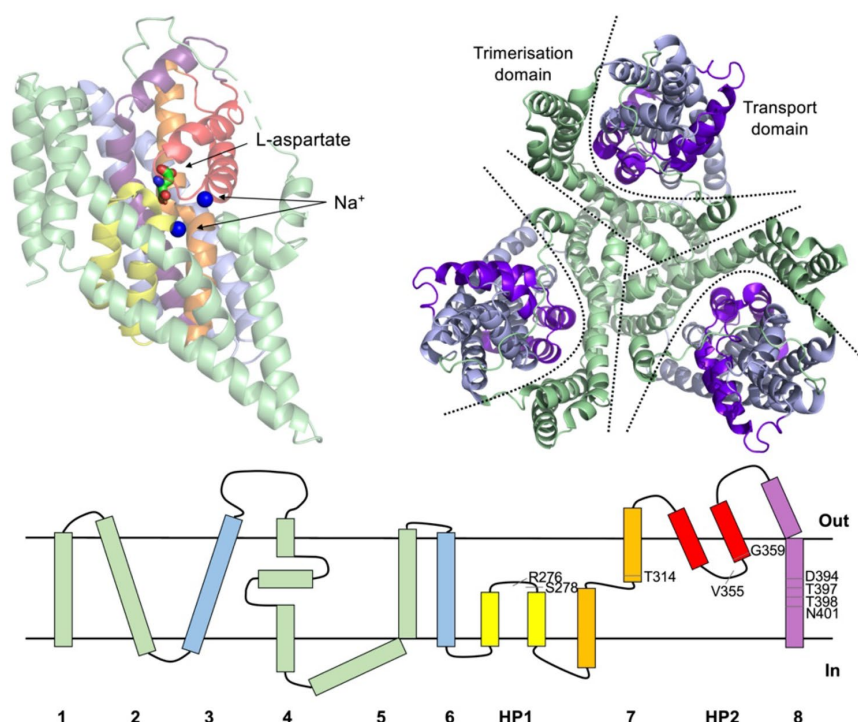


Figure 1. (Top) left: Cartoon representation of a single GltPh protomer with bound aspartate (sticks) and two Na^+ ions (blue spheres). HP1, TM7, HP2 and TM8 containing ligand binding residues are shown in yellow, orange, red and violet respectively. (Top) right: extracellular view of GltPh with trimerisation domains shown in green and transport domains in blue. Individual protomers are separated by straight lines, whilst transport domains are separated from trimerisation domains by curved lines. (Bottom) a cartoon representation of the GltPh fold highlighting the residues involved in substrate binding. The protein images were prepared using PyMOL (The PyMOL Molecular Graphics System; <https://www.pymol.org>).

solved in the outward-open¹⁰, inward-occluded^{11,12}, as well as intermediate¹³ conformations, revealing large-scale translational movement upon substrate binding and transport. The substrate binding site (Fig. 1) is formed by the tips of the two re-entrant loops (HP1 and HP2) and unwound regions of transmembrane (TM) helices 7 and 8¹⁰.

Whilst membrane proteins such as GltPh account for one-third of all proteins encoded by the genomes of most organisms¹⁴, they continue to prove challenging for structural and mechanistic studies. This difficulty extends to include medium to high throughput determination of substrate and inhibitor binding affinities, essential for drug target screening. Until now methods such as fluorescence-based assays^{10,15,16}, isothermal titration calorimetry^{12,15,17} and uptake experiments¹⁸ have been utilised in the determination of substrate K_M and K_D values in both detergent-solubilised and membrane-reconstituted protein environments. Such methods have yielded results that vary considerably with Na^+ concentration, clearly highlighting GltPh's sodium dependency for substrate binding. K_D values for L-aspartate vary from 2–380 nM depending on Na^+ concentration (200 mM and 10 mM respectively), while for the much weaker binding ligand, L-glutamate, they vary between 120–250 μM in 200 mM Na^+ .

Here we present a novel approach for the determination of K_D values for both L-aspartate and L-glutamate using saturation transfer difference (STD) NMR spectroscopy¹⁹. STD NMR is a sensitive method for ligands of medium to weak affinity, relying on the exchange between bound and free states of the ligand (K_D values typically ranging from high nM to low mM). During the STD NMR experiment, resonances of the receptor molecule are selectively saturated, and this saturation rapidly spreads via spin diffusion throughout the protein. A ligand which binds to this protein (for between 0.1 to 100 ms) will receive saturation to some of or all its protons, with those in closest proximity to the protein receiving the most²⁰. When the ligand dissociates, this saturation is transferred to the bulk solution, where a reduction in ligand NMR signal intensity is observed. The difference between a spectrum without saturation (reference spectrum), and the saturated experiment gives rise to the STD spectrum, in which only signals of ligand molecules that interacted with the protein are observed²¹. The work presented here significantly extends the application of STD NMR to membrane proteins and demonstrates its use as an alternative, fast method for determination not only of K_D values for weakly binding ligands but also for very tightly binding ligands, not previously thought accessible through STD NMR. In addition to demonstrating

