Biophysical and structural studies of chlorophyll biosynthetic enzymes magnesium chelatase and protochlorophyllide reductase

David Andrew Farmer

Department of Molecular Biology and Biotechnology

A thesis submitted for the degree of Doctor of Philosophy

September 2019

Abstract

Photosynthetic organisms must generate chlorophyll, the light capturing molecule integral for photosynthesis, whilst also regulating production to adapt to environmental conditions. Chlorophyll is so fundamental for photosynthesis that it is important to understand its biosynthetic pathway; consequently, this sequence of reactions has been studied extensively for several decades, culminating in the engineering of the heterotroph *E. coli* to produce chlorophyll. Two key regulatory points in the pathway are i) at the branchpoint between haem and chlorophyll biosynthesis, where a magnesium ion is inserted into the porphyrin ring, catalysed by the multi-subunit enzyme magnesium chelatase (MgCH); and ii) the light-dependent reduction of the C17-C18 double bond, catalysed by protochlorophyllide oxidoreductase (POR). Despite the extensive work performed to biochemically describe these enzymes, the full reaction mechanisms are still unknown, and the limited amount of structural information for either enzyme has hindered complete characterisation.

The work reported in this thesis has used a range of biochemical, biophysical and structural techniques to study these biologically important and mechanistically unique enzymes: microscale thermophoresis, kinetic analysis and cross-linking mass spectrometry revealed that the driving force behind chelation, the ATPase activity of the AAA+ ChlI subunit, is linked to the chelating ChlH protein via the regulatory ChlD subunit; X-ray crystallography-guided mutagenesis determined the porphyrin binding site in ChlH; (cryo)-electron microscopy was used to initiate structural investigations into the quaternary organisation of MgCH and POR; and 2-dimensional electronic spectroscopy was applied to the light-initiated reaction catalysed by POR, the first use of this complex technique on an enzyme, potentially revealing a novel intermediate in the reaction that is formed on the femtosecond time-scale. The work presented in this thesis aims to develop our biochemical understanding of MgCH and POR, and lay the foundations for further structural and mechanistic studies of these interesting enzymes of chlorophyll biosynthesis.

Acknowledgements

To start I would like to thank my supervisor, Neil, for offering me the opportunity to undertake this PhD, and for his continued support, suggestions and ability to inspire enthusiasm for the project. I have greatly appreciated the Hunter lab environment and have been given several opportunities to work with other lovely people in exciting places thanks to Neil's vast network of collaborators.

I would also like to thank my more hands-on lab supervisors, Nate and Mo for all of their training, guidance and collaboration (especially on enzyme kinetics and molecular biology, respectively). A thanks goes to all of the other Hunter lab members who have helped in and out of the lab, and whose office shenanigans consistently improved the day. The pub trips and Christmas dos ever a source of hilarity, and an enthusiasm for most things sport encouraged me to actually do a bit of moving, which is good!

Next, I would like to thank the people that I've collaborated with: Dr Eli Romero, who I worked with for the data presented in Chapter 6, and was incredibly friendly, helpful and generally good fun to work with. Thanks for helping me enjoy my month in Amsterdam that bit more. I would also like to thank Dr Alistair Siebert, who I undertook my PIPS placement with down at eBIC, despite seemingly being one of the busiest people I know! A thanks too, to Dr Georgio Schiro, whose "Sleep is for the weak" T-shirt aided with the long nights at the ESRF conducting work that, unfortunately, didn't make it into this thesis. I would also like to personally thank any other contributors to work in this thesis not already mentioned: Dr Andy Hitchcock, Dr Phil Jackson, Dr Claudine Bisson, Dr Jim Reid, Dr Svet Tzokov and Dr Dave Armstrong.

Finally, I would like to thank all my other family and friends that have been a part of my life during these past four years. Especially my best mates who I've been lucky enough to live with: Sarah, Alex, Harry, Emma, Josh and Studd; for all the walks, rides, nights out, pub trips, festivals, holidays, BBQs and general chilling – cheers guys. And a last thank you to my parents: your ever-present support and encouragement undoubtedly played a large role in me getting this far.

Table of Contents

Table of Figures

Table of Equations

Table of Tables

List of Abbreviations

2DES - 2-dimensional electronic spectroscopy
AAA+ - ATPases associated with various cellular activities
ADP - Adenosine diphosphate
AFM - Atomic force microscopy
ALA - 5-aminolevulinic acid
ATP - Adenosine triphosphate
AUC - Analytical ultracentrifugation
CCD - Charge-coupled devices
CD - Circular dichroism spectrometry
Chlide - Chlorophyllide
CMC - Critical micelle concentration
CPOX - Coproporphyrinogen III oxidase
CTF - Contrast transfer function
CV - Column volumes
DAD - Diode array detector
DAVA - 4,5-diaminovalerate
DHFR - Dihydrofolate reductase
DLS - Dynamic light scattering
DPOR - Dark-operative POR
DQE - Detective quantum efficiency
DSF - Differential scanning fluorimetry
DTT - Dithiothreitol
DVR - Divinyl(proto)chlorophyllide 4-vinyl reductase
EDC - 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride
EDTA - Ethylenediaminetetraacetic acid
EM - Electron microscopy
FAD - Flavin adenine dinucleotide
FeCH - Ferrochelatase
FMO - Fenna-Matthews-Olson
FRET - Förster resonance energy transfer
FSC - Fourier shell correlation

GGPP - Geranylgeranyl diphosphate
GluRS - Glutamyl-tRNA synthetase
GluTR - Glutamyl-tRNA reductase
GSA - L-glutamic acid 1-semi-aldehyde
GSAT - Glutamate 1-semi aldehyde amino transferase
HMBS - Hydroxymethylbilane synthase
HPLC - High performance liquid chromatography
ICT - Intramolecular charge transfer
IPTG - Isopropyl β-D-1-thiogalactopyranoside
LB - Lysogeny broth
LEW - Lysis-equilibration-wash
LPOR - Light-dependent POR
MBP - Maltose-binding protein
MgCH - Magnesium chelatase
MgP_{IX} - Magnesium protoporphyrin IX
MgPME - Mg-protoporphyrin 13 monomethyl ester
MIDAS - Metal ion dependent adhesion site
MST - Microscale thermophoresis
MTF - Modular transfer function
MX - Macromolecular X-ray crystallography
NADP - Nicotinamide adenine dinucleotide phosphate
NEM - N-ethylmaleimide
NHS - N-hydroxysuccinimide
NIR - Near-infrared
NMR - Nuclear magnetic resonance
NOPA - Non-collinear optical parametric amplifier
PBG - Porphobilinogen
PBGS - Porphobilinogen synthase
Pchlide - Protochlorophyllide
P_{IX} - Protoporphyrinogen IX
PLB - Prolamellar bodies
PLP - Pyridoxal 5'-phosphate
PMF - Proton motive force
PMP - Pyridoxamine 5'-phosphate

POR - Protochlorophyllide oxidoreductase

PPO - Protoporphyrin oxidase

PSII - Photosystem II

RC - Reaction centre

ROS - Reactive oxygen species

SAM - S-adenosyl-L-methonine

SDR - Short-chain dehydrogenase reductase

SDS-PAGE - Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SNR - Signal to noise ratio

TRIC - Temperature-related intensity change

UROD - Uroporphyrinogen III decarboxylase

UROS - Uroporphyrinogen III synthase

XL-MS - Cross-linking mass-spectrometry

B-DDM - Dodecyl-β-maltoside

1 Introduction

1.1 Photosynthesis

Photosynthesis is the key biological process whereby light energy absorbed by a photosynthetic organism is used to fix atmospheric carbon dioxide in the form of carbohydrates. It is responsible for the generation of almost all of the biomass on Earth, resulting in food for chemotrophic organisms including humans, and indirectly as fossil fuels to support our energy usage (Hoffert, 2002). In the case of oxygenic photosynthesis, a by-product of the reaction is gaseous oxygen, which is released into the atmosphere allowing organisms to respire aerobically. It is one of the most important processes for the existence of life on Earth, and a thorough understanding of photosynthesis will surely be key to combatting the pressing issue of man-made climate change. This could be through improved crop security, as well as harnessing the process to improve renewable energy sources and to sequester mounting levels of carbon dioxide.

$$2H_2A + CO_2 \overset{h\nu}{\rightarrow} CH_2O + H_2O + 2A$$

Equation 1.1. The generalised equation for the fixation of carbon dioxide by photosynthesis. H_2A is the reducing agent, whilst A is the oxidation product (van Niel, 1962).

Photosynthetic organisms can be separated by the type of photosynthesis that they perform, namely oxygenic or anoxygenic photosynthesis. Plants, algae and cyanobacteria are all oxygenic phototrophs, and they use water (A is oxygen in Equation 1.1) as the reductant with molecular oxygen as the oxidative product. Anoxygenic phototrophs include the remaining photosynthetic bacteria and these use a range of reductants including organic acids, hydrogen and reduced sulphur. The resulting oxidation products differ depending on the reductant.

Photosynthesis occurs when light energy is absorbed by pigments specifically arranged within proteins located in the membranes of photosynthetic organisms. The majority of pigments are used to capture incoming light energy which is then funnelled via Förster resonance energy transfer (FRET) to two chlorophyll molecules in reaction centre (RC) proteins. These coupled "special-pair" chlorophyll molecules are in an environment that allows their excited state to catalyse the energetically unfavourable oxidation of the reducing agent (H_2A), starting a series of electron transfers between photosystem, cytochrome and iron-sulphur proteins that culminate in reduction

of nicotinamide adenine dinucleotide phosphate ($NADP^+$), the terminal electron acceptor. Along the way, the electron flow is coupled to the generation of a proton motive force (PMF), which is used by the adenosine triphosphate (ATP) synthase enzymes to generate ATP. The resulting reduced $NADP^+$ (NADPH) and ATP molecules are used in the Calvin cycle, the process by which carbon dioxide is fixed (Calvin and Bassham, 1955).

1.2 Chlorophyll and other tetrapyrroles

At the core of photosynthesis is a group of molecules named (bacterio)chlorophyll. These are reduced cyclic tetrapyrroles with chemical properties that allow them to absorb strongly in the UV/visible and near-infrared (NIR) regions of the electromagnetic spectrum. The core macrocycle of a porphyrin is roughly square planar, approximately 9 Å wide with four pyrrole rings connected in a conjugated system (Figure 1.1 **A**). Chlorins distinguish themselves from porphyrin through the presence of a reduced bond at position C17-18. Almost all chlorins in nature also have an isoprenoid tail bound to the D ring and an extra isocyclic ring "E" attached to ring C (Scheer, 1991).

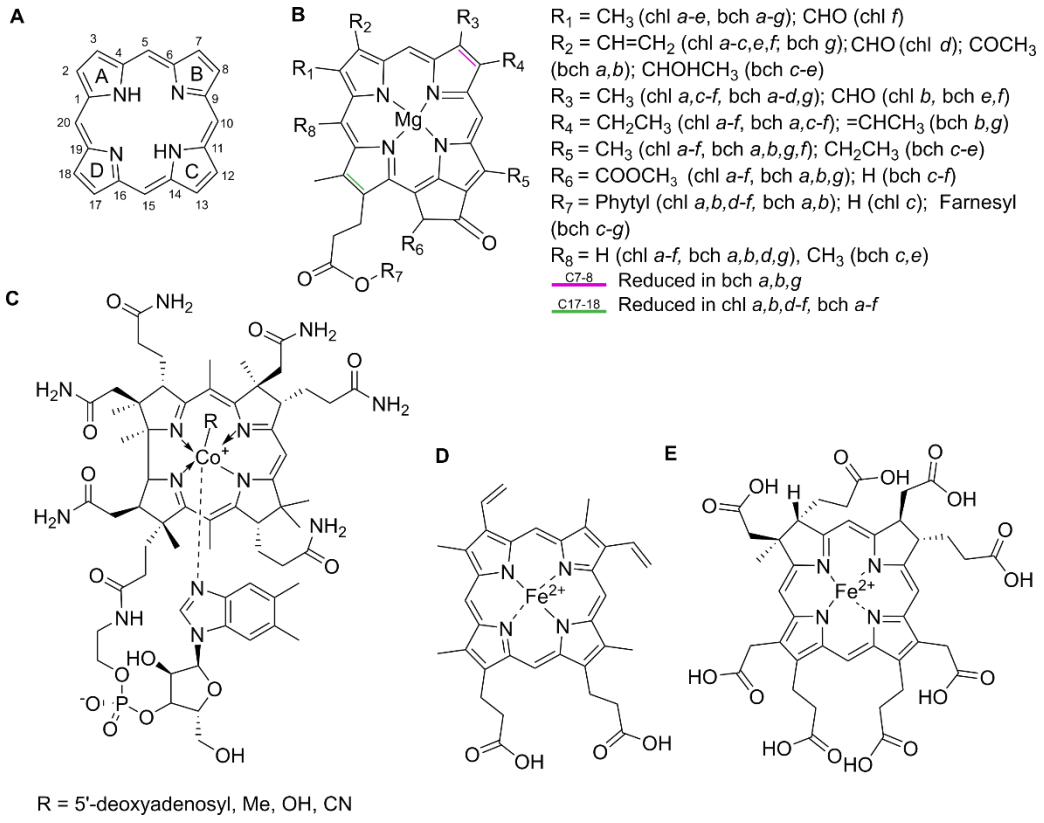


Figure 1.1. The chemical structures of a selection of tetrapyrroles that act as prosthetic groups. **A**, porphyrin, a generic cyclic tetrapyrrole. The carbon atoms and pyrrole rings are labelled according to IUPAC nomenclature. **B**, (bacterio)chlorophyll, **C**, cobalamin (vitamin B_{12}), **D**, haem *b* and **E**, sirohaem.

Tetrapyrroles are used as prosthetic groups by a wide range of proteins, and perform many different functions including electron transfer, oxygen binding, methyl transfer, and light absorption. As shown in Figure 1.1 they all contain a central metal ion, which is often involved in the properties and function of the molecule in biological systems.

(Bacterio)chlorins can be split into two main types in nature: chlorophyll, which is found in oxygenic phototrophs, and bacteriochlorophyll, found in anoxygenic phototrophs. The major difference between the two is whether or not the C7-8 bond is reduced – in bacteriochlorin reduction disrupts the conjugation system of the macrocycle and alters the spectral properties. Both chlorophyll and bacteriochlorophyll have a variety of modifications to the groups labelled R in Figure 1.1 that can further change the spectral and other properties, giving rise to a range of pigments known as *a- f* for chlorophyll and *a-g* for bacteriochlorophyll (Chen et al., 2010; Scheer, 1991).

The variable extended conjugated π system and the reduction state of the macrocycle of chlorins allows them to absorb at near-UV/ blue (380-420 nm) and the red/ near infrared (R/NIR; 600-740 nm) regions, resulting in the high energy B (Soret) band energy transition and the lower excitation states, the Q bands, respectively. The reduced C7-8 bond and presence of different side chains alters the conjugation system, which redshifts the Q band in bacteriochlorophylls. The higher (S2) excitation state is very short lived (ps) and loses energy as heat to drop to the lower S1 state, which is much longer lived (ns) and used to drive the charge separation of the H_2A by the special-pair (bacterio)chlorophyll molecules in RCs (Blankenship et al., 2011).

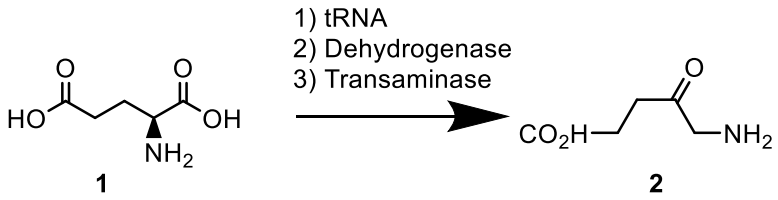
Absorbed energy not harnessed for charge separation is released from (bacterio)chlorins as fluorescence, which provides useful spectral signals; the fluorescence and absorption spectra of (bacterio)chlorins are highly dependent on surrounding chemical environment as well as the overall chemical structure; intermediates in the biosynthetic pathways mostly having distinct 'spectral fingerprints' (Blankenship, 2014).

1.3 Chlorophyll Biosynthesis

1.3.1 5-aminolevulinic acid synthesis

All tetrapyrroles synthesised by organisms share a similar pathway until their biosynthetic pathways diverge at various branch points. In plants, algae and cyanobacteria (i.e. chlorophyll-producing organisms), L-glutamic acid is the precursor for the synthesis of 5-aminolevulinic acid (ALA) (Beale, 1999), which is formed via a three-step reaction that proceeds through L-glutamyl-tRNA

and L-glutamic acid 1-semi-aldehyde (GSA) (Figure 1.2). ALA synthesis is the rate-limiting step of tetrapyrrole biosynthesis (Tanaka and Tanaka, 2007).



1 2

Figure 1.2. L-glutamic acid (**1**) is converted into ALA (**2**) via L-Glutamyl-tRNA and L-glutamic acid 1-Semialdehyde . This is catalysed by glutamyl-tRNA synthetase, glutamyl-tRNA reductase and then glutamate 1-semialdehyde amino transferase.

The first step of this reaction is catalysed by glutamyl-tRNA synthetase (GluRS), a class I aminoacyl-tRNA synthetase that requires $MgATP^{2-}$(Kern et al., 1979; Schimmel, 1987) and uses L-glutamyl-tRNA, also used for protein synthesis, as a substrate (Beale et al., 1975). The activity of GluRS is positively regulated by cellular levels of porphyrin, and downregulated by intracellular hemin (Levican et al., 2007). The crystal structure of GluRS (Perona et al., 1993) shows the presence of a Rossmann fold domain for nucleotide binding and conserved domains for tRNA binding, with specific regions for detecting the anticodons for glutamate (Sekine et al., 2006).

Glutamyl-tRNA reductase (GluTR) catalyses the second step of ALA production from glutamate, the $NADP^+$ dependent dehydrogenation of L-glutamyl-tRNA (Beale, 1990). GluTR is thought to form oligomeric states *in vivo* with three helices of one domain from each monomer intertwining into a six helix bundle (Moser et al., 2001), and to interact with GluRS in the presence of L-glutamyl-tRNA, allowing substrate channelling of the highly reactive molecule from GluRS to GluTR. Following the formation of the GSA product, the catalytic N-terminal domain of GluTR then binds to the subsequent enzyme in the reaction path, GSA amino transferase, again to prevent the escape of the highly reactive GSA (Nogaj and Beale, 2005). Positively charged lysine and arginine residues have been added to this N-terminal catalytic region of GluTR and were found to improve ALA production, demonstrating the importance of this region in regulating activity (Kang et al., 2011; Zhang et al., 2017).

Glutamate 1-semi aldehyde amino transferase (GSAT) catalyses the final step in ALA production, the transamination of GSA (Hennig et al., 1997). Pyridoxamine 5'-phosphate (PMP)/ pyridoxal 5'-phosphate (PLP) (vitamin B_6) provides the additional amine group to be added to the molecule resulting in the intermediate 4,5-diaminovalerate (DAVA); the oxo group of GSA is replaced with

an amine and then the original amine of GSA is replaced with a new oxo group, effectively swapping the amine and oxo groups (Figure 1.3) (Stetefeld et al., 2006). GSAT forms an asymmetric dimer where each of the monomers can be bound to either PLP or PMP. A gating loop closes upon substrate binding, and this has been shown to inhibit the other monomer, conferring negative co-cooperativity (Song et al., 2016; Stetefeld et al., 2006).

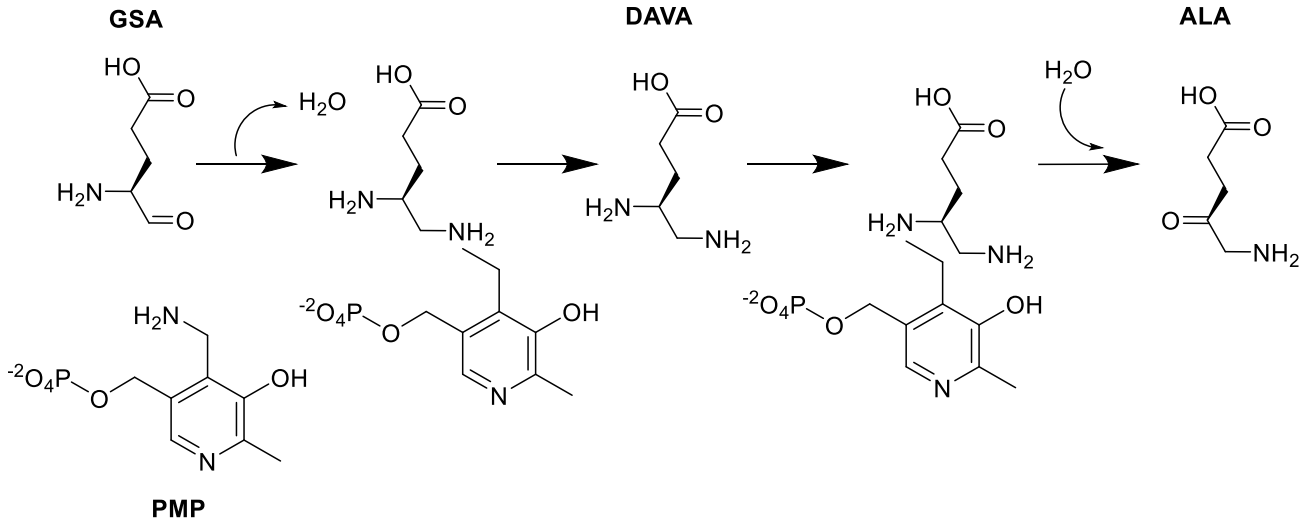


Figure 1.3. Proposed mechanism of transamination of GSA to ALA, catalysed by GSAT with cofactor PMP via a diaminovalerate (DAVA) intermediate.

1.3.2 Three-step condensation of ALA to uroporphyrinogen III

Once ALA has been synthesised, eight molecules of ALA are asymmetrically condensed to form four porphobilinogen (PBG) molecules and then one hydroxymethylbilane molecule. The first reaction is catalysed by the highly conserved porphobilinogen synthase (PBGS) (Figure 1.4) and the second by hydroxymethylbilane synthase (Figure 1.5) (Gibson et al., 1954; Jaffe, 1995).

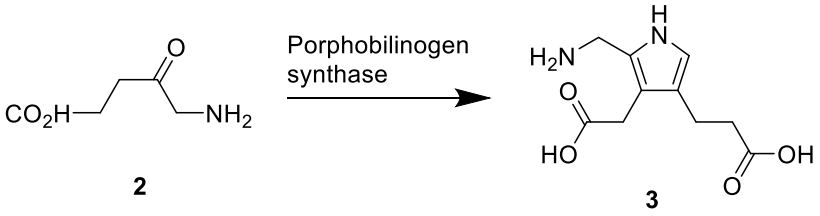
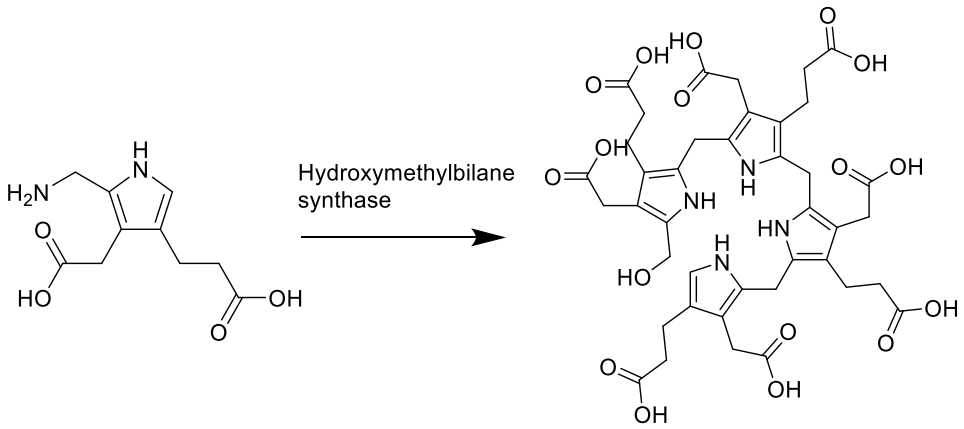


Figure 1.4. The condensation of ALA (**2**) to PBG (**3**) via porphobilinogen synthase.

PBGS requires a highly basic active site environment for the ALA and active site lysine residues to form the Schiff base of the first pyrrole in the biosynthetic pathway. To achieve an active site

environment that differs significantly from the cellular environment PBGS from most species forms a double stack of tetramers which can occlude and gate the active sites (Jaffe, 2016). In the majority of species, PBGS form homooctomers, however, in *Rhodobacter capsulatus*, hexameric, tetrameric and monomeric organisations are used and mediating the oligomeric states regulates PBGS activity (Bollivar et al., 2004). For most species, the octomeric state is allosterically promoted by Mg^{2+} ions, in other species, activity can be regulated through key protonated arginine residues or Zn^{2+} ions, and Zn^{2+} is often essential for catalysis. The *Rhodobacter* genus differs here also; no metal binding has been observed for PBGS from this genus (Jaffe, 2003).



3 **4**

Figure 1.5. Condensation of PBG (**3**) to hydroxymethylbilane (**4**) via hydroxymethylbilane synthase.

Hydroxymethylbilane synthase (HMBS), also known as porphobilinogen deaminase, catalyses the stepwise condensation of four PBG molecules, releasing an ammonia molecule at each step and forming the first tetrapyrrole of the pathway (Bogorad and Granick, 1953; Jordan, 1991). The reaction requires a covalently bound dipyrromethane cofactor which acts as a template to the condensation reaction. A hexapyrrole is produced before the product is cleaved from the cofactor and released to the next enzyme in the pathway, uroporphyrin III synthase (Louie et al., 1996); though the precise mechanism has yet to be elucidated. It is important that the unstable hydroxymethylbilane product reaches the next enzyme swiftly, to avoid spontaneous closure of the macrocycle and inclusion of the incorrect isomer, a toxic molecule that cannot be metabolised (Jordan, 1991).

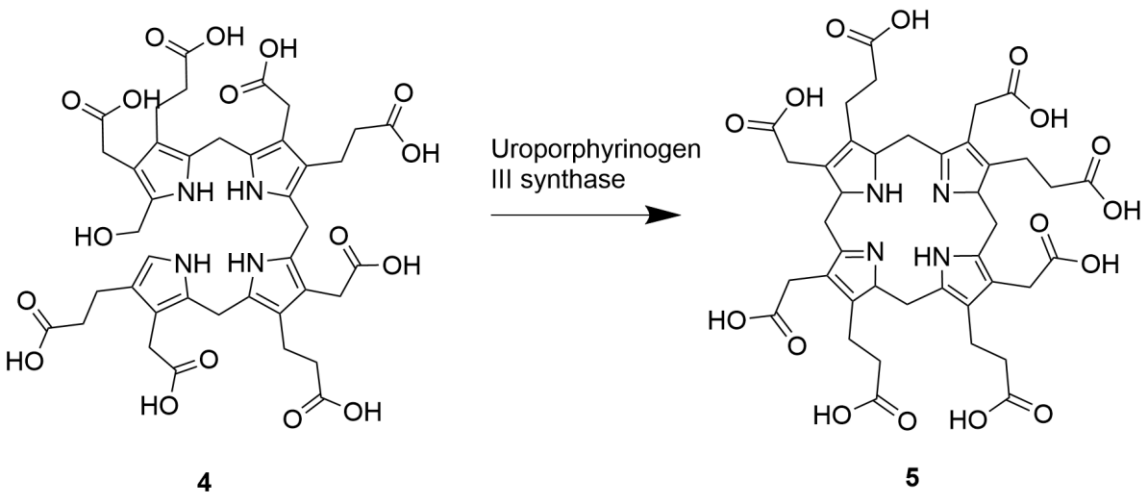


Figure 1.6. The ring closure of hydroxymethylbilane (**4**) by the uroporphyrinogen III synthase resulting in the flipping of ring D and formation of uropophyrinogen III (**5**).

The formation of the last common precursor of the cyclic tetrapyrrole cofactors is catalysed by uroporphyrinogen III synthase (UROS) a small monomeric protein with two similar domains connected by a linker region (Mathews et al., 2001). The cobalamin (vitamin B_{12}) and siroheme biosynthetic pathways split from the common pathway at this stage. Regarding the mechanism of UROS, rings A and B of the substrate interact tightly with the protein whilst ring D is exposed to solvent, allowing for the rotational freedom necessary for ring D flipping (Schubert et al., 2008). A conformational change in UROS occurs after an intermediate is formed from the catalysed loss of water, closing the enzyme around the intermediate and ensuring that ring D is in the correct orientation for product formation (Fortian et al., 2011).

1.3.3 Decarboxylation and oxidation of uroporphyrinogen III to form protoporphyrin IX

On the pathway to forming haem or chlorophyll, uroporphyrinogen III must be specifically decarboxylated by one of two different enzymes. The first, uroporphyrinogen III decarboxylase (UROD), catalyses the decarboxylation of the acetate groups, leaving methyl groups in their place and forming coproporphyrinogen III (**6**). This occurs in a specific order at physiological conditions, starting with ring D and proceeding clockwise (Jackson et al., 1976). UROD forms a homodimer but each monomer acts independently, despite initial proposals that multiple active sites may be necessary to catalyse the multiple decarboxylation steps (Phillips et al., 2009).

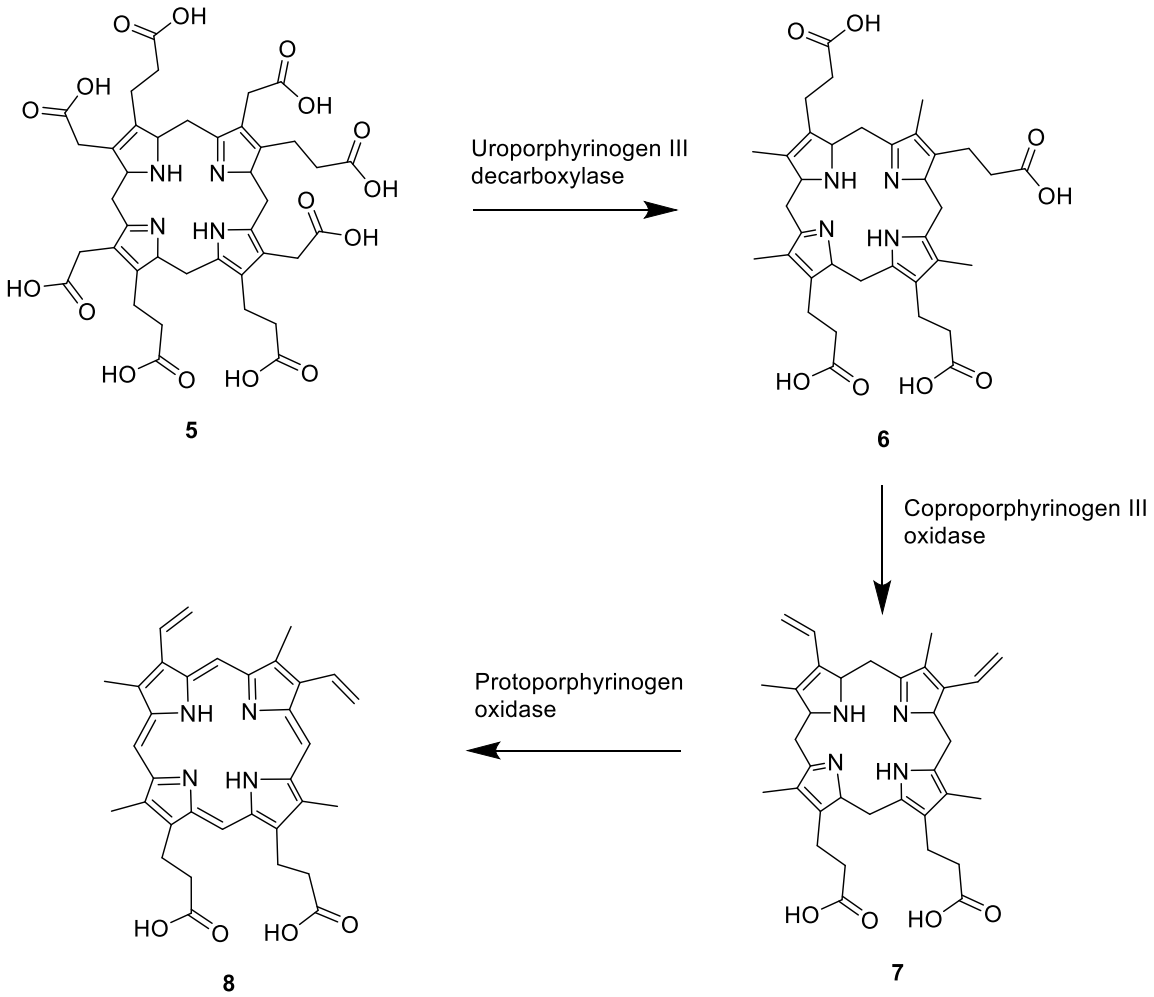
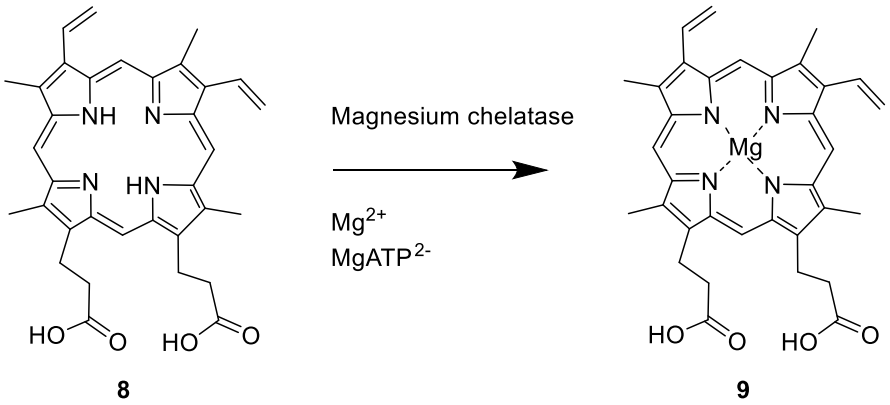


Figure 1.7. The two decarboxylation and subsequent oxidation steps required to form protoporphyrin IX (**8**), the last common precursor of chlorophyll and heme. UROD decarboxylates acetate groups on uroporphyrinogen III (**5**) to form coproporphyrinogen III (**6**). The ring A and B propionate moieties are then oxidised by CPOX, resulting in protoporphyrinogen IX (**7**). An additional oxidation of the macrocycle by PPO produces protoporphyrin IX (**8**).

Further decarboxylation steps then occur on the propionate moieties of rings A and B, catalysed by the enzyme coproporphyrinogen III oxidase (CPOX). The propionates are replaced by vinyl groups resulting in the formation of protoporphyrinogen IX (**7**). There are two distinct and unrelated forms of CPOX: oxygen-dependent and –independent CPOX (Troup et al., 1994, 1995). Oxygen-dependent CPOX uses oxygen as a terminal electron acceptor, and the active site is protected by a lid, to prevent premature exposure of the highly oxygen-sensitive substrate (Phillips et al., 2004). In contrast, oxygen-independent CPOX utilises two s-adenosyl-L-methonine (SAM) molecules as cofactors as well as a 4Fe-4S iron sulphur cluster that promotes formation of the SAM-radical used for catalysis (Ji et al., 2019).

The final step of biosynthesis, shared by both the haem and chlorophyll pathways, is the six-electron oxidation of protoporphyrinogen IX to complete the conjugation of the ring, resulting in protoporphyrin IX (P_{IX}). Three distinct forms of protoporphyrinogen oxidases (PPO) have been discovered so far: PgoX, the most widespread, is an oxygen-dependent PPO with a flavin adenine dinucleotide (FAD) cofactor (Dailey and Dailey, 1996); PgdH1, which is oxygen-independent and also uses a FAD cofactor (Boynton et al., 2009); and PgdH2, most recently discovered, and present in cyanobacteria where the other PPOs are often absent (Kato et al., 2010). PgoX uses oxygen as the terminal electron sink, producing three hydrogen peroxide molecules per reaction. PgdH1 is oxygen independent as it is membrane associated and electrons are passed to respiratory complexes which can then reduce either oxygen or another terminal electron acceptor such as nitrate and fumarate (Mobius et al., 2010). PgdH2 is believed to be a transmembrane protein (Skotnicová et al., 2018), and it has no FAD cofactor but binds haem *b* which is presumed to be catalytically active in the oxidation process.

1.3.4 Chelation of PIX



8 **9**

Figure 1.8. The insertion of a magnesium ion into protoporphyrin IX (**8**) to form magnesium protoporphyrin IX (**9**).

P_{IX} is the last common precursor of haem and chlorophyll. At this branch point either ferrochelatase (FeCH) commits P_{IX} to haem biosynthesis by inserting a ferrous iron into the macrocycle, or magnesium chelatase (MgCH) forms magnesium protoporphyrin IX (MgP_{IX}) (**9**) (Figure 1.8). Magnesium chelatase is a large multi-subunit complex and one of two enzymes investigated in this thesis; a review can be found in section 1.4.

FeCH is a single subunit enzyme, often membrane associated, of *ca*. 40 kDa that forms homodimers (Wu et al., 2001). As well as iron, FeCH catalyses the insertion of many other divalent metal

ions into porphyrin *in vitro*, including zinc, which is often used for *in vitro* assays (Camadros et al., 1984). There is limited sequence similarity between species but consistent fold architecture. However, an iron-sulphur cluster has been found in some variants of the enzyme including human and all other animal FeCH proteins (Wu et al., 2001). There are also variants of FeCH that catalyse the insertion of iron into coproporphyrin III, coproporphyrinogen (**6**) that has been oxidised in much the same way as protoporphyrin IX (Dailey et al., 2015; Hobbs et al., 2017; Lobo et al., 2015).

1.3.5 Methylation and addition of isocyclic ring E

MgP_{IX} is methylated on the carboxyl group of the C13 side chain propionate to form Mg-protoporphyrin 13 monomethyl ester (MgPME) (**10**) (Figure 1.9). This reaction is catalysed by MgP_{IX} methyl transferase (ChlM in chlorophyll producing organisms) and requires methyl group donor SAM (Shepherd and Hunter, 2004). This reaction is thought to be necessary to facilitate the next step of the pathway, the formation of the isocyclic ring E by MgPME cyclase (Beale, 1999). As with many of the chlorophyll biosynthetic enzymes, ChlM could be membrane localised, though not an integral membrane protein (Block et al., 2002). Given the phototoxicity and oxygen sensitivity of the substrates and products, it is unsurprising that substrate channelling from MgCH to ChlM has been proposed (Shepherd et al., 2005). Moreover, the crystal structure of ChlM suggests that two regions of the protein undergo conformational changes to trap each substrate, without disrupting binding of the other (Chen et al., 2014).

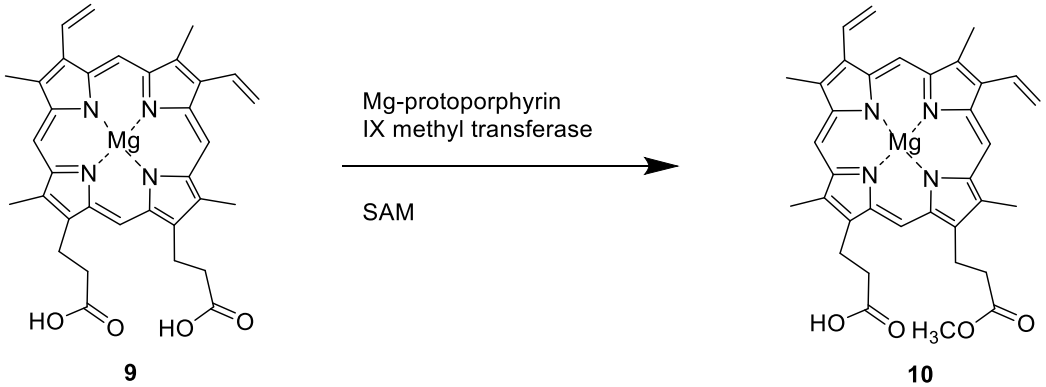


Figure 1.9. The addition of a methyl group to the propanoic acid of C-13 of Mg-protoporphyrin IX (**9**) to form Mg-protoporphyrin IX 13 monomethyl ester (**10**). This reaction uses biological methyl group donor S-adenosyl-L-methionine (SAM).

The next step on the pathway is catalysed by MgPME cyclase, one of the least well characterised enzymes in the pathway. It is the formation of the fifth ring of chlorophyll that confers the shift

from the red colour of the MgPME substrate to the green colour of the divinyl protochlorophyllide (divinyl Pchlide) product, and therefore to chlorophyll. There are two forms of cyclase; BchE is found in anoxygenic phototrophs, and AcsF is found in cyanobacteria, plants, algae and some purple proteobacteria (Chen et al., 2017; Gough et al., 2000). BchE contains an iron-sulphur cluster and operates by a radical-SAM mechanism (Gough et al., 2000). AcsF is an oxygen-dependent single subunit enzyme with a di-iron cofactor (Chen et al., 2018), and it also requires another subunit to function, which varies between organisms, either Ycf54 or BciE (Chen et al., 2017; Gough et al., 2000).