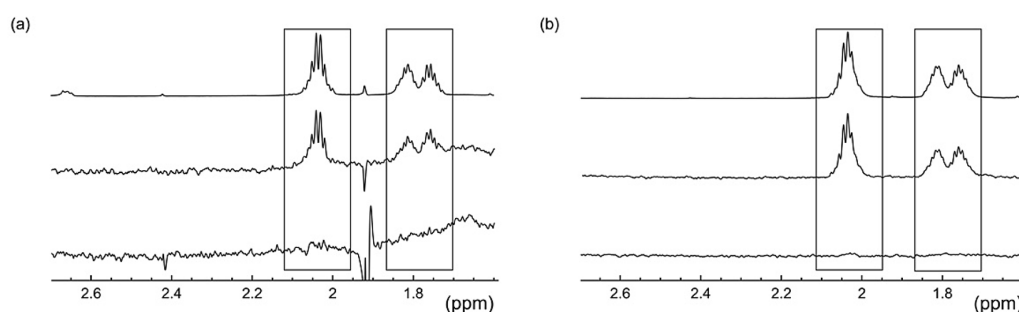


Figure 2. Selected spectral region of the reference ^1H NMR (top) and corresponding STD NMR spectrum of GltPh incubated with 0.5 mM L-glutamate in the presence (middle) and absence (bottom) of 100 mM Na^+ ions; (a) detergent solubilised GltPh ($\sim 20 \mu\text{M}$) and (b) GltPh ($\sim 20 \mu\text{M}$) reconstituted into proteoliposomes (800 MHz, 5°C , 2 s saturation time). Full spectra can be found in the Supporting Information.

that these two ligands share the same binding site, negative cooperativity for the trimeric GltPh in a near native proteoliposome-reconstituted environment is observed.

Current STD NMR literature reporting the direct observation of substrate binding and interaction with membrane proteins is limited to a very small number of systems. Previous examples have included the basic observation of a protein–ligand interaction in whole cells^{22–25}, cell membranes²⁶ and in proteins reconstituted into proteoliposomes²⁷ and nanodiscs²⁸. Only very recently have the first examples of STD NMR-based analysis of ligand binding to detergent-solubilised proteins been reported^{29–32}. There is only one previous report of using STD NMR for a ligand K_D determination to a membrane protein which employed an as-yet unspecified and unpublished fitting routine, whereby an STD NMR effect was measured as a function of ligand concentration for a receptor embedded into a proteoliposome²⁷.

The three protomers within the trimeric GltPh each have one substrate binding site and they have been previously proposed to function independently^{33–36}. Using STD NMR we have directly determined the affinity of the weak binding ligand L-glutamate for GltPh in both detergent micelles and proteoliposomes, whilst binding of the high affinity ligand L-aspartate was identified using a competition-based STD NMR experiment.

Results and discussion

^1H -NMR spectra of GltPh were measured in the presence of L-glutamate both in detergent (Fig. 2a top) and reconstituted into proteoliposomes (Fig. 2b top). Spectra of detergent-solubilised samples are dominated by strong signals from detergent protons (see Supporting Information). The γ -protons of L-glutamate have resonances occurring at 2.05 ppm, whereas β -protons appear at 1.8 ppm, both in a spectral region devoid of interferences from these strong detergent signals (Fig. 2a). STD NMR signals from γ - and β -protons of L-glutamate are observable in the presence (middle) but not in the absence (bottom) of Na^+ . We observe a similar Na^+ dependence for GltPh in detergent micelles and in reconstituted proteoliposomes, which is in accordance with previously published data showing that binding and transport of substrate only occurs in the presence of Na^+ . The STD NMR spectra of detergent-solubilised GltPh also reveal signals arising from detergent molecules as a result of two mechanisms: on–off exchange with the protein/detergent macromolecular assembly in solution, and direct saturation through the on-resonance radio frequency. A lack of ligand STD signal in the absence of Na^+ , however, clearly provides evidence that the ligand only receives its saturation when bound to the protein. Any build-up of saturation is limited by relaxation ($R_1 = 1/T_1$) of saturated ligand protons in the free state; protons with a longer T_1 accumulate more saturation in the bound state and decay more slowly in solution and hence their relative STD intensity is higher. Since proton relaxation rates are sensitive to molecular tumbling in solution which can be affected by the presence of detergent micelles or proteoliposomes, we determined proton T_1 values for L-glutamate under both conditions (Figure S6). The fitted data shown in Figure S6 indicate that the T_1 values remain the same under both conditions (Table S1).

The method for K_D determination requires monitoring of ligand STD NMR signals as a function of the ligand concentration, via ligand titration to a large excess. A GltPh STD NMR spectrum in the presence of Na^+ is recorded initially without L-glutamate. Subsequently, the ligand is titrated incrementally, and further STD NMR spectra are recorded. The experimentally observed STD effect is analysed through conversion to the STD amplification factor (STD-AF) using Eq. (1)²⁰

$$\text{STD} - \text{AF} = \frac{I_0 - I_{\text{SAT}}}{I_0} \times \frac{[L]_{\text{Total}}}{[P]_{\text{Total}}} \quad (1)$$

where $(I_0 - I_{\text{SAT}})$ represents the ligand signal intensity in the STD NMR spectrum, I_0 is the ligand peak intensity in an off-resonance ^1H -NMR spectrum (reference) and $([L]_{\text{Total}}/[P]_{\text{Total}})$ is the ligand excess relative to a fixed and constant protein concentration.

The STD-AF for each titration point was determined for both the γ -protons at 2.028 ppm and β -protons at 1.809 ppm and plotted against the concentration of L-glutamate. The resulting hyperbolic curve (using the γ -protons) has an expected decrease in slope at higher ligand concentrations, eventually reaching a plateau

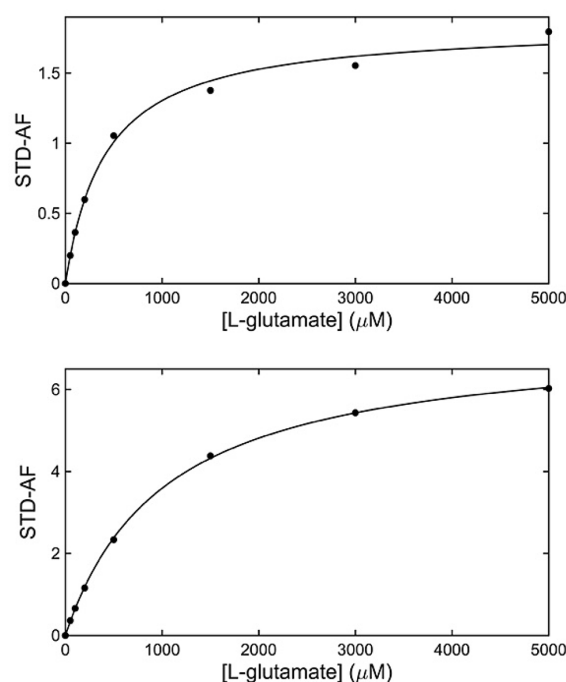


Figure 3. STD amplification factor as a function of L-glutamate concentration for GltPh in detergent micelles (top) and proteoliposomes (bottom). The solid line represents the best fit according to Eq. (3).

caused by saturation of the GltPh binding site (Fig. 3). To utilise this STD-AF for K_D determination, an equation analogous to the Langmuir isotherm is used (Eq. 2). Here, α_{STD} is the maximum STD-AF and $[L]$ is the total concentration of free ligand.

$$STD - AF = \frac{\alpha_{STD}[L]}{K_D + [L]} \quad (2)$$

Determination of the K_D is simplified when $[L] = [L]_{Total}$, which is true under the experimental conditions used in STD NMR experiments since $[L]_{Total} \gg [P]_{Total}$. The STD-AF will increase with increasing $[L]_{Total}$ until a maximum value is reached. At this point, $[L]_{Total} \gg K_D$ and the protein binding sites are fully occupied³⁷. Strictly speaking in the case of GltPh, since the ligand can bind from both sides of the transporter and STD is more sensitive to the off rates, then our data is reporting on an apparent K_D , instead of a K_M , because the glutamate on/off rates are still an order of magnitude faster than transporter isomerization which is linked to the transport event (i.e. K_M).

α_{STD} and K_D for L-glutamate are determined by fitting the observed experimental curve using Eq. (2), and the resulting K_D values for L-glutamate are 414 and 1034 μM for detergent-solubilised and proteoliposome-reconstituted GltPh respectively (Fig. 3) using the γ -protons (very similar results of 458 μM and 984 μM respectively were obtained if the β -protons were used). These values are higher than those previously observed in detergent micelles (between 120–250 μM)^{10,16}, however it must be noted that the sodium ion concentration used here is half of that used in those previous studies. It is also known that due to fast protein–ligand rebinding in solution, K_D values obtained from STD NMR are always greater than, or equal to, the true thermodynamic value³⁸. In order to combat such over-estimations, a more involved protocol can be employed in which initial growth rates of each STD amplification factor are determined and used for calculation of binding isotherms for K_D determination; this application however is not suitable for somewhat unstable membrane proteins³⁸. It is therefore expected that the values determined here likely reflect the upper limits of ligand K_D 's.

STD NMR is of limited use for the direct determination of K_D values of high-affinity ligands (i.e. K_D 's < μM) since tight ligand binding results in low k_{off} rates, thus precluding the saturation of the bound ligand being effectively transferred into solution and thus no STD NMR effect is observed. No STD NMR effect from L-aspartate (in the presence of Na^+) is observed for GltPh which is clearly consistent with its expected much higher binding affinity.

To overcome such a limitation, a competition experiment may be performed where detection of high-affinity ligand binding is identified via observation of a reduction in the STD NMR signal of a weakly binding so-called 'reporter' ligand³⁹. Using such a competitive displacement approach we determined an upper limit for the K_D values of the biologically more relevant, high-affinity GltPh substrate L-aspartate in both detergent-solubilised and proteoliposome-reconstituted environments. In contrast to the initial experiment a titration of L-aspartate