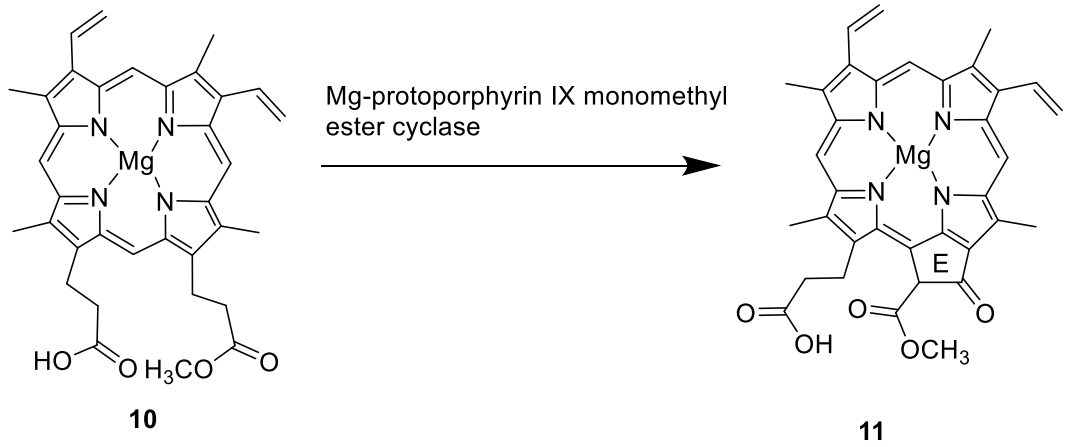


Figure 1.10. The formation of ring E (as labelled in molecule **11**) by Mg-protoporphyrin IX monomethyl ester cyclase, converting MgPME (**10**) to divinyl-Pchlide (**11**).

1.3.6 The formation of chlorophyllide

Divinyl Pchlide is a substrate for either of the subsequent two enzymes, protochlorophyllide oxidoreductase (POR) or divinyl(proto)chlorophyllide 4-vinyl reductase (DVR) (Canniffe et al., 2014). POR catalyses the reduction of the C17-C18 double bond on ring D, disrupting the conjugation system and red-shifting the Q_Y band, resulting in a distinct spectral change corresponding to the product (divinyl)chlorophyllide (**12/14**). There are two types of POR, an ATP dependent oxygen-sensitive POR (DPOR) and a photoactivated form: light-dependent POR (LPOR). In LPOR, light energy is absorbed by the substrate, providing the driving force for the reaction. LPOR is the second of the two enzymes investigated in this thesis, and both light and dark POR are reviewed in section 1.5.

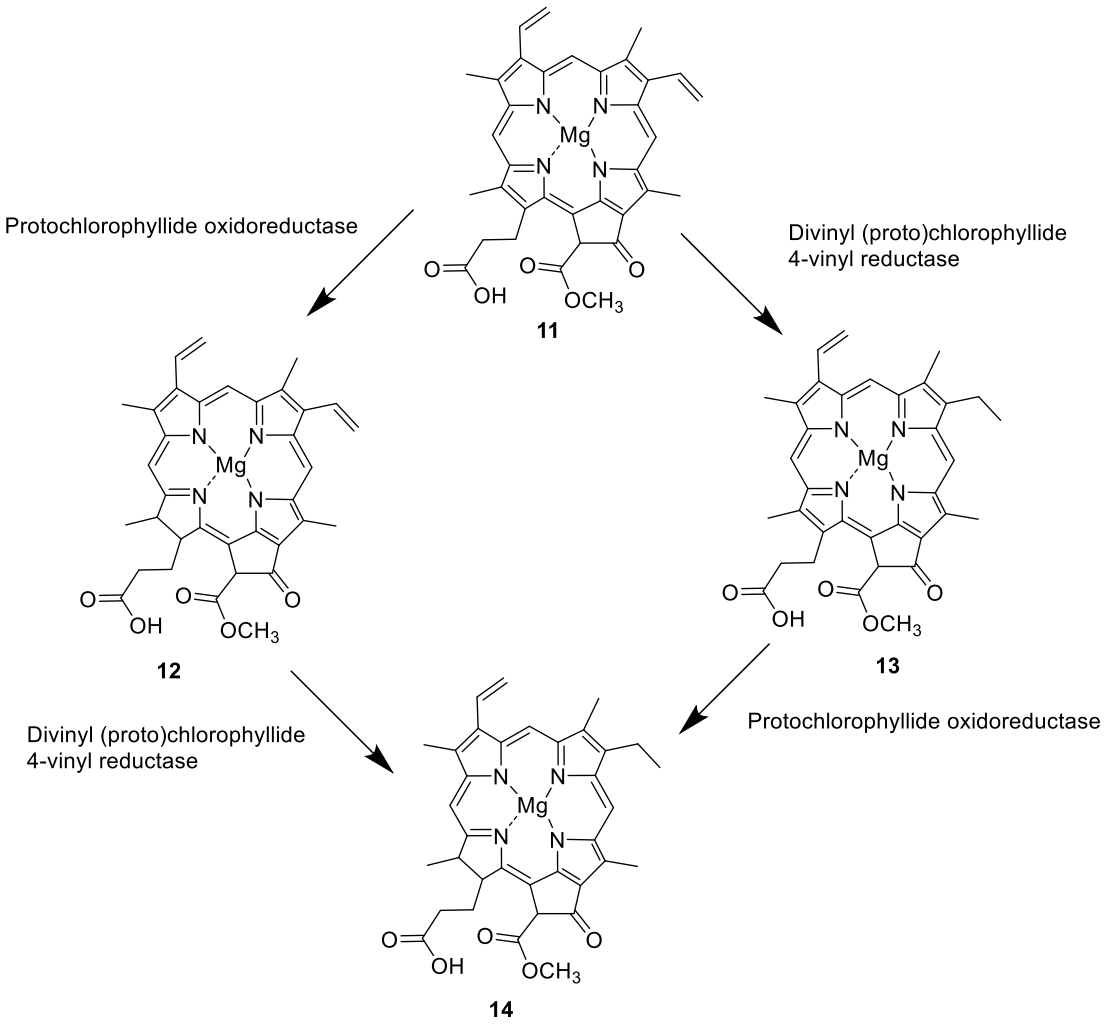


Figure 1.11. The reduction of divinyl protochlorophyllide (**11**) to chlorophyllide (**14**) is achieved via divinyl chlorophyllide (**12**) or protochlorophyllide (**13**).

DVR catalyses the reduction of the C8 vinyl group of divinyl(proto)chlorophyllide to form monovinyl(proto)chlorophyllide (**13**/**14**) (Figure 1.11). Again, there are two main forms of DVR, BciA, a short-chain dehydrogenase reductase (SDR) requiring NADPH for catalysis used by many eukarya, and BciB, a ferredoxin-dependent hydrogenase utilised by cyanobacteria and red algae (Islam et al., 2008; Nagata et al., 2005). Homologues to both BciA and BciB are present in *R. sphaeroides* and cyanobacterium *Acaryochloris marina* (Canniffe et al., 2013; Chen et al., 2016). BciB is thought to have a broad specificity for substrates, implying that this reaction can happen earlier in the pathway as well as at this stage (Wang et al., 2013).

1.3.7 The addition of the phytyl tail to form chlorophyll

The final step of the chlorophyll biosynthesis pathway is the addition of a phytyl tail to the propionate group on C17, catalysed by chlorophyll synthase (ChlG) (Rüdiger et al., 1980). It is thought to proceed by esterification of geranylgeranyl diphosphate (GGPP), followed by three successive hydrogenations of the alcohol by geranylgeranyl diphosphate reductase (ChlP) (Oster et al., 1997). ChlG is poorly characterised but is an integral membrane protein, and is thought to interact with proteins that assist in light harvesting protein assembly (Chidgey et al., 2014). These interactions seem to be evolutionarily conserved, as when non-native ChlG proteins from other chlorophyll producing organisms were introduced into ΔChlG *Synechocystis* the organism functioned as WT (Proctor et al., 2018).

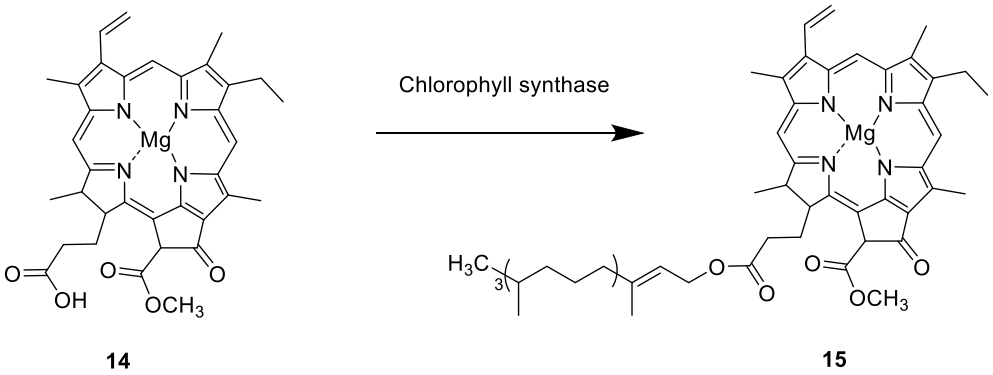


Figure 1.12. The addition of a phytol to monovinyl chlorophyllide (**14**) via esterification to produce the final product, chlorophyll (**15**). This reaction is catalysed by chlorophyll synthase.

1.4 Magnesium chelatase

One of two enzymes of this important biosynthetic pathway studied in this thesis is magnesium chelatase, a large multimeric complex that catalyses the first committed step of the chlorophyll biosynthesis pathway (Figure 1.8). Because of this position in the pathway, MgCH is at a key regulatory point in the production of chlorophyll, and thus the assembly of photosynthetic complexes. MgCH catalyses the ATP-dependent insertion of a magnesium ion into protoporphyrin IX. This substrate is shared with the biosynthesis pathway of haem so MgCH competes with the ATP independent ferrochelatase. Ferrochelatase inserts a ferrous (Fe^{2+}) ion into protoporphyrin IX to produce haem, another key tetrapyrrole cofactor. A photosynthetic cell must tightly control this branch point so that chlorophyll and haem production are balanced according to the needs of the cell in an environment (Chidgey et al., 2017). Variable levels of light and oxygen require photosynthetic cells to adjust the MgCH/ferrochelatase branchpoint to optimise chlorophyll and haem

levels, and the assembly of complexes involved in photosynthesis and respiration, two key energy generating processes for a photosynthetic cell.

Magnesium chelation was first reported by Gorchein in *Rhodopseudomonas spheroides* when protoporphyrin was fed to whole cells and MgPME was produced (Gorchein, 1972). It was postulated that the chelation and methyl transfer steps were obligatorily coupled as part of a multi-enzyme complex (complex formation between MgCH and ChlM has been modelled since, though not confirmed experimentally) (Chen et al., 2015a). However, the unmethylated and methylated products were subsequently separated (Fuesler et al., 1982). After the discovery of MgCH activity Gorchein (1973) found a requirement for ATP when uncouplers of the PMF were shown to inhibit the chelatase reaction. MgCH activity was later observed in higher plants, in etioplasts prepared from developing cucumber cells (Smith and Rebeiz, 1977). Castelfranco et al., (1979) found that L-glutamate increased the reaction rate of MgCH in cucumber etioplasts, revealing a potential regulatory role for the precursor to ALA. Until the development of recombinant expression systems in 1995-6, plant etioplasts continued to be the source of MgCH activity; it was found that at least three different components were necessary for catalysis and the porphyrin specificity of the enzyme was examined (Guo et al., 1998; Walker and Weinstein, 1991, 1994).

MgCH consists of three main subunits: protoporpyhrin binding H (*ca.* 140-150 kDa), ATPase I (*ca.* 40 kDa) and regulatory D (*ca.* 60-75 kDa). There are bacteriochlorophyll and chlorophyll homologues of MgCH produced by anaerobic phototrophs (purple sulphur bacteria and green sulphur bacteria) and aerobic phototrophs (cyanobacteria, algae and higher plants), respectively. Three bacteriochlorophyll MgCH subunits were identified as essential for production of magnesium protoporphyrin IX: BchH, BchI and BchD (Gibson et al., 1995). The chlorophyll producing homologues are ChlH, ChlI and ChlD, respectively (Jensen et al., 1996). A fourth subunit, Gun4, is required *in vivo* and greatly enhances catalysis *in vitro*. It is a small, porphyrin binding protein thought to be involved in an array of regulatory networks (Larkin et al., 2003).

The MgCH subunits are closely related to the subunits of the cobaltochelatase utilised in the aerobic synthesis of cobalamin (Heldt et al., 2005). These are grouped as Class I chelatases, while a group of small monomeric/ homodimeric chelatases such as ferrochelatase are grouped as Class II chelatases. Another class of homodimeric chelatases unrelated to Class II are involved in siroheme biosynthesis (Brindley et al., 2003). CobN shares a 31 % identity with BchH; CobT shares 14 % with BchD; and CobS 15 % with BchI. There are also homologies between the subunits; the N-

terminus of ChlD has *ca.* 40 % identity to ChlI and BchI, perhaps the result of a past gene duplication event (Jensen et al., 1996). CobT has similar N-terminal homology with CobS.

Chelation of cobalt into the tetrapyrrole precursor to cobalamin can occur via two different pathways, the aerobic or anaerobic (Mattes et al., 2017). Depending on the pathway, the chelation of cobalt happens at different points catalysed by two distinct enzymes, CbiK/X for anaerobic, CobNST for aerobic. CbiK/X is a small ATP-independent Class II chelatase similar to ferrochelatase (Brindley et al., 2003; Schubert et al., 1999). Similarities between CobNST and MgCH were observed when first purified and CobST are thought to adopt an oligomerisation similar to that of BchID (Debussche et al., 1992; Lundqvist et al., 2009)

1.4.1 The reaction

The individual subunits of MgCH from several different species have been recombinantly expressed and purified (Gibson et al., 1995; Jensen et al., 1996; Petersen et al., 1998; Willows et al., 1996). Assays have been developed and conditions include a Mg^{2+} concentration above that of ATP, showing the need for free magnesium as well as that bound to ATP in solution; an indication that the magnesium is not directly handed over from the $MgATP^{2-}$ molecule. As well as this, the *Synechocystis* enzyme showed a positive cooperative response to magnesium. (Jensen et al., 1998). Magnesium appears to have a much more complex role than just that of a substrate. These assays were simplified by the use of a more water-soluble analogue of P_{IX}, deuteroporphyrin IX (D_{IX}), which reacts with MgCH in much the same way (Karger et al., 2001).

The optimal ratios of the subunits for chelatase activity were 4:2:1 for ChlH:ChlI:ChlD; higher ChlD concentrations inhibit the reaction (Jensen et al., 1998). When substrates are added to the subunits, with a low protein to substrate ratio, there is a lag period prior to onset of chelation, indicating a slow complex formation upon substrate binding. This lag period was eliminated if ChlH was pre-incubated with porphyrin, $MgCl_2$ and ATP while ChlI and ChlD were pre-incubated with ATP and $MgCl_2$, leading to the conclusion that ChlH associates with porphyrin and ChlI with ChlD both in a $MgATP^{2-}$ dependent manner. ChlH was therefore proposed as the subunit with the site of chelation (Jensen et al., 1998). The lag period can also be eliminated if the enzymes alone are pre-incubated under high protein to substrate ratio (Viney et al., 2007).

ATP hydrolysis is coupled to magnesium insertion, with the maximal rate of chelation roughly equal to the maximal rate of ATP hydrolysis. ATP hydrolysis also differs with conditions that would affect the chelation rate such as porphyrin and ChlH concentration (Jensen et al., 1999).

The rate of ATP hydrolysis, when coupled to chelation by the MgCH complex, is much lower than that of ChlI alone (Reid and Hunter, 2004). Around 15 ATP molecules are hydrolysed per chelation reaction, resulting in a ΔG of between 25 and 33 $kJmol^{-1}$, chelation operating far from equilibrium (Reid and Hunter, 2004). The product of ATP hydrolysis, adenosine diphosphate (ADP), inhibits chelation, a potential regulatory feature (Jensen et al., 1999).

The chelation reaction was directly monitored with enzyme in excess of porphyrin and the rate-limiting formation of an intermediate state was attributed to either chelation of porphyrin within enzyme-porphyrin complex or to an isomerisation of the enzyme-porphyrin complex. The formation of this intermediate state was reversible, with an equilibrium constant in favour of the unchelated/ unisomerised enzyme-porphyrin complex; the forward reaction (i.e.: chelation/ isomerisation) is made more favourable by $MgATP^{2-}$ binding with the backwards rate of this slow step basically eliminated. Thus, MgCH may have two distinct states depending on the enzyme to substrate ratios (Viney et al., 2007). The specific mechanism linking ATP hydrolysis and magnesium chelation is still unclear; it is thought to be a thermodynamic driving force. When ATP is available, MgCH can maintain an equilibrium constant shifted by a factor of 10^8 in favour of the product state in comparison to the uncatalysed reaction (Adams et al., 2016a).

The mechanism of magnesium insertion has not been confirmed in MgCH but there has been extensive work on the iron chelation of protoporphyrin catalysed by ferrochelatase, involving enzymatic and non-enzymatic metal chelation kinetics and thermodynamics, information from crystal structures and computer modelling (Dailey et al., 2000). The ring of the porphyrin is distorted, and protons are removed to expose core nitrogen atoms to the incoming metal ion. The degree of distortion has been proposed to select for metal ion incorporation (Al-Karadaghi et al., 2006) and it seems probable that MgCH would follow a similar mechanism; ring distortion is observed when ChlH is bound to DIX (Karger et al., 2001).

1.4.2 ChlI and ChlD

Powering the chelation is the ID complex, which is presumed to assemble in a nucleotide-dependent manner from its constituent Bch/ChlI and Bch/ChlD proteins and then interact with Bch/ChlH-porphyrin molecules. Structurally, ChlD can be split into three different domains, an N-terminal ATPase domain, a C-terminal integrin I-like domain with the two connected by an acidic proline-rich domain. ChlI is homologous to the N-terminal domain of ChlD. The crystal structure of BchI has been solved (Fodje et al., 2001) (Figure 1.13), confirming that this N-terminal region belongs

to the superfamily of ATPases associated with various cellular activities (AAA+) (Section 1.4.2.2) (Neuwald et al., 1999).

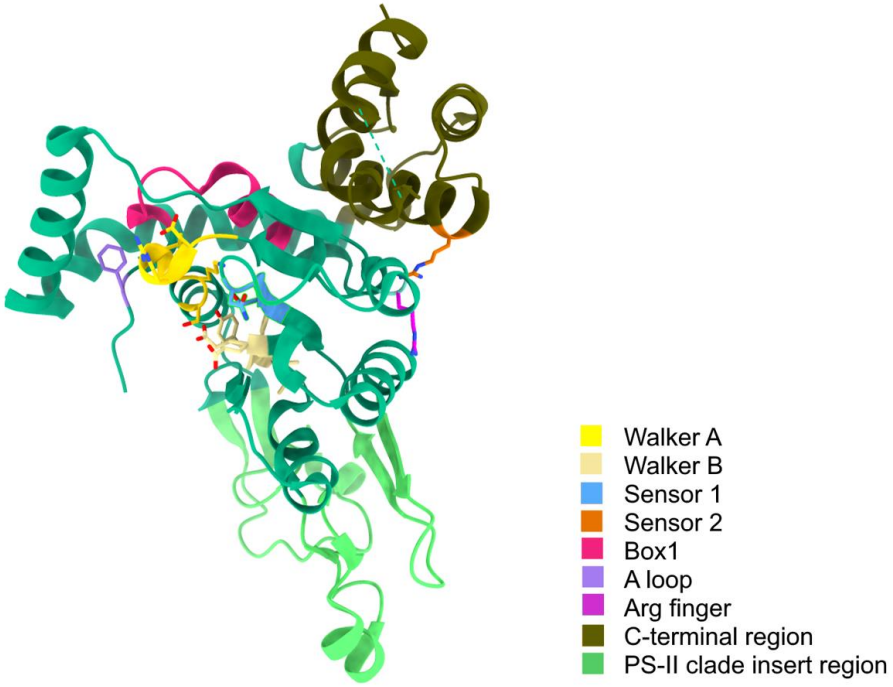


Figure 1.13. The 2.1 Å resolution crystal structure of BchI (PDB: 1g8p), coloured as in the key to highlight conserved features of the AAA+ proteins. Residues shown as sticks are involved in substrate binding. The C-terminal region is unusual in its position relative to the rest of the structure when compared to other AAA+ proteins (Fodje et al., 2001). It is unknown whether this is biologically relevant or a crystal packing artefact.

Because the structure of a whole MgCH complex has yet to be determined, the stoichiometry of the subunits is still unknown. Multiple different protein interaction events can occur between the subunits, not all of which are productive for chelation. *In vitro* studies with purified subunits show that ChlI and ChlD interact both with themselves and each other; dimeric ChlD has been observed for tobacco MgCH (Gräfe et al., 1999) and multimeric BchI/ChlI complexes have been observed *in vitro* in the presence of nucleotide (Hansson et al., 2002; Reid et al., 2003). Both hexameric and heptameric rings of ChlI have been seen under negative stain EM (Reid et al., 2003; Willows et al., 2004). Oligomerisation may not require ATPase activity; mutations that abolish ATPase activity in BchI can still form oligomers but are catalytically inactive (Hansson et al., 2002).

ChlI and ChlD interact with each other in a nucleotide dependent manner via the AAA+ domain (Adams et al., 2016b; Gräfe et al., 1999; Jensen et al., 1999, 2000; Papenbrock et al., 1997). Conserved residues normally involved in protein-protein interaction between AAA+ proteins, the sensor II and arginine finger regions, influence this interaction (Adams et al., 2016b), suggesting that ChlI interacts with ChlD in a similar manner to other proteins of the AAA+ family. As with the formation of ChlI oligomers, the ChlI-ChlD interaction does not require active ATP hydrolysis, only the presence of nucleotide (Jensen et al., 1999), and it still occurs when ChlI ATPase activity is abolished by N-ethylmaleimide (NEM) (Jensen et al., 2000). Mutations that would disrupt potential ATPase activity in ChlD do not disrupt ChlID assembly (Adams and Reid, 2013).

It has been suggested that BchI may interact with BchD via the acidic proline rich region, though with little experimental evidence (Fodje et al., 2001). A low resolution electron microscopy (EM) model of two stacked rings of *Rba. capsulatus* BchI and BchD proposed an interaction via the integrin I-like domain of BchD (Lundqvist et al., 2010). However, the EM map is of questionable quality and they were unable to fit the AAA+ domain of BchD inside the density (Figure 1.14). The BchID complex may well differ from the ChlID complex; the AAA+ domain of BchD lacks several inserted regions that are present in BchI that may prevent formation of a heterogeneous BchID ring. However, the AAA+ domain of ChlD is very similar to both BchI and ChlI, suggesting that a heterogeneous ChlID ring could be more likely.

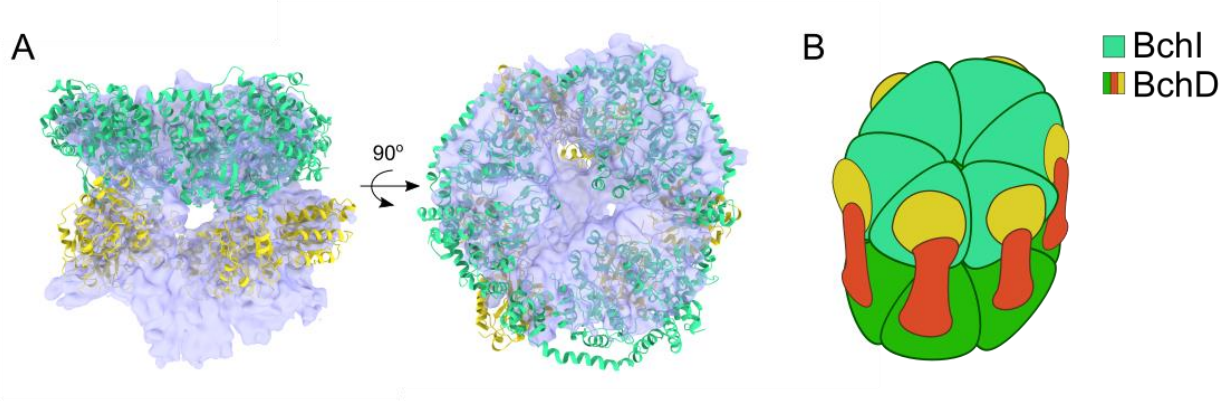


Figure 1.14. The BchID complex proposed by Fodje et al., (2001). **A**, EM map (EMDB: 1676; lilac) with crystal structure of BchI (cyan-green) and homology model of the C-terminal integrin I-like domain of BchD (yellow) fit into the map (PDB: 2x31). The extra density is believed to be for the AAA+ domain of BchD but there is insufficient room to fit six such domains, amounting to approximately 240 kDa. **B**, a schematic model of the whole BchID complex, coloured as in key, with the integrin I-like domain of BchD interacting with BchI. It is unclear how this proposed model would interact with BchH.

ChlI is an active ATPase that proceeds via an enzyme-phosphate complex (Adams and Reid, 2012). In the presence of ChlD the rate of hydrolysis is decreased (Jensen et al., 1999). Indeed, ChlD seems to play a regulatory role in the MgCH complex. It has not been observed to hydrolyse ATP but the binding of nucleotide is essential for chelation to occur (Adams and Reid, 2013; Jensen et al., 1999). Mutations in ChlD of conserved residues normally involved in ATP hydrolysis do not interfere with MgCH activity, only hamper ID complex formation (Adams and Reid, 2013). Phosphorylation of a histidine within ChlD by ChlI potentially plays a role in stimulating MgCH activity in oxygenic phototrophs (Sawicki et al., 2017). ChlD from *Synechocystis* has key glutamate residues in the C-terminal region of the protein that confer the cooperative response to magnesium found in this cyanobacterial enzyme (Brindley et al., 2015). As well as these key residues, arginine finger, sensor II and Walker B mutants of ChlD showed a decreased cooperative response to magnesium by MgCH (Brindley et al., 2015).

The C-terminus of ChlD is homologous to the Integrin I-like (or von Willebrand A) family of proteins (Fodje et al., 2001) (Section 1.4.2.1). Usually these are membrane proteins involved in cell to cell adhesion; a conserved metal ion dependent adhesion site (MIDAS) regulates protein-protein interactions (Luo et al., 2007). These proteins often bind specific sequences on their target protein. Both ChlI and ChlH have examples of these target sequences, RGD/RGE and LDV, respectively (Fodje et al., 2001). The architecture of (Chl/Bch)D is unusual as the combination of the AAA+ and integrin I-like domains has only been observed in one other protein: midasin, distantly related to the large subunit of dynein and involved in the assembly of the ribosome (Garbarino and Gibbons, 2002; Romes et al., 2016).

1.4.2.1 The integrin family

Integrins are large multi-domain transmembrane proteins that interact with the extracellular matrix of cells and promote a response to their environment; they participate in a wide range of signalling and membrane trafficking pathways (De Franceschi et al., 2015) . The integrin α I domain interacts with a vast array of ligands, and is homologous to the C-terminal domain of ChlD (Humphries et al., 2006). Crystal structures have been solved for several α I domain proteins both with and without the rest of the integrin, revealing a conserved Rossmann-like fold, though with no evidence of nucleotide binding. A different class of integrin I proteins (β I domains) do bind nucleotide in this region as well as the divalent metal (Luo et al., 2007).

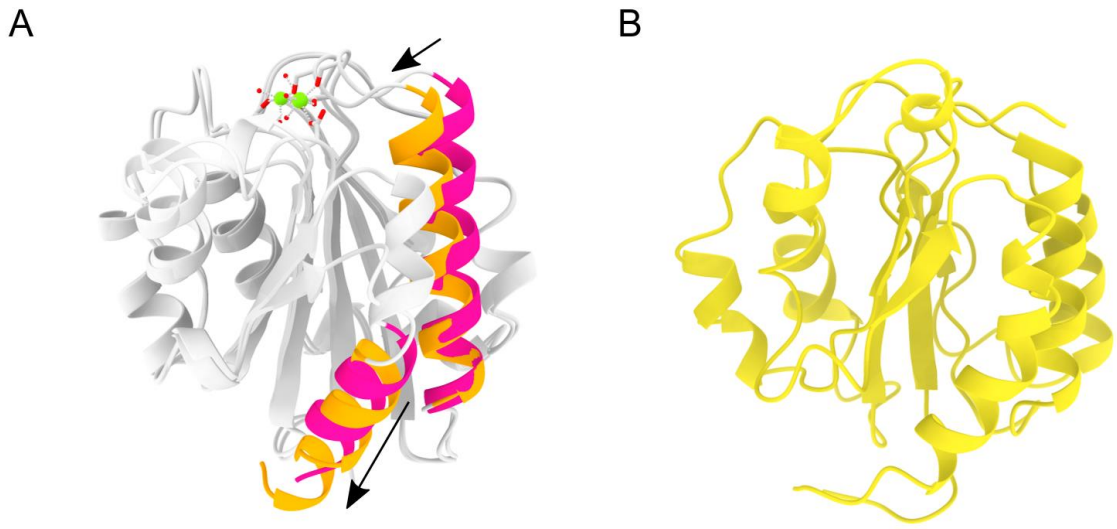


Figure 1.15. Integrin I domains. **A**, Example of a conformational change induced by different ligand coordination of the MIDAS region. Aligned structures of integrin CR3 in a closed (PDB: 1JLM; magenta) and open (PDB 1IDO; orange) conformation. Arrows show direction of movement from closed to open helices. Whether this conformational change is relevant to BchD/ChlD is unknown. **B**, Homology model of the C-terminal domain of BchD, showing the structural similarity (PDB: 2x31).

Central to the interaction with their ligands is the MIDAS region (DXSXS, where X is any amino acid), with a bound divalent metal ion cofactor (often magnesium). The magnesium is coordinated by three to five residues as well as water molecules in ligand free situations; however, this water is displaced upon ligand binding, which subtly changes the conformation of several helices in the α I domain. This subtle change is amplified over the rest of the integrin protein, spanning the cell membrane and causing a response within the cell (Lee et al., 1995).

1.4.2.2 The AAA+ superfamily

The AAA+ superfamily is very large with a diverse range of important functions in cell biology that includes cell cycle control, proteolysis, cell trafficking and transcription (Erzberger and Berger, 2006). AAA+ proteins often adopt hexameric ring structures in the presence of nucleotide, and cycle through ATP binding, hydrolysis and then release of ADP; each of these steps is coupled to structural changes within the protein linked to rearrangements elsewhere, transducing a physical force from the ATP hydrolysis cycle (Hanson and Whiteheart, 2005).

Several structural features of AAA+ proteins have various degrees of conservation, including essential Walker A and B motifs, involved with nucleotide binding. The aspartate of the hhhhDE

sequence of Walker B (where h is any hydrophobic residue) coordinates a magnesium ion required for catalysis. The glutamate residue of the Walker B motif can be mutated to inactivate AAA+ proteins, whilst not disrupting ATP binding, creating a substrate trap (Wendler et al., 2012). Three other structural elements are also involved in catalysis and substrate binding, the Arginine finger and Sensor 1 and 2 regions. Sensor 1 and 2 are involved in coordinating the γ-phosphate of ATP, though sensor 2 coordinates the ATP in a neighbouring subunit when in the oligomeric state. The arginine finger is also required for intersubunit substrate coordination. In addition to the core AAA+ component there are several additional features that can be used to subdivide the super-family, in the form of insertions and differences in the C-terminal region that may have functions related to the biological role of the specific protein (Figure 1.13) (Snider et al., 2008).

1.4.3 ChlH

ChlH is a large (*ca.* 150 kDa) porphyrin binding protein that is presumably the site of chelation (Jensen et al., 1998). Structural information was limited to a low resolution negative stain EM and SAXS model (Qian et al., 2012) until the recently solved crystal structure of the *Synechocystis* enzyme (Chen et al., 2015a); both models show a similar overall architecture with a large globular body connected to a smaller N-terminal "head" by a "neck". It was thought from the EM structure that the internal cavity for porphyrin binding was large but the crystal structure suggests that it is instead quite small. However, the crystal structure has no bound porphyrin so this is just a putative active site. Seemingly large structural changes were observed by negative stain EM in the presence of porphyrin (Qian et al., 2012; Sirijovski et al., 2008). The low resolution of the structures makes it difficult to interpret the functional significance however. The "head" of H can be removed and the rest of the protein folds well and still binds porphyrin (Adams et al., 2014); this truncated protein is severely impaired catalytically but rescuable to an extent by Gun4.

For such a large protein, there are several point mutations remote from proposed active site that inactivate MgCH. For example, A942 and P595 (*Synechocystis* numbering) are important for ChlH catalytic activity, but not for substrate binding, and the effects can be relieved by addition of Gun4 (Davison and Hunter, 2010). It is not understood why H is so large in comparison to its substrate, but it must interact with the ID motor complex, Gun4, and possibly other proteins in the biosynthetic pathway such as ChlM.

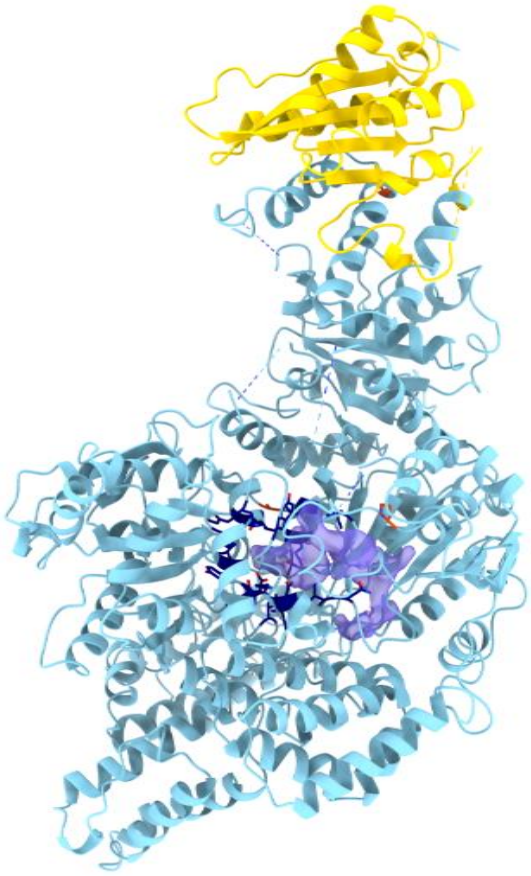


Figure 1.16. The 2.5 Å resolution crystal structure of the ChlH subunit of MgCH represented in ribbon format. In yellow is the N-terminal region that can be removed without disrupting folding, although removal severely impairs activity. Navy residues represented as sticks show the proposed active site, with the cavity shown as a purple surface. A942, P595 and G195, residues distal to the active site that affect activity when mutated are shown in red. There are many unstructured loop regions (dashed lines).

The gene encoding ChlH is expressed in a wide variety of plant tissues, only some of which make chlorophyll. In conflicting reports, several studies have reported that ChlH is a receptor of abscisic acid, potentially positively regulating stomatal opening, post-germination growth and germination in *Arabidopsis thaliana* (Cutler et al., 2010; Shen et al., 2006). It has also been suggested that ChlH is an anti-sigma factor in cyanobacteria, helping to regulate sugar catabolism (Osanai et al., 2009). Within the chlorophyll biosynthetic pathway, ChlH interacts with Magnesium-protopor-

phyrin O-methyltransferase (Bch/ChlM), the following enzyme in the (bacterio)chlorophyll pathway, stimulating the activity of ChlM, and being potentially involved directly in the reaction (Shepherd et al., 2005).

1.4.4 Gun4

Gun4 is a small (*ca.* 26 kDa) soluble protein essential for growth in many chlorophyll producing organisms, with the ability to bind various porphyrin-like molecules (Larkin et al., 2003); it enhances the catalytic activity of MgCH, reducing the magnesium concentration needed for activity at low D_{IX} concentrations (Davison et al., 2005). Gun 4 is thought to increase catalytic efficiency of the enzyme by aiding in the supply and removal of the substrate/product pigment from ChlH, potentially trafficking it from/to the next enzyme in the pathway, PPO/ChlM (Shepherd et al., 2005; Verdecia et al., 2005). The propionate group of MgD_{IX} that is methylated by ChlM has been observed to protrude from the binding cleft of Gun4, potentially exposing it to ChlM (Chen et al., 2015b). The binding of Gun4 to porphyrin causes Gun4 and ChlH to associate with membranes in *A. thaliana*, potentially improving chelation rates by associating ChlH with the membrane bound PPO (Adhikari et al., 2009).

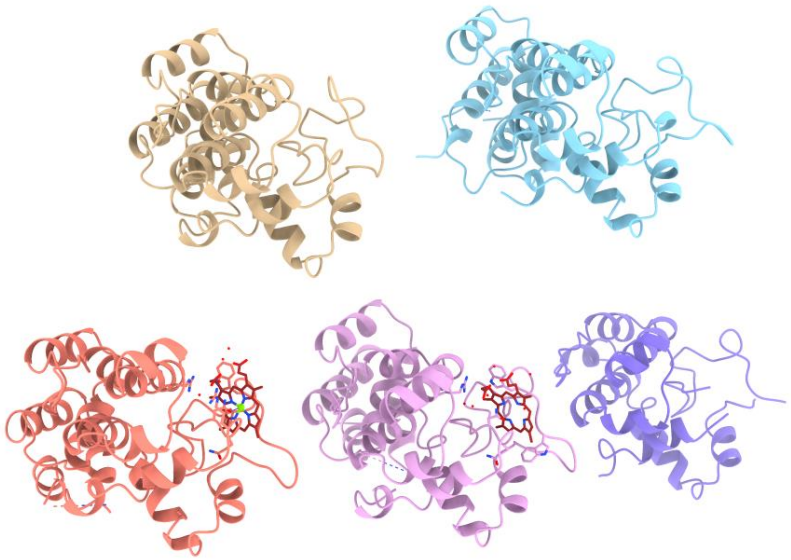


Figure 1.17. The crystal structures of Gun4 from various organisms. *Synechocystis* apo-Gun4 (cream, PDB: 1Y6I); *Thermosynechococcus elongatus* apo-Gun4 (cyan, PDB: 1Z3Y); *Synechocystis* MgD_{IX}-Gun4 (red, PDB: 4XKC); *Synechocystis* D_{IX}-Gun4 (pink, PDB: 4XKB); *Chlamydomonas reinhardtii* apo-Gun4 (purple, PDB: 4YKB). The *C. reinhardtii* protein has a significant N-terminal truncation in comparison to the other organisms.

The crystal structures of apo Gun4 have been solved for *T. elongatus*, *Synechocystis* and *Chlamydomonas reinhardtii* (Davison et al., 2005; Tarahi Tabrizi et al., 2016; Verdecia et al., 2005). Structures of *Syn*Gun4 bound to both D_{IX} and MgD_{IX} have also been solved, confirming the binding site proposed earlier (Chen et al., 2015b). This binding site is at the C-terminal end of Gun4, with flexible loops between α2/3 and α6/7 forming a hydrophobic "greasy palm" to bind the hydrophobic porphyrin (Verdecia et al., 2005). Porphyrin binding by Gun4 could be disrupted by phosphorylation of the C-terminus in angiosperms, reducing the stimulation of MgCH (Richter et al., 2016).

Gun 4 plays a key regulatory role in chlorophyll biosynthesis by controlling the efficiency of ChlH. This perturbation in the biosynthesis pathway was observed to create a plastid to nucleus signal (Mochizuki et al., 2001). Specifically, an accumulation of MgP_{IX} represses nuclear transcription of various plastid genes (Larkin et al., 2003). It is unclear if this signalling is due to regulation of MgCH alone, or if it helps to traffic MgP_{IX} to the nucleus.