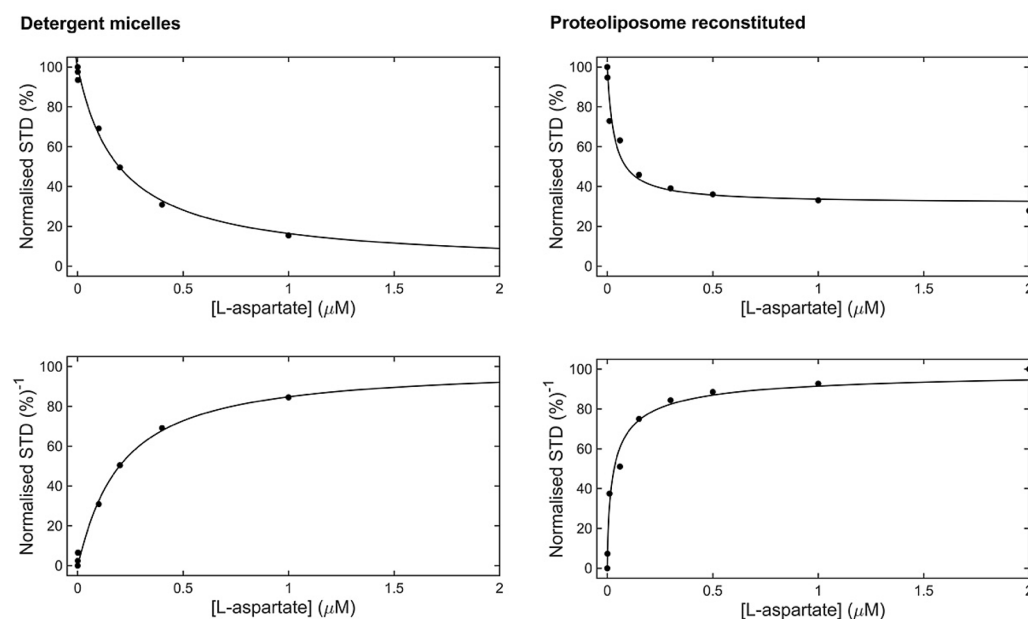


Figure 4. Competition binding of L-aspartate in the presence of excess L-glutamate (1.5 mM) in GltPh. (Top) Normalised STD effect (%) as a function of L-aspartate concentration for GltPh in detergent micelles (left) and reconstituted proteoliposomes (right). The solid line represents the best fit according to Eq. (3). (Bottom) Normalised data fitted to a standard Hill equation (Eq. 4), resulting in Hill coefficients of 1.06 and 0.69 in detergent micelles (left), and reconstituted proteoliposomes (right) respectively.

is performed on a sample containing a large excess of L-glutamate (1.5 mM) relative to both its previously determined K_D and the protein concentration. Upon subsequent titration with L-aspartate, a reduction of the STD NMR effect arising from the L-glutamate reporter ligand is observed. The reduction implies the specific displacement of L-glutamate, most probably from the same binding pocket, by the higher affinity ligand L-aspartate. In Fig. 4 the STD effect for L-glutamate, determined by comparison of the reference and STD spectra, and then normalised (%), is recorded against a titration of L-aspartate resulting in a hyperbolic curve which can be fitted directly using Eq. (3)^{40,41}. Here, A is a proportionality constant, $K_{D_{glu}}$ the dissociation constant measured for L-glutamate and L_0 its fixed concentration, I_0 is the total concentration of L-aspartate and $K_{D_{asp}}$ is its dissociation constant.

$$\%bound \propto -A \times \frac{\left[\frac{K_{D_{glu}}}{L_0}\right] \left[\frac{I_0}{K_{D_{asp}}}\right]}{1 + \left[\frac{K_{D_{glu}}}{L_0}\right] \left[1 + \left(\frac{I_0}{K_{D_{asp}}}\right)\right]} + 100 \quad (3)$$

By fitting the experimental titration data (Fig. 4) K_D values for L-aspartate of 42 nM and 13 nM for detergent-solubilised and proteoliposome-reconstituted GltPh were determined respectively. These are comparable to previously estimated values, also in the nM range^{10,12,15–18}. A comparison of the K_D values determined in detergent and proteoliposomes reveals interesting insight into the protein function. The observed values for the weakly binding ligand, L-glutamate, are comparable in magnitude but vary slightly from those observed previously. The concentration of sodium ions used here is lower than used in most previous cases^{10,16}, which might explain the slightly higher K_D values determined here. Indeed, to test this hypothesis a further set of STD NMR measurements were carried out as a function of Na^+ concentration, using detergent-solubilised WT GltPh in the presence of a fixed, excess concentration of L-glutamate. The STD effect increases with excess Na^+ concentration (Figure S7) which can be explained by an decrease in K_D in accordance with previous biochemical data¹⁵, further highlighting the sensitivity of this particular technique to the system at hand. This serves as an additional control for the Na^+ dependency of substrate binding.

For the tightly binding L-aspartate ligand there is little difference in the observed K_D values between protein environments, however upon closer inspection of the titration data in Fig. 4 a distinct difference can be seen. The value of A used in Eq. 3 to fit the experimental data is different for detergent and proteoliposomes. Whereas in detergent the tightly binding ligand fully displaces the L-glutamate ($A = 100$), in proteoliposomes this value is only 67 suggesting that L-aspartate is not able to fully displace the L-glutamate in GltPh reconstituted into proteoliposomes. To further understand this observation, data from both competition experiments were normalized and subsequently fitted to a standard Hill equation (Fig. 4 bottom) as described by Eq. (4), where n is the Hill coefficient.

$$y = \frac{1}{\left[\left(\frac{K_D}{[L]}\right)^n + 1\right]} \quad (4)$$

The Hill coefficient calculated in detergent micelles was 1.063, indicative of typical non-cooperative, independent ligand binding in the trimeric GltPh. In proteoliposomes, however the Hill coefficient obtained was only 0.69. A Hill coefficient significantly less than 1 is typical of binding events which exhibit negative cooperativity, where binding of substrate to one binding site makes binding of another substrate less likely. This result suggests some interplay between the protomers within the trimeric GltPh when it is reconstituted into a proteoliposome. In this case ligand binding could be associated with an actual transport process across the proteoliposome membrane leading to a difference in environment on either side of the membrane, as would be expected under physiological conditions.