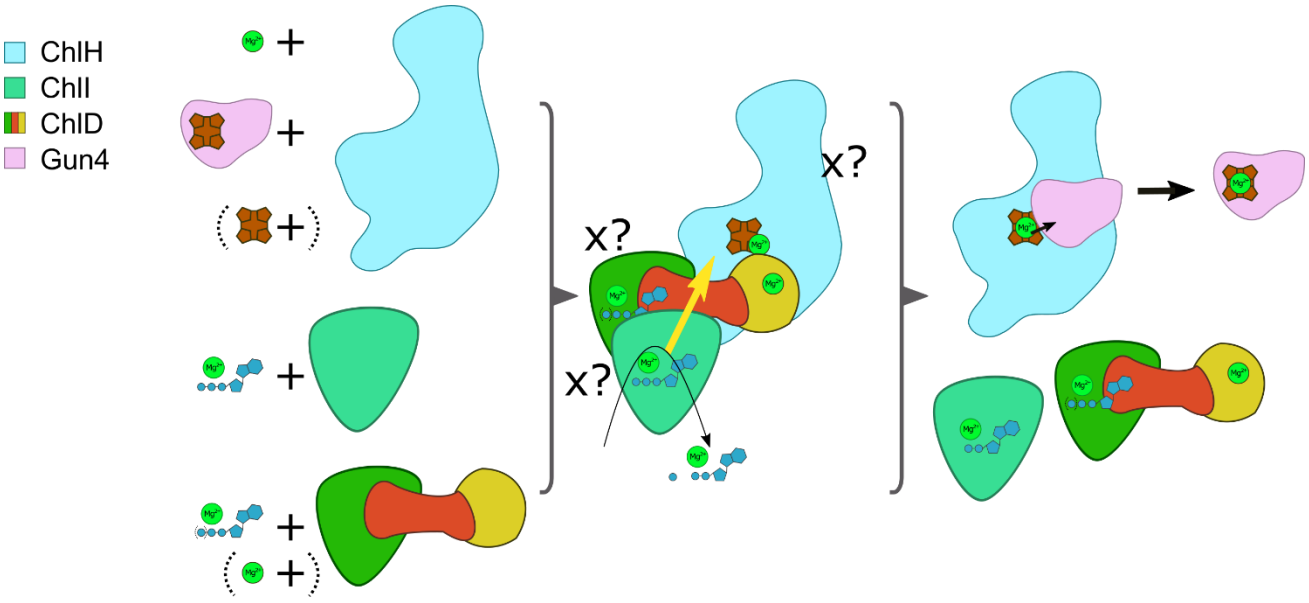


Figure 1.18. The overall reaction of magnesium chelatase. Proteins coloured as in key; ChlD AAA+ domain in green; acidic proline-rich region in red; integrin I-like domain in yellow. ChlH binds magnesium and P_{IX} (brown). The porphyrin is likely delivered by Gun4 *in vivo*, however this is not necessary with *in vitro* assays. ChlI binds $MgATP^{2-}$ and ChlD binds nucleotide ($MgADP^{2-}$ or $MgATP^{2-}$, it is unknown which promotes catalysis). ChlH, ChlI and ChlD come together to form a complex of currently unknown stoichiometry where ATP hydrolysis in ChlI is somehow coupled to chelation of P_{IX} in ChlH (yellow arrow). ChlD has no observed ATPase activity. Gun4 could also play a role in increasing product release rate. It is also unknown if the complex disassembles after each chelation event, or whether the ID complex may stay together and activate other ChlH molecules. It has been suggested that BchID complex may disassemble between chelation steps (Lundqvist et al., 2013).

It has been suggested that binding of Gun4 to porphyrin also plays a photoprotective role. Porphyrin molecules are photosensitive and can generate reactive singlet oxygen species (ROS) when exposed to high light levels. It was suggested that the "greasy palm" shields the molecules from oxygen and water (Verdecia et al., 2005). However, it has recently been observed *in vitro* that a Gun4-P_{IX} complex actually increases the rate of ROS generation compared to P_{IX} alone. ROS are also retrograde signalling molecules, so Gun4 may play a role in this signalling pathway too (Tarahi Tabrizi et al., 2016).

1.5 Protochlorophyllide oxidoreductase

Protochlorophyllide oxidoreductase (POR) catalyses the penultimate step of the chlorophyll biosynthetic pathway, the reduction of the C17-C18 double bond in protochlorophyllide (Pchlide) to form chlorophyllide (Chlide) (Figure 1.11). There are two distinct types of POR, light-dependent (LPOR) and light-independent, or dark-operative (DPOR). Most photosynthetic organisms contain copies of both LPOR and DPOR with the exceptions being anoxygenic photosynthetic bacteria which, for the most part, contain only DPOR, and angiosperms (flowering plants) which contain only LPOR (Kaschner et al., 2014; Schoefs and Franck, 2004).

1.5.1 DPOR

DPOR is encoded by three genes, *chlL*, *chlN* and *chlB* (in chlorophyll producing organisms), which are related to nitrogenase enzymes. DPOR, like the related nitrogenase, is oxygen sensitive so having both LPOR and DPOR has been suggested as an adaptation to light and oxygen levels in the environment, especially for photosynthetic bacteria (Yamazaki et al., 2006).

A crystal structure of DPOR from *R. capsulatus* has been solved (Muraki et al., 2010), revealing a heterotetrameric arrangment of N and B components of DPOR form the site of Pchlide reduction, while a dimer of L proteins supplies electrons for the reaction sourced from ATP hydrolysis (Nomata et al., 2016). This cluster is uniquely bound by three cysteine and one aspartate residues, the aspartate being essential for catalysis. The proposed mechanism of catalysis is as follows: electrons are transferred from an [4Fe-4S] cluster bound by the NB complex to Pchlide, generating an anion radical which is then protonated by the C17 propionate group of Pchlide; the group is distorted over the C17=C18 so that the carboxyl is in close proximity (Figure 1.19) (Muraki et al., 2010; Nomata et al., 2014). Another electron is supplied by the [4Fe-4S] cluster and a final proton is added by an aspartate residue. The donation of a proton from C17 propionate is suggested as no other proton donating residues are near to the Pchlide substrate.

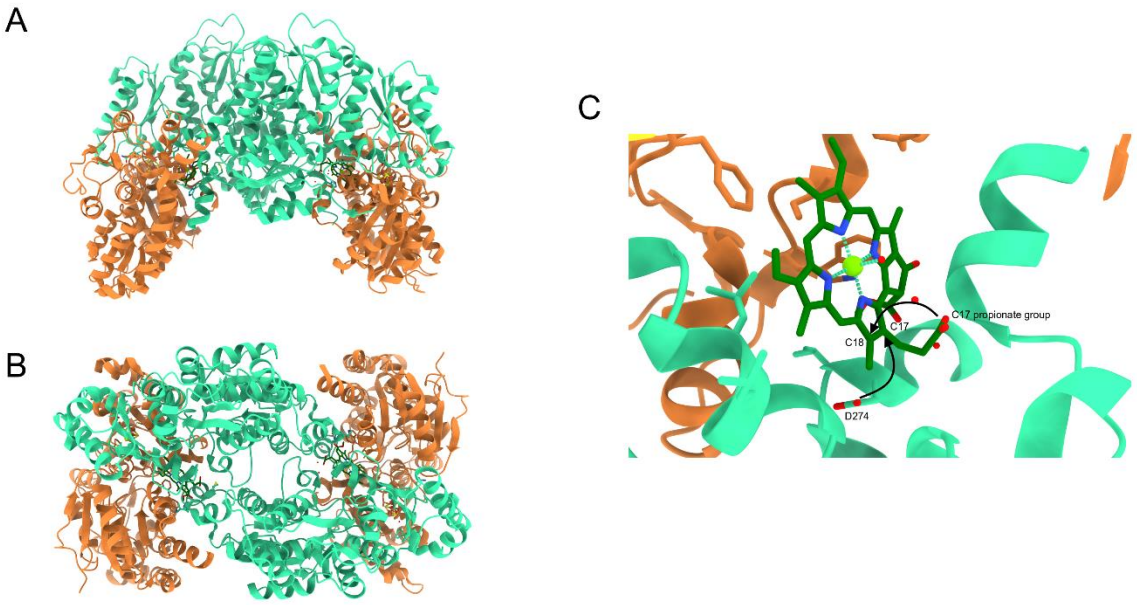


Figure 1.19. The 2.3 Å resolution crystal structure of *Rhodobacter capsulatus* DPOR NB component in complex with Pchlide (PDB: 3AEQ). **A**, side view of the complex (orange, N; cyan-green, B). **B**, top view of the complex. **C**, one of the two active sites, showing proposed mechanism of reduction. C17 is hydrogenated by D274, then C18 by the C17 propionate group distorted over ring D.

1.5.2 LPOR

LPOR is a member of the short-chain dehydrogenase/ reductase (SDR) family and therefore requires nicotinamide nucleotide, specifically NADPH, as a cofactor for the reaction (Baker, 1994; Kavanagh et al., 2008). It is an important regulatory step in the chlorophyll biosynthetic pathway due to the reaction being driven by light, and this is especially relevant in angiosperms; in newly developing chloroplasts hidden away from light (known as etioplasts) within a developing plant, POR molecules bound to their substrates form prolamellar bodies (PLB), large crystalline membrane arrays also containing lipids and carotenoids. Upon exposure to light, Pchlide is reduced to Chlide; these arrays break down and chlorophyll production along with thylakoid development is promoted (Boddi et al., 1989). PLBs are also thought to have a photoprotective and light harvesting role for developing etioplasts (Reinbothe et al., 1999).

As well as biological significance, the light dependence of the reaction allows for study of the ultrafast catalytic mechanisms, for example involving oxidation of NADPH, not normally available for reductases or indeed any enzyme. POR and its substrates can be assembled in the dark to form a ternary complex and the reaction can only be initiated by a femtosecond pulse of light, allowing for ultrafast measurements of the catalytic cycle. In contrast, the “dead time” for mixing

the reactants limits access to ultrafast (picoseconds or less) processes. Moreover, Pchlide reduction gives rise to distinct and readily measured spectral signals of the substrate, product, substrate bound and intermediate states. Thus, POR can be regarded as a generally useful model enzyme for dehydrogenation/ reduction and proton/ hydride transfer reactions (Heyes and Hunter, 2005).

1.5.3 SDR proteins

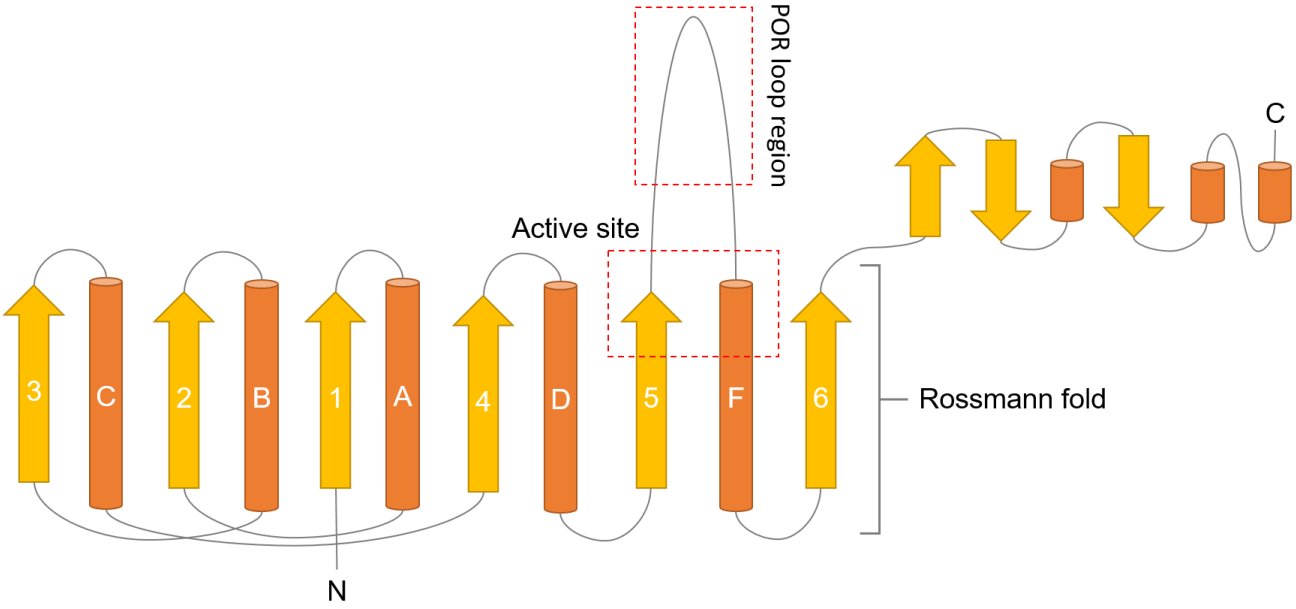


Figure 1.20. Schematic of the predicted overall architecture of POR proteins, highlighting the conserved Rossman fold of SDR proteins. β-sheets as yellow arrows, α-helices as orange cylinders. The large loop insert in POR between strand 5 and helix F is localised close to the active site, highlighted by the dashed-line box and label. POR has an extra C-terminal region, potentially involved in regulating oligomerisation (Ghosh et al., 2001).

POR belongs to the short-chain dehydrogenase/reductase superfamily of enzymes, which catalyse the reduction of several different moieties, mainly carbonyl and hydroxyl groups but also alkene (such as POR) and imine groups (Kavanagh et al., 2008). The unifying feature of SDR proteins is the presence of a Rossmann fold which binds a nucleotide cofactor required for reduction (Lesk, 1995). The nucleotide binds in a 4-pro S configuration, in contrast with medium chain dehydrogenase/reductases which uses the 4-pro R hydride for hydride transfer (Begley and Young, 1989; Kavanagh et al., 2008). The active site also contains a catalytic tyrosine, with a nearby lysine which lowers the pKa of the tyrosine in order to facilitate catalysis (Benach et al., 1999; Shafqat et al., 1996). A conserved asparagine facilitates proton relay via a water molecule coordinated nearby the aforementioned lysine which can be transferred to the catalytic tyrosine (Filling et al.,

2002). A Gly-rich motif specifies whether NADPH or NADH is the cofactor, and an acidic residue coordinates the 2' and 3' hydroxyls of the ribose component of the cofactor.

1.5.4 Reaction of LPOR

The reaction of LPOR was initially studied by isolating etioplast membranes, resulting in identification of the substrates for the reaction. A distinct spectral shift is observed upon reduction of the C17-C18 double bond of Pchlide, altering the conjugation system of the macrocycle. Pchlide has a prominent Q_y peak at *ca.* 630 nm, while in Chlide this is red shifted to *ca.* 665 nm. This spectral change allowed preliminary characterisation of the reaction catalysed by POR, including the necessity of cysteine residues, and determination of K_M values for Pchlide and NADPH. The substrate specificity of POR was also investigated, including the stereospecificity for the *pro-S* hydrogen of NADPH; Pchlide analogues esterified on the D-ring, along with analogues containing no central Mg^{2+} ion (although a central Zn^{2+} ion can be sufficient to restore activity), are not substrates for the enzyme (Griffiths, 1975, 1978, 1980; Valera et al., 1987). Purification of POR from oat (*Avena sativa*) allowed more information to be gained on substrate specificity, showing the importance of rings D and E, whilst modifications of the A and B rings are more tolerated. This study suggested that bound Pchlide is orientated with rings D and E forming more interactions with the POR active site residues, as may be expected for reduction of the D ring (Klement et al., 1999).

Initially the gene encoding POR from pea (*Pisum sativum*) was expressed in *Rhodobacter sphaeroides,* but the recombinant protein was in low abundance; the more conventional *Escherichia coli* expression systems were then used, producing POR as a fusion with maltose-binding protein (MBP) (Martin et al., 1997; Wilks and Timko, 1995). Improvements in recombinant expression systems allowed large scale production of POR from the cyanobacteria *Synechocystis* sp. PC6803 and *Thermosynechococcus elongates,* which provided the protein for thorough characterisations of the reaction and photochemistry (Heyes et al., 2000; McFarlane et al., 2005; Townley et al., 1998).

Michaelis-Menton kinetics were observed for the *Synechocystis* C-terminally tagged POR_{His6} with V_{max} values on the order of μM for both substrates. The K_{Dapp} value for NADPH was very low, only 11 nM, indicating a very tight binding, with a higher value for Pchlide of *ca.* 7 μM. Similar experiments were conducted with the *T. elongatus* enzyme, which has subsequently been used for most further experiments into the reaction mechanism (McFarlane et al., 2005). NEM treatment

reacts with reduced cysteine residues, allowing insights into their functional roles (Heyes et al., 2000, 2002).

Further investigation into these cysteine residues identified C226 (*T. elongatus* numbering) as important in Pchlide binding while C37 was shown to be involved in NADPH binding (Menon et al., 2010). It was later discovered that the C226 mutant may bind Pchlide in a different orientation to the WT, as time-resolved absorption techniques revealed spectral changes consistent with a hydride being added to the C18 first then to C17, the reverse order of hydrogenation in comparison with WT (Archipowa et al., 2018).

Low-temperature fluorescence spectroscopy has been repeatedly used with POR to observe independent steps in the reaction mechanism. The MBP-*P. sativum* PORb fusion was used to show that there is an initial intermediate generated by a light pulse, followed by a light-independent conversion to Chlide, both of which can occur at sub-zero temperatures, implying limited protein motions (Lebedev and Timko, 1999). This was repeated at sub-200 K with the MBP-free *Synechocystis* enzyme, showing that the light-dependent reaction can still occur. A non-fluorescent intermediate with absorption at *ca.* 695 nm was observed after the light pulse, which was suggested to be an ion radical, though it was then reappraised as an excited state with a intramolecular charge transfer (ICT) character (Heyes et al., 2006, 2014). Recently, however, upon further probing of the hydride/ proton transfer by time-resolved absorption spectroscopy no evidence of an ICT state was found, only excited singlet and triplet states (Archipowa et al., 2018).

By studying the thermophilic *T. elongatus* enzyme using low-temperature fluorescence spectroscopy, two further dark steps were observed that allowed for a more complete picture of the product release and binding steps. This work showed that $NADP^+$ must be replaced by NADPH before the Chlide will dissociate from the enzyme to be replaced with fresh Pchlide (Heyes and Hunter, 2004). It was also found that $NADP^+$ release is biphasic, with a fast structural rearrangement, before release of $NADP^+$ (Heyes et al., 2007). Substrate binding also relies on conformational alterations, three distinct phases were observed for NADPH binding, and two for Pchlide binding. The large structural alterations observed could reflect a switch in the state of oligomerisation (Heyes et al., 2008a).

The dark reactions were found to proceed only after a rise in temperature. Activation energies and binding energies were calculated from the temperature dependencies of these reactions and it was suggested that some of the absorbed energy from the photon could be used to drive the dark reactions (Heyes et al., 2002).

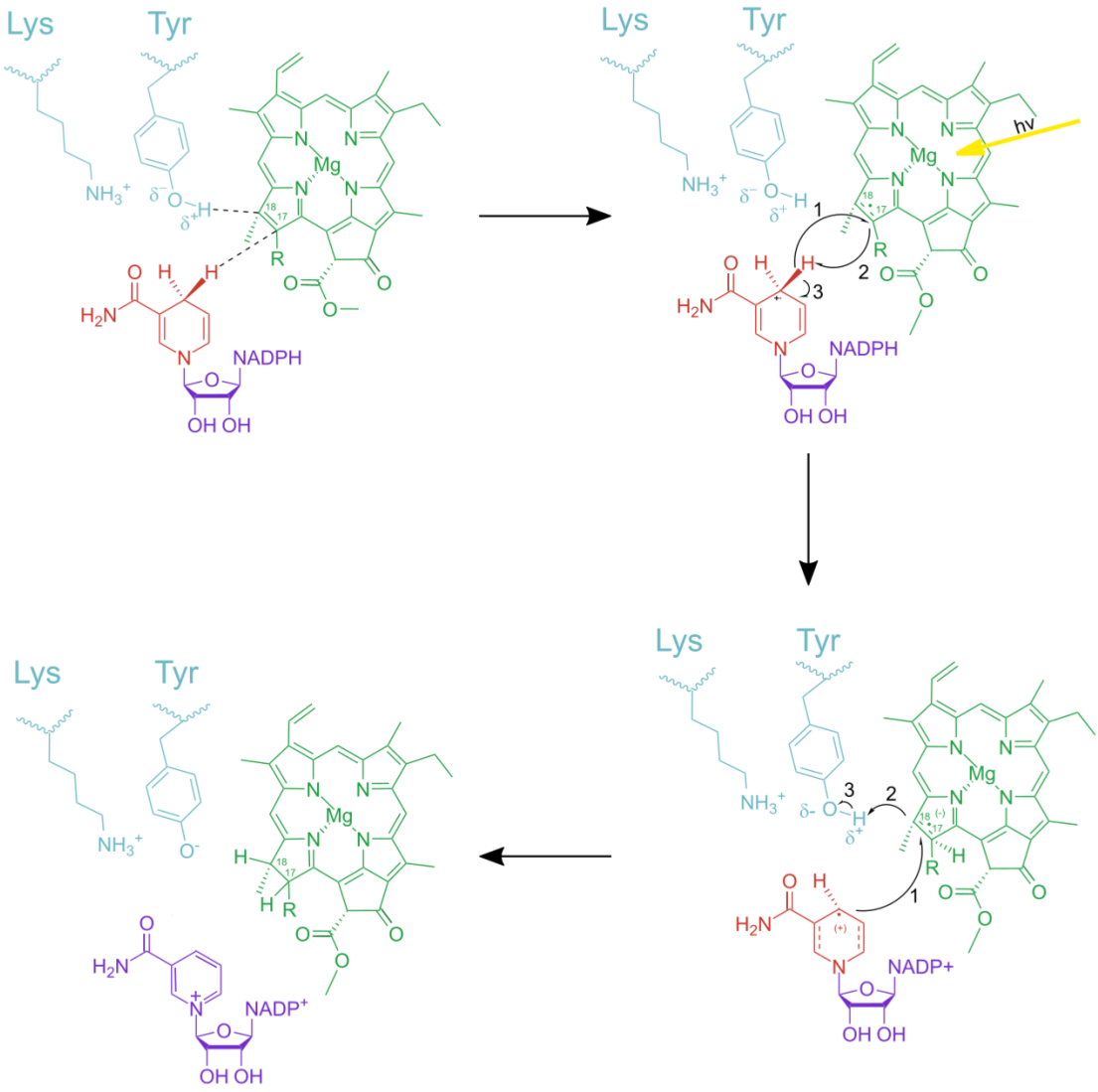


Figure 1.21. The POR active site and most recently proposed reaction mechanism (Archipowa et al., 2018). A step-wise proton/hydride transfer occurs first by electron transfer from NADPH to generate an anionic radical Pchlide species, promoting subsequent NADPH provided proton donation. After this, a further electron is supplied to the Pchlide-Chlide intermediate by the NADP species, which then receives the second proton from the catalytic tyrosine residue.

Like other SDR proteins, POR has several highly conserved residues in the proposed active site that are thought to be important for catalysis. Tyrosine 189 and lysine 279 mutations inactivate the *Synechocystis* protein as K279 is thought to lower the pK_a of the Y189 to facilitate deprotonation of Pchlide. It is also thought to help co-ordinate the NADPH for *pro-S* hydride transfer (Heyes and Hunter, 2002). More mutations of these residues in the *T. elongatus* enzyme revealed that the

tyrosine residue is not completely essential for catalysis, with the Y193F mutation showing *ca.* 20% activity of WT. However, this work also revealed more than just a catalytic necessity of these residues. It was suggested that they are also involved in substrate interaction and Pchlide excited state stabilisation, allowing for a more efficient hydride transfer (though hydride transfer rate was unaffected by the mutation) (Menon et al., 2009).

The exact reaction mechanism of reduction catalysed by POR has remained elusive, as suggested above, with multiple proposed theories of hydride/ proton transfer. Chlide formation has been suggested to proceed by both/ either concerted and sequential hydride transfer, via an ion radical, ICT state and by electron transfer to a singlet excited state of Pchlide (Archipowa et al., 2018; Heyes et al., 2003, 2006, 2014). However, most evidence suggests that interaction of the C18 carbonyl group of Pchlide with the conserved lysine 279 is key in enabling the excited state of Pchlide to be reduced by POR with NADPH (Archipowa et al., 2018; Heyes et al., 2014; Sytina et al., 2010). The most recent work using time-resolved absorption and emission spectroscopy suggests a stepwise hydride transfer where an electron is transferred to the excited singlet Pchlide molecule, followed by a proton coupled electron transfer and then the final proton (Archipowa et al., 2018).

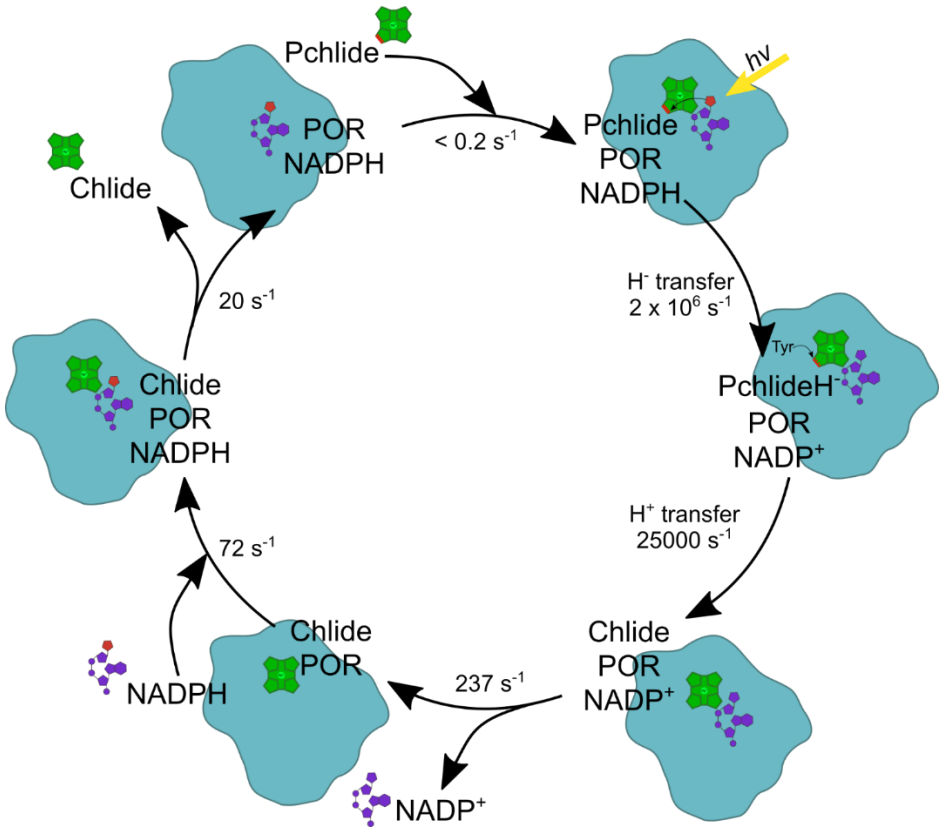


Figure 1.22. Overall reaction scheme of POR, rate constants derived for the *T. elongatus* enzyme (Heyes et al., 2007, 2008a, 2008b). Adapted from Scrutton et al. (2012).

1.5.5 Photochemistry of Pchlide

The substrate of POR, Pchlide, has been studied in the absence of POR in an attempt to understand the photochemistry behind the enzyme-catalysed reaction. It has been found that the excited state of Pchlide differs in solvents of different polarity; in polar solvents such as methanol and acetonitrile an intermediate of charge transfer character can form after initial relaxation, but not in the non-polar solvent cyclohexane (Dietzek et al., 2006). Indeed, various aspects of the fluorescence and absorption spectra of Pchlide can be attributed to the hydrogen bonding network solvating the molecule. These bonds are strengthened in the excited state of Pchlide (Zhao and Han, 2008). TR-IR characterisation showed that Pchlide enter a long-lived low-lying excited triplet state, highlighting the importance of POR to performing catalysis quickly to avoid this highly reactive and potentially damaging excited state (Dietzek et al., 2009).

1.5.6 Hydrogen tunnelling in POR

The role of quantum effects in biological systems has garnered some attention recently; evidence has emerged for a variety of processes utilising quantum mechanical properties. These include magnetoreception that birds use for navigation and electron transport in photosynthesis to improve light-energy harvesting efficiency (Lambert et al., 2013). In addition, hydrogen tunnelling events may be a driving force in several proton/hydride transfer enzymes, challenging the transition state theory of enzyme catalysis (Masgrau et al., 2004, 2006, 2007). The reduction of Pchlide catalysed by POR may be grouped with such hydrogen tunnelling enzymes, on the basis of kinetic isotope effects with temperature and solvent dependences consistent with proton tunnelling reactions. The sub-picosecond promoting motions needed for effective quantum tunnelling are potentially key for efficient catalysis (Heyes et al., 2008b, 2009). Comparison of POR from three species showed that while hydride transfer seemed to be conserved, there were differences in proton transfer between cyanobacterial and non-cyanobacterial POR proteins. Non-cyanobacterial POR seem less dependent on the temperature and solvent, potentially being more optimally configured for proton transfer (Heyes et al., 2011).

1.5.7 Structural information on POR

At the time of writing, there are no reports of the structure of POR, although a homology model was produced some time ago that aided with mutagenesis efforts to understand the enzyme (Townley et al., 2001). Since then a new updated model has been suggested based on refinements guided by simulated IR and visible spectra of the active site (Gholami et al., 2018). An updated

homology model with input from nuclear magnetic resonance (NMR) spectroscopy attempted to account for regions in the structure that have limited homology to other solved SDR proteins (Armstrong, 2014). These regions include a loop insert, presumed to be involved in substrate binding, and a C-terminal helix. The loop is proposed to operate in a similar manner to dihydrofolate reductase (DHFR), another SDR protein, in that it allows the protein to adopt an open and occluded state. Thus, the binding and release of substrates/ products could be controlled by conformational switching in the presence of its ligands (Heyes and Scrutton, 2009).

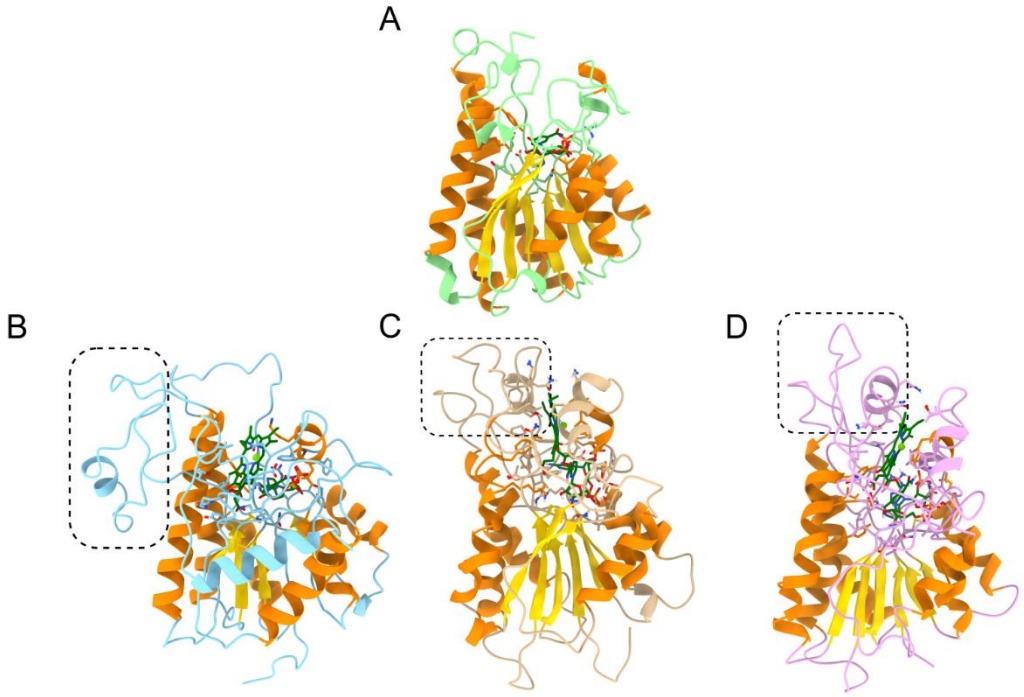


Figure 1.23. Structural models of POR. (A) A classic example of an SDR protein, 3α,20β-hydroxysteroid dehydrogenase (PDB: 2hsd; green) with core SDR β-strands and α-helices shown in yellow and orange, respectively. (B-D) Structural models of POR from (B, D) *T. elongatus* (Armstrong, 2014; Gholami et al., 2018) and (C) *Synechocystis* (Townley et al., 2001). The dashed box shows the loop region involved in substrate binding. The loop in (B) is based on carbonyl reductase enzymes where the movement of this loop upon substrate binding promotes oligomerisation. There is no consensus on the correct model.

The apo-POR protein can be monomeric but the oligomerisation upon substrate binding makes structural work on the ternary complex challenging. Cross-linking mass-spectrometry (XL-MS) has been used to try and elucidate some information about the interactions between POR proteins in *A. thaliana*. The results suggested oligomerisation can actually occur in both substrate-free and

bound POR and that active sites between POR monomers become closer upon substrate binding (Gabruk et al., 2016).

1.6 Techniques

1.6.1 Cryo-electron microscopy

There are several methods used to study biological molecules at the atomic level: macromolecular X-ray crystallography (MX), NMR spectroscopy and EM. EM has been used to visualise and calculate models of biological molecules for over fifty years (De Rosier and Klug, 1968). These models were limited in the resolution of structural information to tens of Angstroms apart from for molecules with significant symmetry or the propensity to form 2D crystalline arrays (Böttcher et al., 1997; Henderson and Unwin, 1975; Kühlbrandt et al., 1994; Miyazawa et al., 2003; Nogales et al., 1998). However, in recent years EM under cryogenic conditions (cryo-EM) has seen a "resolution revolution" due to various technological advances in both hardware and software. It has become, in the past 3-5 years, almost routine to determine structures of biological particles as small as 150 kDa. Cryo-EM is now an extremely valuable tool for studying the supramolecular organisation of large, multimeric protein complexes in a near-native state.

In an electron microscope, analogies can be drawn with conventional light microscopes in terms of optics. Electrons, however, can be energised to have a wavelength much smaller than visible light, so they can interact with a sample on an atomic level, in contrast with the roughly 200-400 nm resolution of diffraction-limited optical microscopy. High energy electrons can be focused by magnetic lenses, but they are significantly less precise than glass lenses, so aberrations are more of an issue. As focused electrons pass through the sample, they can be scattered, refocused and then detected. This results in information about the Coulomb potential of the sample at various points in space (Glaeser, 2016).

The organic molecules of life mostly use elements which are weakly scattering, so contrast is a significant issue; it has been and will continue to be a major hurdle for the EM community. This issue is compounded by the fragility of the samples; a low dose of electrons has to be used otherwise damage to the sample occurs, ruining high resolution structural information. Staining biological samples with heavy metals helps in overcoming these issues because heavy metal atoms are strongly scattering, so contrast is much higher. They also act as a protective layer so a higher dose can be applied. However, by staining with heavy metal salts all high resolution information is lost; models resulting from negative stain EM are generally unable to go beyond 20 Å resolution.

Photographic film was first used to detect electrons, which was mostly replaced with charge-coupled devices (CCDs), which are easier to work with. However, these detectors have a low detective quantum efficiency (DQE), so signal to noise is more of an issue (Murata and Wolf, 2018).

As part of the so-called "resolution revolution" traditional CCD detectors have been replaced with complementary metal–oxide–semiconductor (CMOS) detectors that are radiation protected and ultrathin to reduce indirect detection of electrons. The devices directly detect electrons and can be operated in a low-dose collection mode where single electron events can be recorded over a longer exposure, greatly improving DQE. CMOS detectors can also collect at a high frame rate which means that "movies" of the sample can be collected, to allow correction for beam-induced sample motion and improve blurred images (McMullan et al., 2016).

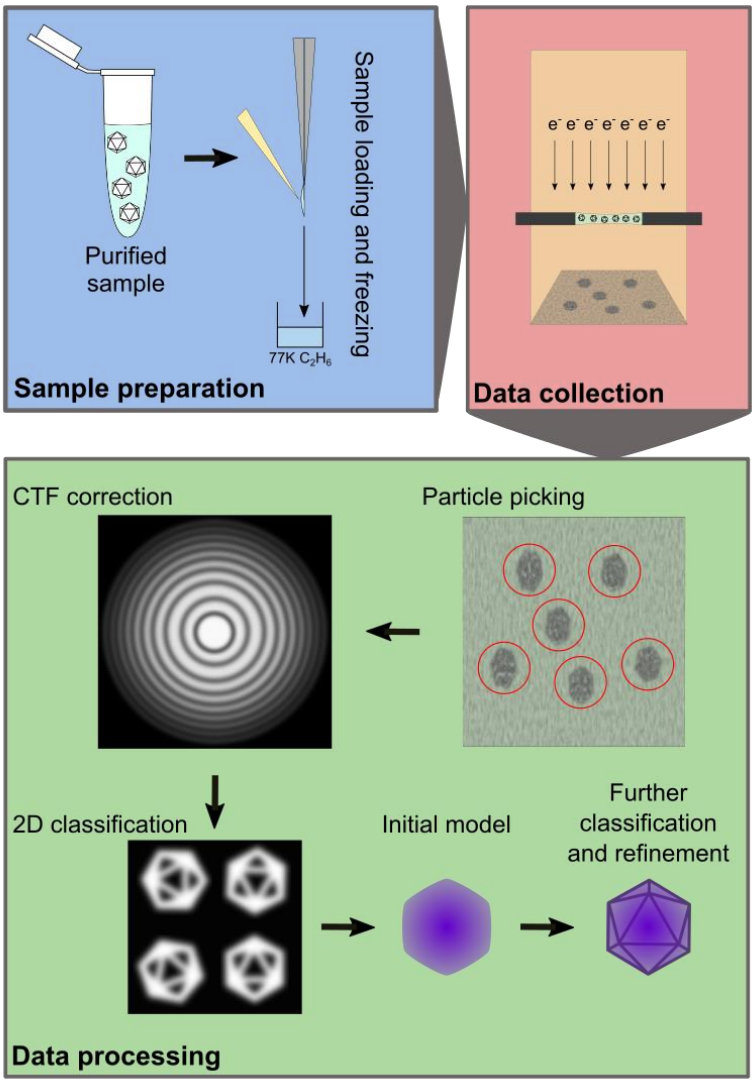


Figure 1.24. Simplified workflow of single particle cryo-EM.

Cryo-EM samples can be of several types: 2D crystalline arrays of single particles; long tubular filaments of single particles; monodisperse single particles; or larger biological structures that require the use of tomography techniques. Single particle cryo-EM (in the form of filaments or monomers) has shown the most recent development and focus as it can be used to solve 3D structures of proteins and complexes at resolutions similar to structures determined by X-ray crystallography.

In single-particle cryo-EM, highly purified biological samples are rapidly frozen in liquid ethane, hopefully resulting in an ultrathin (10-70 nm) film of vitreous ice with sample suspended within. In an ideal sample, molecules will be monodisperse and in a range of different orientations so that different projections can be combined to form a 3-D reconstruction. This frozen grid must be kept at liquid nitrogen temperature while handling and imaging to avoid disrupting the ice and sample. Once loaded into the microscope, images can be automatically collected over several days after an initial time setting up. This automation step has greatly reduced the direct work of data collection for EM scientists. Images must subsequently be subjected to a myriad of processing steps in order to calculate a 3D EM density map, which is used for fitting a model of the protein.

The images (or micrographs) are first motion corrected (as mentioned briefly earlier), by aligning frames from the movie to reduce blurring induced by the beam interacting with the stage. At this step, dose-weighting can also be applied, where frames recorded earlier in the movie are given more weight as they will have been exposed to less radiation damage and will have retained more of the high resolution information. Motioncor2 is the now commonly used motion correction software, and Unblur has also been used (Grant and Grigorieff, 2015; Zheng et al., 2017).

After motion correction, contrast transfer function (CTF) estimation is performed. The CTF is a mathematical description of how images taken by EM are modified by various aberrations in the system. These aberrations affect the amplitudes of Fourier spatial frequencies that make up an image. The CTF gives the characteristic “power spectrum” of Fourier transformed micrographs, as shown in Figure 1.24. Progressing through higher spatial frequencies, the amplitudes oscillate positively and negatively, with some having a zero value, and contributing nothing to the image. The affected spatial frequencies can be altered by adjusting the defocus of the image, so it is common practice to collect data at a range of defocus values so that the full range of spatial frequencies have high amplitudes. As well as the oscillations, an envelope function dampens higher frequencies, reducing contribution of this high resolution information to the image (Wade, 1992). CTF estimation is performed on each micrograph to find the values of certain parameters that describe the CTF. These can then be applied later on in the data processing to correct for the aberrations

introduced during imaging. Currently used popular CTF estimation software are CTFfind and the GPU accelerated GCTF (Rohou and Grigorieff, 2015; Zhang, 2016).

Next, particles are picked from the micrographs, either manually or automatically. Automatic particle picking algorithms can use template based or machine learning approaches (Scheres, 2015; Wagner et al., 2018). The averaging of many particles is essential to achieving a high resolution structural model as it greatly improves the signal to noise ratio (SNR) of the images. Once particles are picked, they can be aligned and classified into groups of similar particles to make averaging more effective. There are many different classification algorithms, utilising an array of statistical methods (Grigorieff, 2016; Ludtke, 2016; Scheres, 2016). The resultant classes represent relatively detailed projection views of the sample from various angles, which are then used to generate an initial model of the protein, where 'common-lines' of correlation in Fourier space between 2D classes reveal the relative positioning of each class in 3D space.