Using the STD NMR method presented here, a constant protein concentration is maintained throughout the entire experiment. Determination of the apparent K_D does thus not rely on accurately knowing the protein concentration, which is often difficult to obtain for membrane proteins. The STD NMR approach therefore presents a simpler route for K_D estimation as compared with other methods commonly used^{10,12,15–18}. Indeed, if we deliberately use a wrong estimation of the protein concentration (by a factor of two) to perform the calculations there is no difference in the K_D determined using this method. The two essential factors are to maintain a protein concentration below the expected range of the K_D of the reporter ligand (weak binder) and also to ensure that the protein concentration is kept constant. Thus, this is the first simple and quick analytical estimation of a ligand K_D to a membrane protein using STD NMR. Further, this method was expanded to include an accurate K_D estimation for a tightly binding ligand normally not accessible with STD-NMR, as well as the observation of important functional differences of a membrane protein reconstituted in different membrane mimetics, making this technique much more universally applicable.

This simplified approach is now being applied to other membrane proteins for both ligand and inhibitor studies. In less favourable cases where interferences from the strong detergent signals are present, perdeuterated detergent molecules could be used or double difference techniques (STDD NMR) applied²³. One further alternative is to perform similar studies on membrane proteins encapsulated in novel membrane protein scaffolds (SMALPS) or indeed in whole cells⁴². Such experiments are currently being performed in our lab.

In summary, we have demonstrated a fast, STD NMR-based protocol for the determination of ligand binding constants across at least four orders of magnitude for membrane proteins. This experimental approach can be applied to both detergent-solubilised and proteoliposome-reconstituted proteins leading to important physiologically relevant information. Further, with improved spectrometer performance and sensitivity such competition STD NMR techniques could lead to important new applications such as lead discovery in which compound libraries are screened against a range of drug targets to identify both high- and low-affinity ligands. Analogues may then be rapidly rank-ordered, and structure–activity relationships derived that can be used to optimise such NMR “hits” into viable drug leads. Hence, the STD NMR approach described here could ultimately be considered as a powerful additional screening tool for future structure-based and fragment-based drug design (SBDD and FBDD respectively) especially in these and other key membrane protein systems.

Materials and methods

Plasmids. GltPh (wild type) inserted into pBAD24 plasmid was created by Eric Gouaux (Oregon Health and Science University, USA) and supplied by Harald H. Sitte and Thomas Stockner (Medical University of Vienna, Vienna, Austria)¹⁸.

Expression and purification of GltPh in detergent. Plasmids including wild type (WT) GltPh were used to transform chemically competent TOP10 *Escherichia coli* bacterial cells ($1 \text{ ng } \mu\text{L}^{-1}$, 100 mL). Cells were grown in LB medium up to an optical density of 0.7–0.8 at 600 nm (37 °C, 190 rpm). Protein production was induced by addition of 0.2% L-arabinose and cell harvesting was carried out after 3 h of incubation (30 °C, 190 rpm). Cells were pelleted by centrifugation (Thermo Scientific Sorvall RC6 Plus) at 6000 rcf for 8 min. The cells were re-suspended in lysis buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 5 mM MgCl_2) with DNase, PMSE, lysozyme and glycerol, and lysis was performed using a high pressure EmulsiFlex system at 10,000 psi. Removal of all cellular debris was carried out via centrifugation (6000 rcf, 10 min). The resulting supernatant was then subjected to ultracentrifugation (176,000 rcf, 2 h) and the membrane pellet was re-suspended in lysis buffer. n-Dodecyl- β -D-Maltoside Detergent (DDM) was added to a final concentration of 1%. The solution was left for 90 min at 4 °C with gentle agitation, after which the total detergent concentration was diluted to 0.5% and insoluble protein/protein aggregates were isolated via ultracentrifugation (176,000 rcf, 15 min).

For protein binding and purification, imidazole (pH 8.0) was added to a final concentration of 10 mM, and the solution was incubated with TALON Metal Affinity Resin overnight at 4 °C, under mild agitation. Washes 1 and 2 were carried out in 40 and 5 column volumes of buffer A (20 mM HEPES pH 7.5, 200 mM KCl, 0.075% DDM) supplemented with 10 mM and 20 mM imidazole respectively. Protein was eluted in 3 column volumes of buffer A supplemented with 500 mM imidazole. The purity of the eluted protein was analysed via SDS-PAGE gel electrophoresis. The eluted protein was exchanged to buffer B (20 mM HEPES pH 7.5, 100 mM KCl, 0.075% DDM) and afterwards concentrated using Amicon Ultra-15 Centrifugal Filter Units (100 kDa MWCO). The protein concentration was calculated by measuring the optical absorbance at 280 nm and using the molar extinction coefficient of WT GltPh. The purified protein was kept in 20% glycerol, aliquoted, flash frozen and stored at –80 °C.

WT GltPh in buffer B with 20% glycerol was washed using a PD-10 desalting column prior to NMR experiments, to remove glycerol and other small molecule impurities. Buffers C (20 mM potassium phosphate pH 7.5, 100 mM NaCl, 0.075% DDM) and D (20 mM potassium phosphate pH 7.5, 100 mM KCl, 0.075% DDM) were prepared in D₂O. Washed WT GltPh was prepared in buffer C via buffer exchange using an Amicon Ultra-4 Centrifugal Filter Unit (100 kDa MWCO), with a final protein concentration of 19.5 μ M. Buffer exchange was repeated 4 times for maximum efficiency. Washed WT GltPh was additionally prepared in buffer D, to a concentration of 22.7 μ M, ready for control NMR experiments in the absence of Na⁺ ions.