Heterogeneity within the model can be isolated by further classification in 3D. Eventually a refinement procedure is performed on the subset of particles that correspond to the model where all information out to the highest resolution (Nyquist frequency, which is double the pixel size used) is used. To restrict overfitting, two subsets of data are independently refined until they both converge, where they are then combined to produce the final model; this can be masked to reduce solvent noise, and then sharpened to account for other blurring effects. With an excellent dataset, a high quality density map is produced into which a structural model of the protein can be fitted at high resolution, in some cases below 3.0 Å.

However, there are many factors that can produce a poor data set, and it is essential to produce a good quality purified sample and grids. Even if the sample is very pure, there may be significant heterogeneities in its conformation and aggregation state, which compromise the quality of the final structural model Nevertheless, the rapid advances in grid preparation, microscopy and image analysis make this an exciting time for cryo-EM research.

1.6.2 2-dimensional electronic spectroscopy

2-dimensional electronic spectroscopy (2DES) is a technique used to determine information about aggregates of electronically excitable molecules. The spectral properties of the interacting pigments are difficult to disentangle without the use of a 2D technique and many coupling events are invisible. 2DES can resolve excitation and emission events occurring on the femtosecond timescale; it operates in a manner analogous to 2D NMR spectroscopy, where a quick succession of

laser pulses excite and polarise the sample in such a way as to promote the release of a "photon echo", which contains information about resonant interactions within the pigment or pigment complex as well as energy transfer. If two electronic states are coupled, then an indicative “cross-peak” is observed in the 2D spectrum and the pigments must be close in space (Ginsberg et al., 2009).

Because the experiment is performed at multiple population times, kinetic information can be obtained about the evolution of specific resonant frequencies. Amplitude oscillations, known as quantum beats, can be observed as a function of population time, the result of a superposition of excited states (excitons) that can persist for hundreds of femtoseconds, significantly longer than expected. The independent states that form these excitons are said to be coherent with one another. Fourier transforms of the data yield the frequencies of the amplitude oscillations, which can be linked to vibrational frequencies of intra/inter molecular interactions with the exciton. Display of specific frequencies of interest within the 2D frequency maps can be used to identify which parts of the spectra are coherent at that frequency (Schlau-Cohen et al., 2012). The long-lasting nature of these excitons has been attributed to the observed vibrational modes that act in a way to stabilise the inherently unstable exciton state. The coherence of excitons with vibrational modes led to the term vibronic coherence (Romero et al., 2017).

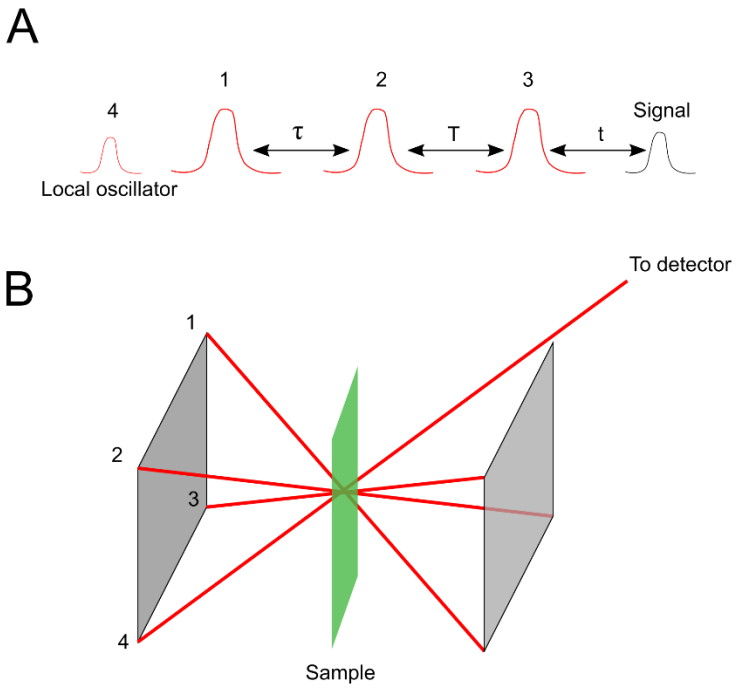


Figure 1.25. 2DES pulse sequence (**A**) and non-colinear geometrical configuration (**B**). The first three pulses cause excitation and generate a signal which is heterodyne-detected using the local oscillator. The time between pulses 1-3 and signal generation are described by τ, the coherence time; T, the population time; and t, the rephasing time.

2DES has been used on photosynthetic systems to aid in characterising multi-pigment complexes (Zigmantas et al., 2006); for example by calculating the flow of excitation energy in Fenna-Matthews-Olson (FMO) photosynthetic pigment-protein complex, which links light harvesting by chlorosomes to charge-separating RCs in green sulphur bacteria (Brixner et al., 2005). Quantum coherence effects have been observed to occur in photosynthetic systems, potentially aiding efficient energy transfer (Cheng and Silbey, 2006). More recently the charge separation catalysed by the RC of photosystem II (PSII) was probed by 2DES and electronic coherence was proposed to aid in this efficient and ultrafast process (Romero et al., 2014). These recent findings suggest that quantum effects are able to contribute to overall efficiency of some ultrafast biological reactions, despite the highly disordered state (in terms of quantum phenomena) of biological molecules at physiological temperatures. Given that the great majority of such work has been performed on photosynthetic systems, mainly light harvesting complexes, it is of considerable interest to find out if coherence effects are involved in the ultrafast photochemistry of POR and other photoenzymes.

1.6.3 Microscale thermophoresis

Microscale thermophoresis (MST) is a technique developed recently that can characterise the interactions between two molecules. A fluorescent label is covalently bound to one of the binding partners (or a fluorescent small molecule ligand can be used) and the other partner is titrated in; the reaction is in a chamber exposed to an infrared (IR) laser beam for rapid and highly localised micro-heating and the response of the fluorophore is measured as a function of titrated ligand concentration (Seidel et al., 2013).

The fluorescence of a molecule changes at different temperatures due to temperature-related intensity change (TRIC) and thermophoresis, which is the movement of molecules within a thermal gradient (either away from, or toward the area of increased temperature) (Gupta et al., 2018). TRIC is caused by the alteration of the immediate environment of the fluorophore under different temperatures, which affects fluorescence intensity. TRIC can be caused by a binding partner which directly alters the environment of the fluorophore (Baaske et al., 2010). For thermophoresis, the movement is based on three main properties: the size of the particle, the surface charge, and the hydration entropy. All of these properties are affected by interaction events: the size is an obvious change - if two molecules are interacting, the overall molecule size is larger; the surface charge may alter if binding interfaces are covered, or if the added binding partner contributes to surface

charge; and the same can be said for solvation layer, which will be displaced and rearranged upon interaction of two binding partners (Duhr and Braun, 2006).

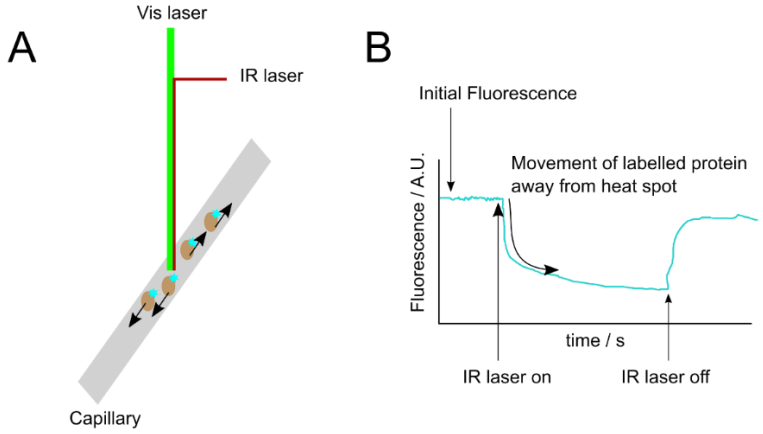


Figure 1.26. Principles of MST. **A**, laser and capillary schematic, where proteins tagged with a fluorescent dye are exposed to micro-heating by the IR laser, and move along the temperature gradient; the fluorescence change is monitored by the visible light laser. **B**, representation of a MST curve following the changes in fluorescence.

For simple systems of 1:1 binding, MST is a rapid technique for quantifying interaction strengths in the form of K_d values (Seidel et al., 2013). More complicated binding events can also be determined by altering the time for collecting the fluorescence arising from the IR beam. The temperature dependence of interactions can also be determined by controlling the strength of the IR laser. As well as a technique for measuring binding, MST has been used to follow equilibrium unfolding, by using intrinsic fluorescence of proteins and titrating in a chemical agent used to unfold proteins such as guanidinium hydrochloride (Jerabek-Willemsen et al., 2014).

There are many interactions involved in the MgCH complex and POR is known to form oligomers. MST can provide a fast and simple method to monitor the steady-state interaction strengths between the various subunits of these complexes. Determining the factors that affect these binding events may allow further understanding of the roles of subunits in the reaction cycle. MST has previously been used on MgCH to help characterise the ChlI-ChlD interaction (Adams et al., 2016b).

1.7 Aims and objectives

MgCH and POR have been studied for many years and are well characterised in many ways; however, a complete understanding of the reaction mechanisms and structural organisation of them is still lacking, hindering a more thorough understanding. This thesis aims to further our knowledge

of the biochemistry that these enzymes perform; by using a variety of structural, biochemical and biophysical techniques, it aims to characterise some of the protein-protein interactions occurring within the POR and MgCH complexes, reveal their overall quaternary structural arrangements, and increase our understanding of how these enzymes perform the complex and unique chemistry on their substrates.

2 Materials and Methods

2.1 Materials

All chemicals and reagents were purchased from Sigma-Aldrich Company Ltd. (Dorset, UK) or Fisher Scientific UK Ltd (Loughborough, UK) unless otherwise stated. Most chromatographic resins were bought from GE Healthcare Life Sciences (Little Chalfont, UK). The Protino® Ni-TED Resin was supplied by Macherey-Nagel (Düren, Germany). Milli-Q® Integral ultrapure water (Milli-Q water) was the standard water source used at a purity of 18.2 mΩ cm at 25 ºC produced by a Milli-Q® Integral Water Purification System (Millipore (UK) Ltd., Watford, UK) unless otherwise stated.

2.2 Bacterial strains and plasmids

Chemically competent *E. coli* strains BL21(DE3)pLysS and Rosetta™(DE3)pLysS were used for protein over-expression whilst strains JM109 and XLI Blue were used for cloning transformations.

The Δ*bchJ* strain of *R. sphaeroides* used for production of Pchlide was originally generated by Dr Sarah Hollingshead (Hollingshead, 2013).

Synechocystis sp. PCC6803 strains used are described in Table 2.1.

Table 2.1. Strains of *Synechocystis sp.* PCC6803 used.

Strain	Description	Source
WT	Glucose tolerant wild-type strain of *Synechocystis* sp. PCC 6803.	Dr Roman Sobotka, Centre Algatech, Czech Republic
FLAG-ChlD	Gene encoding 3×FLAG tagged ChlD expressed from *psbAII* promoter. Kanamycin resistant.	Dr Andrew Hitchcock, University of Sheffield
FLAG-ChlD Δ*chlD*	Native *chlD* (slr1777) deleted from FLAG-ChlD strain. Kanamycin and zeocin resistant.	Dr Andrew Hitchcock, University of Sheffield

A table of plasmids used is shown in Table 2.2. The genes cloned into these plasmids encode POR from *T. elongatus* (*T.*POR) and magnesium chelatase subunits from *Synechocystis sp*. PCC6803 (*Syn*I/D/H).

Table 2.2. A list of plasmids used.

Plasmid	Features/ Purpose
pET9aHis	Kanamycin resistance, N-terminal His_6-tag
pET9a	Kanamycin resistance
pET21a	Ampicillin resistance, N-terminal His_6-tag
pET28a	Ampicillin resitance, N-terminal His_6-tag
pETcoco-2	Ampicillin resistance, N-terminal His_6-tag, single copy plasmid

2.3 Buffers, reagents and media

2.3.1 Media

Lysogeny broth (LB) and LB Agar were supplied by Formedium™ (Kings Lynn, UK) and made up as instructed. 'TY' broth was made up of 1.6 % tryptone, 1 % yeast extract and 5 % NaCl. Autoinducing medium (Studier, 2005) was made using 1 % tryptone, 0.5 % yeast extract. These were all autoclaved at 15 psi, 121 ºC for 20 minutes to sterilise. Immediately prior to use, 1x final concentration of M component and 5052 and trace elements (Table 2.3), as well as 2 mM $MgSO_4$ were added to autoinducing medium.

Table 2.3. Autoinducing media additional components.

50x M component	To make 1 l
Disodium phosphate	335 g
Monopotassium phosphate	170 g
Ammonium chloride	134 g
Disodium sulphate	35.5 g
50x 5052	To make 1 l
Glycerol	250 g
Glucose	25 g
α-lactose	100 g
1000x Trace elements	To make 100 ml
Iron(III) chloride	2.7 g
Calcium chloride	15.8 g
Manganese (II) chloride	19.8 g
Zinc sulphate	28.8 g

Cobalt (II) chloride	4.76 g
Copper (II) chloride	1.70 g
Nickel (II) chloride	4.76 g
Sodium molybdate	2.42 g
Sodium selenite	1.73 g
Borinic acid	0.62 g

R. sphaeroides was grown on M22 medium, recipe detailed Table 2.4, and autoclaved as above. 1.5 % Bacto Agar (Becton Dickinson UK Ltd, Oxford) was added to 1x M22 without casein hydrolysate to produce M22 agar. M22 vitamins (detailed Table 2.4) were added to the medium immediately prior to use.

Table 2.4. M22 media components.

Solution C	To make 4 l
Nitrilotriacetic acid	40 g
Magnesium chloride	96 g
Calcium chloride	13.36 g
EDTA	0.5 g
Zinc chloride	1.044 g
Ferrous chloride	1.0 g
Manganese chloride	0.36 g
Ammonium heptamolybdate	0.037 g
Cupric chloride	0.031 g
Cobaltous nitrate	0.0496 g
Boric acid (orthoboric acid)	0.0228 g
10X M22 stock	To make 4 l
Potassium dihydrogen orthophosphate	122.4 g
Dipotassium hydrogen orthophosphate	120.0 g
DL – lactic acid	100.0 g
Ammonium sulphate	20 g

Sodium chloride	20 g
Sodium succinate	173.7 g
Sodium glutamate	10.8 g
Aspartic acid	1.6 g
Solution c	800 ml
1X M22	To make 1 l
10 x M22 stock	100 ml
Casein hydrolysate	2 g
M22 vitamins (10000x)	To make 100 ml
Nicotinic acid	1 g
Thiamine	0.5 g
4-aminobenzoic acid	0.1 g
Biotin	10 mg

Synechocystis strains were grown by Dr Andrew Hitchcock on BG11 medium (Rippka et al., 1979) buffered with 10 mM *N*-Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES)-KOH pH 8.2 (BG11-TES).

2.3.2 Antibiotics

Chloramphenicol stock solutions were prepared by dissolving 35 mg ml^{-1} of chloramphenicol in 100 % ethanol then stored at -20 ℃. Ampicillin and kanamycin stocks were made by dissolving 100 mg ml^{-1} ampicillin or 30 mg ml^{-1} kanamycin in Milli-Q water. The antibiotic-water solutions were filter sterilised, dispensed into 1 ml aliquots and stored at -20 ℃.

2.3.3 Reagents

1 M stocks of isopropyl β-D-1-thiogalactopyranoside (IPTG), arabinose and dithiothreitol (DTT) as well as 100 mM of $NADP^+/H$ were made by dissolving the required mass in Milli-Q water. The IPTG and arabinose solutions were filter sterilised and all solutions were stored at -20 ºC.

2.3.4 Buffers

All buffers were made up with Milli-Q water and pH adjusted with the appropriate acid/base. Buffers used in chromatography were vacuum sterilised using a 0.2 μm filter and degassed.

2.4 Site directed mutagenesis

Primers were designed (for list of primers, see Table 2.5) to incorporate the required codon into the appropriate site in the sequence using Agilent Technologies QuikChange primer design tool. The primers were preferably shortened to end with 3' AT and 5' GC; melting temperatures were verified using a method produced by Dr Andrew Hitchcock, and ordered from Sigma-Aldrich. The mutagenesis was performed using the Agilent Technologies QuikChange kit as per the manufacturer supplied protocol. Plasmid DNA was extracted from cells by using a plasmid mini prep kit (Nippon Genetics Co., Ltd. or Qiagen) according to the instructions provided by the manufacturer. The success of the site directed mutagenesis was verified by DNA sequencing (GATC Biotech GmbH, Cologne, Germany).

Table 2.5. A list of primers used.

Primer name	Forward sequence	Reverse
ChlH active site		
S1103A	TTGACGTTGGAAGCATAGGAACCGGAGGCG	AACGCCTCCGGTTCCTATGCTTCCAACGTC
S1178A	TTGGTGGGGTCGGCGTCGAAGTAGTGG	CACTACTTCGACGCCGACCCCACCAAG
W985A	TTTGATGTTATCCGTACCCGCGAGCACAGAAGCAATGG	AACCATTGCTTCTGTGCTCGCGGGTACGGATAACATC
W985P	TTGATGTTATCCGTACCCGGGAGCACAGAAGCAATGG	AACCATTGCTTCTGTGCTCCCGGGTACGGATAACATC
W985F	TGTTATCCGTACCGAAGAGCACAGAAGCAATGGTTTCGG	CCGAAACCATTGCTTCTGTGCTCTTCGGTACGGATAACA
K1129A	TTGAAAGCGAAGGACGCGCGTTTGAGGTACATTTCC	ATGTACCTCAAACGCGCGTCCTTCGCTTTCAACTCC
S659A	ATGGTAGCTTCAGCAGGATTGTTGGCGGC	CCGCCAACAATCCTGCTGAAGCTACCATCG
D1177A	TAAACAGATCCCGGAATGCCCCGGAACAATTCACC	TCCCACTACTTCGCCTCCGACCCCACC
V1041A	TGGGGTCGGAGGCGAAGTAGTGGG	AATTGTTCCGGGGCATTCCGGGATCTG

E625Q	CTGTTTACCGGGCATAAACTGCAAAGAAC-CGTGGGTGCC	GGCACCCACGGTTCTTTGCAG-TTTATGCCCGGTAAACAG
E625D	TGTTTACCGGGCATAAAATCCAAAGAAC-CGTGGGT	ACCCACGGTTCTTTGGAT-TTTATGCCCGGTAAACA
E625H	CTGTTTACCGGGCATAAAATGCAAAGAAC-CGTGGGTGCC	GGCACCCACGGTTCTTT-GCATTTTATGCCCGGTAAACAG
E625K	GCACCCACGGTTCTTTGAAATTTATGCCCGG-TAAA	TTTACCGGGCATAAATTTCAAAGAAC-CGTGGGTGC
Y653F	CCATTCCCAACTTGTACTACTTCGCCGCCAACA	TGTTGGCGGCGAAGTAGTACAAGTT-GGGAATGG
E660A	AACAATCCTTCTGCAGCTACCATCGCCAAACG	TTTGGCGATGGTAGCTG-CAGAAGGATTGTTGG
E660W	AACAATCCTTCTTGGGCTACCATCGCCAAACG	TTTGGCGATGGTAGCCCAA-GAAGGATTGTTGG
E660D	AACAATCCTTCTGATGCTACCATCGCCAAACG	TTTGGCGATGGTAGCTCGA-GAAGGATTGTTGG
E660H	AACAATCCTTCTCATGCTACCATCGCCAAACG	TTTGGCGATGGTAGCATGA-GAAGGATTGTTGG
E660Q	AACAATCCTTCTCAAGCTACCATCGCCAAACG	TTTGGCGATGGTAGCTTGA-GAAGGATTGTTGG
E660R	AACAATCCTTCTCGAGCTACCATCGCCAAACG	TTTGGCGATGGTAG-CATCAGAAGGATTGTTGG
D988K	TCTGTGCTCTGGGGTACGAAGAACATCAAAAC-CTACGGC	GCCGTAGGTTTTGATGTTCTTCG-TACCCCAGAGCACAGA
T987A	TTCTGTGCTCTGGGGTGCGGA-TAACATCAAAACC	TAGGTTTTGATGTTATCCG-CACCCCAGAGCACAGAAGC
D988A	TTCTGTGCTCTGGGGTACGGCTAACATCAAAACC	TTTTGATGTTAGCCGTACCCCAGAG-CACAGAAGC
T987AD988A	TTCTGTGCTCTGGGGTGCGGCTAACATCAAAACC	TAGGTTTTGATGTTAGCCG-CACCCCAGAGCACAGAAGC
CHLD MIDAS		
D487E	TTGTGTTTTTGGTGGAAGCGTCGGGTTCCATGG	ATGGAACCCGACGCTTCCAC-CAAAAACAC
S489A	TTGGTGGATGCGGCGGGTTCCATGG	AACGCCATGGAACCCGCCGCATCCACC

S491A	ATGCGTCGGGTGCCATGGCGTTGAATCG	ATTCAACGCCATGGCACCCGAC-GCATCC
S554R	TTGCCCTGTGGCGGTGGTCGTCCCCTTTCC	AAGCCGTGGGAAAGGGGACGACCAC-CGCCACAGG

2.5 Transformation of chemically competent *E. coli*

A 50 µl aliquot of competent cells for each transformation was thawed on ice for 10 min. Approximately 10 ng of plasmid DNA was added to each tube and left on ice for 15 min. The tubes were then held in a 42 ℃ water bath for 50 s, followed by 2 min on ice to induce heat shock. 400 µl of LB medium was added to the tubes which were then left to incubate for 40 min at 37 ℃, shaking at 230 rpm. After this incubation, the cells were spread on LB agar plates containing appropriate antibiotics and incubated at 37 ℃ for approximately 16 hr. Transformed Rosetta(DE3)pLysS cells were grown in medium with 34 µg ml^{-1} chloramphenicol.

2.6 Overproduction of proteins in *E. coli*

2.6.1 Starter cultures

A single colony from a successful transformation was picked using a sterile loop and added to LB with appropriate antibiotics to grow for approximately 16 hr at 37 ℃, shaking at 230 rpm. The volume of LB needed depends on the volume of the large scale growth being used: 10 ml starter per 500 ml large scale growth.

2.6.2 Large scale growth for overexpression

The starter cultures were used to inoculate baffled flasks containing growth medium. ChlH was grown in 1 l autoinducing medium and pET9aHis-*T.*POR was grown in 1.5 l LB. These were both grown at 28 ℃, shaking at 180 rpm for approximately 20 hr. ChlI were grown in 1 l LB whilst ChlD was grown in 1 l TY broth. They were all grown at 37 ℃, shaking at 230 rpm and the OD_{600} monitored. At an OD_{600} of *ca.* 0.6 the flasks were transferred to an 18 ℃ incubator shaking at 180 rpm for approximately 10 min. After this time, IPTG was added to a final concentration of 0.4 mM and the flasks were returned to the cooler incubator for 16 hr.

2.6.3 Harvesting cells

Cells were harvested by centrifugation at 14000 x *g*, 4 ℃ for 15 min. The pellet was resuspended in either 25 mM HEPES, 0.5 M NaCl, 5 mM imidazole, pH 7.4 (for MgCH; NaCl omitted for co-

purification experiments) or 25 mM HEPES, 0.5 M NaCl, pH 7.4 (for POR) and stored at -20 °C for future use.

2.7 Growth of *Synechocystis* strains

Starter cultures of 35 ml and then 500 ml were grown in a rotary shaker (150 rpm) at a constant illumination of 40 μmol photons m^{-2} s^{-1}. For purification of FLAG-tagged proteins 8 l cultures of *Synechocystis* were grown with vigorous bubbling with sterile air at a constant illumination of 150 μmol photons m^{-2} s^{-1}. Growth was monitored as the optical density at 750 (OD_{750}). For growing on solid medium, BG11-TES was supplemented with 1.5 % (w/v) agar, 0.3 % (w/v) sodium thiosulphate and appropriate antibiotics and plates were incubated at 30 °C with 40 μmol photons m^{-2} s^{-1} constant illumination.

2.8 Purification of proteins

2.8.1 POR

2.8.1.1 Ammonium sulphate precipitation

Frozen cell pellets were thawed and transferred to a round bottomed glass sonicating vessel. A ROCHE cOmplete™ mini EDTA-free protease inhibitor cocktail was added and the solution sonicated on ice for 12 cycles (30 s sonication, 30 s rest, 80% amplitude). The lysed cells were heat treated at 42 °C in a 50 ml Beckmann centrifuge tube for 12 min, which was clarified by centrifugation at 41 000 *xg* for 15 min, 4 °C and the supernatant reserved. Ammonium sulphate was added to the supernatant at a concentration of 0.523 g ml^{-1} and incubated with mixing for 1 hr at 4 °C. The tubes were then centrifuged as above. This pellet could be stored for later use at -20 °C or resuspended in lysis-equilibration-wash (LEW) buffer: 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0 buffer for immediate use with the Protino® Ni-TED column.

2.8.1.2 Immobilised nickel affinity chromatography

The resuspended pellet was further clarified by ultrafiltration through a 0.45 and 0.2 μm filter and applied to a 10 ml column packed with 1.5 g Protino® Ni-TED resin pre-equilibrated with 15 ml LEW buffer. The column was washed with 50 ml LEW buffer. Protein was eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole) and concentrated to approximately 10 ml using a VivaSpin (GE Healthcare) spin concentrator.

2.8.1.3 Cation exchange chromatography

The following was carried out at 4 ℃. A 5 ml HiTrap SP HP column was attached to an ÄKTA FPLC system (GE Healthcare) and equilibrated with 3 column volumes (CV) of 95 % Buffer A (25 mM Sodium Phosphate buffer, 1 mM dithiothreitol (DTT), pH 6.5), 5 % Buffer B (25 mM Sodium Phosphate buffer, 1 M NaCl, 1 mM DTT, pH 6.5). The sample (diluted to 50 ml with buffer A) applied the column was washed with 2 CV of 95 % Buffer A, 5 % Buffer B followed by a 1 CV length gradient to 38 % buffer B. Protein was eluted over a 15 CV linear gradient to 55 % buffer B. Eluting protein was detected by absorption at 280 nm and purity assessed by SDS-PAGE. Fractions with pure POR were pooled, concentrated to 0.2 – 1 ml snap frozen in liquid nitrogen and stored at -80ºC.

2.8.2 Magnesium chelatase

2.8.2.1 Immobilised nickel affinity chromatography

Frozen cell pellets were sonicated as in 2.8.1.1 and clarified by centrifugation at 53343 x *g* for 15 min. The supernatant was further clarified by ultrafiltration using 0.45 μm and 0.2 μm filters and applied to a 2.5 ml Ni(II) Fast flow chelating Sepharose resin column, pre-equilibrated with 10 CV of binding buffer (25 mM HEPES, 0.5 M NaCl, 5 mM imidazole, pH 7.4). The column was then washed with 10 CV binding buffer, 10 CV wash 1 buffer (25 mM HEPES, 0.5 M NaCl, 20 mM imidazole, pH 7.4) and 10 CV wash 2 buffer (25 mM HEPES, 0.5 M NaCl, 50 mM imidazole, pH 7.4). Protein was eluted with elution buffer (25 mM HEPES, 0.1 M NaCl, 400 mM imidazole, pH 7.4). If ChlD was being purified, then the use of wash 2 was omitted.

2.8.2.2 Anion exchange chromatography

Protein diluted 1:1 with Buffer A (50 mM Tricine, 0.3 M glycerol, pH 7.9) was loaded onto a ResourceQ column pre-equilibrated with anion exchange Buffer A attached to an ÄKTA FPLC. The column was washed with 2 CV Buffer A before protein was eluted with a gradient to 100% Buffer B (50 mM Tricine, 0.3 M glycerol, 1 M NaCl, pH 7.9) over 20 CV. If ChlD was being purified DTT was added to buffers to a concentration of 1 mM. Eluting protein was detected by UV absorption at 280 nm and purity assessed by SDS-PAGE. Fractions with pure MgCH subunit were pooled, concentrated to 500 μl and used for size exclusion chromatography.

2.8.2.3 Size exclusion chromatography

Protein was loaded onto a S200 Increase size exclusion column pre-equilibrated with size exclusion buffer (50 mM Tricine, 0.3 M Glycerol, 200 mM NaCl, pH 7.4). Eluting protein was detected

by UV absorption at 280 nm with pure protein eluting between 0.3 CV and 1 CV. Purity was assessed by SDS-PAGE. Fractions with pure MgCH subunit were pooled, concentrated to 0.2 – 1 ml snap frozen in liquid nitrogen and stored at -80 ℃. When purifying larger complexes, a Sephracyl S300 column was used under the same conditions.

2.8.2.4 Purification of StrepII tagged protein

Cells were suspended in binding buffer (100 mM Tris, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8) and lysed by sonication as above, and cell debris cleared by centrifugation (53343 x *g*, 15 min). The supernatant was applied to a StrepTrap HP column (GE Healthcare) pre-equilibrated in binding buffer and unbound proteins removed with 15 CV of binding buffer. ChlH-StrepII was eluted with binding buffer supplemented with 2.5 mM desthiobiotin.

2.8.2.5 Purification of FLAG-tagged proteins

Synechocystis strain FLAG-ChlD was grown to an OD_{750} of 0.9. Cells were pelleted and resuspended in buffer A (25 mM sodium phosphate, 50 mM NaCl, 10 mM $MgCl_2$, 10 % (w/v) glycerol pH 7.4 and EDTA-free protease inhibitors). Cells were mixed with an equal volume of 0.1 mm glass beads and broken in a Mini-Beadbeater-16 (BioSpec) (50 s beating, 2 min cooling on ice, 10 cycles). Soluble proteins and membranes were separated by centrifugation (65000 × *g*, 30 min, 4 ℃). The membrane fraction was washed with an excess of buffer A, then resuspended in buffer A with 1.5 % (w/v) dodecyl-β-maltoside (β-DDM), Applichem) and solubilised for 2 h at 4 ℃. The solubilised membranes were separated from insoluble material by centrifugation (65000 ×*g*, 30 min, 4℃).

FLAG-ChlD and associated proteins were purified from the soluble and membrane fractions using a 200 ml anti-FLAG-M2 agarose columns (Sigma–Aldrich) equilibrated in buffer A (+0.25 % (w/v) β-DDM for membrane fraction). FLAG-ChlD was eluted with 500 ml of buffer A (+0.25 % β-DDM for membrane fraction) containing 150 mg ml^{-1} 3xFLAG peptide (Sigma–Aldrich).

2.9 Protochlorophyllide preparation

2.9.1 Growth of the Δ*bchJ R. sphaeroides* mutant

The substrate for POR, Pchlide, is not commercially available so it is produced using the Δ*bchJ* strain of *Rhodobacter sphaeroides*. A glycerol stock of Δ*bchJ* was streaked onto M22 agar with vitamins (detailed 2.3.1) and grown at 30 ℃ in darkness for two days. A single colony was picked to inoculate 10 ml of M22 medium with vitamins. The culture was grown at 34 ℃, shaking at 180

rpm in the dark for 1-2 days, then used to inoculate 80 ml of M22 medium with vitamins, grown as before. After two days, 20 ml was used to inoculate a flask with 2 l of M22 medium with vitamins and 0.02 % Tween 20, which was incubated for 2 days.

2.9.2 Protochlorophyllide extraction

After centrifugation at 4424 x *g* for 30 minutes at 4 °C the supernatant was saved and kept in the dark. Under green light, a 2:1 mix of acetonitrile:diethylether was added to the medium in a ratio of 2 volumes medium:1 volume solvent solution and thoroughly mixed. The pigment containing solvent was separated using separation funnel and concentrated using a rotary evaporator. Highly concentrated pigment solution was dried in a vacuum centrifuge before storage in the dark at -20 °C.

2.9.3 Protochlorophyllide purification by high performance liquid chromatography

Preparative high performance liquid chromatography (HPLC), performed on an Agilent 1200 instrument, was used to purify the pigment obtained from solvent extraction. The dried material was dissolved in a minimal amount of 100 % methanol and clarified by centrifugation for 5 minutes at 1700 x *g*. The supernatant was applied to a reverse phase C18 column, pre-equilibrated with 35 % Buffer A (350 mM ammonium acetate and 30 % (v/v) MeOH - filtered through a 0.22 μm membrane and degassed) and 65 % Buffer B (100 % (v/v) MeOH). Pigment was eluted over a linear gradient of 65 % to 75 % Buffer B over 35 minutes. Elution of pigments was monitored using a multichannel diode array detector (DAD), recording absorbance at 440 nm and 632 nm, and a fluorescence detector using an excitation wavelength of 440 nm and emission wavelength at 632 nm.

Eluted pigment was desalted using a 1 ml C18 reverse phase gravity column (Sigma-Aldrich) pre-equilibrated with 50 % (v/v) MeOH. The column was washed with 5 CV of 50 % (v/v) MeOH before eluting with 100 % (v/v) MeOH. Eluted pigment was dried in a vacuum centrifuge and stored in the dark at -20 °C.

2.10 Protein analysis

2.10.1 Determination of protein concentration

Protein concentration was determined by absorbance at 280 nm and converted into an approximate protein concentration using Equation 2.1, where n_{Trp}, n_{Tyr}, and n_{Cys} are the number of tryptophan,

tyrosine and cysteine residues in the protein, and M_r is the predicted molecular weight of the protein.

$$A_{280} = \frac{(5960n_{Trp} + 1280n_{Tyr} + 120n_{Cys})}{M_r}$$

Equation 2.1. Determining protein concentration.

2.10.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein solutions were separated using 12 % Bis-Tris precast gels (Life technologies, Carlsbad, CA, USA) and cross-linked MgCH complexes were separated using 3-8 % Tris-Acetate precast gels. Samples were loaded with Laemmli sample buffer alongside Precision plus unstained protein standards (BioRad, Hemel Hempstead, UK). Bis-Tris gels were run in MES Running Buffer (Life technologies) at 180 V for 55 min at room temperature. Tris-Acetate gels were run in Tris-Acetate Running Buffer (Life technologies) at 150 V for 1.5 hr. Protein bands were visualised with Coomassie Brilliant Blue G250.

2.10.3 Native PAGE

Proteins were separated in non-denaturing conditions by 3-12 % bis-acrylamide Native PAGE precast gels (Life technologies). Samples were loaded with sample buffer (125 mM Tris-HCl pH 6.8, 0.004 % (v/v) Bromophenol blue, 20 % glycerol) and run at 150 V, 4 ℃ for 2 hr. For blue native PAGE, 1 ml of NativePAGE cathode buffer (Life technologies) was added to the upper buffer chamber.

2.10.4 Immunoblotting

Polyvinylidene fluoride membranes (Novex), were soaked in methanol for 30 s and then, along with filter papers and porous pads, washed in transfer buffer (10 mM $NaHCO_3$, 10 mM $NaCO_3$, 10 % methanol). The gel from the PAGE was sandwiched between the membrane, filter paper and pads and placed in a transfer tank. The separated proteins were transferred at 30 mA overnight or 350 mA for 1 hr at 4 ℃.

The membrane was washed in Tris-buffered saline (TBS; 50 mM Tris/HCl, 150 mM NaCl, pH 7.6) before being blocked in TBS containing 5 % milk powder and 0.2 % (w/v) Tween 20 for 1-3 hr at room temperature. Diluted (1:10000 with TBS and 0.05 % Tween 20) primary antibodies were incubated with the blocked membrane for 4 hr at room temperature, or overnight at 4 ℃. The

membrane was washed with TBS and 0.05 % Tween 20 and incubated with appropriate secondary antibodies conjugated with horseradish peroxidase (Sigma-Aldrich) (diluted as above) for 1 hr at room temperature. After further washing in TBS, the membrane was soaked in WESTAR ETA C 2.0 chemiluminescent substrate (Cyanagen) and imaged on an Amersham Imager 600 (GE Healthcare).

2.10.5 Complex stabilisation by cross-linking

20 μM of MgCH complex (in buffer from previous purification step; immediately after nickel affinity or size exclusion chromatography - section 2.8.2) was incubated at 34 ℃ with 0.02 % glutaraldehyde, 3 mM ADP, 13 mM $MgCl_2$ for 30 min before quenching with the same buffer containing 1 M glycine. 8 μM D_{IX} was occasionally used when ChlH was present (see relevant results section). The complex was purified by SEC as in 2.8.2.3 with the addition of 3 mM ADP and 13 mM $MgCl_2$ as well as other components in different experiments (see results).

50 μM POR was incubated (50 ℃, 5 min) with 1 mM $NADP^+$, 100 μM Pchlide solubilised as in 2.13.1 in POR assay buffer (detailed 2.13.1). The ternary complex was diluted 2x and incubated with 0.02 % glutaraldehyde for 30 min at 50 ℃ before quenching with POR assay buffer containing 1 M glycine. The cross-linked solution was centrifuged at 1500 x *g* for 10 min to clarify any aggregate and loaded onto a Sephacryl S300 or Superdex S200 increase size exclusion column pre-equilibrated with POR assay buffer for purification.

2.11 Electron microscopy

2.11.1 Negative stain EM

1 μM protein was diluted to a suitable concentration with appropriate buffer. 5 μl of the protein solution was applied to a carbon-coated copper 400-mesh grid (Agar Scientific) that had been glow-discharged for 25 s. Sample was blotted with filter paper and the grid washed twice with distilled water. 5 μl of 0.75 % (w/v) uranyl formate was used to stain the grid for 20 s with excess stain removed by blotting and drying with air. Grids were imaged on a Philips CM100 microscope with a 1K x 1K Gatan Multiscan 794 CCD.

Automated data collection was carried out on a negatively stained grid using a Technai Arctica equipped with a Falcon 3EC (ThermoFisher) direct electron detector after the grid was clipped and loaded at room temperature. The grid was brought to cryogenic temperatures by the cooling of the microscope. For imaging parameters and processing details, see relevant results section.

2.11.2 Cryo-EM

5 µl of purified sample was applied to freshly glow discharged holey carbon grids (Quantifoil R1.2/1.3, 400 mesh Cu; glow discharging was omitted for graphene oxide and gold grids). Grids were blotted from the front or behind for 2-8 s using a Leica EM GP (70 % humidity, 20 ℃) and plunge frozen in liquid ethane. Grids were clipped and loaded into either a Technai Arctica, Talos Arctica, or Titan Krios microscope (ThermoFisher) for imaging using a Falcon 3EC or K3 (Gatan) detector. For imaging parameters, see relevant results sections.

Beam-induced motion correction and dose-weighting was conducted using MotionCor2 (Zheng et al., 2017). CTF estimation was carried out by either GCTF or CTFFIND4 (Rohou and Grigorieff, 2015; Zhang, 2016). Particle picking (manually or automatically), classification, initial model generation, and model refinement were carried out within RELION 2.1 or 3.0 (Fernandez-Leiro and Scheres, 2017; Zivanov et al., 2018). For project specific processing details, see relevant results section.

2.12 Microscale thermophoresis

Protein was labelled with NHS-NT647 dye (NanoTemper Technologies) according to the manufacturer's instructions before being exchanged into MST buffer (50 mM Tris/NaOH, 0.2 % pluronic, with or without 10 mM $MgCl_2$, pH 7.8). MST experiments were performed in triplicate, using 10 nM labelled protein with a serial dilution of partner protein. Samples were loaded into hydrophilic capillaries (NanoTemper) and MST conducted on a Monolith NT.115 (NanoTemper). Samples were illuminated with 20 % RED LED power and thermophoresis induced with 40 % IR laser power. Data was analysed using the MO.affinity analysis software (NanoTemper) where K_d values were calculated using the single site binding model (Equation 2.2).

$$f([P]) = U + \frac{(B-U) \times ([P] + [L_{tot}] + K_d - \sqrt{([P] + [L_{tot}] + K_d)^2 - 4[P][L_{tot}]}}{2[L_{tot}]}$$

Equation 2.2. Single site binding model used in the MO.affinity analysis software. [P] is the non-fluorescent partner concentration; $[L_{tot}]$ is the total concentration of fluorescent ligand; U and B are normalised fluorescence values from the ligand unbound and fully bound state, respectively; and K_d is the dissociation constant.

2.13 Kinetic assays

2.13.1 POR

A roughly 100 mM NADPH stock was made up with MilliQ water. The concentration was calculated by the absorption at 340 nm using a Cary 50 spectrophotometer with ε_{340} 6200 M^{-1} cm^{-1}. Purified dry Pchlide was resuspended in minimal Pchlide solubilisation buffer (50 mM Tris pH 7.4, 200 mM NaCl, 5 mM DTT, 50 mM octyl-β-glucoside (βOG)) by shaking for 10 min. This was then clarified by centrifugation at 1500 x *g* for 10 min at 4 ℃. The concentration of Pchlide was calculated by absorption at 630 nm in 80 % acetone, 20 % water with ε_{630} 30400 M^{-1} cm^{-1}.

POR assay mixtures were assembled in low light conditions. In a standard assay, 1 μM POR, 100 μM NADPH and 20 μM Pchlide were mixed in POR assay buffer (50 mM Tris pH 7.4, 200 mM NaCl, 5 mM DTT, 20 mM βOG) and incubated at 55 ℃ in a temperature controlled cuvette holder for 2 min before illumination with a 400 – 500 nm filtered KL 1500 electronic fibre optic light source. Spectra were collected at a rate of 150 nm s^{-1}.

2.13.2 MgCH

2.13.2.1 Reaction

MgCH assays were conducted in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, *I* (molar ionic strength) 0.1 and pH 7.7 at 34 °C. ATP was kept in the $MgATP^{2-}$ form by keeping at least 1 mM excess of $MgCl_2$ over ATP. Standard substrate concentrations were 15 mM $MgCl_2$, 5 mM ATP and 8 μM D_{IX}. MgCH subunits were added to initiate the reaction, generally to final concentrations of 0.4 μM ChlH, 0.2 μM ChlI and 0.1 μM ChlD.

D_{IX} and MgD_{IX} solutions were prepared in assay buffer, shaking for 5 min to dissolve *ca.* 0.5 g pigment. This was clarified by centrifugation at 1500 x *g* for 10 min at 4 °C. 100 mM stocks of nucleotide were prepared by weighing the appropriate amount, dissolving in MiliQ-H_2O and aliquoting into 1 ml tubes stored at -20°C. The concentrations of pigment and nucleotide were determined spectrophotometrically: ε_{398} 433 000 M^{-1} cm^{-1} in 0.1 M HCl for D_{IX} and MgD_{IX}; ε_{260} 15 400 M^{-1} cm^{-1} for nucleotide.

100 μl reactions were prepared in 96-well plates and followed in a FLUO-Star Optima plate-reader equppied with a xenon flash lamp (BMG Labtech, Aylesbury, Buckinghamshire, UK). Excitation light passed through a 420 nm (10 nm bandpass) filter and emission light through a 580 nm (10 nm bandpass) filter.

A 2-D calibration for ChlH and MgD_{IX} concentrations was conducted to account for changes in fluorescence with increasing ChlH concentration.

2.13.2.2 Data analysis

Steady-state rates were estimated using MARS Data Analysis Software Version 1.x R2 (BMG Labtech) and nonlinear regression performed using Igor Pro Version 7.0.6.1 (Wavemetrics, Lake Oswego, OR). Maximal rates were fit to the Michaelis-Menten equation (Equation 2.3), with errors obtained from least squares analysis of the fit.

$$v_{ss} = \frac{V_{max}[ATP]}{K_M + [ATP]}$$

Equation 2.3. The Michaelis Menton equation

Where v_{ss} is the steady state rate, V_{max} is the maximal rate of the system and K_M is the Michaelis constant.

V_{max} values obtained were plotted as a function of relevant protein concentration and the data were fitted to a single site binding model (Equation 2.4) to obtain apparent K_d values.

$$V_{max} = v_{lim}\frac{([E] + n_d + K_{app}) - \sqrt{([E] + n_d + K_{app})^2 - 4n_d[E]}}{2n_d}$$

Equation 2.4. Single site binding model.

Where v_{lim} is the V_{max} at saturating enzyme concentration, E is the enzyme concentration, n_d is the concentration of binding sites for enzyme and K_{app} is the apparent dissociation constant for the enzyme-protein complex.

Cooperative responses to substrate were fit to the Hill equation with errors obtained from least squares analysis of the fit. Where $S_{0.5}$ is the ligand concentration of half V_{max} and n is the Hill coefficient.

$$v_{ss} = \frac{V_{max}[S]^n}{S_{0.5} + [S]^n}$$

Equation 2.5. The Hill equation.

2.14 Differential scanning fluorimetry

Each assay contained 5 µM ChlD, 10 mM $MgCl_2$ 1x SYPRO Orange and 1x protein buffer to give a final volume of 50 µl. The QPCR machine (Agilent) was programmed to generate a temperature increase of 1 °C per cycle per minute. A fluorescence reading was taken each cycle. Melting temperatures (T_m) were assigned at the point where 50 % of the protein is unfolded, as determined by fitting melting data to a sigmoid Boltzmann distribution (Equation 2.6) where F_0 is initial fluorescence, F_{max} is maximal fluorescence, T is the temperature, T_m is the melting temperature and T_0 is the initial temperature.

$$F = F_0 + \left(\frac{F_{max}}{1 + \exp^{\frac{T_m - T_0}{\frac{\delta T}{\delta F}}}} \right)$$

Equation 2.6. Sigmoid Boltzmann distribution for DSF

2.15 Circular dichroism spectrometry (CD)

Spectra were recorded with a JASCO-810 spectrometer (JASCO, Great Dunmow, UK). Protein (0.05 mg ml^{-1}) was in buffer exchanged into 5 mM sodium phosphate, 1 mM β-mercaptoethanol, pH 7.5 using a Zeba Spin column according to the manufacturer's instructions. Spectra were recorded from 260 to 200 nm (1 nm steps, 4 s nm^{-1}, 4 accumulations) and background subtracted.

2.16 Cross-linking mass spectrometry

Purified ChlH was desalted into activation buffer (100 mM MES, 0.5 M NaCL, pH 6) using a Zeba Spin column according to the manufacturer's instructions. 1 mg ml^{-1} of ChlH in 100 µl activation buffer was incubated with 2 mM 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC; Thermo Fisher), and 5 mM N-hydroxysuccinimide (NHS; Sigma Aldrich) at room temperature for 15 min and then buffer exchanged into PBS. 1 mg ml^{-1} activated ChlH was mixed with 1 mg ml^{-1} ChlD and incubated at 34 °C for two hours prior to flash-freezing samples.

Cross-linked samples were analysed by SDS-PAGE using NuPAGE 3-8 % Tris-acetate gels in denaturing buffer (Tris-glycine-SDS), and cross-linked protein detected by immunoblotting for *Synechocystis* ChlD and ChlH, using the methods reported above.

2.16.1 In-gel digestion

Bands corresponding to putative ChlH-ChlD complexes were excised and subjected to in-gel digestion with trypsin as in (Pandey et al., 2000). Finely diced bands were incubated in 50 mM ammonium bicarbonate for 5 min, before incubating for 10 min after a 1:1 addition of acetonitrile. This process was repeated three times, replacing the solvent each time. The gel pieces were then reduced with 10 mM DTT, at 56 °C for 30 min, after which the solvent was removed and acetonitrile added and incubated at room temperature for 10 min. Solvent was then discarded and freshly prepared 55 mM iodoacetamide, 50 mM ammonium bicarbonate was added. This was incubated at room temperature in the dark for 20 min. Solvent was then discarded and the initial ammonium bicarbonate and acetonitrile additions were repeated as above. The gel pieces were then dried in a vaccuum centrifuge, rehydrated in trypsin digestion solution (12.5 ng μl^{-1} trypsin (Promega), 5 mM $CaCl_2$, 50 mM ammonium bicarbonate) and incubated overnight at 37 °C. The resultant peptide solution was transferred to a LoBind tube. For the remaining gel pieces 2 volumes acetonitrile were added to 1 volume gel and the mixture was incubated for 15 min at 37 °C. This solution was added to the previously removed peptide solution. 50 µl 5 % formic acid was added to the gel pieces which were then incubated at 37 °C for 15 min before transfer to the peptide solution. Acetonitrile was again incubated with the gel pieces as above which was then transferred to the peptide solution. The pooled peptide solution was dried in a vacuum centrifuge and stored at -20 °C.

2.16.2 Mass spectrometry and analysis

Mass spectrometry was conducted by Dr. Philip Jackson. The peptide extracts were redissolved in 10 µL 0.1 % (v/v) TFA, 3 % (v/v) acetonitrile and 2 µL analysed by nano-flow liquid chromatography (Ultimate 3000 RSLCnano system, Thermo Scientific) coupled to a mass spectrometer (Q Exactive HF, Thermo Scientific) according to (Hollingshead et al., 2016). The mass spectra were processed with Byonic v. 2.9.38 (Protein Metrics, CA, USA) with parameters set to default except that carbamidomethyl-Cys and Met oxidation were specified as fixed and variable (common) modifications respectively. Protein identifications were made by searching the Cyanobase *Synechocystis* sp. PCC 6803 proteomic database (http:// www.uniprot.org/uniprot/?query=organism:1111708) using the Byonic search engine (Protein Metrics). The identification of crosslinked peptides was enabled by manually creating a rule for EDC with the following syntax: ’EDC / -18.010565 @ K,D,E | xlink’. Crosslinked peptides were

identified in Byonic using a two-dimensional posterior error probability (2-D PEP) < 0.01 as a threshold. All crosslinked peptides identified in Byonic were manually verified.

2.17 2D electronic spectroscopy

Room temperature 2D electronic spectra were recorded with a diffractive optic-based inherently phase-stabilized four-wave mixing set-up, described in detail previously (Brixner et al., 2004). The repetition rate of the laser system (PHAROS, Light Conversion) was set to 1 kHz. The pulse generated by a home-built non-collinear optical parametric amplifier was centred at 685 nm with a full width half maximum of 90 nm and a duration of 15 fs. The excitation intensity was 12 nJ per pulse with all pulse polarisation set parallel to each other. The coherence time (τ) was scanned with 1 fs steps in the range $-85 < \tau < 100$ fs (up to 2 ps) and $-85 < \tau < 90$ fs (up to 1 ns) for Pchlide and $-140 < \tau < 160$ fs (up to 2 ps) and $-100 < \tau < 110$ fs (up to 1 ns) for the POR protein samples by employing movable fused silica wedges. The population time (T) was scanned in the range -24 fs $< T <$ 2000 fs with 8 fs steps to sample quantum beats with frequencies of up to 2100 cm^{-1}. Several 2D spectra were taken before time zero (-1000 and -100 fs) to check the noise levels of the measurement and at longer times: 5, 10, 20, 50, 1000, 250, 500 ps and 1 ns to investigate the spectral evolution up to 1 ns after excitation.

3 The ChlH-ChlD interaction

3.1 Introduction

Magnesium chelatase is a large multimeric enzyme, and multiple protein-protein interactions between the different subunits of the complex are involved in catalysing the chelation of a magnesium ion into the porphyrin macrocycle. The ChlI-ChlD interaction has been observed in multiple previous studies (Adams and Reid, 2013; Jensen et al., 1999); biochemical characterisation by Adams *et al.* (2016b) showed that interactions between the two subunits occurs via the AAA+ domains of ChlI and ChlD. Because mutations in key regions of both AAA+ domains show that they are involved in oligomerisation, it is reasonable to believe that ChlI and ChlD may form a heteromeric ring in a way not dissimilar to other AAA+ proteins (Adams and Reid, 2013; Gribun et al., 2008). However, there is no direct structural evidence for a ChlI-ChlD complex, and an EM study of the magnesium chelatase from *Rhodobacter* proposed that there are stacked rings of BchI and BchD but the EM map is of poor quality (Figure 1.14) (Lundqvist et al., 2010). As noted in the introduction (Section 1.4.2), the AAA+ domains of BchI and BchD show differences with respect to each other but in the case of ChlI and ChlD the homology between them is higher, which could explain the discrepancies between the two the enzymes of these different species.

While there has been some research on the interaction between I and D, it is still unknown how the energy produced by the ATPase activity of I is transferred to H in order to drive the chelation reaction; there has been no evidence of a direct interaction between I, the site of ATP hydrolysis, and the site of chelation on H. As the AAA+ domain of ChlD interacts with ChlI, ChlD could interact with ChlH via the proline-rich region or, perhaps more likely given that this domain mediates protein-protein interactions in other systems, the integrin I-like domain. The only previous work on the ChlH-ChlD interaction was a yeast two-hybrid study on the rice (*Oryza sativa*) chelatase (Luo et al. (2018), where they found that the C-terminal domain of ChlD was not necessary for chelatase activity; an interaction was still observed between the truncated ChlD and ChlH. However, this result contradicts other studies using mutagenesis and biophysical techniques with purified subunits that show that an intact ChlD complex is necessary for chelation (Adams et al., 2016b; Brindley et al., 2015).

In this chapter, the ChlH-ChlD interaction in the *Synechocystis* enzyme is probed with the aim of clarifying the discrepancies between the work of Luo *et al.* (2018) and other studies. It is revealed

that ChlH and ChlD do interact, and it is most likely via this C-terminal integrin I-like domain. This work is published as Farmer *et al.* (Farmer et al., 2019).

3.2 The ChlH-ChlD interaction can be observed *in vitro* and *in vivo*

3.2.1 FLAG-ChlD can pull down ChlH when expressed in native *Synechocystis*

To investigate the *in vivo* complex formation of ChlD with the other chelatase subunits, a *Synechocystis* strain was constructed by Dr Andrew Hitchcock in which ChlD has a 3xFLAG-tag encoded at the N-terminus (FLAG-ChlD). The *chlD* gene was expressed under the control of the *psbAII* promoter, and subsequently the native *chlD* gene was removed (FLAG-ChlD Δ*chlD*) (Figure 3.1, Table 2.1).

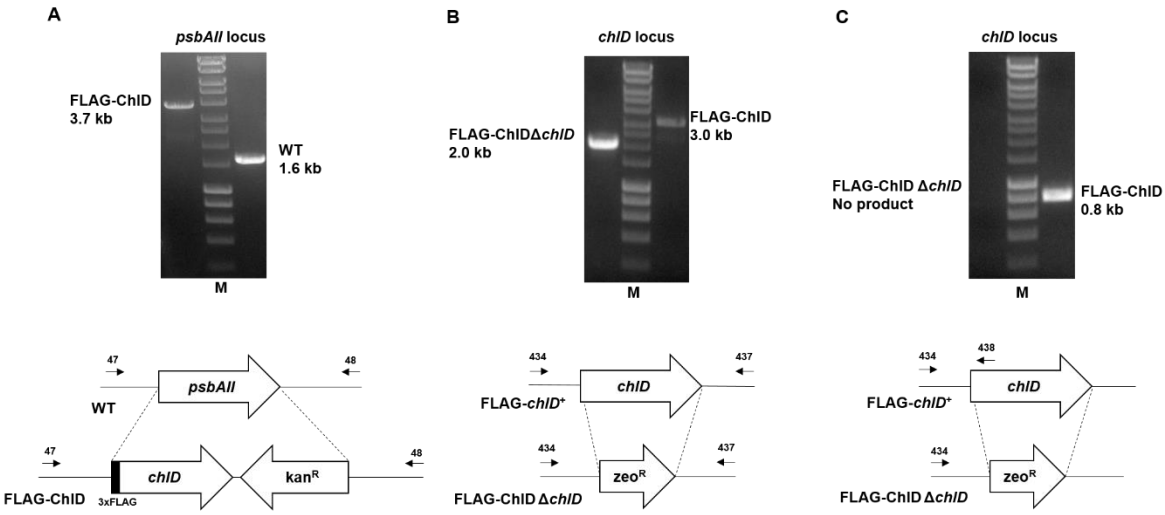


Figure 3.1. Generation of *Synechocystis* strains used in this study. **A**, integration of gene encoding 3×FLAG-tagged ChlD at the *psbAII* (slr1311) locus in strain FLAG-ChlD. Agarose gel analysis with primers flanking the integration site shows that a single larger PCR product is amplified from the FLAG-ChlD genomic (g)DNA compared to that from the wild-type (WT) due to insertion of the tagged gene and the kanamycin resistance cassette (kan^R) in place of *psbII*. **B**-**C**, deletion of the native *chlD* (slr1777) gene in strain FLAG-ChlD Δ*chlD*. **B**, agarose gel analysis with primers flanking the deletion site shows that a smaller PCR product is amplified from FLAG-ChlD Δ*chlD* gDNA compared to FLAG-ChlD gDNA due to replacement of the *chlD* gene with the smaller zeocin resistance cassette (zeo^R). **C**, to confirm that the FLAG-ChlD Δ*chlD* has fully segregated at the *chlD* locus PCR with a reverse primer internal to the deleted portion of the gene was performed. A product is present using FLAG-ChlD gDNA as template but absent when using gDNA from the FLAG-ChlD Δ*chlD* strain. In all panels the position of primers used in the PCR analysis are indicated with small black arrows and lane M contains Hyperladder™ I (Bioline, UK).

FLAG-ChlD Δ*chlD Synechocystis* PCC6803 was grown as described in (Section 2.7), membranes were solubilised in FLAG buffer (Section 2.8.2.5) containing 1.5 % β-DDM and material from

both soluble and membrane fractions were separated by SDS-PAGE (Section 2.10.2). The gel was then probed by immunoblotting (Section 2.10.4) for ChlD and ChlH (Figure 3.2 **A**). Both soluble and membrane fractions contained ChlH and ChlD, though more was present in the membrane fractions. The chlorophyll biosynthetic enzymes are believed to be membrane associated (Masuda, 2008) so it is perhaps unsurprising that the membrane fraction contained more of the MgCH proteins.

The solubilised membrane fraction was applied to an anti-FLAG column (Section 2.8.2.5); eluates were separated by SDS-PAGE and probed for the constituents of the MgCH enzyme, revealing the presence of ChlH, ChlD and ChlI, as well as a *ca.* 50 kDa breakdown product of ChlD that appears in most preparations of ChlD (Figure 3.2 B). The epitope of the anti-ChlD antibody is the N-terminus of ChlD so this breakdown product must be the N-terminal region.

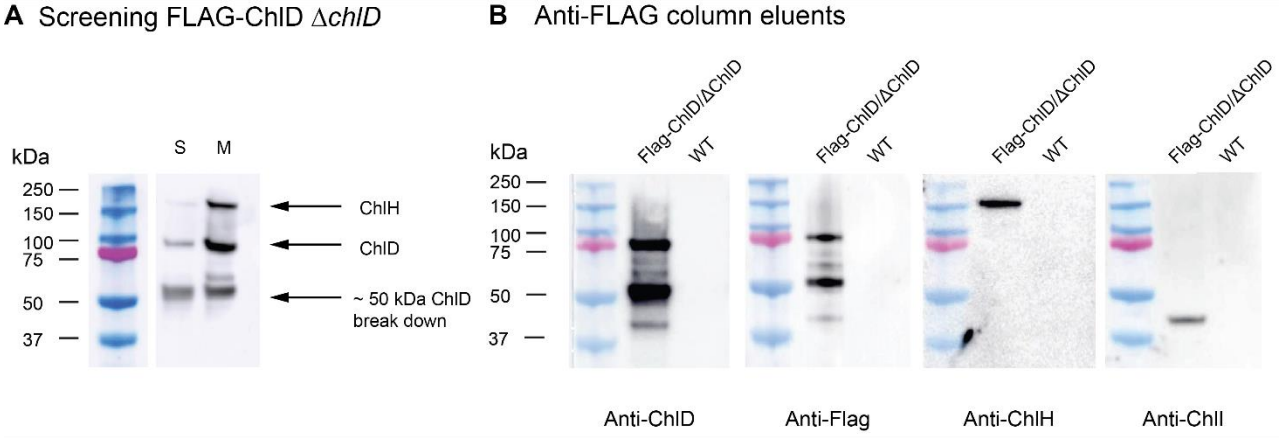


Figure 3.2. Purification of FLAG-ChlD-ChlH complex from *Synechocystis* cells. **A**, FLAG-ChlD Δ*chlD Synechocystis* soluble and detergent solubilised extracts were applied to an anti-FLAG column to optimise purification conditions. Proteins were separated by SDS-PAGE as in Section 2.10.2 and immunoblotted as in Section 2.10.4. Lanes: 1, markers; S, soluble fraction; M, membrane fraction. The ChlD-ChlH complex detected in both soluble and membrane fractions, with larger amounts of both proteins present in the membrane fraction. **B**, immunoblot probed with antibodies for each chelatase subunit in FLAG-ChlD Δ*chlD Synechocystis* and WT *Synechocystis* membrane fractions eluted from a FLAG column. The three constituent proteins of magnesium chelatase can all be co-purified from solubilised membranes when there is a N-terminal FLAG tag on ChlD.

3.2.2 Strep-ChlH can pull down ChlD when expressed in non-native *E. coli*

It was important to see if this interaction persists in non-native organisms and therefore does not rely on other external factors in *Synechocystis*. A plasmid containing both an N-terminal StrepII-tagged ChlH and an untagged ChlD was designed by Dr Amanda Brindley and the respective genes

were co-expressed in *E. coli* (Section 2.6) for co-purification by StrepTactin affinity chromatography (Section 2.8.2.4). Eluates were separated by SDS-PAGE and bands of the correct molecular mass for ChlH (150 kDa) and ChlD (75 kDa) were observed (Figure 3.3 **A**). The gel was analysed by immunoblotting to confirm the presence of both ChlH and ChlD (Figure 3.3 **B**).

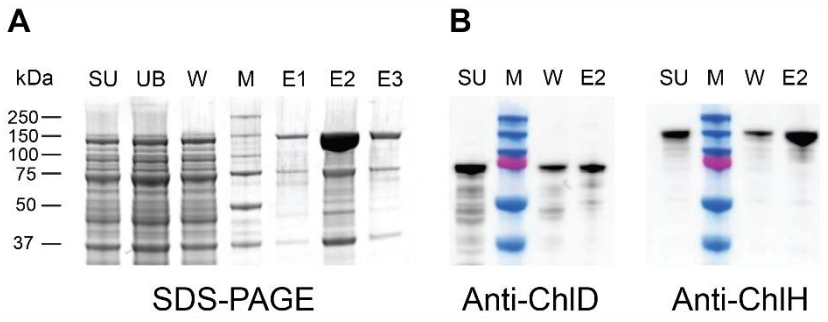


Figure 3.3. Copurifying recombinant ChlD and ChlH. **A**, *Synechocystis* pET21aChlH-STREP was co-produced with non-tagged pET3aChlD and purified on a STREP HP column. Lanes: Su, cell supernatant; M, molecular weight markers as indicated on right hand side; UB, unbound fraction; W1, 5 ml wash 1 (binding buffer (20 mM Sodium phosphate, 10 mM MgCl2, 50 mM NaCl, 10 % glycerol, pH 7.4); W2, 10 ml buffer; W3, 15 ml binding buffer, Eluents (E 1 – 6): Binding buffer with 2.5 mM biotin. **B**, immunoblot analysis of ChlH-STREP and non-tagged ChlD fractions from **A**, showing the presence of ChlD in both the supernatant and elution fraction.

3.2.3 Separately purified ChlH and ChlD form a complex observable on a native gel

As a method to try to capture the ChlH-ChlD complex directly, ChlH and ChlD were expressed and purified independently (Section 2.8.2) before being combined at 15 μM and 12 μM concentrations, respectively. The mixture was incubated at 34 ℃ for 10 mins, and then purified by native-PAGE (Section 2.10.3). Complexes of *ca.* 350 kDa and 450 kDa were observed, consistent with a ChlH-ChlD complex in a ratio of 1:2/ 2:1 and 2:2, respectively (Figure 3.4). Electro-elution was attempted in order to recover some material for negative stain EM studies, but no protein was observed, and the electro-elution deemed unsuccessful.

3.3 Characterising the ChlH-ChlD interaction

The above work on *Synechocystis* demonstrates that ChlH and ChlD interact in a manner strong enough to withstand purification procedures, and independent of other native proteins. Further biochemical characterisation of the interaction could help to elucidate its role in the function of the MgCH complex. MST and kinetic assays were applied to the complex to reveal information

about the strength of the interaction in the active MgCH complex. As well as this, cross-linking mass-spectrometry (XLMS) was used to identify the regions crucial for the ChlH-D interaction.

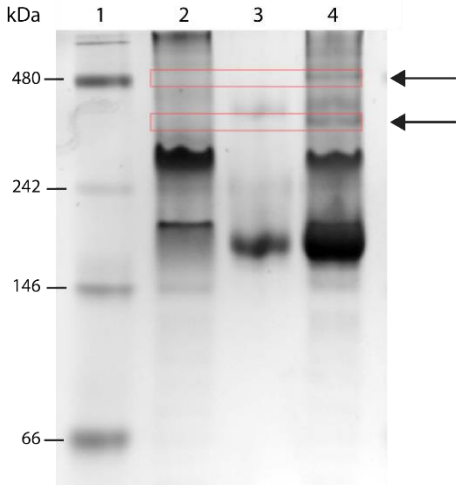


Figure 3.4. Capturing the ChlD-ChlH complex via Native PAGE. 15 µM ChlH was mixed with 12 µM ChlD and incubated at 34 °C prior to separation by 7 % bis-acrylamide Native PAGE. Lanes: 1, NativeMark (ThermoFisher) protein size marker (kDa); 2, 12 μM ChlD; 3, 15 μM ChlH; 4, ChlD-ChlH mix. Red boxes and arrows highlight new bands in lane 4 not seen in lane 2 or 3. The band at *ca*. 150 kDa in lane 2 is presumed to be a ChlD dimer.

3.3.1 Quantitative analysis of the ChlH-ChlD interaction by MST

MST was conducted on independently purified ChlH and ChlD, to quantify the strength of the interaction independent of ChlI and substrates and to determine what may affect the interaction. Magnesium is one of the substrates of the reaction, but has been found to have additional roles; this is especially true of the *Synechocystis* enzyme, which has a cooperative response to magnesium (Reid and Hunter, 2004). MST was thus conducted in both the presence and absence of magnesium (Section 2.12). ChlH and ChlD were purified as in (Section 2.8.2) and ChlH labelled with NHS-NT647 (Nanotemper) according to the manufacturer's instructions. ChlD was titrated into a constant concentration of labelled ChlH. Dissociation constants (K_d) values were calculated using a single-site binding model (Equation 2.4) and the MO.Affinity Software (Nanotemper) supplied with the instrument. The ChlH-ChlD interaction proved to be magnesium dependent; a K_d of 332 ± 58 nM was obtained in the presence of 10 mM Mg^{2+}, with an order of magnitude increase to 2089 ± 387 nM in the absence of Mg^{2+} (i.e.: the interaction is stronger when magnesium is present) (Figure 3.5).

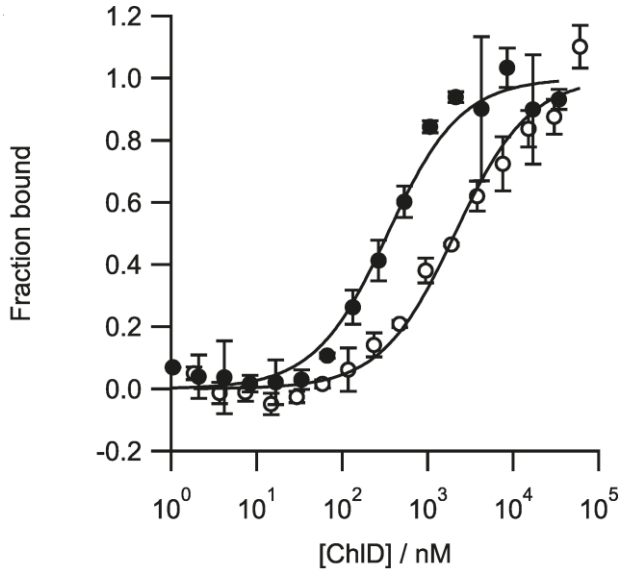


Figure 3.5. Quantifying the assembly of the ChlD-ChlH complex with MST. The presence of magnesium increases strength of the ChlD-ChlH interaction. Thermophoresis was performed with ChlD titrated into 20 nM labelled ChlH with or without 10 mM Mg^{2+}. The interaction shows Mg^{2+} dependency, with a higher affinity shown at 10 mM Mg^{2+} (filled circles) and Mg^{2+} free (open circles) conditions. Fitting of the resultant binding isotherms revealed K_d values of 332 ± 58 nM and 2089 ± 387 nM for Mg^{2+} present and Mg^{2+} free, respectively

The assembly states of ChlH and ChlD were assessed to confirm that the observed MST trace did correspond to the ChlH-ChlD interaction. Unlabelled ChlD was titrated into ChlD that had been labelled with NHS-NT647 (Nanotemper) according to the manufacturer's instructions; this analysis showed that a single-site binding event occurs between two ChlD molecules with a K_d of 296 nM (Figure 3.6 **A**). ChlH was titrated against itself and it was found that it is monomeric at the 20 nM concentration used in the ChlH-ChlD MST experiments (Figure 3.6 **B**). However, at concentrations higher than 100 nM there is a binding event between ChlH molecules. Whether the oligomerisation of ChlH has a role in the function of MgCH, or if it still occurs in the presence of the other subunits could be an avenue for further research. These results suggest that in the ChlH-ChlD MST experiment a dimer of ChlD may be interacting with a monomer of ChlH, though the data for the ChlH-ChlD interaction fit well with a single-site binding curve so more complex assemblies are unlikely.

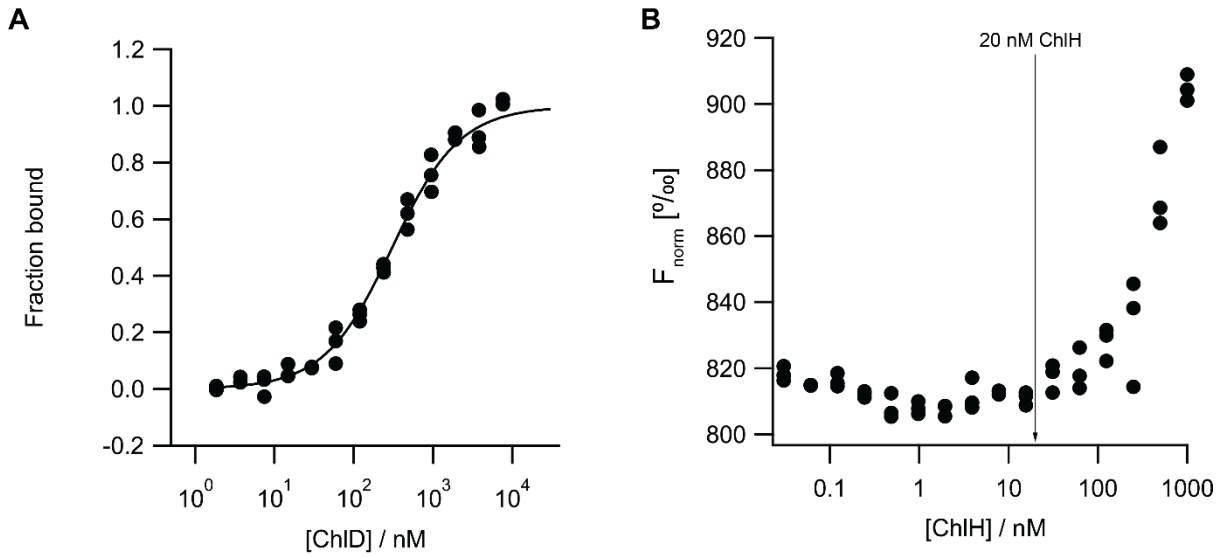


Figure 3.6. Monitoring self-assembly of **(A)** ChlD and **(B)** ChlH. Proteins were labelled with NHS-NT-647 and MST was conducted. **A**, a single site binding model fits the data, resulting in a K_d = 296 nM, suggesting that ChlD may bind to ChlH as a dimer. **B**, as indicated by the arrow, there is no observed interaction of ChlH with itself at the concentrations used in ChlD titrations (20 nM). It does seem that ChlH will undergo dimerisation at higher concentrations, however. Each chelatase subunit was titrated into 20 nM of the same NHS-NT647 labelled subunit and MST performed as in Section 2.12.

Previously, Brindley et al. (2015) showed that five glutamate residues in the C-terminal region of ChlD (E510, E513, E600, E603 and E605; *Synechocystis* numbering) regulate the cooperative response to Mg^{2+} observed in chelatase assays in the *Synechocystis* enzyme. Binding studies were performed on ChlH and a quintuplet mutant of ChlD with these five glutamate residues mutated to test if the Mg^{2+} dependent decrease in K_d for the ChlH-ChlD association observed in Figure 3.5 was related to the magnesium cooperativity. The five glutamate residues were previously mutated with the corresponding residues found in the non-cooperative *T. elongatus* ChlD subunit (E510Q/E513Q/E600T/E603P/E605T) (Brindley et al., 2015). The mutant (ChlD QuinE) was produced and purified as in Sections 2.6 and 2.8.2, then MST conducted as before. In the presence of Mg^{2+} the interaction between ChlD QuinE and ChlH has a K_d similar to that of wild-type protein, which is essentially unchanged in the absence of Mg^{2+} (Figure 3.7). These results suggest that Mg^{2+} dependent regulation of the ChlH-ChlD interaction is the source of the magnesium cooperativity observed for the *Synechocystis* enzyme.

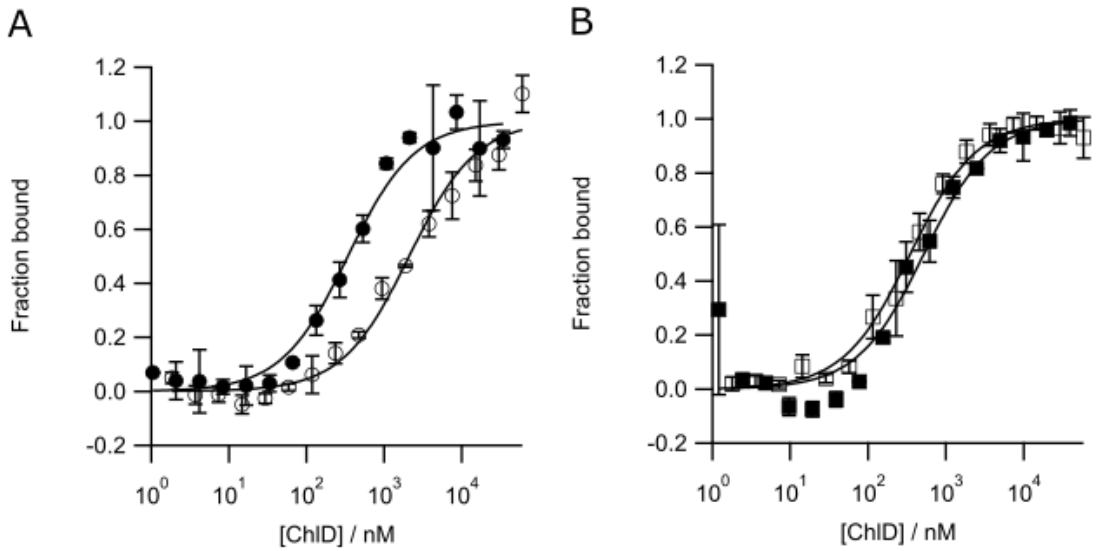


Figure 3.7. The Quintuple mutant (QuinE) of ChlD does not have a Mg^{2+} dependent K_d for ChlH. Unlike WT ChlD, the interaction between ChlH and the non-cooperative QuinE mutant of ChlD is not dependent on Mg^{2+}. MST traces of WT **(A)** and QuinE **(B)** in the presence and absence of Mg^{2+} (filled and unfilled symbols respectively). The fitted curves reveal a K_d of 338.5 ± 39.7 nM and 512.5 ± 161.31 nM for the QuinE mutant in the presence and absence of Mg^{2+}, respectively.

3.3.2 Quantitative analysis of the ChlH-ChlD interaction in the active chelatase by kinetic titrations

To investigate the binding interaction between ChlD and ChlH in the context of the active chelatase (i.e.: ChlIDH and a full set of substrates), ChlH was titrated against a constant concentration of ChlID and the formation of magnesium deuteroporphyrin IX was monitored (Section 2.13.2). Steady state rates (v_{ss}) were obtained at a range of $MgATP^{2-}$ concentrations to calculate saturating kinetic values (Figure 3.8 **A**) which can be described by Equation 2.3, the Michaelis-Menten relationship. The V_{max} values obtained from this fit were plotted as a function of ChlH concentration and can be described by a binding model (Equation 2.4 with parameters: v_{lim} = 0.87±0.04 nM min^{-1}; K_{app} = 13.4±0.05 nM; *nd* = 1.92±0.13 μM (Figure 3.8 **B**)).

The coefficient *nd* provides information about the number of binding sites for ChlD i.e.: [ChlD]× (number of sites per ChlD). The number of binding sites is an order of magnitude higher than the number of ChlD molecules which can either describe approximately ten ChlD proteins binding to one ChlH protein, or the more realistic prospect that one ChlID complex is able to bind to and supply approximately ten ChlH proteins with energy per chelation cycle. The previous results from the MST also suggest that ChlH has only one binding site for ChlD, making the latter explanation

more likely. It is perhaps more productive for the ChlID complex to dissociate from ChlH once the driving force is provided for the reaction and then move onto another ChlH molecule. However, some of these results could be biochemical artefacts related to the large excess of ChlH, and lack of additional chlorophyll biosynthesis proteins (e.g. Gun4, the methyltranferase or cyclase) to assist ChlH in product release.

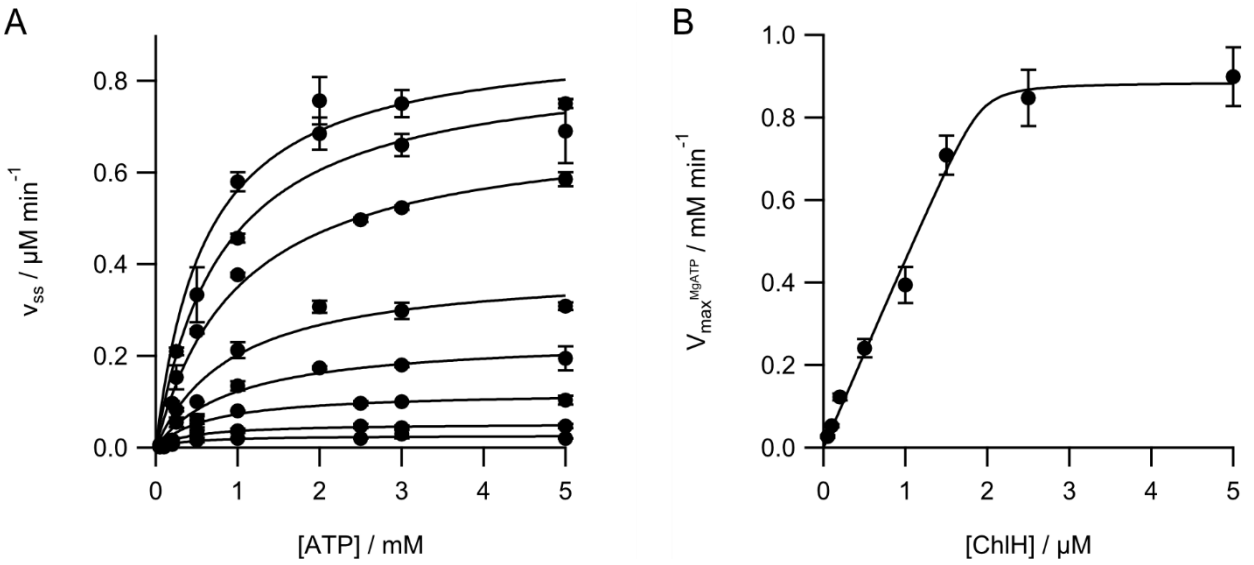


Figure 3.8. Quantifying the ChlH-ChlD interaction within the context of the whole chelatase. **A**, secondary plot of the ChlD–ChlH assembly titration. Steady state rates estimated and plotted as a function of ATP concentration. Assays conducted as in Section 2.13.2. Each plot represents a different ChlH concentration (5 µM to 0.02 µM ChlH) The curves can be described by Equation 2.3 to calculate V_{max} and K_M values for each ChlH concentration. **B**, tertiary plot of resultant V_{max} values as a function of ChlH concentration. This curve can be described by Equation 2.4 with parameters: v_{lim} 0.87 ± 0.044 mM min^{-1} K_{app} 13.4 ± 0.045 nM *nd* 1.92 ± 0.131 µM.

3.3.3 Cross-linking-MS reveals that the C-terminal domain of ChlD mediates the interaction with ChlH

Chemical cross-linking was employed to gain an insight into the physical location of the interaction site between ChlH and ChlD. The water-soluble zero length 1-ethyl-3-(3-dimethylaminopropyl) carbodimide hydrochloride (EDC) cross-linker was used (Section 2.16). EDC reacts with primary amines in the proteins, primarily lysine residues. The resultant covalently linked complexes were separated by SDS-PAGE. After immunoblotting for ChlH and ChlD, two high molecular weight bands with both proteins present (marked by arrows in Figure 3.9 **A**) were excised and subjected to tryptic digest and mass spectrometry analysis (Section 2.16). The bands corresponded to *ca.* 175 kDa and 275 kDa, which do not correlate with multiples of ChlH and ChlD. However, due to the introduction of the covalent bonds by the cross-linking reagent, the complexes may not

be completely linearized by the SDS, which could alter the characteristics of migration down the gel.