Reconstitution of GltPh into liposomes. For reconstitution experiments WT GltPh was exchanged into buffer E (20 mM potassium phosphate buffer pH 7.5, 100 mM KCl, 0.075% DDM) and concentrated to approximately 50 μ M. *Escherichia coli* polar lipid extract (Avanti Polar Lipids Inc.) and 1,2-dioleoyl-*sn*-glycero-phosphocholine (DOPC) (Avanti Polar Lipids Inc.) were mixed in a 1:3 ratio. The mixture was dried under a constant steady stream of argon to form a thin lipid film. The dried lipid mixture was then resuspended in degassed buffer F (20 mM potassium phosphate buffer pH 7.5, 100 mM KCl) to a final concentration of 20 mg.mL⁻¹. The lipid suspension was subjected to 5 freeze-thaw cycles using liquid nitrogen, thawing at room temperature, resulting in the formation of large multilamellar vesicles (LMVs). The mixture was extruded through polycarbonate membranes with a pore size of 400 nm (Avanti Polar Lipids Inc.) to form large unilamellar vesicles (LUVs). The liposomes were diluted to a lipid concentration of 4 mg.mL⁻¹ using buffer G (20 mM potassium phosphate buffer pH 7.5, 100 mM KCl, 25% glycerol). Liposomes were destabilised by dropwise addition of Triton X-100 (10% w/v), adding 150 μ L in total per 5 mL of liposomes. WT GltPh in buffer E was added to the destabilised liposomal mixture at 20 μ g protein per mg of lipid and left to agitate gently for 30 min. Detergent was removed by stepwise addition of Bio-Beads SM-2. The proteoliposome mixture was incubated and gently mixed with four successive additions of 200 mg Bio-Beads SM-2 per 5 mL of proteoliposome mixture, which were then removed by filtration. Liposomes were concentrated by centrifugation at 140,000 $\times$ g for 2 h at 4 °C. Buffers H (20 mM potassium phosphate pH 7.5, 100 mM NaCl) and I (20 mM potassium phosphate pH 7.5, 100 mM KCl) were prepared in D₂O. Liposomes were resuspended in either buffer H or I to a final lipid concentration of approximately 40 mg.mL⁻¹. Proteoliposomes (PL-WT GltPh) were used immediately for NMR experiments, and not flash frozen between times.

Empty liposomes were also prepared in the same manner and at the same lipid concentration, however steps involving liposome destabilising, protein addition and removal of detergent using Bio-Beads SM-2 were removed in this protocol. These served as control samples. The external buffer H was used for resuspension of the liposomes in the final step.

NMR spectroscopy. The NMR spectra were recorded at 278 K with D₂O as solvent, on an Ultra-Compact 800 MHz Bruker Avance III NMR spectrometer equipped with an inverse triple resonance (H/C/N) z-gradient probe head. STD NMR experiments were carried out by a pseudo 2D pulse sequence including spoil pulses to destroy residual magnetization during the relaxation delay (two trim pulses of 2.5 and 5 ms, followed by a 3 ms gradient pulse on the Z-axis). For selective saturation of GltPh, cascades of 49 ms Gaussian-shaped pulses (field strength of 90 Hz) were used with a 1 ms delay between successive pulses. Total saturation time for STD measurements was 2 s with a recycling delay of 2 s in experiments consisting of 512 scans. The saturation time of 2 s was selected as a compromise between sensitivity and accuracy on K_D determination, as a long full initial-slope STD NMR approach⁴⁸ is not appropriate for relatively fast-degrading membrane proteins. Selective saturation of the protein was achieved by setting the on-resonance frequency at 0.6 ppm, to produce efficient saturation of the aliphatic side chains of the protein. This saturating radio frequency indeed also directly hit on the most up-field proton signal of the detergent molecules, contributing to a very efficient saturation of the whole system. Potential saturation transfer effects arising from non-specific interactions of the ligands with the detergent could be ruled out by negative results of STD NMR experiments of the protein-ligand system in the absence of Na⁺. For the reference (off-resonance) spectrum, the irradiation frequency was shifted to 40 ppm.

In the case of detergent solubilised WT GltPh, a series of STD NMR experiments were first carried out using WT protein in buffer C and increasing L-glutamate concentrations ranging from 0 to 5 mM (example NMR spectra in Figure S2). A control STD NMR measurement was then undertaken using detergent solubilised WT GltPh in buffer D (without Na⁺ ions) supplemented with 0.5 mM L-glutamate (Figure S1). Finally, a series of STD NMR experiments were taken following a competition-based titration of L-aspartate (ranging from 0 to 1 μ M) into WT GltPh in buffer C supplemented with 1.5 mM L-glutamate.

For PL-WT GltPh, the first set of STD NMR experiments were carried out using proteoliposomes in external buffer H and increasing L-glutamate concentrations ranging from 0 to 5 mM (Figure S5). Two control STD NMR experiments were also undertaken, using (1) empty liposomes in external buffer H (Figure S3) and (2) PL-WT GltPh in external buffer I (Figure S4), each supplemented with 1.5 mM L-glutamate. Finally, a series of STD NMR experiments were taken following a competition-based titration of L-aspartate (ranging from 0 to 2 μ M) into PL-WT GltPh in buffer H supplemented with 1.5 mM L-glutamate. Bruker Topspin (version 3.5) was employed for processing the NMR spectra and determination of peak integrals for STD-AF calculations as described in reference 21. All figures, except Fig. 1, were prepared using Matlab. The K_D determination was performed using the curve fitting tool in Matlab.

Spin-lattice relaxation times (T_1) were determined using the inversion-recovery method⁴³ for (a) 1.5 mM L-glutamate in buffer C, and (b) 1.5 mM L-glutamate added to empty liposomes in external buffer H. For these experiments, spectra of 16 K data points were obtained for 16 and 12 delay time (τ) values in cases (a) and (b) respectively, ranging from 1 ms to 5 s. Initial quick T_1 estimation was performed for all samples in order to set

the appropriate relaxation delay, which was set to 5 s in all samples. T_1 values were obtained by fitting the data (Figure S6, table S1) to the following equation:

$$I = I_{\max} \left(1 - 2e^{-\tau/T_1} \right) \quad (5)$$

Data availability

The data sets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 6 July 2020; Accepted: 4 September 2020

Published online: 05 October 2020

References

- Danbolt, N. C. Glutamate uptake. *Prog. Neurobiol.* **65**, 1–105 (2001).
- Owe, S. G., Marcaggi, P. & Attwell, D. The ionic stoichiometry of the GLAST glutamate transporter in salamander retinal glia. *J. Physiol.* **577**, 591–599 (2006).
- Zerangue, N. & Kavanaugh, M. P. Flux coupling in a neuronal glutamate transporter. *Nature* **383**, 634–637 (1996).
- Fairman, W. A., Vandenberg, R. J., Arriza, J. L., Kavanaugh, M. P. & Amara, S. G. An excitatory amino-acid transporter with properties of a ligand-gated chloride channel. *Nature* **375**, 599–603 (1995).
- Canul-Tec, J. C. *et al.* Structure and allosteric inhibition of excitatory amino acid transporter 1. *Nature* **544**, 446–451 (2017).
- Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* **431**, 811–818 (2004).
- Jensen, S., Guskov, A., Rempel, S., Hänel, I. & Slotboom, D. J. Crystal structure of a substrate-free aspartate transporter. *Nat. Struct. Mol. Biol.* **20**, 1224–1227 (2013).
- Garaeva, A. A. *et al.* Cryo-EM structure of the human neutral amino acid transporter ASCT2. *Nat. Struct. Mol. Biol.* **25**, 515–521 (2018).
- Ji, Y., Postis, V. L. G., Wang, Y., Bartlam, M. & Goldman, A. Transport mechanism of a glutamate transporter homologue GltPh. *Biochem. Soc. Trans.* **44**, 898–904 (2016).
- Boudker, O., Ryan, R. M., Yernool, D., Shimamoto, K. & Gouaux, E. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature* **445**, 387–393 (2007).
- Reyes, N., Ginter, C. & Boudker, O. Transport mechanism of a bacterial homologue of glutamate transporters. *Nature* **462**, 880–885 (2009).
- Akyuz, N. *et al.* Transport domain unlocking sets the uptake rate of an aspartate transporter. *Nature* **518**, 68–73 (2015).
- Verdon, G. & Boudker, O. Crystal structure of an asymmetric trimer of a bacterial glutamate transporter homolog. *Nat. Struct. Mol. Biol.* **19**, 355–357 (2012).
- Psakis, G. *et al.* Expression screening of integral membrane proteins from *Helicobacter pylori* 26695. *Protein Sci.* **16**, 2667–2676 (2007).
- Hänel, I., Jensen, S., Wunnice, D. & Slotboom, D. J. Low affinity and slow Na⁺ binding precedes high affinity aspartate binding in the secondary-active transporter GltPh. *J. Biol. Chem.* **290**, 15962–15972 (2015).
- Silverstein, N., Ewers, D., Forrest, L. R., Fahlke, C. & Kanner, B. I. Molecular determinants of substrate specificity in sodium-coupled glutamate transporters. *J. Biol. Chem.* **290**, 28988–28996 (2015).
- Reyes, N., Oh, S. & Boudker, O. Binding thermodynamics of a glutamate transporter homolog. *Nat. Struct. Mol. Biol.* **20**, 634–640 (2013).
- Venkatesan, S. K. *et al.* Refinement of the central steps of substrate transport by the aspartate transporter GltPh: Elucidating the role of the Na² sodium binding site. *PLoS Comput. Biol.* **11**, 2 (2015).
- Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chemie - Int. Ed.* **38**, 1784–1788 (1999).
- 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. *J. Am. Chem. Soc.* **123**, 6108–6117 (2001).
- 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. *J. Chem. Educ.* **88**, 990–994 (2011).
- Mari, S., Serrano-Gómez, D., Cañada, F. J., Corbí, A. L. & Jiménez-Barbero, J. 1D saturation transfer difference NMR experiments on living cells: The DC-SIGN/oligomannose interaction. *Angew. Chemie Int. Ed.* **44**, 296–298 (2004).
- Claassen, B., Axmann, M., Meinecke, R. & Meyer, B. Direct observation of ligand binding to membrane proteins in living cells by a saturation transfer double difference (STDD) NMR spectroscopy method shows a significantly higher affinity of integrin $\alpha\text{IIb}\beta_3$ in native platelets than in liposomes. *J. Am. Chem. Soc.* **127**, 916–919 (2005).
- Airoldi, C., Giovannardi, S., Laferla, B., Jiménez-Barbero, J. & Nicotra, F. Saturation transfer difference NMR experiments of membrane proteins in living cells under HR-MAS conditions: The interaction of the SGLT1 co-transporter with its ligands. *Chem. A Eur. J.* **17**, 13395–13399 (2011).
- Cox, B. D. *et al.* Structural analysis of CXCR4—Antagonist interactions using saturation-transfer double-difference NMR. *Biochem. Biophys. Res. Commun.* **466**, 28–32 (2015).
- Venkitakrishnan, R. P., Benard, O., Max, M., Markley, J. L. & Assadi-Porter, F. M. Use of NMR Saturation Transfer Difference Spectroscopy to Study Ligand Binding to Membrane Proteins. In *Membrane Protein Structure and Dynamics* 47–63 (Humana Press, Totowa, 2012).
- Meinecke, R. & Meyer, B. Determination of the binding specificity of an integral membrane protein by saturation transfer difference NMR: RGD peptide ligands binding to integrin $\alpha\text{IIb}\beta_3$. *J. Med. Chem.* **44**, 3059–3065 (2001).
- Fredriksson, K. *et al.* Nanodiscs for INPHARMA NMR characterization of GPCRs: Ligand binding to the Human A2A Adenosine Receptor. *Angew. Chemie Int. Ed.* **56**, 5750–5754 (2017).
- Igonet, S. *et al.* Enabling STD-NMR fragment screening using stabilized native GPCR: A case study of adenosine receptor. *Sci. Rep.* **8**, 1–14 (2018).
- Yong, K. J. *et al.* Determinants of ligand subtype-selectivity at a 1A-Adrenoceptor revealed using saturation transfer difference (STD) NMR. *ACS Chem. Biol.* **13**, 1090–1102 (2018).
- Vaid, T. M., Chalmers, D. K., Scott, D. J. & Gooley, P. INPHARMA based determination of ligand binding modes at α_1 -adrenergic receptors explains the molecular basis of subtype selectivity. *Chem. A Eur. J.* <https://doi.org/10.1002/chem.202000642> (2020).
- Bumbak, F. *et al.* Conformational changes in tyrosine 11 of neurotensin are required to activate the neurotensin receptor 1. *ACS Pharmacol. Transl. Sci.* <https://doi.org/10.1021/acspstsci.0c00026> (2020).