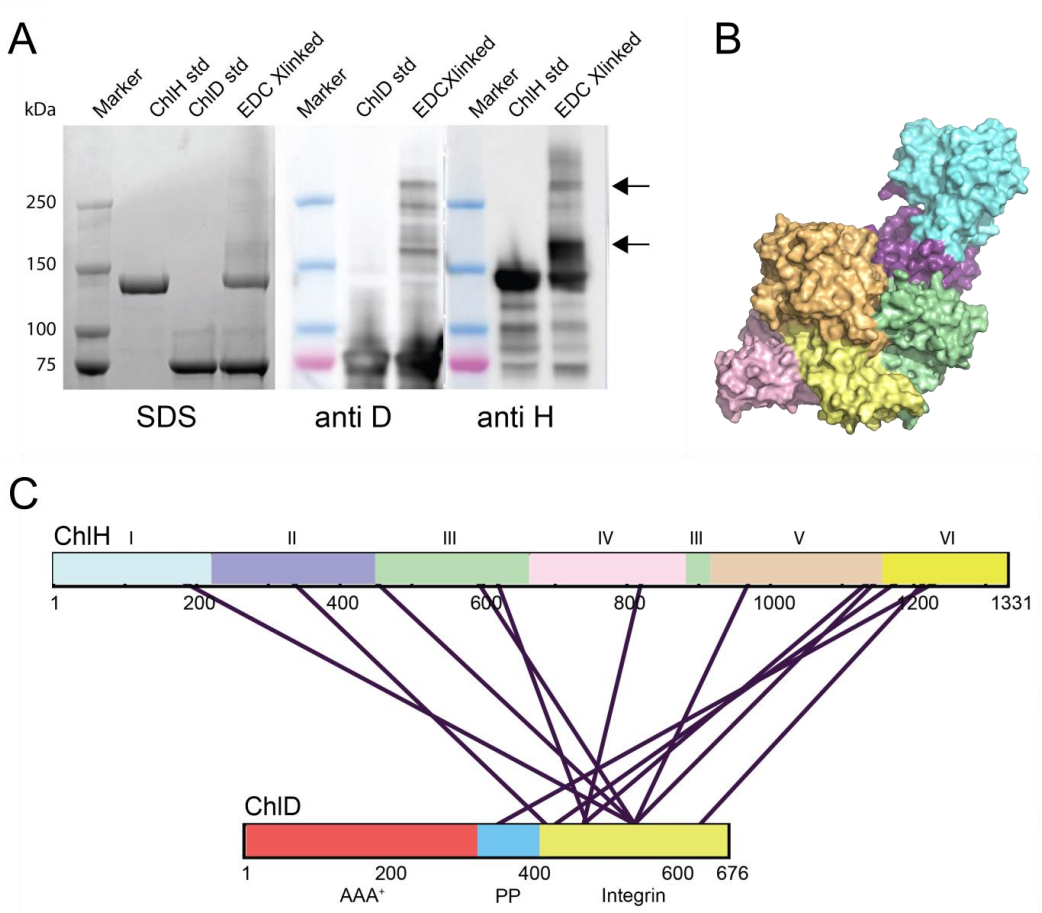


Figure 3.9. Capturing the ChlD-ChlH complex. **A**, SDS-PAGE analysis of chemically cross-linked proteins. 1 mg ml^{-1} ChlH was cross-linked with 1 mg ml^{-1} ChlD using EDC (see Section 2.16); immunoblotting shows high molecular weight bands containing both ChlH and ChlD (indicated by arrows). Lanes as labelled above. **B**, atomic structure of ChlH, domains coloured as in **C**. **C**, bands indicated by arrows in **A** were subjected to in-gel digestion with trypsin and analysed for EDC-crosslinked residues by mass spectrometry; the primary sequence of the ChlH and ChlD proteins are presented as rectangles with residues as marked. For ChlH, the six domains are numbered as identified by (Chen et al., 2015a). ChlD has three identified domains as labelled below, AAA+, acidic poly proline region (PP) and the C-terminal Integrin domain. Purple lines denote chemical crosslinks between ChlH and ChlD.

Mass spectrometry analysis was kindly conducted by Dr Nathan Adams and Dr Philip Jackson. Crosslinks were located between peptides that corresponded to the acidic proline-rich domain and the C-terminal integrin I-like domain of ChlD (blue, yellow, respectively; Figure 3.9 **B**). These were covalently linked with, for the most part, the "body" of ChlH, as well as one fragment corresponding with the "head" region. These results are at odds with those of Luo *et al.* (2018), but

consistent with other studies implicating the importance of fully intact ChlD (Adams et al., 2016b; Brindley et al., 2015).

3.3.4 C-terminal truncations of ChlD display weakened affinity for ChlH

The cross-linking results show a clear affinity of the C-terminal domain of ChlD for ChlH. However, there are cross-links between the acidic proline-rich region also, which suggests that ChlD may be forming a complex with ChlH in a manner that positions the N-terminal domain of ChlD in close proximity to ChlH. To determine whether the N-terminus of ChlD is involved in interactions with ChlH, MST was performed on previously designed C-terminally truncated ChlD proteins (Figure 3.10 **A**) (Adams et al., 2016b). The two ChlD mutants were truncated to remove both the integrin I-like domain and the proline-rich region, or the integrin I-like domain alone.

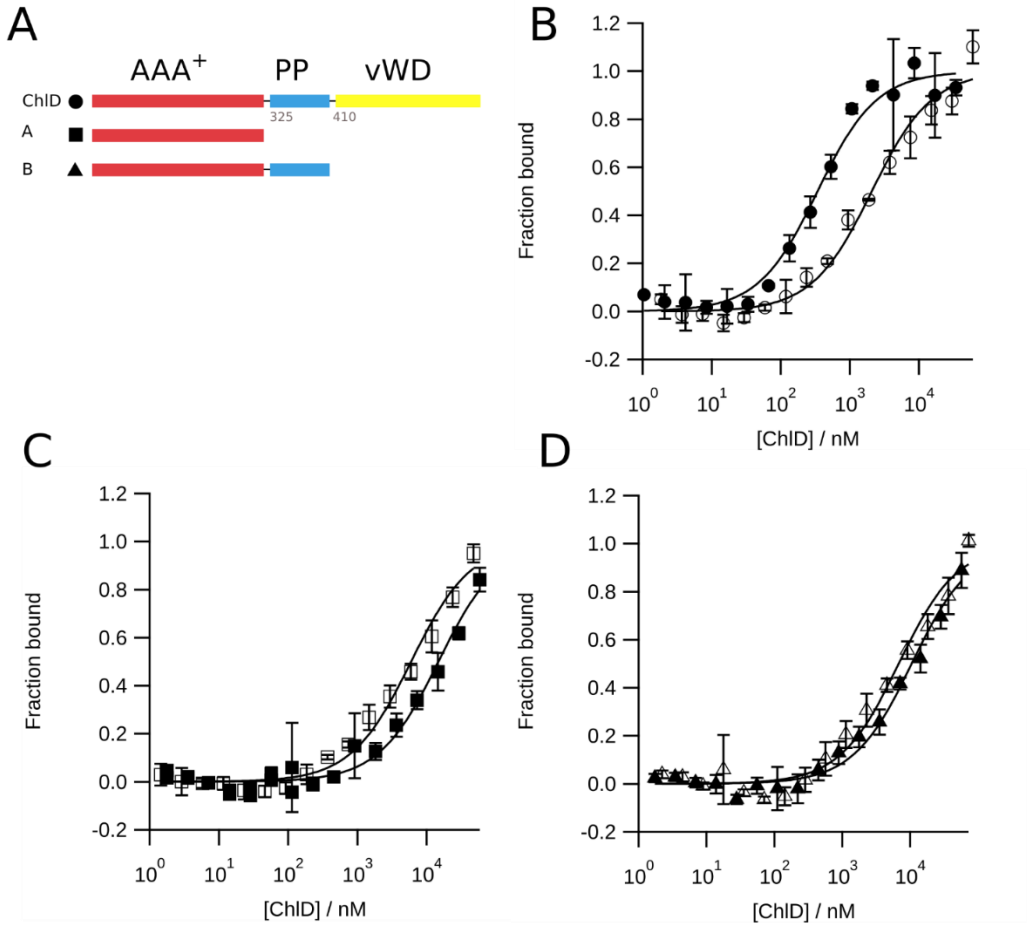


Figure 3.10. Quantifying the strength of the ChlH-ChlD interaction in the absence of the C-terminal domain of ChlD. **A**, Schematic illustrating the truncation mutants. **B**, WT MST trace for comparison. **C-D**, MST assays to measure binding of C-terminal truncations of ChlD to ChlH. Truncations consisted of either the AAA+ domain alone (squares), or the AAA+ domain with the polyproline region (triangles). Closed markers, 10 mM free Mg^{2+}; open markers, no free Mg^{2+}.

The proteins were overproduced and purified as in (Sections 2.6 and 2.8.2); ChlH was labelled and MST conducted as before. The two mutants had an almost 50-fold increase in K_d highlighting the importance of the C-terminal domains of ChlD in the ChlH-ChlD interaction (Figure 3.10 **B-D**). However, that these mutants can still bind ChlH at all suggests that the interaction may not be wholly driven by the integrin I-like domain of ChlD.

3.4 The role of the MIDAS motif in the ChlH-ChlD interaction

The above results demonstrate that the majority of the interactions with ChlH involve the C-terminal integrin I-like domain of ChlD occurs. Integrin I domains contain a conserved metal ion dependent adhesion site (MIDAS) which regulate the interaction strengths with the binding partners and often coordinate a conformational change upon ligand binding. A MIDAS motif is present in subunit D of MgCH and the importance of this motif has been demonstrated previously in BchD (Axelsson et al., 2006). There are also common amino acid motifs within the binding partner protein that are complementary to the MIDAS motif, which include LDV and RDG, amongst others (Humphries et al., 2006; Plow et al., 2000). An LDV motif is present in BchH and has previously been suggested as a possible site of interaction for the integrin I-like domain of BchD (Fodje et al., 2001). The same LDV motif is also present in ChlH, potentially revealing some functional conservation; the motif is located near the surface of ChlH, though is not completely exposed to solvent. The MIDAS and the associated LDV motifs were probed by mutagenesis to assess their role in the ChlH-ChlD interaction.

3.4.1 MST of the LDV motif in ChlH

The LDV motif (residues 474-476, *Synechocystis* numbering) of ChlH was probed by mutagenesis and MST. Double mutations were introduced by Quickchange (Agilent) as described in the manufacturer's instructions; the LDV residues were mutated to alanine residues to disrupt the motif (Section 2.4). The resultant ChlH proteins were overproduced and purified as for the wild-type (Section 2.6 and 2.8.2) and labelled with NHS-NT647 as described in the manufacturer's instructions. MST was conducted as before to probe for disruptions of the ChlH-ChlD interaction. As shown in Figure 3.11, all mutants have a K_d value approximately equal to that of wild-type when in the presence of Mg^{2+}, suggesting that this motif is not essential for the interaction when Mg^{2+} is present. However, in the absence of Mg^{2+} the LDV mutants have decreased affinity for ChlD when compared to wild-type (Table 3.2). While there is no clear dependency on the LDV motif

for the ChlH-ChlD interaction in the presence of Mg^{2+}, the motif may play more of a role of stabilising the interaction in low Mg^{2+} conditions.

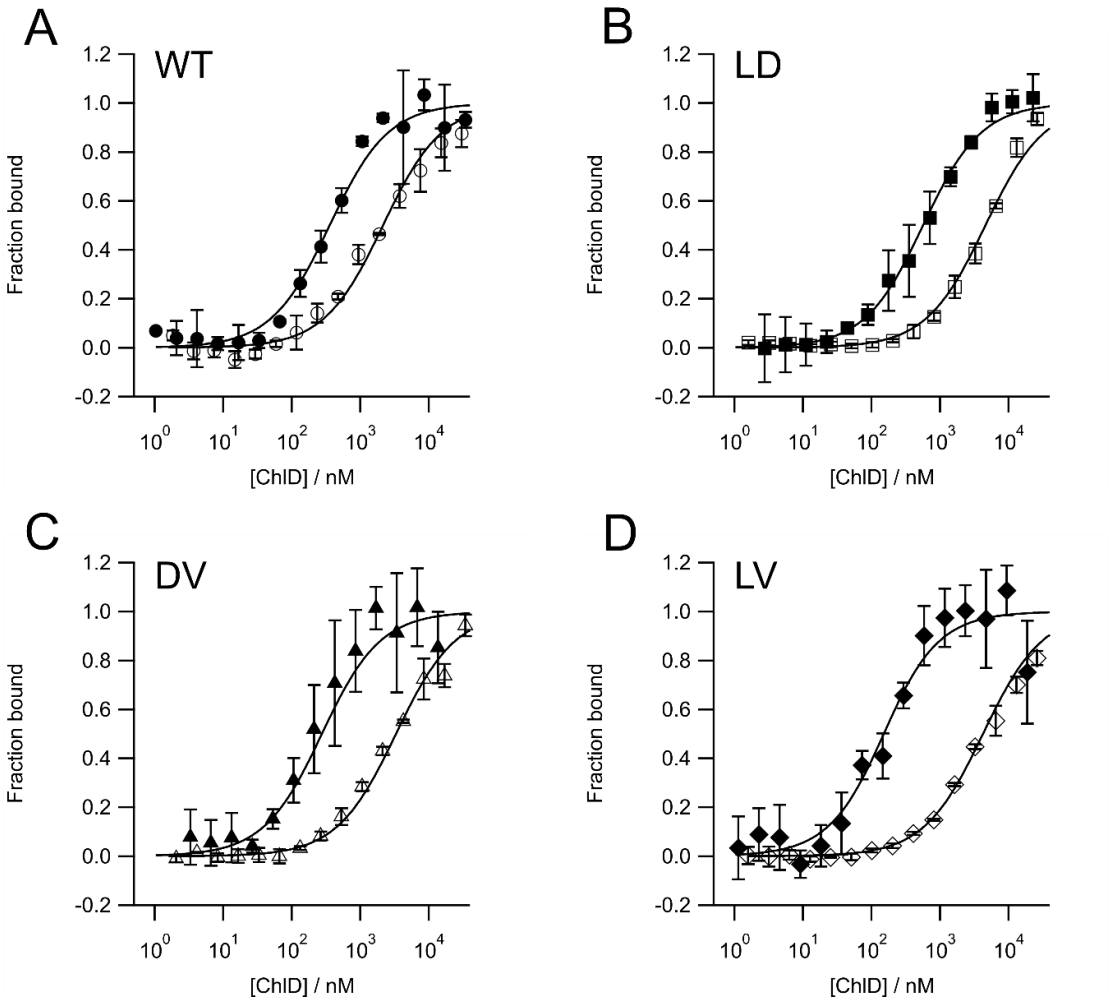


Figure 3.11. Quantifying the strength of the ChlH-ChlD interaction in ChlH LDV mutants. **A**, WT MST trace for comparison. **B-D**, MST traces for the ChlH LDV mutants (labels representing the residues in the LDV motif mutated to alanine). All mutants have similar K_d values to WT in the presence of Mg^{2+} (closed markers). K_d values in the absence of Mg^{2+} (open markers) are higher for all of the LDV mutants (for specific values, see Table 3.2).

3.4.2 Producing a homology model of the C-terminal domain of ChlD

A homology model of the integrin I-like domain of ChlD was built to aid the design of mutations in the MIDAS region. Two classes of mutants were selected that would disrupt the magnesium binding of the MIDAS motif (Type I) and disrupt metal binding while providing a replacement positive charge (Type II). A search of the protein data bank for homology to the integrin I-like

domain of ChlD using the HHPRED (Söding et al., 2005) sequence alignment server resulted in a match of 21 % sequence identity with the extracellular domain of Tumour Endothelial Marker 8 (TEM8, PDB accession code: 3N2N). The alignment between residues 481-672 of *Synechocystis* ChlD and TEM8 was used with Modeller (version 9.21) (Webb and Sali, 2002) to generate ten homology models, with the best one chosen by Modeller's inbuilt DOPE score (Shen and Sali, 2006).

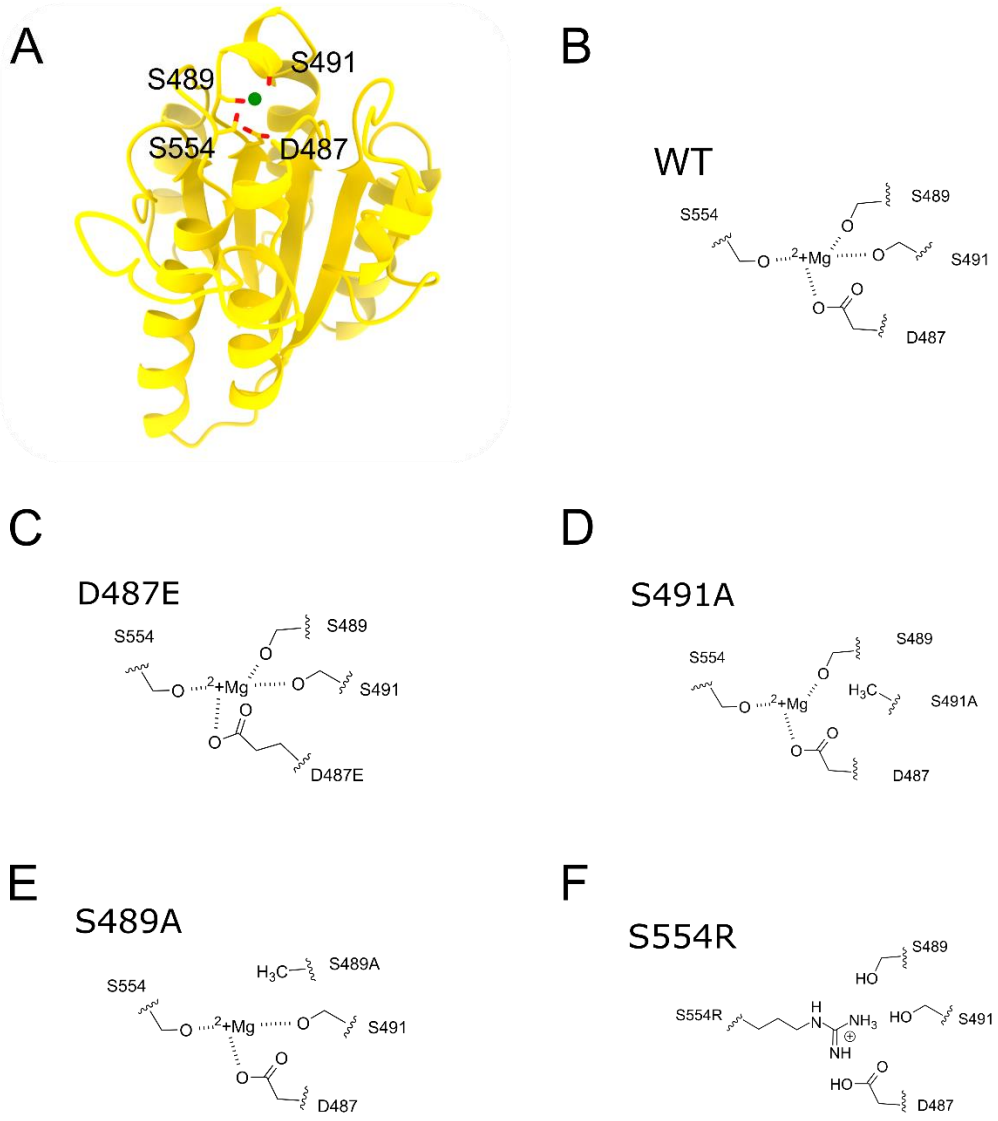


Figure 3.12. The MIDAS motif and the mutants used to probe its importance for the ChlH-ChlD interaction. **A**, homology model of the C-terminal integrin I-like domain of ChlD. Resides involved in the MIDAS motif that were probed by mutagenesis are shown as sticks and labelled. Mg^{2+} atom modelled in based on the homology with TEM8 (PDB: 3N2N). **B-F**, schematic representations of WT **(A)** and mutant **(C-F)** MIDAS motif highlighting the predicted disruption to Mg^{2+} binding.

The residues involved in the MIDAS motif of ChlD were assessed visually in PyMOL and virtual mutagenesis allowed for the design of the aforementioned classes of mutations. The mutants selected for Type I (metal binding perturbation) were D487E, S489A and S491A while for Type II (metal binding perturbation and replacement positive charge) the mutation was S554R; the positively charged guanidine group of the arginine residue predicted to be positioned in place of the Mg^{2+} (Figure 3.12 **F**).

3.4.3 Mutations in the MIDAS region are impaired catalytically

The selected mutations of ChlD were produced using the Quickchange kit (Agilent) according to the manufacturer's instructions, and the successful introduction of the mutations confirmed by DNA sequencing (GATC). All recombinant proteins were expressed and purified as wild-type (Section 2.6 and 2.8.2). The mutant ChlD proteins were used in MgCH assays (Section 2.13.2) to determine the effect of the mutations on chelation rate. All Type I mutants showed no activity (Figure 3.13 **B**-**D**), confirming the importance of the MIDAS motif for chelation. The Type II S554R mutant was found to have *ca.* 5 % of the activity of wild-type (Figure 3.13 **E**), perhaps suggesting that despite the importance of the motif highlighted by the above results, a potential replacement positive charge provided by the arginine residue may allow for roughly correct ChlH-ChlD association and power transfer from the ChlID complex. However, the reduced chelation rate of the S554R mutant reiterates the importance of this motif for efficient catalysis. It is presumed that the mutations created disrupt the MIDAS motif in such a way as to reduce metal binding in this region, however, it is difficult to confirm this experimentally.

3.4.4 Assessing the stability of the MIDAS mutants

Circular dichroism (CD) spectroscopy (Section 2.15) and differential scanning fluorimetry (DSF) (Section 2.14) were used to verify the foldedness and stability of these mutants under assay conditions. DSF curves were fit to the sigmoid Boltzmann distribution equation (Equation 2.6) and T_m values for the mutants (Table 3.1) are all above the 34 °C that the chelatase assay is conducted at, as well as being roughly as stable as wild-type (Figure 3.14 **A**). All mutants displayed broadly similar CD traces to wild-type, indicating no significant changes in secondary structure (Figure 3.14 **B**-**C**).

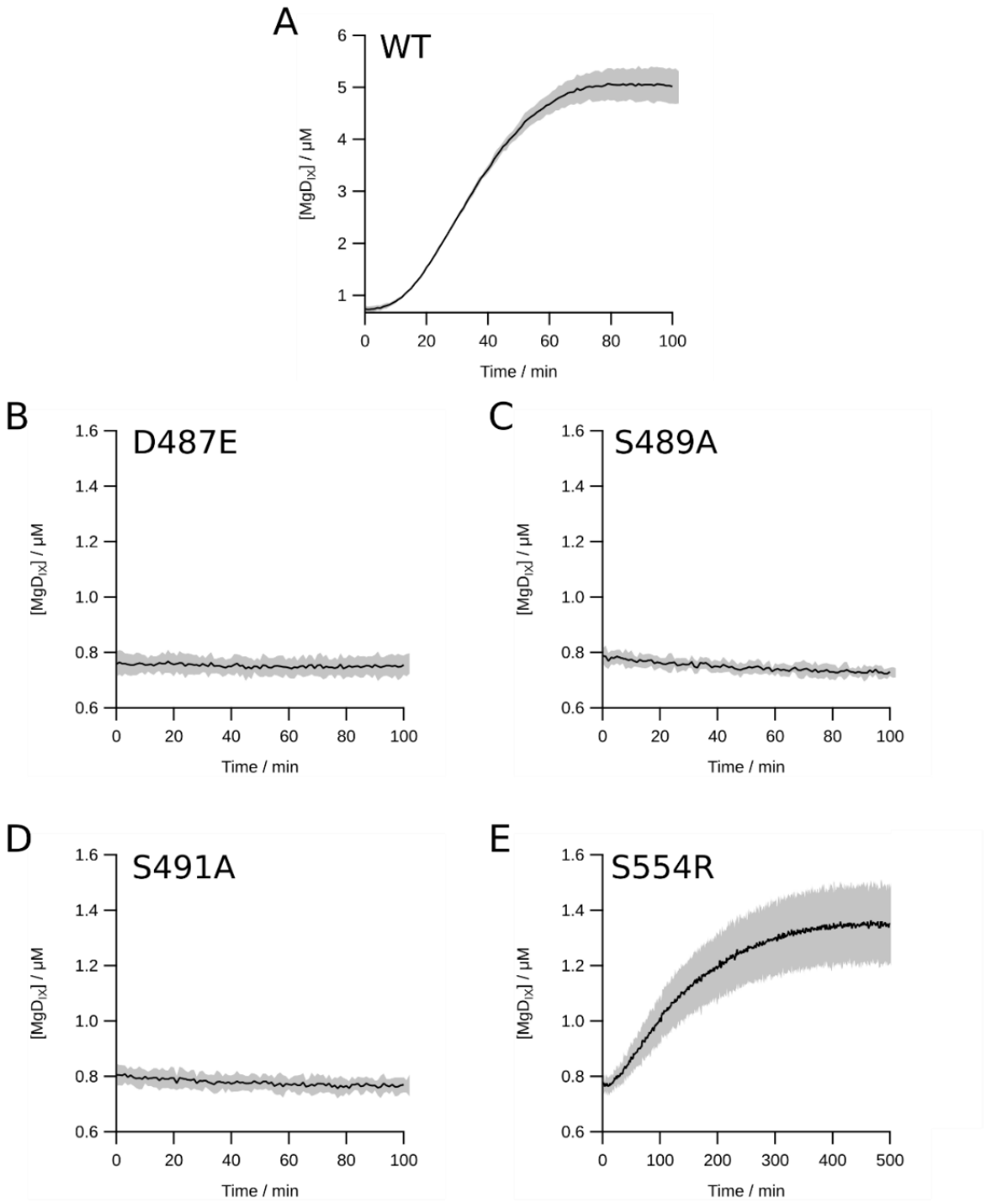


Figure 3.13. Primary plots from chelatase assays of WT **(A)** compared to MIDAS mutants **(B-E)**. Only mutant S554R shows any activity, though this is *ca.* 5 % of the activity of WT. Standard deviation (n = 3) is represented as shading. Conditions are as in Section 2.13.2.

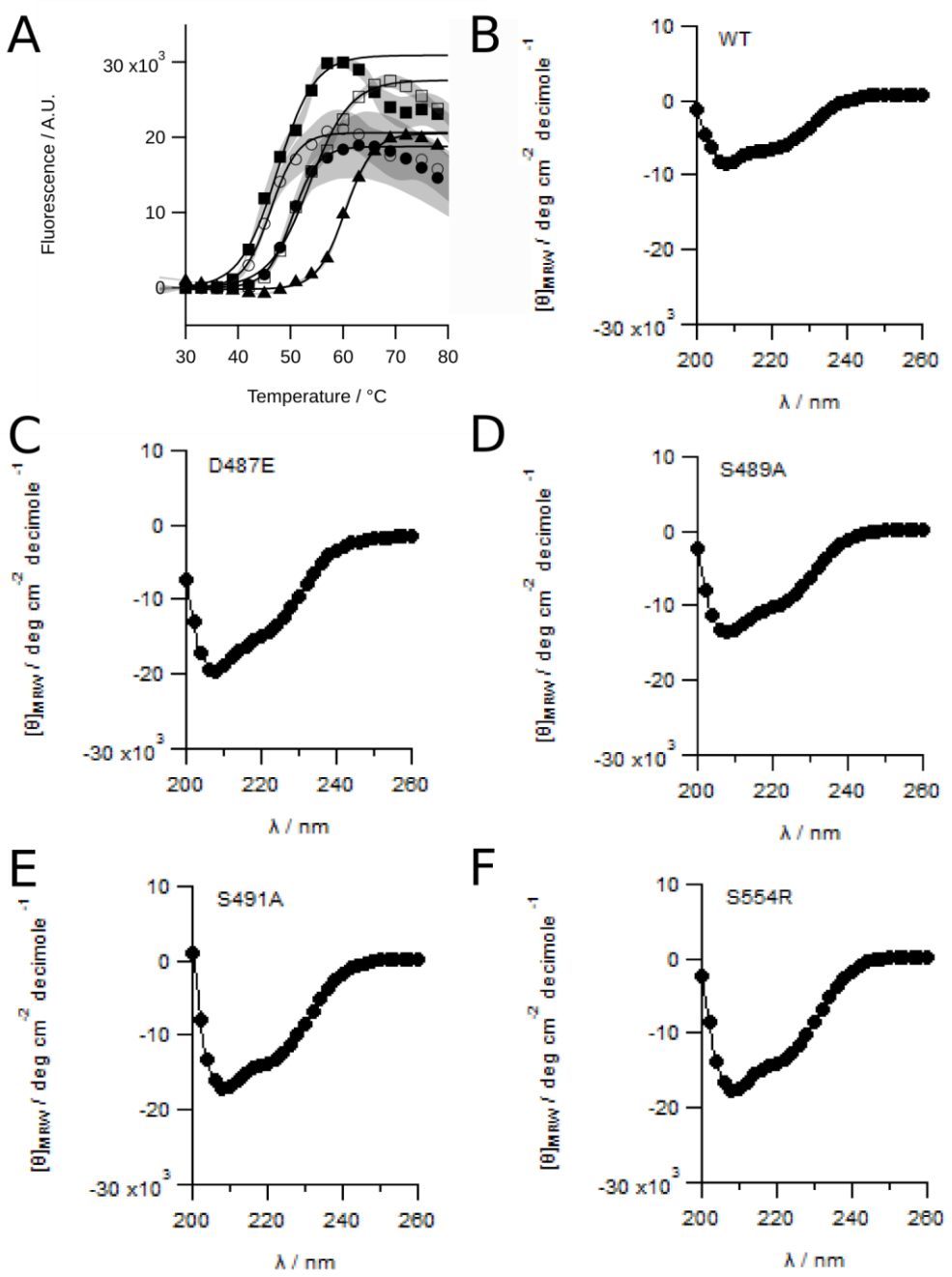


Figure 3.14. Temperature stability and extent of folding of the MIDAS mutants**. A**, DSF melting curves for WT and MIDAS mutants. All mutants (D487E, filled square; S489A, filled triangle; S491A, open circle; S554R, open square) have a thermal stability comparable or above WT (filled circles), with all calculated T_m values greater than the MgCH activity assay temperature of 34 °C. T_m values are shown in Table 3.1. Assay conducted as in Section 2.14. Every three values are shown for clarity, and can be described by equation Equation 2.6 to calculate T_m values. **B-F**, CD spectra of the ChlD MIDAS mutants. All four mutants have CD spectra broadly similar to the WT, suggesting that secondary structure of these mutants has not been disrupted.

Table 3.1. T_m values calculated from fits of the DSF melting curves of ChlD MIDAS mutants.

ChlD MIDAS mutant	T_m / °C
WT	50
D487E	47
S489A	60
S491A	46
S554R	54

3.4.5 The effect of the MIDAS mutations on the ChlH-ChlD interaction

The impact of the single residue mutations on chelatase activity reveals the importance of the MIDAS motif for chelation efficiency. Given that the motif is involved in regulating protein-protein interactions in other systems, the effect of these mutations on the ChlH-ChlD interaction was investigated by MST. MST was performed on each of the MIDAS mutants in the presence or absence of Mg^{2+} (Figure 3.15). The K_d values (Table 3.2) show that all mutants are able to bind ChlH at least as well as the wild-type in the absence of Mg^{2+}, with ChlD S491A appearing to have a greater affinity for ChlH than wild-type, especially in Mg^{2+} free conditions. The only active mutant, S554R, has K_d values for ChlH roughly similar to that of WT protein. The MIDAS motif appears to be important for more than regulating the ChlH-ChlD interaction; it perhaps plays a more active role in the catalytic cycle, although it is unclear what that may be from these results. Experimental structural information about this complex and the whole MgCH complex would undoubtedly yield more answers.

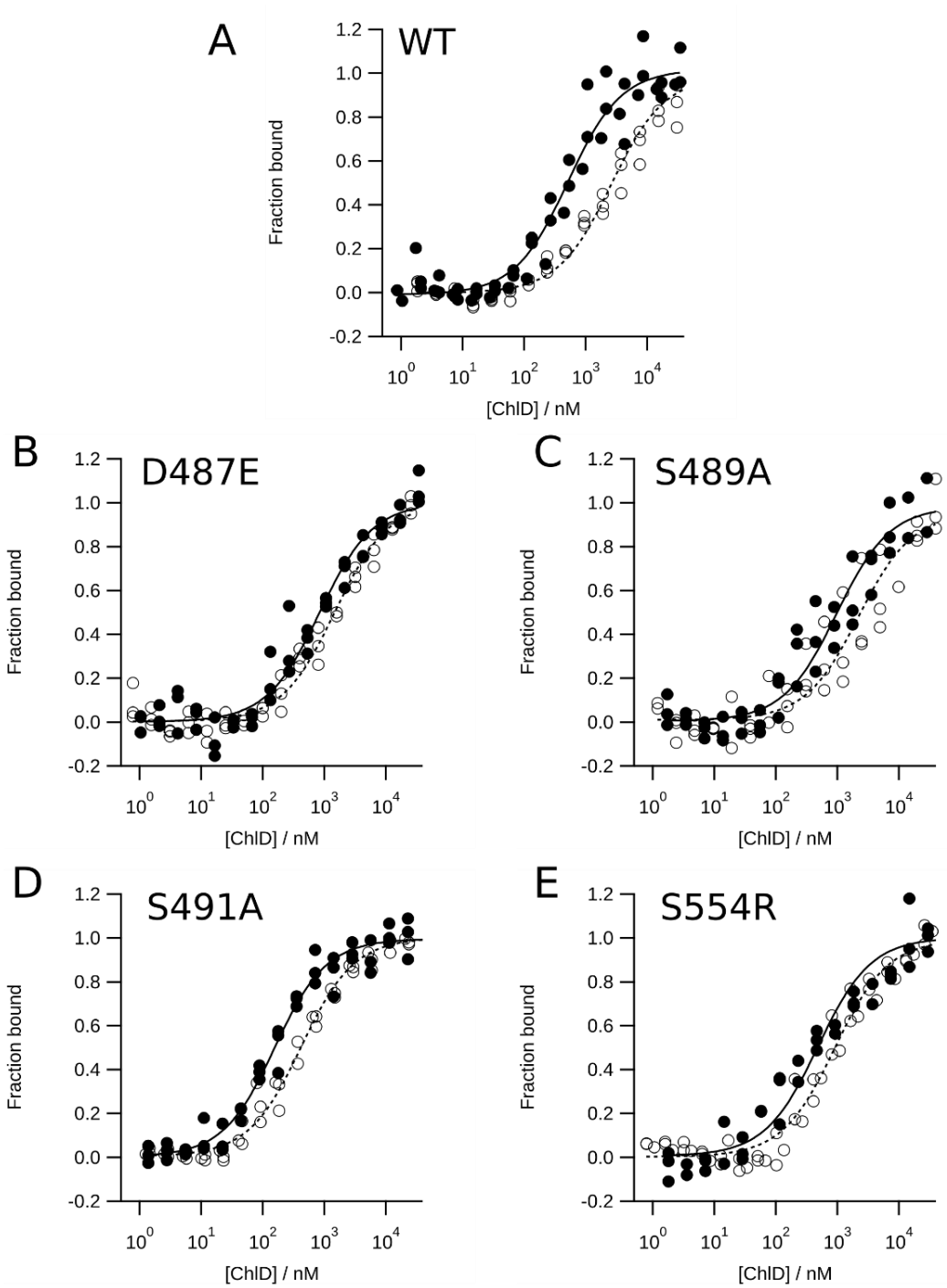


Figure 3.15. MIDAS mutants of ChlD continue to form a ChlD-ChlH complex. **A**, WT and **B-E**, MIDAS mutants. In the model structure, ChlD is represented as a cartoon with predicted Mg^{2+} binding residues represented as sticks. MST assays were performed in the presence (closed circles) and absence (open circles) of 10 mM Mg^{2+}. MST experiments were performed in triplicate from independent batches of protein. K_d values shown in Table 3.2.

Table 3.2. Dissociation constant (K_d) values of the ChlH-ChlD interaction for wild-type and all ChlD and ChlH mutants. Mutations are in ChlD unless stated within the table.

Mutant	Mutation type	Mg^{2+} present K_d / µM	Mg^{2+} free K_d / µM
Wild type		0.331 ± 0.05	2.10 ± 0.387
QuinE	Allosteric response to Mg^{2+}	0.339 ± 0.04	0.512 ± 0.161
Truncation A	N-terminal truncations	14.6 ± 2.99	5.93 ± 1.08
Truncation B		10.5 ± 1.90	6.76 ± 1.52
ChlH LD	ChlH LDV double mutants	0.546 ± 0.06	4.58 ± 0.408
ChlH DV		0.266 ± 0.05	3.32 ± 0.756
ChlH LV		0.129 ± 0.02	4.23 ± 0.784
D487E	Type I MIDAS mutants	0.842 ± 0.162	1.489 ± 0.221
S489A		1.98 ± 0.52	3.21 ± 0.76
S491A		0.147 ± 0.026	0.351 ± 0.097
S554R	Type II MIDAS mutant	0.481 ± 0.140	0.988 ± 0.126

3.5 Discussion

The work within this chapter presents a biochemical characterisation of the C-terminal integrin I-like domain of ChlD; there is limited previous work to describe the role this region plays in the catalysis carried out by MgCH. The D subunit of MgCH uniquely comprises an AAA+ module tethered to an integrin I unit, and an understanding of this arrangement may be crucial to discovering the mechanism of energy transduction from the ATPase activity of ChlI to the chelation site of ChlH.

Through *in vivo* and *in vitro* co-purification it was revealed that ChlH and ChlD interact, and with the aid of MST and XLMS techniques it was found that the C-terminal integrin I-like domain is the major contributor to the interaction. Mutagenesis of the conserved MIDAS motif in the C-terminal domain severely disrupts chelation, which can be somewhat saved by a replacement positive charge provided by an arginine residue. The precise role of the MIDAS motif still remains unclear, however, as the ChlH-ChlD interaction is only mildly affected by the mutations that impair or eliminate chelation activity.

Recent work by Brindley et al. (2015) discovered the amino acids responsible for the cooperativity of the *Synechocystis* magnesium chelatase, and the presented work builds upon that study, showing that magnesium binding by these residues regulates the strength of the interaction between ChlH and ChlD, conferring the cooperativity.

The work presented in this chapter provides a biochemical and kinetic counterpart to the results of the recent study by Luo et al. (2018), which showed that ChlH and ChlD interact in a manner that is not dependent on the C-terminus of ChlD. It has previously been shown that C-terminal truncations of ChlD from *Synechocystis* are not able to produce active chelatases (Adams et al., 2016b) and the results of the work presented here agree with that finding. It could be possible that MgCH from higher plants operate via a different mechanism but with a 51% sequence similarity between the ChlD subunits of *Synechocystis* and *Oryza sativa* (rice) this seems unlikely. Thorough biophysical characterisation of the protein-protein interactions in rice MgCH could further resolve the discrepancies.

4 Structural studies of ChlH

4.1 Introduction

The central metal atoms in the cyclic tetrapyrrole-derived cofactors, magnesium in chlorophyll, cobalt in cobalamin, iron in haem, and nickel in F_{430}, are collectively important for most living systems. The steps required to insert each metal ion into its cognate macrocyclic ring can require surprisingly complex and poorly understood enzymes. These metal ion chelatases can also play a regulatory role in directing and controlling flux down various branches of tetrapyrrole metabolism. A reasonably good structural, evolutionary, and mechanistic understanding exists for the relatively simple Class II chelatases (Brindley et al., 2003; Dailey et al., 2017; Hoggins et al., 2007; Medlock et al., 2007, 2009; Romão et al., 2011), but our knowledge of the complex multisubunit Class I chelatases, typified by magnesium chelatase, is more limited.

Much of the mechanistic work on this class of chelatases has focused on the more tractable magnesium chelatases from bacteriochlorophyll and chlorophyll producing organisms. The current mechanistic and structural data suggest a model for the MgCH mechanism where the two AAA+ subunits form a ChlID complex (Adams and Reid, 2013; Adams et al., 2016a). This complex transiently interacts with the ChlH protein (Chapter 3; Farmer et al., 2019), and hydrolyses ATP, which drives a conformational change in the ChlH-porphyrin complex that promotes insertion of the Mg^{2+} ion into the protoporphyrin IX (P_{IX}) ring (Reid and Hunter, 2002).