33. Koch, H. P. & Larsson, H. P. Small-scale molecular motions accomplish glutamate uptake in human glutamate transporters. *J. Neurosci.* **25**, 1730–1736 (2005).
34. Grewer, C. *et al.* The individual subunits of the glutamate transporter EAAC1 homotrimer function independently of each other. *Biochemistry* **44**, 11913–11923 (2005).
35. Erkens, G. B., Hänel, I., Goudsmits, J. M. H., Slotboom, D. J. & Van Oijen, A. M. Unsynchronised subunit motion in single trimeric sodium-coupled aspartate transporters. *Nature* **502**, 119–123 (2013).
36. Ruan, Y. *et al.* Direct visualization of glutamate transporter elevator mechanism by high-speed AFM. *Proc. Natl. Acad. Sci. USA* **114**, 1584–1588 (2017).
37. Claridge, T. D. W. High-Resolution NMR Techniques in Organic Chemistry. In *Protein-Ligand Screening by NMR* 421–455 (Elsevier Inc., Amsterdam, 2016).
38. Angulo, J., Enriquez-Navas, P. M. & Nieto, P. M. Ligand-receptor binding affinities from saturation transfer difference (STD) NMR. *Chem. A Eur. J.* **16**, 7803–7812 (2010).
39. Wang, Y., Sen Liu, D. & Wyss, D. F. Competition STD NMR for the detection of high-affinity ligands and NMR-based screening. *Magn. Reson. Chem.* **42**, 485–489 (2004).
40. Cantor, C. R. & Schimmel, P. R. Biophysical Chemistry: Part III. In *Ligand interactions at equilibrium* 849–1371 (WH Freeman and Co., Oxford, 1980).
41. Cheng, Y. & Prusoff, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
42. Postis, V. *et al.* The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochim. Biophys. Acta Biomembr.* **1848**, 496–501 (2015).
43. Vold, R. L., Waugh, J. S., Klein, M. P. & Phelps, D. E. Measurement of spin relaxation in complex systems. *J. Chem. Phys.* **48**, 3831–3832 (1968).

Acknowledgements

This work was supported in part by the Royal Society (FM was a Wolfson Research Merit Award Holder); the EU (COST Action CM1306 Understanding Movement and Mechanism in Molecular Machines) to FM, EJC and TS; the Austrian Science Fund (FWF) for the Special Research Program F35, subproject F3506 to HHS and F3524 to TS; Applied Molecular Biosciences Unit- UCIBIO financed by national funds from FCT/MCTES/Portugal (UID/Multi/04378/2019) to EJC and a BBSRC New Investigator Award (Grant no. BB/P010660/1) to JA. JH was supported by the BBSRC Norwich Research Park Biosciences Doctoral Training Partnership grant number BB/M011216/1.

Author contributions

J.H. performed production of biological material as well as all NMR data curation and analysis, Wrote the original draft and contributed to reviewing & editing final submission A.S. supported production of biological material E.J.C. supported NMR data analysis C.M. supported NMR Data curation T.S. supported production of biological material and NMR data interpretation H.H.S. supported production of biological material J.A. supported NMR data curation and analysis F.M. Designed, administered and supervised the project Wrote original draft and final submission All Authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-73443-z>.

Correspondence and requests for materials should be addressed to F.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020