This sequence of events accounts for the kinetic data, but a complete model requires the structures of the H, I and D subunits, in particular the porphyrin binding subunit where insertion of Mg^{2+} takes place. SAXS, and electron microscopy of negatively stained ChlH, provided low resolution views of the ChlH protein, both free and bound to porphyrin (Qian et al., 2012), and the recently published crystallographic structure of apo-ChlH from *Synechocystis* at 2.5 Å represented a significant advance in structural studies of this subunit (Chen et al., 2015a), showing the domain-level organization of the protein in detail. ChlH protein can be viewed as consisting of five domains, the “head” (I), “neck” (II) and “body” (III - V); an internal pocket was also identified at the interface between domains III and V, postulated as the porphyrin binding site. Given the mechanistic importance of the porphyrin binding site, and the lack of direct structural or mechanistic evidence for its location, this Chapter reports crystallographic, cryo-EM, computational, kinetic

and mutagenesis approaches used to identify the active site of magnesium chelatase. Acknowledgements to Dr Claudine Bisson, who conducted the X-ray crystallography work in this chapter, Dr Amanda Brindley, who helped with molecular cloning, and Dr Nathan Adams, who helped with the modelling and kinetic work.

4.2 X-ray crystallography guided mutagenesis of ChlH

4.2.1 Locating part of the substrate binding site in ChlH

An inactive, stable truncated form of ChlH from *Thermosynechococcus elongatus* (T_ChlH_929) (residues D938-E1326) was designed based on a structural study of a corresponding fragment of the cobaltochelatase CobN subunit (Prof. M Warren, University of Kent, personal communication) and expressed in the presence of selenomethionine prior to crystallization. Experimentally derived phases were used to build a backbone model of the protein, which was subsequently used to determine a 1.6 Å resolution structure of the ChlH fragment. T_ChlH_929 consists of two lobes (Figure 4.1 **A**), which share a split seven-stranded beta-sheet that underlies a prominent surface groove (Figure 4.1 **B**).

T_ChlH_929 has 78 % sequence identity with the corresponding region of full-length ChlH from *Synechocystis* sp. PCC 6803 and comparison with this structure (PDB:4ZHJ) shows that T_ChlH_929 represents part of "body" domain, including one side of the putative active site (Chen et al., 2015a) that corresponds to the surface groove on T_ChlH_929. After refinement of the polypeptide structure, inspection of the difference density map revealed two peaks that could not be accounted for by the protein or by water molecules. Both peaks were positioned within the surface groove of T_ChlH_929, at the interface of two of the beta-strands and the N-terminal ends of two bent helices, which pack against them. One of the peaks corresponded to a 6-coordinate metal ion, which based on ligand distances and size of the peak in comparison to the surrounding protein, was modelled as a potassium ion. The ligands surrounding this site include three water molecules, the sidechain of K200 and the main chain amides of G111 and A174. In the structure of a similarly truncated form of ChlH from *A. thaliana* (PDB: 5EWU) (72% sequence identity to T_ChlH_929) (Figure 4.1 **C**), the same binding site is occupied by a magnesium ion.

The second difference peak (Figure 4.1 **D**) sits in a slight depression between the sidechains of W56, T58, S110 and V112, indicating the binding of a small molecule that either exogenously co-purified with the protein or that originated from the crystallization conditions. A molecule of benzoate occupies the same position in the *A. thaliana* ChlH truncate structure, with the carboxyl

group of the benzoate forming hydrogen bonds with the main chain amide of G1034 and T1035, and sidechains of T1035 and S1087 (equivalent to G57, T58 and S110, respectively, in T_ChlH_929), indicating the location of a putative carboxyl binding site. Despite the high resolution of the T_ChlH_929 structure, it was not possible to identify the bound ligand, which may be partially disordered. Nevertheless, the conservation of both the architecture of this region and the surrounding residues (Figure 4.1 **D**), suggests that the groove may be important for substrate binding. Thus, based on these observations, it is hypothesized that this is the binding site of one of the two pendant propionate groups of the ChlH substrate, P_{IX}, assigning this position as 'site 1'.

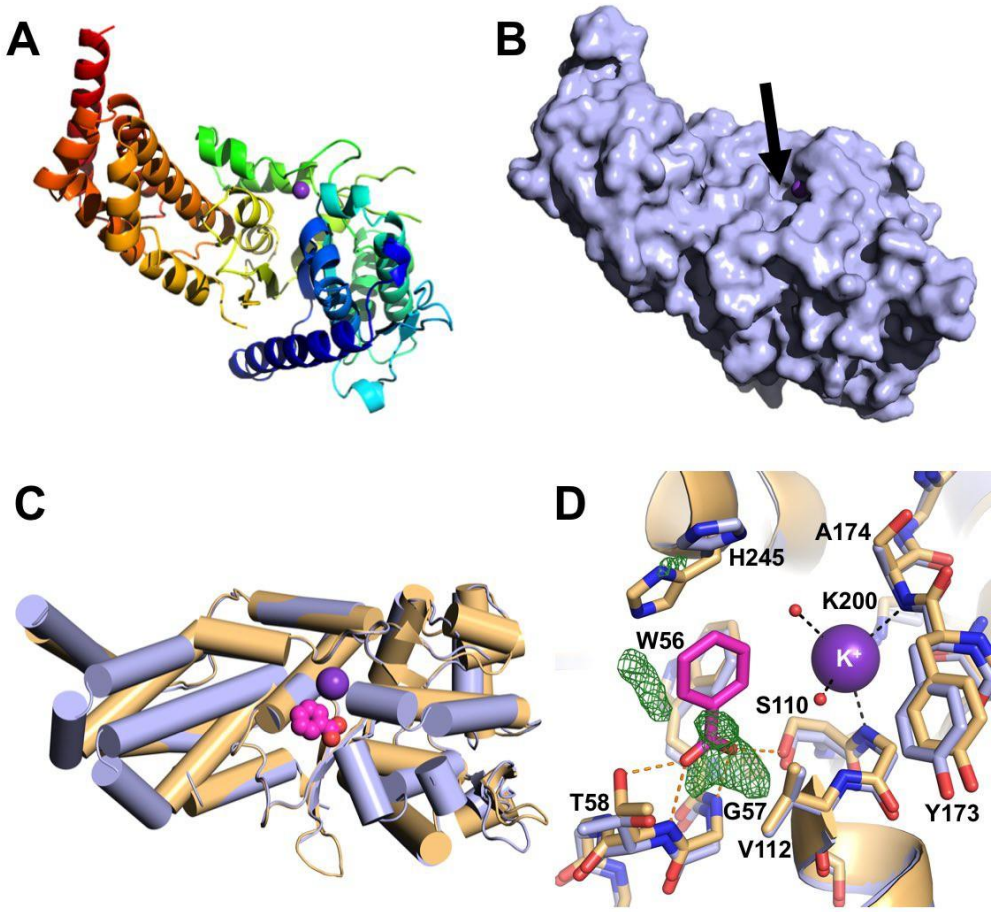


Figure 4.1. Crystal structure of the ChlH truncation. **A**, a cartoon representation of the *Thermosynechoccus elongatus* ChlH truncate, T_ChlH_929, coloured from the N-terminus (blue) to the C-terminus (red). The bound potassium ion is shown as a purple sphere. **B**, a surface representation (blue) of T_ChlH_929 in the same view, showing the surface groove (black arrow). **C**, a superposition of T_ChlH_929 (blue) with a similar truncate of ChlH from *A. thaliana* (PDB: 5EWU). The potassium ion from T_ChlH_929 is shown as a purple sphere and the benzoate molecule bound in *A. thaliana* ChlH is shown for reference (magenta spheres). Helices are represented as cylinders. **D**, a detailed view of the binding groove of T_ChlH_929 (blue) superimposed on that of *A. thaliana* ChlH showing the conservation of amino acids in this region and the interactions that they make with the ligands in each structure. Also shown is a section of the Fo-Fc difference map from the T_ChlH_929 structure (green mesh, contoured at 3 σ) indicating the binding of an unknown small molecule, the position of which corresponds to the benzoate in the *A. thaliana* ChlH structure, suggesting a conserved carboxyl site.

Having identified site 1 in the structure of the truncate, equivalent alanine mutations of the sidechains (Table 4.1, site 1) were made in full-length *Synechocystis* ChlH (Section 2.4, 2.6, 2.8.2) and assayed (Section2.13.2) for chelatase activity. Unsurprisingly, given the probability of multivalent interactions between the substrate and the enzyme at this site, including those to mainchain amides, the mutants remained active, enabling analysis of their kinetic profiles. While S1039A and V1041A had a higher k_{cat} than wild-type (WT) magnesium chelatase, the K_M was also increased, suggesting that the mutation significantly decreases the enzyme's affinity for D_{IX} (Figure 4.2 **D**, Table 4.1). Due to the complexity of the reaction taking place, and K_M in this system being a composite of many steps in the reaction, a better kinetic constant to evaluate the mutated enzymes is the apparent second-order rate constant (or specificity constant, k_{cat}/K_M) for the reaction of porphyrin substrate compared to wild-type protein. The S1039A mutant has an almost 3-fold decrease in efficiency of converting substrate to product.

The T987A mutant had almost double the activity of WT protein at 5 µM D_{IX}, however, it had a lower apparent K_M for D_{IX} than the wild-type enzyme. The kinetic profile (Figure 4.2 **D**) is reminiscent of substrate inhibition, but attempts to fit the data to a substrate inhibition equation produces physiologically implausible results (V_{max} = 3.6 µM min^{-1}, K_M = 32.4 µM, K_I = 4.9 µM). However, if T987 is involved in a hydrogen bonding network to one of the propionate groups of the porphyrin, then abolishing these interactions could conceivably allow for faster product release. In addition, the T987A mutant has a two-fold higher turnover rate compared to WT enzyme, close to the maximal steady state rates observed for WT magnesium chelatase in complex with the Gun4 accessory protein (Davison et al., 2005; Kopečná et al., 2015). Binding of Gun4 could potentially alter the conformation of ChlH to disengage this threonine from the substrate, suggesting a mechanism for Gun4-mediated product release.

While the other site 1 mutants appeared to increase k_{cat} while decreasing affinity for the substrate, S659A displays opposite behaviour, with half the k_{cat} of WT but also a 4x lower K_M value, which increases the specificity constant to double that of WT. This is a particularly interesting result as the S659 is located in a position to form a hydrogen bond with the carboxyl of the substrate in site 1 so one would predict that by disrupting this bond affinity for substrate would decrease, rather than increase. Perhaps disrupting the potential interaction here could allow for porphyrin to bind in an alternate conformation that is more stable; alternatively, the model may be incorrect.

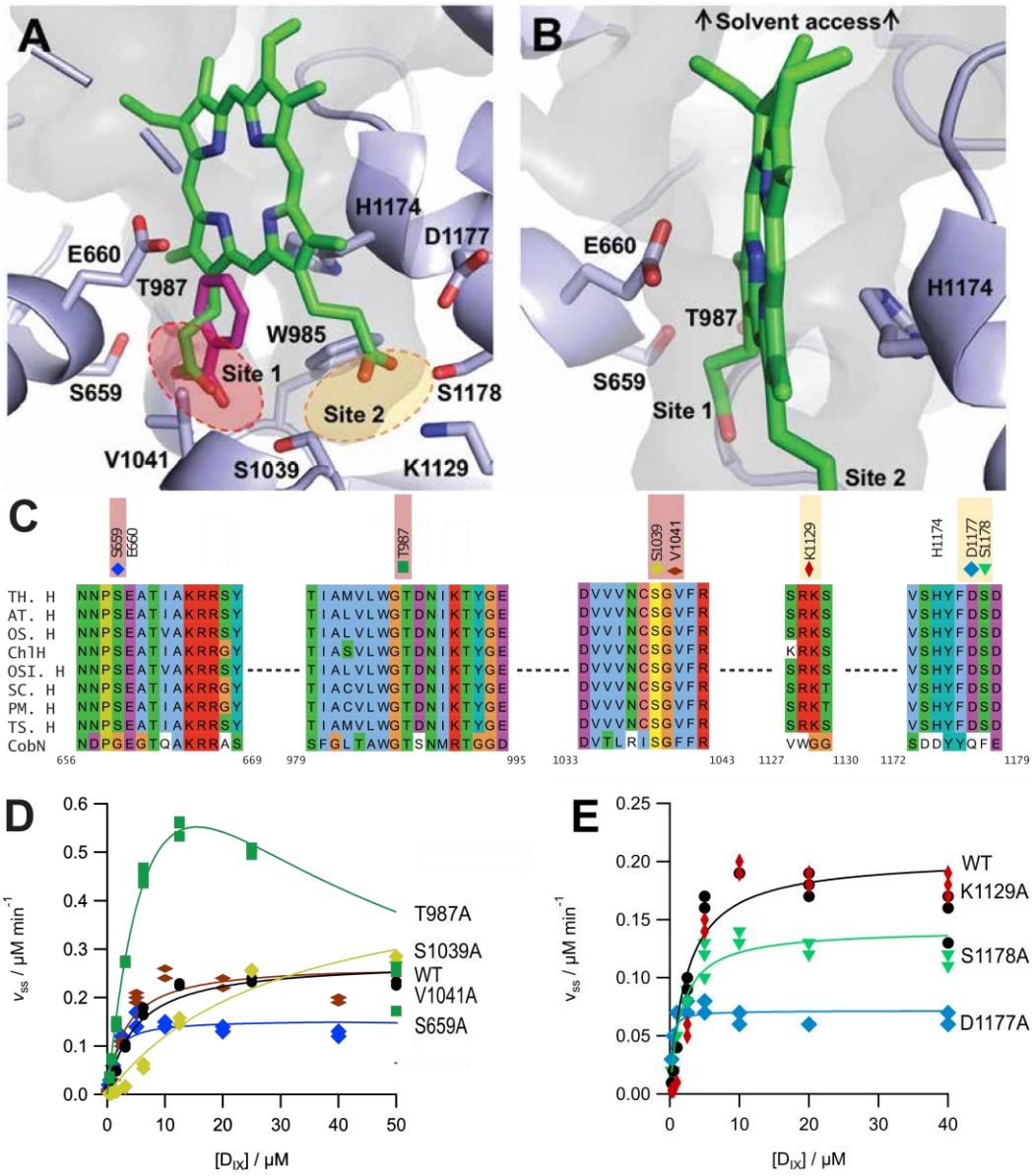


Figure 4.2. A model of P_{IX} binding in Syn_ChlH. **A**, P_{IX} (green sticks) sits within the bent disk-shaped central cavity of Syn_ChlH (partially transparent grey surface) with each of its two propionate groups sat in either site 1 or site 2, highlighted with a red or orange ellipse, respectively. The residues that surround each site are shown as blue sticks and the mainchain is represented as cartoon. **B**, viewed along the plane of the substrate, the positions of E660 and H1174 are shown on opposite sides of the porphyrin ring, which itself is slightly bent. The positions of flexible loop regions, which provide access to the active site, are indicated with black arrows. **C**, sequence alignment and secondary structure representations of the sequence highlighting the level of conservation amongst ChlH proteins and CobN for these core residues. TH. H, *Thermosynechococcus elongatus;* AT. H, *Arabidopsis thaliana;* OS. H, *Oryza sativa subsp. japonica*; ChlH, *Synechocystis sp.* 6903; SC. H, *Synechococcus sp*. (strain WH9102); PM. H, *Prochlorococcus marinus* (strain MIT 9301); TS. H, *Thermosynechococcus sp.* NK55a; CobN, *Brucella melitensis*. Numbering represents Syn_ChlH residue numbers. **D** and **E**, steady state magnesium chelatase assays with respect to [D_{IX}]. Assays conducted as in Section 2.13.2. Steady state rates were fitted to the Michaelis-Menten equation (Equation 2.3. The Michaelis Menton equationEquation 2.3) except T987A where the line is fitted to a substrate inhibition equation, but gives physiologically impossible fit coefficients. Fit coefficients are listed in Table 4.1.

4.2.2 Solving the full-length structure of Syn_ChlH for substrate binding

The assays confirmed that the propionate binding site formed between T987 and S1039 (site 1) seems to be important for substrate binding in ChlH. However, as P_{IX} has two pendant propionate groups it would be expected that a second propionate binding site would be present within the active site. Therefore, to get a global overview of the structure of the putative active site of ChlH and identify the second propionate binding site, WT ChlH from *Synechocystis* sp. PCC 6803 (Syn_ChlH) was overproduced, purified as in Section 2.6 and 2.8.2, then crystallised by Dr Amanda Brindley, the structure of which was determined to 2.5 Å by molecular replacement using a previously published 3 Å model of the enzyme (PDB:4ZHJ) (Chen et al., 2015a). The putative active site sits within a bent, disk-like cavity at the centre of the "body" domain of Syn_ChlH (Figure 4.3 **A**-**B**). The structure shows that the "body" is actually comprised of two halves which are connected by a loop region (A942 - T950) (Figure 4.3 **D**). The C-terminal portion comprises the part of the structure represented by T_ChlH_929 and the N-terminal portion of the "body" domain generates a C-shaped domain that cradles the C-terminal portion of the "body" within a cleft (Figure 4.3 **C**). The active site cavity sits at the interface between the two halves of the "body" and is enclosed on all sides, with several loops forming an access point to the external solvent at the top of the cavity. These loops have varying degrees of order across the multiple copies of Syn_ChlH within the asymmetric unit of the structures, suggesting that they are flexible. This implies that they may adopt different conformations in the presence and absence of substrate and, by extension, that the conformation modelled in the apo-crystal structure may not be equivalent to the conformation of these loops when Syn_ChlH is in complex with the substrate.

In the structure of full-length Syn_ChlH, as in T_ChlH_929, site 1 also contains a significant (5 sigma) buried electron density feature, suggesting a bound ligand. Although the 2.5 Å resolution of the structure prevents direct identification of the ligand, it is likely to be partially ordered citrate from the crystallization conditions (2.4 M Na citrate). Numerous attempts to crystallize Syn_ChlH with various substrate analogues, both by soaks and co-crystallization, in these and alternate conditions, were unsuccessful. It is therefore likely that the presence of citrate stabilizes a conformation that only crystallizes in this particular crystal form. ChlH is known to undergo a conformational change upon substrate binding (Qian et al., 2012) and comparison of the conformation of Syn_ChlH in this crystal form with the previously determined low resolution EM map does not show which conformation is represented by the crystal structure.

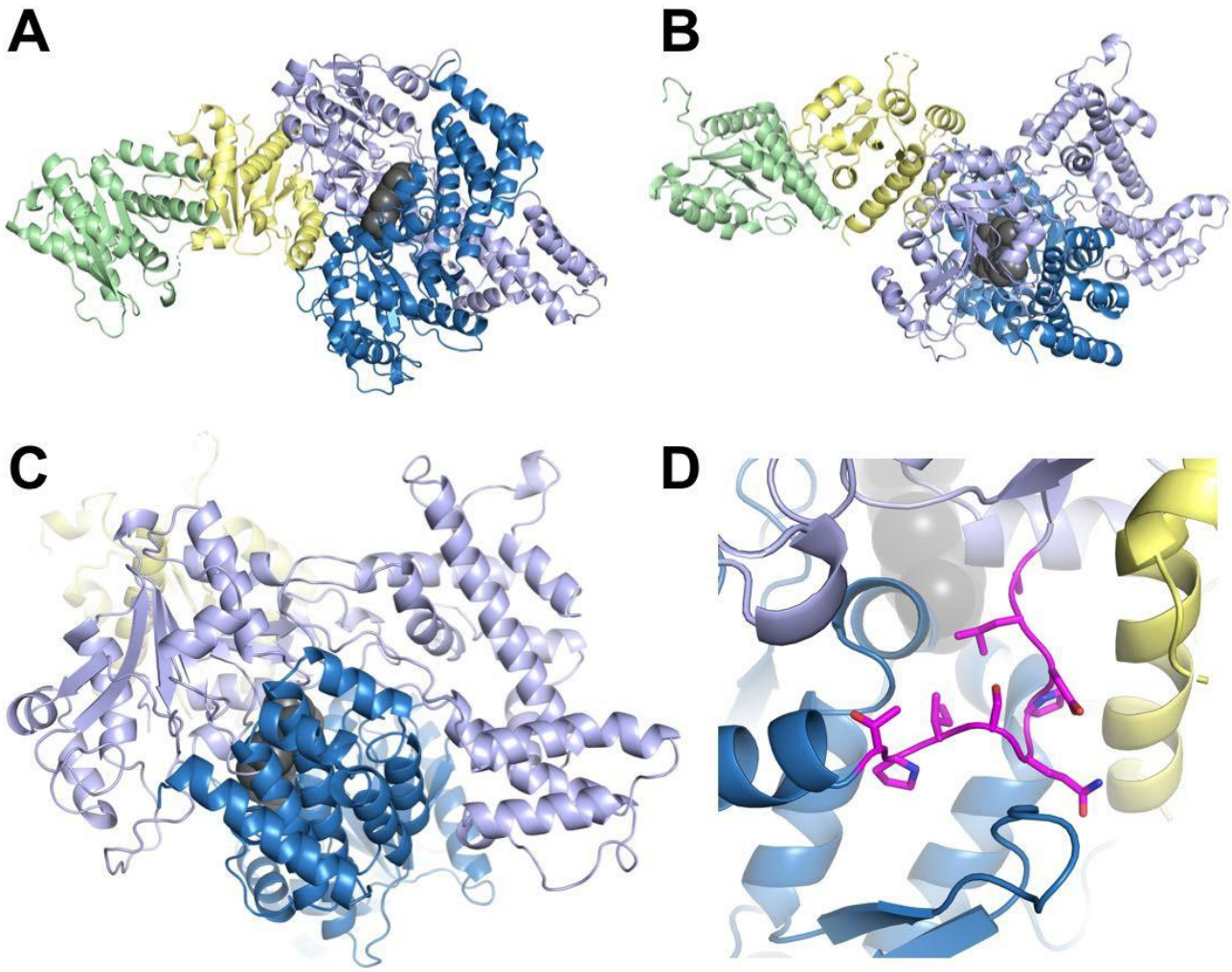


Figure 4.3. **A** and **B**, the three-dimensional fold of Syn_ChlH represented as a cartoon. The "head" and "neck" domains are coloured green and yellow, respectively. The "body" domain is coloured according to the C-terminal portion (dark blue), which is equivalent to T_ChlH_929, and the N-terminal portion (light blue), which forms a C-shaped domain that cradles the C-terminal portion (highlighted in **C**). The active site, indicated with a dark grey volume, is at the interface of the two halves of the "body" domain. **D**, a loop (magenta) connects the two halves of the "body" domain (loop sequence: A942-LDPQSIP-T950).

Next, the active site of Syn_ChlH was probed for P_{IX} binding by computational modelling using FlexAID and AutoDock Vina (Gaudreault and Najmanovich, 2015; Trott and Olson, 2009). The top 5 hits produced from FlexAid place the C ring propionate group of P_{IX} in site 1, identified from the T_ChlH_929 truncate structure (Figure 4.2 **A**-**B**). The computational modelling identified a second propionate binding site (site 2) within the central cavity of Syn_ChlH, positioned between the side chains of S1178, K1129 and N1097 and main chain amide of S1103. Site 2 is within a clear pocket of space at the base of active site cavity, thus a series of alanine mutants was made, to probe the importance of this site.

Whereas the mutants in site 1 mostly increase the K_M of D_{IX} binding, S1178A in site 2 appears to decrease the K_M for D_{IX}. This does not lead to a marked effect on the specificity constant k_{cat}/K_M, as the values for k_{cat} also decrease. However, a similar effect observed with the S659A site 1 mutant also occurs in mutant D1177A but to a greater extent. D1177A has the most decreased turnover rate of 0.17 min^{-1} compared to 0.69 min^{-1} for WT but the value for K_M is 0.2 μM, an order of magnitude lower than WT (4.8 μM), leading to a vastly increased apparent second order rate constant. The D1177A mutant is interesting, as it shows particularly tight affinity for the substrate, leading to a six-fold increase in catalytic efficiency.

The K1129A mutant was unusual in that the steady state rates are similar to WT individually, but when looked at as a function of [D_{IX}] they display what appears to be a cooperative response to D_{IX}, which seems implausible considering the size of the active site and the location of the mutant. Ideally, this assay would be repeated to confirm that the cooperativity is real and not a result of experimental error but time constraints prevented this.

Table 4.1. List of site 1 and 2 mutants with kinetic parameters.

Protein	Binding site	k_{cat} / min^{-1}	K_M / μM	k_{cat}/K_M / min^{-1} μM^{-1}
WT	N/A	0.69 ± 0.03	4.8 ± 0.8	0.14 ± 0.02
S1039A	1	1.19 ± 0.17	29.9 ± 8.4	0.04 ± 0.01
V1041A	1	1.26 ± 0.32	9.24 ± 4.0	0.13 ± 0.07
T987A	1	N/A	N/A	N/A
S659A	1	0.37 ± 0.02	1.20 ± 0.2	0.32 ± 0.07
S1178A	2	0.29 ± 0.07	1.45 ± 0.43	0.20 ± 0.08
K1129A	2	N/A	N/A	N/A
D1177A	2	0.17 ± 0.01	0.20 ± 0.09	0.87 ± 0.39

4.2.3 Locating critical catalytic residues in the active site of Syn_ChlH

Mutating key residues surrounding the two putative propionate binding sites showed that they fitted the kinetic profiles associated with reduced efficiency of substrate binding, providing evidence that the computational model of substrate binding in Syn_ChlH is biologically relevant. With the two propionate groups of the substrate docked at the base of the active site, the slightly bent porphyrin ring of the substrate is sandwiched closely between the two walls of the cavity. On one wall H1174 sits adjacent to the centre of the porphyrin ring and on the opposite wall, projecting toward the opposite face of the porphyrin ring, is a strictly conserved glutamate residue (E660). This generates a Glu-His pair that form an axis (Glu-O to His-N = 6 Å) near to the magnesium binding site of the porphyrin ring.

An axial histidine would be well placed to either interact with the central Mg^{2+} ion of the product, or to act as a base that abstracts protons from the porphyrin substrate, and histidine residues have such a role in class 2 chelatases such as CbiK (an anaerobic cobalt chelatase from *Salmonella typhimurium*; (Schubert et al., 1999) and *Bacillus subtilis*), and human and *Saccharomyces cerevisiae* protoporphyrin IX ferrochelatases (Al-Karadaghi et al., 1997; Karlberg et al., 2002; Wu et al., 2001). H1174 is conserved across the entire magnesium chelatase family (Figure 4.2 **C**). In the closely related late stage cobalt insertion chelatase, CobN, this histidine is replaced by an aspartate, perhaps reflecting the difference in metal specificity in the two enzymes. Mutation of H1174 to a valine in full-length Syn_ChlH showed no decrease in chelatase activity in the enzyme with respect to concentrations of D_{IX} and magnesium (Figure 4.4 **A-B,** Table 4.2), demonstrating that it is not critical for catalysis. Additionally, as a valine sidechain cannot act as a hydrogen bond donor or acceptor, any water molecules in the vicinity that could act as a general base in the reaction must be positioned in a separate hydrogen bonding network. Density functional theory calculations for Mg^{2+} insertion into porphyrin show a steep increase in the energetic challenge of deprotonation of the porphyrin ring by a water as opposed to a histidine base (Shen and Ryde, 2005). We conclude, therefore, that H1174 is not catalytic, but as it is in a convenient position to interact with the central magnesium it may have a stabilizing role during the reaction, but more evidence is required to confirm this.

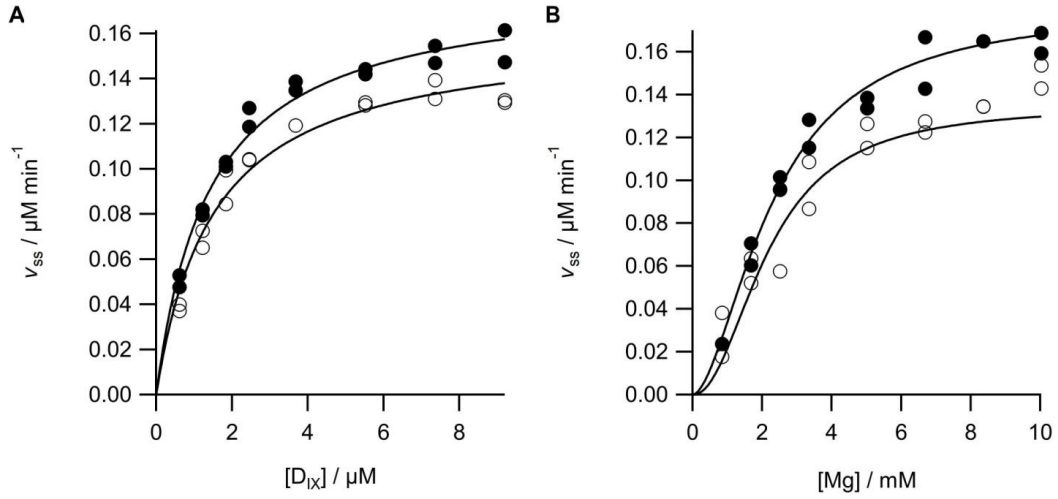


Figure 4.4. **A**, D_{IX} and **B**, Mg^{2+} dependence of the steady state rate of Mg^{2+} chelation for ChlH WT (closed circles) and H1174V (open circles). Assays conducted as in Section 2.13.2. Lines are theoretical, where steady state rates (v_{ss}) were fitted to the Michaelis-Menten equation (Equation 2.3) in panel **A** or the Hill equation (Equation 2.5) in panel **B**. Kinetic coefficients are reported in Table 4.2.

Table 4.2. Steady state parameters of ChlH WT and H1174V proteins.

Protein	k_{cat} / min^{-1}		K_M / μM	$S_{0.5}$ / mM	k_{cat}/K_M / min^{-1} μM^{-1}	k_{cat} / $S_{0.5}$	*n*
	D_{IX}	Mg^{2+}	D_{IX}	Mg^{2+}	D_{IX}	Mg^{2+}	Mg^{2+}
WT	0.45 ± 0.01	0.45 ± 0.03	1.42 ± 0.12	2.32 ± 0.21	0.32 ± 0.03	0.19 ± 0.02	1.7 ± 0.3
H1174V	0.42 ± 0.01	0.42 ± 0.06	1.52 ± 0.17	2.76 ± 0.68	0.28 ± 0.03	0.15 ± 0.04	1.4 ± 0.3

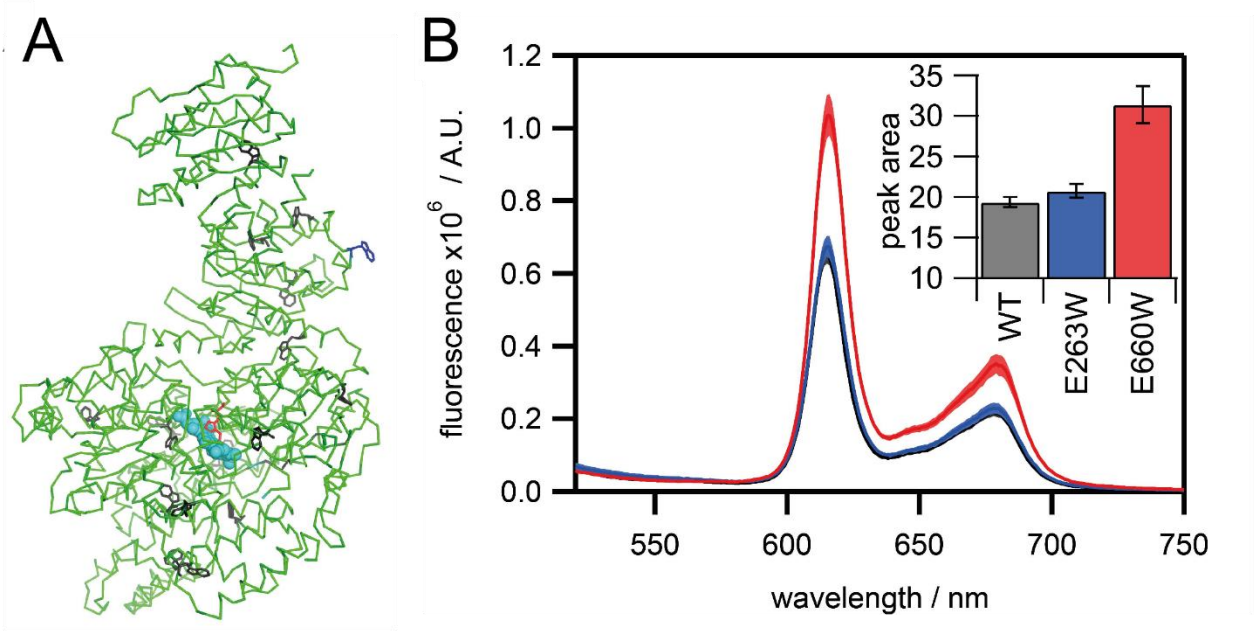


Figure 4.5. Observing energy transfer from tryptophan residues to porphyrin in the active site. **A,** the structure of ChlH represented as green ribbon with all native tryptophan residues represented as black sticks. The control E263W is modelled as blue sticks, and active site tryptophan E660W is shown as red sticks, with protoporphyrin docked in the active site represented as cyan spheres. **B,** fluorescence emission of D_{IX} after excitation of tryptophan residues at 280 nm. Black, WT ChlH; blue, E263W; Red, E660W. Averages of three independent biological repeats are represented with a central solid line and shading representing standard deviation.

Residue E660 is in close proximity to the modelled porphyrin substrate on the opposite side to H1174. To support the proposed location of adjacent to the substrate, E660 was mutated to a tryptophan (E660W) and Förster Resonance Energy Transfer (FRET) assays were carried out, monitoring emission from D_{IX} following excitation of tryptophan at 280 nm (Viney et al., 2007). The D_{IX} fluorescence emission from the E660W mutant was 47 % greater than from the wild-type protein, while a control mutation, E263W, where the tryptophan is ca. 50 Å from the proposed porphyrin binding site, only showed a 7 % increase in fluorescence (Figure 4.5). Thus, it was concluded that the substrate binds in close proximity to E660W which, being buried in the core of the body of the protein, provides direct evidence that the porphyrin binds within the central cavity, adjacent to

E660 in wild-type Syn_ChlH. In addition, it was observed that the active site cavity of Syn_ChlH is clearly bent, such that a bowed conformation of the substrate fits within the cavity better than a planar molecule. Non-planar porphyrins have been observed in several crystal structures of unrelated proteins, including PDB:4UC1 and PDB:5DUO (*Rhodobacter sphaeroides* translocator proteins) and tryptophan fluorescence quenching experiments have directly suggested binding of a slightly distorted non-planar conformation of substrate in ChlH (Karger et al., 2001). In Syn_ChlH, the convex face of the porphyrin ring points towards E660, presenting the four central nitrogen atoms of the porphyrin to the carboxyl side chain, which suggests that bowing of the substrate within the Syn_ChlH cavity might facilitate metal insertion.

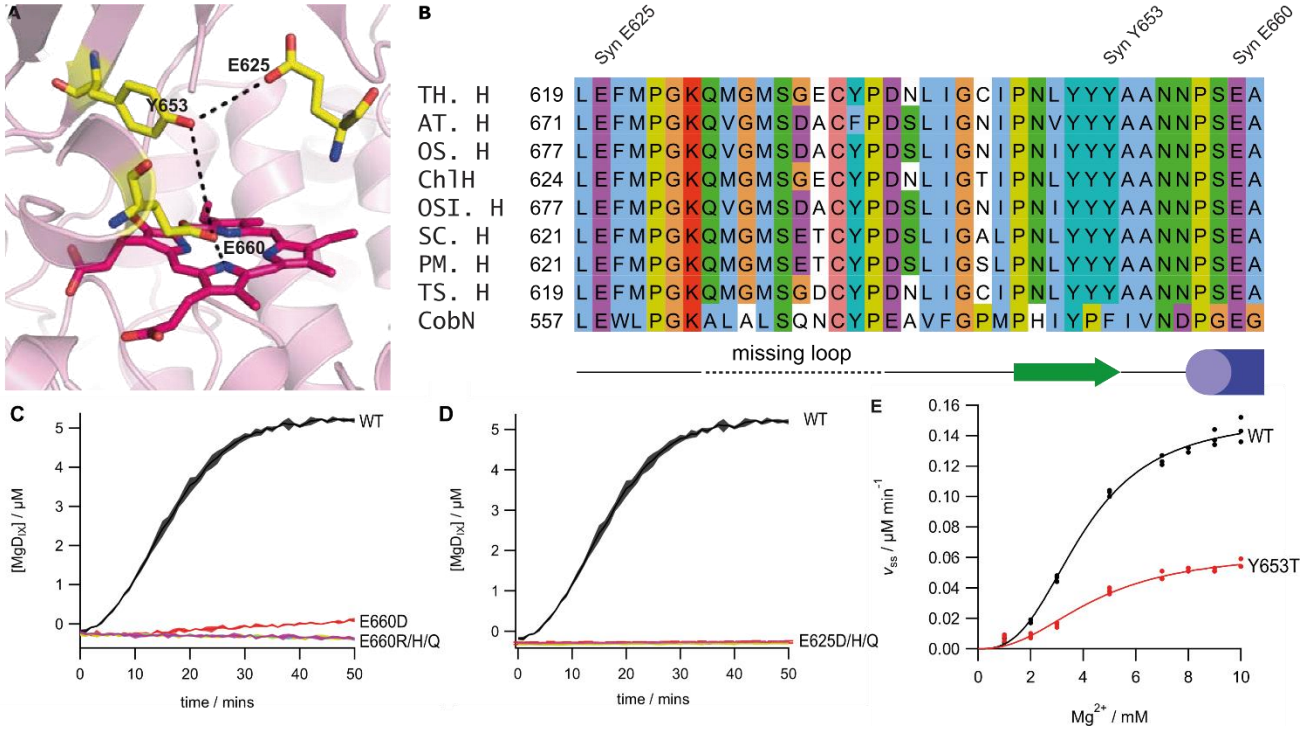


Figure 4.6. Conserved negatively charged residues within ChlH. **A**, view of the potential metal ion delivery residues, with important residues highlighted in yellow sticks, with ChlH represented as cartoon and docked PIX in magenta. **B**, sequence alignment of ChlH and CobN, naming as in Figure 4.2. **C**, **D**, Time courses of MgCH assays with WT, E660 and E625 mutants of ChlH; assays were performed in triplicate and are displayed as the mean with stand deviations shown in shading. Only E660D had any activity, v_{ss} = 0.012 ± 0.001; 5 % of WT v_{ss} = 0.230 ± 0.009. **E**, Mg^{2+} dependence of the steady state rate of Mg^{2+} chelation for WT and Y653T ChlH, lines are theoretical and described by Equation 2.5, with parameters: WT, k_{cat}^{Mg} = 0.38 ± 0.01 µM min^{-1}, $S_{0.5}^{Mg}$ = 3.99 ± 0.12, n = 2.8 ± 0.2 µM; Y653T, k_{cat}^{Mg} = 0.16 ± 0.01 µM min^{-1}, $S_{0.5}^{Mg}$ = 4.41 ± 0.38, n = 2.4 ± 0.3 µM.

The position of E660 on the other side of the porphyrin ring to H1174 suggests that it could also interact with the central magnesium ion during catalysis. To probe this, several mutations were made: E660R, E660H, E660D and E660Q. Assays showed (Figure 4.6 **C**) that only by retaining a

carboxyl group at position 660, the glutamate to aspartate mutation, could any activity be measured (*ca.* 5 % compared to wild-type). This suggests that the carboxyl group of E660 is essential for catalysis in Syn_ChlH and its position, adjacent to the centre of the porphyrin ring, also suggests that E660 has a role in inserting the Mg^{2+} ion into P_{IX}.

4.2.4 Insights into magnesium ion binding in ChlH

The involvement of E660 in chelation of magnesium by P_{IX} requires a delivery mechanism to supply magnesium ions to the active site. The resolution of the full-length Syn_ChlH structure is not sufficient to directly observe bound magnesium ions, a problem compounded by the complexity of the difference map features in the active site arising from non-specific binding of citrate. However, examination of the section of difference map adjacent to E660 reveals an intriguing chain of four features sitting to one side of the active site, indicating a solvent-filled buried channel that leads from the surface of the protein to the active site next to E660.

The closest peak in the chain to the active site is ~4 Å away from the nearest E660 oxygen atom. The distance from this peak to the next in the chain is 3 Å, increasing to 3.7 Å to the next peak and finally, 5.2 Å to the final peak. The amino acids surrounding these features include two tyrosine residues (Y653 and Y651), which are stacked on top of each other to one side of the putative channel, and a glutamate (E625), which is positioned between two phenylalanine residues (F459 and F626). A number of main chain carbonyls and also main chain amides also project into the solvent channel. To explore whether any of these residues plays a role in metal binding, residue E625 was mutated to form E625D, E625H and E625Q enzyme variants, producing normally folded protein with abolished chelatase activity (Figure 4.6 **D**). Due to this complete loss of activity, it is difficult to be certain what mechanistic role E625 plays in chelation, but as the residue points away from the active site and is ~8 Å away from E660 (closest O to O distance), E625 is unlikely to directly interact with the substrate. Given that the E625D mutant is also completely inactive, this suggests that there is no tolerance for even a slight movement in the position of the carboxyl group. Mutation of Y653 to a phenylalanine reduces k_{cat}^{Mg} by ~40 % (Figure 4.6 **E**), suggesting that Y653 is not essential for supplying magnesium to the active site, but may be involved in the process. The clear involvement of E625 and E660 is, however, consistent with the notion of vectoral delivery of Mg^{2+} to the porphyrin substrate; more work is required to ascertain the critical role of E625.

4.3 Electron microscopy of ChlH

While ligand modelling that guided mutagenesis, as well as tryptophan FRET assays, resulted in compelling evidence that the porphyrin binding site is in the proposed location, some structural information for a ChlH-porphyrin complex was still desired. Conformational changes within ChlH upon ligand binding have been observed previously in a low resolution EM study of negatively stained ChlH (Qian et al., 2012), and it is not clear which conformation ChlH is in from the X-ray crystallography structure. ChlH has only crystallised under specific conditions that are quite different from the physiological environment, so observing the structure of the molecule under more biologically relevant conditions could also reveal information that has been lost by using crystallography. It would be important to investigate both substrate bound and substrate free forms of ChlH.

As ChlH is 150 kDa, monomeric ChlH is a small molecule to study using cryo-EM; larger multimeric complexes or bigger molecules like ribosomes and virus particles are more frequently investigated using this technique. However, there have been successful reports of using cryo-EM to elucidate information about proteins and complexes of this size (Merk et al., 2016).

4.3.1 Initial data collection

ChlH was purified as in Section 2.8.2. A highly concentrated D_{IX} solution was made as in Section 2.13.2.1 and incubated with half of the purified ChlH sample at 34 °C for 10 min. The ChlH-D_{IX} solution was then purified by SEC on a HPLC system as in Section 2.8.2.3 so that absorbance and fluorescence of both protein and porphyrin could be observed, and so that excess porphyrin was removed. HPLC SEC was also performed on ChlH with no porphyrin.

Table 4.3. Imaging parameters for automated data collection of ChlH-D_{IX}.

Parameter	**Value**
Magnification	120 000
Voltage (kV)	200
Electron exposure (e^-/ Å^2)	1.4 per frame, (56 e^- on 40 frames)
Defocus range (µm)	-1.0 to -3.0
Pixel size (Å)	1.15

ChlH-D_{IX} sample was applied to Quantifoil 1.2/1.3 Cu 400 mesh grids, plunge frozen as in Section 2.11.2 and preliminary automated data collection was performed using the EPU software (ThermoFisher) on a Talos Arctica 200 kV microscope equipped with a Falcon 3EC direct electron detector. For imaging conditions, see Table 4.3.

Dose-fractionated motion correction of collected images was performed using MotionCor2 (Zheng et al., 2017). The CTF of images was estimated using CTFFIND4.1 (Rohou and Grigorieff, 2015). Further processing was conducted in RELION 2.0 (Scheres, 2016). Particles were picked using the autopicking software within RELION 2.0 and reference-free 2D classification of the particles was performed. A preferred orientation issue was revealed; most particles were positioned in the ice such that only one view of the molecule was present (Figure 4.7). With only a limited selection of orientations of particles, accurate map generation becomes impossible.

While the data was not sufficient to generate a 3D map, the classes that were generated were ~ 20 Å shorter in length than the crystal structures for ChlH. This could be due to a conformational change upon binding D_{IX}. It was therefore essential to also characterise apo ChlH by cryo-EM.

Figure 4.7. Best 2D classes generated from 3858 particles after data collection on a Talos Arctica 200 kV microscope. While promising secondary structure features are present, the limited variation in orientation of the classes hampered further data processing steps. The particles are only ~100 Å in length, ~20 Å shorter than the crystal structure.

4.3.2 Optimising grid preparation for ChlH

As discussed above, preferred particle orientation within the ice was an issue with initial grid preparation. However, subsequent attempts to produce grids of both ChlH-D_{IX} and apo-ChlH proved more challenging; the preparations resulted in aggregation of molecules within the ice and on the carbon layer of the grid, with no observable monodisperse particles. Several strategies for grid preparation were conducted to address the issue: varying grid properties, multi-blotting sample and addition of detergent, as suggested by Drulyte *et al.* (2018).

The use of gold grids in cryo-EM has been explored as a more stable alternative to carbon-coated copper grids; the stability reduces motion of the sample upon exposure to the electron beam (Russo and Passmore, 2016). In addition to this, gold grids can be functionalised using thiol chemistry, as demonstrated by Meyerson et al. (2014). Using this chemistry, it is possible to control the surface environment, which can encourage monodisperse protein particles within the ice layer (Li et al., 2017). PEGylated UltrAuFoil (Quantifoil) grids were produced following the methods in the study by Meyerson et al. (2014) and both apo-ChlH and ChlH-D_{IX} applied before blotting and plunge freezing (Section 2.11.2). These grids were then imaged on a Technai Arctica 200 kV microscope equipped with a Falcon 3EC detector. Unfortunately, this method appeared not to resolve the issue and no particles were observed within the ice.

The plunge-freezing process for preparing cryo-grids can be particularly harsh on biological samples as once the loaded sample is blotted, the water film has a very high surface area to volume ratio. This means that even in the short time taken to go from blotting to freezing, samples can collide with the air-water interface thousands of times (Drulyte et al., 2018). Air is hydrophobic, so these collisions can promote unfolding of proteins, potentially leading to aggregated states observable on the microscope. It has been suggested that one way to lessen the detrimental effect of plunge-freezing grids is to add surfactants to the sample, even for soluble protein complexes (Glaeser et al., 2016). To see if the addition of a surfactant would aid in resolving the aggregation of ChlH upon freezing, β-DDM was added to the apo-ChlH and ChlH-D_{IX} solution a final concentration of 0.02% (w/v) and grid preparation using Quantifoil 1.2/1.3 Cu 400 mesh grids conducted as before (Section 2.11.2). The grids were screened on a Technai Arctica as above and again, no particles were observed.

As the sample clearly had a preference for the carbon layer on the surface of the grid, rather than in the ice-filled holes that are used for imaging, Dr Svetomir Tzokov suggested multiple blotting and loading steps as conducted by Snijder et al. (2017). The repeated addition and blotting of sample on the grid was proposed to saturate the carbon with protein so that protein would be forced into the holes. Apo-ChlH and ChlH-D_{IX} were loaded independently onto grids, blotted manually with filter paper, then more sample was quickly added. This was repeated so that 15 µl of sample was added in total (5 µl per loading) and upon final addition the grid was blotted and plunge frozen as in Section 2.11.2. The grids were then imaged on a Technai Arctica as above, where particles were observed within the ice (Figure 4.8). While particles were within the ice, there was still an obvious preference for the edge of the holes, where multiple molecules gathered together. The

ChlH-D_{IX} grid had better particle distribution, but the ice of the hole imaged was non-vitreous as indicated by the mottled appearance of the background.

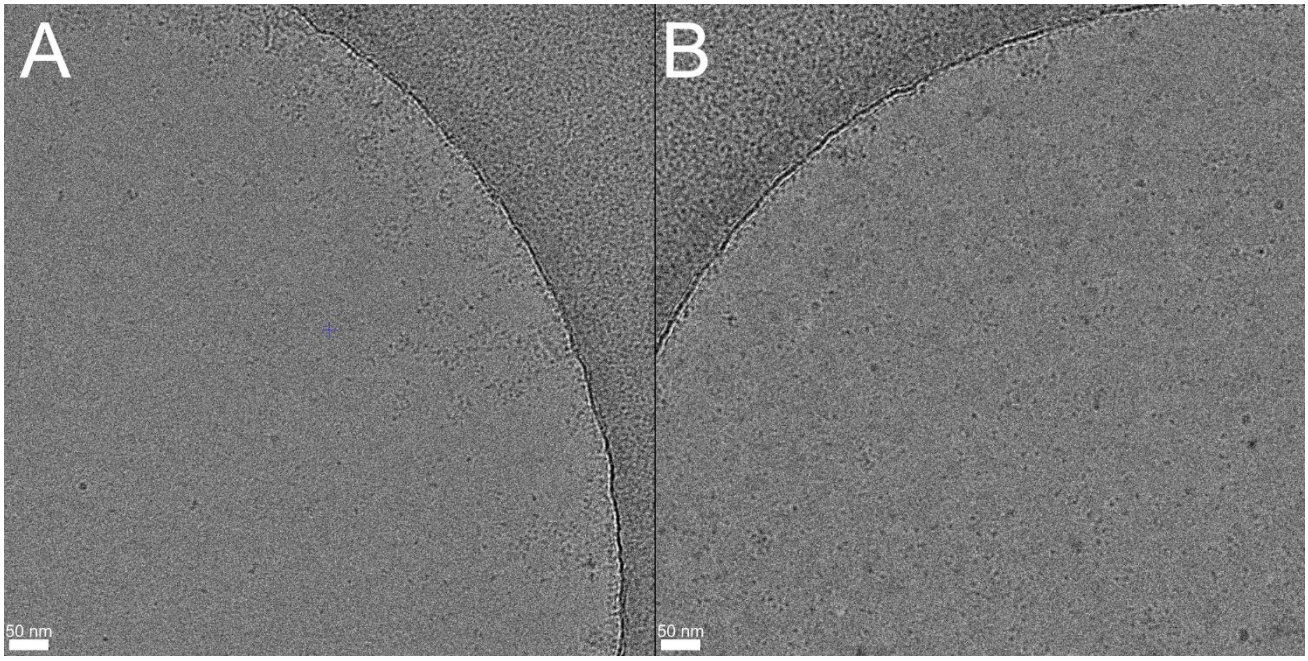


Figure 4.8. Micrographs of apo-ChlH (**A**) and ChlH-D_{IX} (**B**) produced using the multiple loading method. Particles can be seen within the hole, but tend to cluster together near the edges, especially for apo-ChlH. The best ChlH-D_{IX} grid prepared in this method has good particle distribution, but non-vitreous ice (indicated by the mottling of the hole). The presence of non-vitreous ice can disrupt protein structure and make data processing more challenging, but while it is present in this location on the grid, elsewhere may be more suitable for data collection. Imaged at -3 μm defocus.

Altering the grid properties to improve affinity of the protein for the hole was desired, as overloading the carbon was not completely successful. One common way of overcoming sample affinity for carbon is to use grids with an ultra-thin layer of carbon over the holes. However, this comes at the cost of reduced contrast, an issue not too pressing for large complexes, but incompatible when studying a protein as small as monomeric ChlH. An alternative to carbon coating is to use the much more electron-transparent graphene oxide. Coating holey-carbon copper grids with ultra-thin, ideally only one molecule thick, graphene oxide has been found to improve particle distribution within the ice with minimal contribution to background noise (Palovcak et al., 2018). Graphene oxide monolayers are deposited onto the grids and lay over the holes, providing a support for molecules to bind to and promoting the localisation of the protein to the holes.

Graphene oxide grids were produced following the protocol used by Palovcak *et al.* (2018) whereby graphene oxide films are lowered onto glow discharge treated Quantifoil 1.2/1.3 400

mesh Cu grids by pumping out a subphase solution. Sample was loaded onto these grids after a period of drying and plunge frozen as previously (Section 2.11.2). The grids were imaged on a Technai Arctica as before and the presence of graphene oxide assessed by imaging in the diffraction plane and looking for the typical hexagonal lattice of graphene oxide monolayers. It appeared that there was limited coverage of graphene oxide; despite this, monodisperse molecules of ChlH appeared within the ice (Figure 4.9). The presence of particles even in the absence of graphene oxide could be explained by the altered hydrophobicity of the grid – grids were not glow discharged immediately before sample application as the glow discharge process can disrupt the monolayer of graphene oxide. Perhaps the increased hydrophobicity of the carbon edges decreased the affinity of the ChlH molecules for the carbon edges. Unfortunately, this was not the case for the apo-ChlH so it is unclear how reliable this effect is.

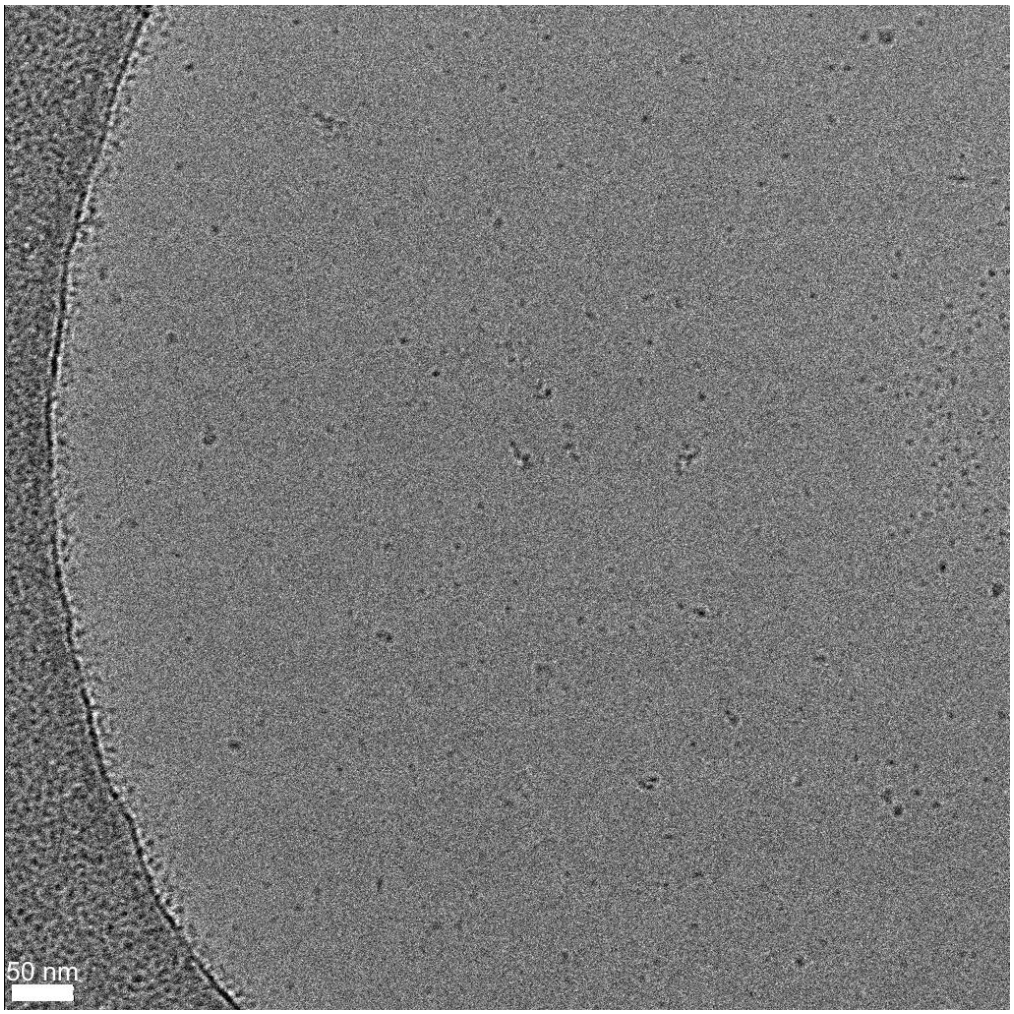


Figure 4.9. Micrograph of ChlH-D_{IX} applied to homemade graphene oxide grid. While there is no evidence for the monolayer of graphene oxide, particles appear within the hole. Imaged at -3 μm defocus.

The graphene oxide grid was deemed suitable for data collection so it and, due to time pressures limiting further optimisations, the previous two grids prepared by the multi-loading technique were sent to eBIC (Harwell Campus) for data collection on a 300 kV Titan Krios equipped with a K3 detector (Gatan) operating in super-resolution mode with a Volta phase-plate. The phase-plate was used on the initial data collection due to concerns about lack of contrast.

Unfortunately, automated data collection on the graphene oxide grid encountered a technical error and the grid squares were exhausted without producing any suitable images for data processing. The grids produced by the multi-loading technique were then imaged automatically using the EPU software (ThermoFisher), in order to get some data for processing (for imaging conditions, see Table 4.4). Both of these grids appeared to have been damaged at some point during transit or handling and whilst data was collected, the micrographs were of poor quality – there was significant ice contamination and non-vitreous ice present.

Table 4.4. Imaging parameters for automated data collection on ChlH grids produced by the multi-loading technique.

Parameter	**ChlH-D_{IX}**	**Apo-ChlH**
Nominal magnification	92000	92000
Voltage (kV)	300	300
Electron exposure (e^-/ Å^2)	1.0 per frame (50 total dose)	1.0 per frame (50 total dose)
Defocus range (μm)	-2 to -3.2 (300 nm steps)	-2 to -3.2 (300 nm steps)
Pixel size (Å)	0.42 (super resolution)	0.42 (super resolution)

4.3.3 ChlH-D_{IX}

In spite of the low quality data, processing was conducted: micrograph movies were corrected for beam-induced motion and dose-weighted with MotionCor2; CTF estimation was performed with GCTF; and further processing within RELION 3.0 (Zhang, 2016; Zheng et al., 2017; Zivanov et al., 2018). Particles were automatically picked with the autopicking algorithm in RELION, and the coordinates manually refined, removing and adding particles as appropriate. The resultant 20260 particles were extracted from the micrographs and subjected to reference-free 2D classification. The classes seemed to display more variety of orientations in comparison to the initial collection and the ten most well defined 2D classes were selected, which contained a total of 7710 particles (Figure 4.10).

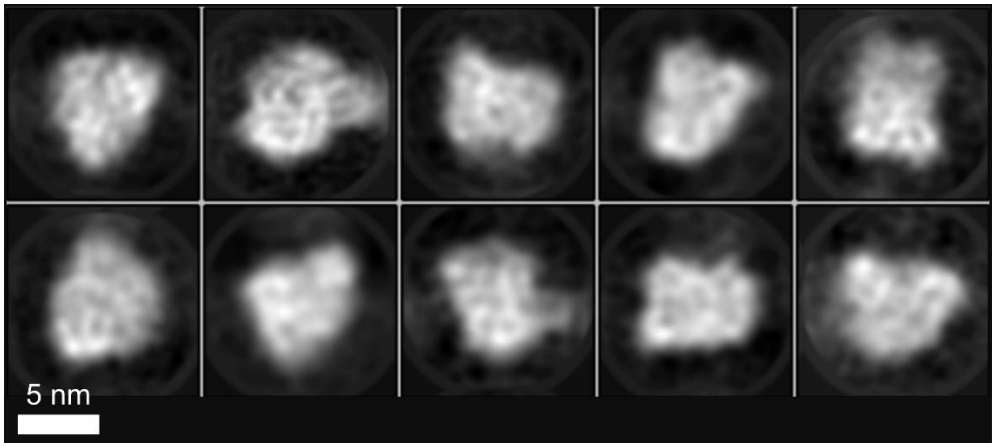


Figure 4.10. Ten best 2D classes resulting from reference-free 2D classification within RELION. Some secondary structural elements are visible. These classes contained 7710 particles and further processing was done using these particles. As in Figure 4.7, the classes are about 20 Å shorter in length than the crystal structure.

The particles were used to generate an *ab initio* 3D initial map which was used as a template for 3D classification. The low quantity of particles limited the resolution of the 3D classification and meant that overfitting was an issue. Selecting an appropriate map was therefore challenging, but the chosen class contained 4825 particles which were then subjected to further 3D classification with parameters altered to reduce overfitting. These particles were then used for high-resolution refinement using the so-called "gold-standard" Fourier shell correlation (FSC) calculation of two independently refined half-sets. A mask was generated within RELION from the 3D map low-pass filtered to 15 Å, with a 12.6 Å soft-edge which was used along with modular transfer function (MTF) detector correction, and an estimated temperature factor of -1000 to sharpen the map to a final resolution of 8.4 Å as calculated by the FSC cut off at 0.143 (Figure 4.11).

The resultant map is a roughly similar shape to that of the crystal structure, with a wider "body" region and more narrow "head" region, however it is more compressed than the crystal structure, in common with the results from the previous data collection (Section 4.3.1). When the crystal structure is fit within the map, one could imagine the long head and neck region folding inwards. Some secondary structure can be seen in the map but it is too low resolution to confidently model (Figure 4.13).

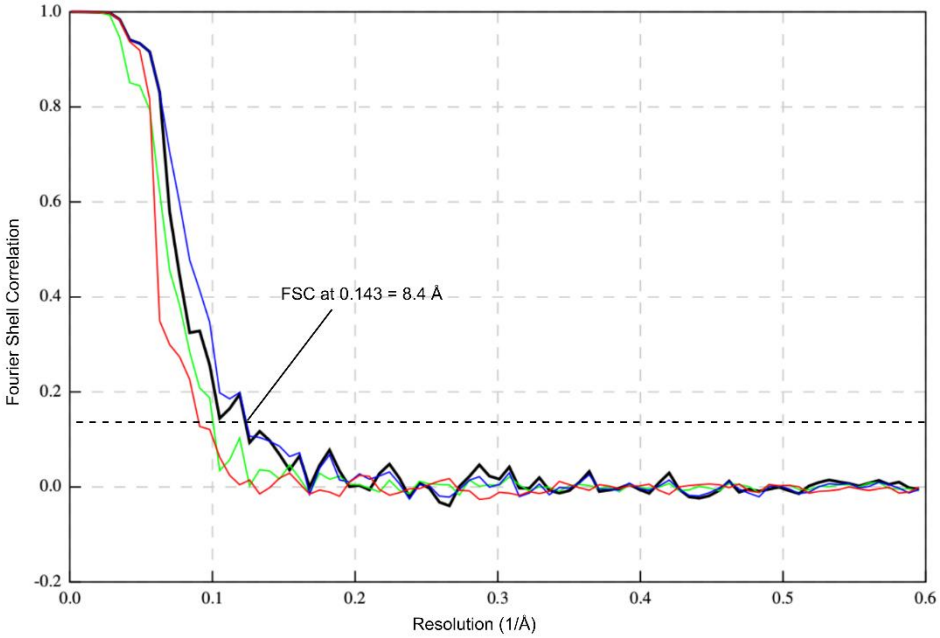


Figure 4.11. FSC calculated within RELION. Black, MTF corrected and masked map; green, unmasked map; blue, masked map; red, phase randomised masked maps, used as validation.

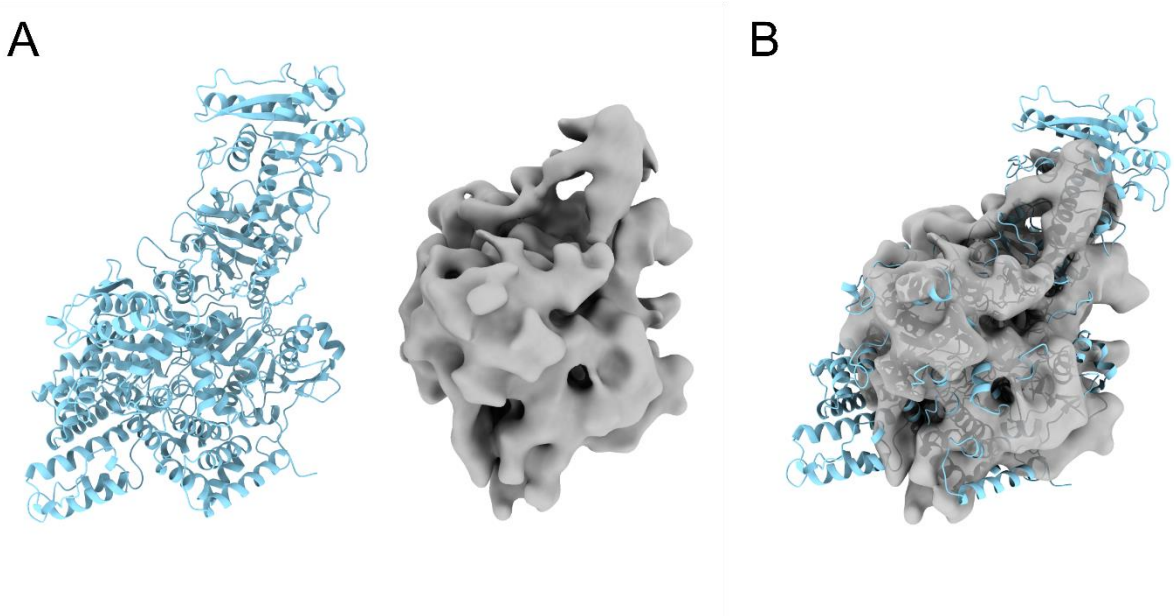


Figure 4.12. ChlH crystal structure compared with 8.4 Å resolution ChlH-D_{IX} map produced. **A**, comparison between the ChlH crystal structure. **B**, crystal structure fit in the map using the "fit in map" function within USCF Chimera. The map reveals a more compressed structure in comparison to the crystal structure, consistent with earlier data collection (Figure 4.7).

4.3.4 Apo-ChlH

Overall, the quality of the data collected for the apo-ChlH was lower than for ChlH-D_{IX} so processing was more challenging. Processing was performed in much the same as for the ChlH-D_{IX} sample: micrograph movies were dose-fractionated and corrected for beam-induced motion with MotionCor2; CTF estimation was performed with GCTF; and further processing conducted within RELION 3.0 (Zhang, 2016; Zheng et al., 2017; Zivanov et al., 2018). Attempts to use the autopicking function within RELION were ultimately unsuccessful so 17855 particles were manually picked from the motion corrected micrographs. The extracted particles were subjected to reference-free 2D classification and the resultant classes show some signs of secondary structure, but also show evidence of overfitting so it is unclear how real the apparent secondary structure elements are (Figure 4.13). When data is noisy or low quality, and particles small in size, the classification algorithms can struggle to accurately align and classify. If the classes are real, however, then it seems that this particle is also smaller than the crystal structure, to a similar extent as the ChlH-D_{IX} particle. This could suggest that the elongated architecture present in the crystal structure may be a crystal packing artefact. The unit cell comprises two ChlH monomers, interacting over a large interface which extends from the base of the "body" up to the tip of the "head". This dimerisation may be what is driving the elongated conformation. There does not seem to be any evidence of dimerisation from the EM micrographs, but dimerisation of ChlH has been observed from MST data in Chapter 3; Figure 3.6. Whether these different conformational states have any role in the reaction mechanism of ChlH is unclear.

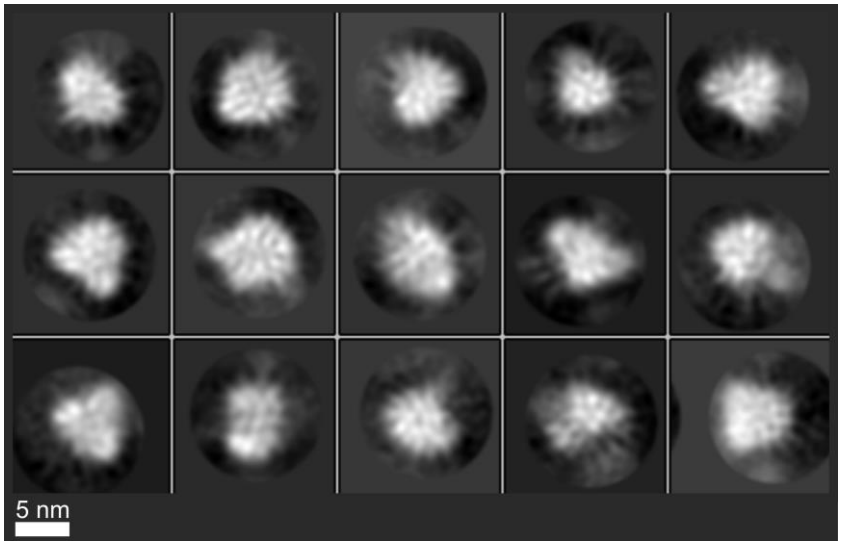


Figure 4.13. The 2D classes generated from the apo-ChlH particles. It looks like there may be some secondary structure features present, but the streaking of white extending out of the classes is a sign of overfitting. Overfitting may result in features that look like higher resolution information, but are not real.

Producing a reliable 3D map for this sample was unsuccessful, as the 3D classification and refinement algorithms could not align the particles and enable the calculation of an accurate FSC. Unfortunately, further work will be necessary to obtain direct structural evidence of substrate binding in ChlH.

4.4 Discussion

The entire biosynthetic pathway for chlorophyll has been elucidated, and heterologously expressed in *E. coli* (Chen et al., 2018). While we now know all the enzymes required for this most vital pathway we still lack a detailed mechanistic understanding of the individual reactions. The most pivotal, in terms of initiation and regulation of the chlorophyll pathway is the insertion of the Mg^{2+} ion into protoporphyrin, catalysed by MgCH. The insertion of magnesium into protoporphyrin is a particularly challenging reaction. Synthetic routes to Mg-porphyrin products usually require anhydrous and oxygen free environments, with high temperature and long reaction times (Smith and Goff, 1986), in stark contrast to the constraints that operate in nature, which require insertion of magnesium into protoporphyrin at ambient temperatures and therefore the enzyme must sequester porphyrin within a specialized, sub-yoctoliter compartment, controlling the location of water molecules.

The presented work has probed the structure of ChlH, through a combination of x-ray crystallography, electron microscopy, *in silico* modelling of ligand binding and rational mutagenesis. Despite no direct structural evidence, the binding site of the porphyrin substrate has been located, buried deep within the ChlH protein, and located close to the catalytically essential E660. Magnesium is likely delivered vectoral to the protoporphyrin along a delivery channel via a series of interactions with oxygen rich amino acid side chains.

Nature has evolved only one enzyme to catalyse this reaction. The bacteriochlorophyll and chlorophyll producing enzymes share remarkable subunit composition and sequence identity. The structural conservation displayed between the *Synechocystis*, *Thermosynechococcus* and *A. thaliana* ChlH proteins indicate this is the way to build a magnesium insertion machine.

The Mg^{2+} dependency of the ChlH protein within the MgCH complex is difficult to assess, given the multiple roles played by this ion. $MgATP^{2-}$ is required for assembly of the active MgCH complex, and for activity. The MgCH response to Mg^{2+} is dominated by the ChlD protein, where a series of acidic residues in the C-terminal region conveys cooperativity (Adams and Reid, 2013; Brindley et al., 2015). The $S_{0.5}^{Mg}$ for magnesium chelatase is in the mM range, but binding of the

Mg^{2+} to ChlH for chelation, might be expected be orders of magnitude tighter, in the same range measured for porphyrin binding.

Whilst the EM work within this chapter did not result in a high-resolution density map, it showed promise that with further optimisations an appropriate grid could be produced for large data collection that would allow the construction of an atomic model of ChlH-D_{IX}. That preliminary data seems to suggest an alternate compressed conformation of ChlH in comparison to the crystal structure highlights the necessity of solving the structure using this technique; the more native-like conditions of cryo-EM may give further insight into the reaction mechanism of this magnesium insertion machine.

5 The quaternary organisation of POR and MgCH

5.1 Introduction

When in their active, substrate bound state POR and MgCH are multi-subunit complexes of currently undiscovered stoichiometry. POR undergoes oligomerisation in the presence of both Pchlide and nucleotide, and this process is believed to play a role in membrane formation within developing angiosperm chloroplasts. ChlH-P_{IX}, ChlI-$MgATP^{2-}$, and ChlD-$Mg(ATP^{2-}/ADP^{-})$ associate to form the active MgCH complex and drive chelation of Mg^{2+} into the porphyrin ring. Elucidating the architecture of the overall active MgCH complex would lead to a greater understanding of the reaction mechanism, specifically how ATPase activity in ChlI is utilised by ChlH to drive chelation. With the recent advances in cryo-electron microscopy, as well as the utility of this approach for studying unstable complexes, it was chosen as the main technique to be used for these structural investigations.

Previous work showed that ChlI and ChlD interact through their AAA+ domains (Adams and Reid, 2013; Adams et al., 2016b), and Chapter 3 of this thesis shows that ChlD interacts with the "body" of ChlH, primarily through the C-terminal integrin I-like domain of ChlD. The strength of each of these interactions is mediated by substrates of the chelation reaction; magnesium-nucleotide in the case of ChlI-ChlD, and magnesium ions in the case of ChlD-ChlH. Despite this knowledge of interaction partners, the overall stoichiometry of the complex is unknown. Purified ChlI and BchI both form heptameric/ hexameric rings in the presence of substrate (Reid et al., 2003; Willows et al., 2004), and will hydrolyse $MgATP^{2-}$ in an unregulated manner, but it is unclear whether this state is biologically relevant as this would waste valuable ATP reserves of a photosynthetic cell. A stack of rings of BchI and BchD has been suggested from EM studies, but the density map is of low quality and unreliable (Lundqvist et al., 2010). An alternative arrangement for the ID component of the MgCH machine could be a heterohexameric ring, typical of some AAA+ motors (Monroe et al., 2017; Schmidt et al., 2012). This ring could then associate with ChlH to provide the work required to catalyse the chelation reaction.

POR is believed to form a light-sensitive stable ternary complex, which then dissembles at least to some extent upon exposure to light. Previous work on *T. elongatus* POR has revealed that an inactive ternary complex can also be formed with the addition of $NADP^{+}$ instead of NADPH

(Armstrong, 2014). As part of this work, some progress was made towards understanding the overall structure of the POR oligomer; atomic force microscopy (AFM), EM, analytical ultracentrifugation (AUC) and dynamic light scattering (DLS) data revealed that megadalton POR-ternary aggregates can be formed, which dissipate after exposure to light, but also that more regular rings of 4 – 7 POR monomers can be present. In another study, XLMS was used to probe which regions of PORA from *Arabidopsis thaliana* are potentially involved in the oligomerisation process in the presence and absence of substrates; all solvent exposed regions of POR not including the substrate binding pocket were found to form cross-links with neighbouring POR molecules (Gabruk et al., 2016). POR is a member of the SDR family, which also have a propensity to form multimeric complexes (Kavanagh et al., 2008). Information from other members of this family could be used to guide efforts in understanding the oligomerisation of POR.

5.2 Magnesium chelatase

5.2.1 Capturing the whole chelatase for structural studies

Several approaches were undertaken in an attempt to capture the MgCH complex for structural investigation using electron microscopy. The initial method was to assemble the components of the reaction as for preparing an MgCH assay. This could be done lacking certain substrates, to reduce heterogeneity of the sample; the complex would not be turning over, so a single inactive complex could be isolated. With the full set of substrates, the active chelating complex would be present, but if complex formation is transient then assembly intermediates could complicate interpretation of the micrographs. If these different components were sufficiently distinctive then the complex could be "purified *in silico*" during the classification stages following cryo-EM data collection. Identification of all assembly intermediates could reveal the assembly process that ultimately leads to chelation.

5.2.1.1 Establishing a MgCH assay for preparing sample grids to image the active complex

In the first instance, ADP was used instead of ATP, as it has previously been demonstrated that ChlI and ChlD interact tightly in the presence of ADP (Adams et al., 2016b) and this would potentially lower the heterogeneity within the sample as this avoids an ADP/ATP mixture of substrates. D_{IX} was also omitted. MgCH subunits were produced as described in Sections 2.6 and 2.8.2. 3 μM ChlH, 2 μM ChlI and 1 μM ChlD were incubated at 34 °C in 50 mM MOPS/KOH buffer, 10 mM free Mg^{2+}, 5 mM ADP, 0.3 M glycerol, 1 mM DTT, *I* 0.1, pH 7.7 for 5 min before

being diluted 10-fold and applied to a freshly glow discharge treated carbon-coated 400 mesh Cu grid which was then stained with uranyl acetate as in Section 2.11.1. The grid was imaged on a Philips CM100 microscope with a 1K x 1K Gatan Multiscan 794 CCD (Figure 5.1 **A**). Particles are clearly visible but are smaller than expected for a multimeric complex. No particles much bigger than monomeric ChlH (*ca.* 12 nm long) are visible which could suggest that the presence of ATP and/or porphyrin are required for formation of a functional chelating complex.

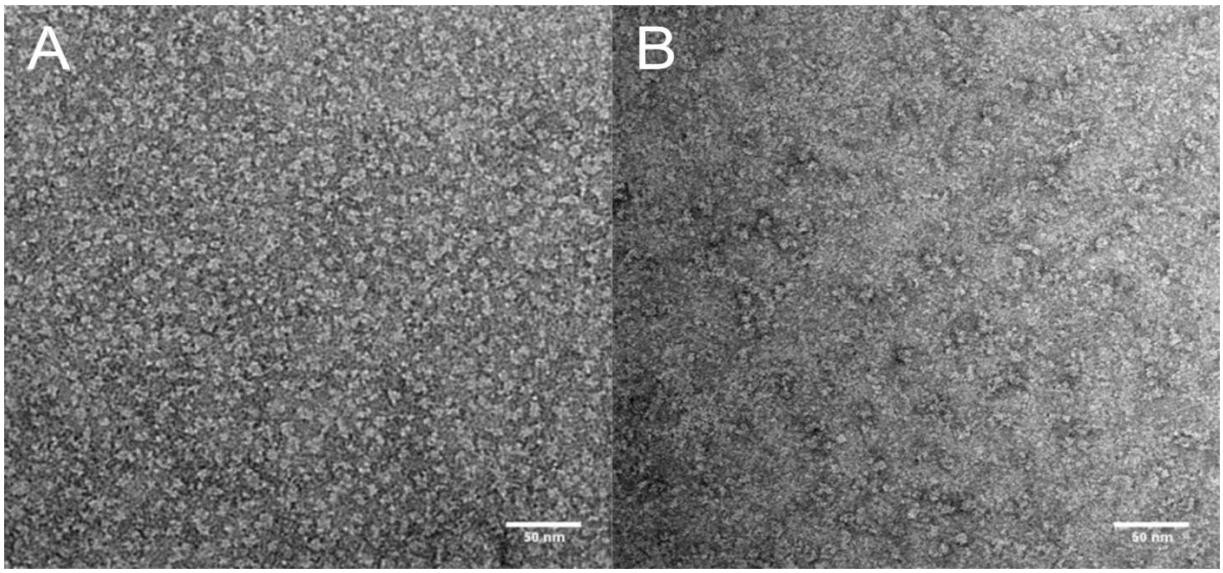


Figure 5.1. EM micrographs of negatively stained MgCH assembled as for conducting a chelatase assay. **A**, a non-active assay containing ADP and no D_{IX}. **B**, an active assay containing all substrates required for chelation.

Next, the full set of substrates was used so that a chelating complex would definitely be present. ChlH, ChlI and ChlD were incubated as above, but with 5 mM ATP instead of ADP, as well as 8 μM D_{IX}. After the incubation, negatively stained grids were produced and imaged as before (Figure 5.1 **B**) revealing a situation similar to the previous attempt. The contrast in this micrograph is less strong so it is difficult to interpret but it could be argued that there are more particles that are larger than 12 nm. However, there is still a large amount of heterogeneity and it is not clear enough to easily determine different specific complexes. It was clear that optimisation and purification steps would be necessary to isolate the complex and reduce heterogeneity.

5.2.1.2 Purification of the native MgCH from *Synechocystis*

The grids imaged from assembling an assay contained a large variety of particles, and it was impossible to identify particles corresponding to MgCH complexes; some purification of the complex was required. Purification from the native organism, *Synechocystis* sp. 6803, was attempted as it

has been shown that the whole complex can be purified from a FLAG-tagged ChlD strain of *Synechocystis* (Farmer et al., 2019; Chapter 3). Purification from the native environment could also improve stability of the complex, potentially leading to clearer EM micrographs.

The FLAG-ChlD Δ*chlD* strain of *Synechocystis* constructed in Chapter 3 was grown as in Section 2.7, before cells were lysed, solubilised and applied to an anti-FLAG column for purification (Section 2.8.2.5). The resultant eluates from the applied solubilised membrane (as in Chapter 3, MgCH is present in higher quantities within the membrane fraction) were applied to a freshly glow-discharge treated carbon-coated 400 mesh Cu grid and stained with uranyl formate as in Section 2.11.1. The grid was then imaged on a Philips CM100 microscope with a 1K x 1K Gatan Multiscan 794 CCD.

The micrograph (Figure 5.2 **A**) had a high level of background noise, and no particles could be identified. The presence of 150 mg ml^{-1} FLAG peptide present in the elution buffer was believed to be the source of the noise so the eluate was buffer exchanged using a Zeba™ Spin desalting column (ThermoFisher) into the same buffer without FLAG peptide according to the instructions provided by the manufacturer. This process was repeated twice to thoroughly remove FLAG peptide and the resultant solution applied to a grid, stained and imaged as before.

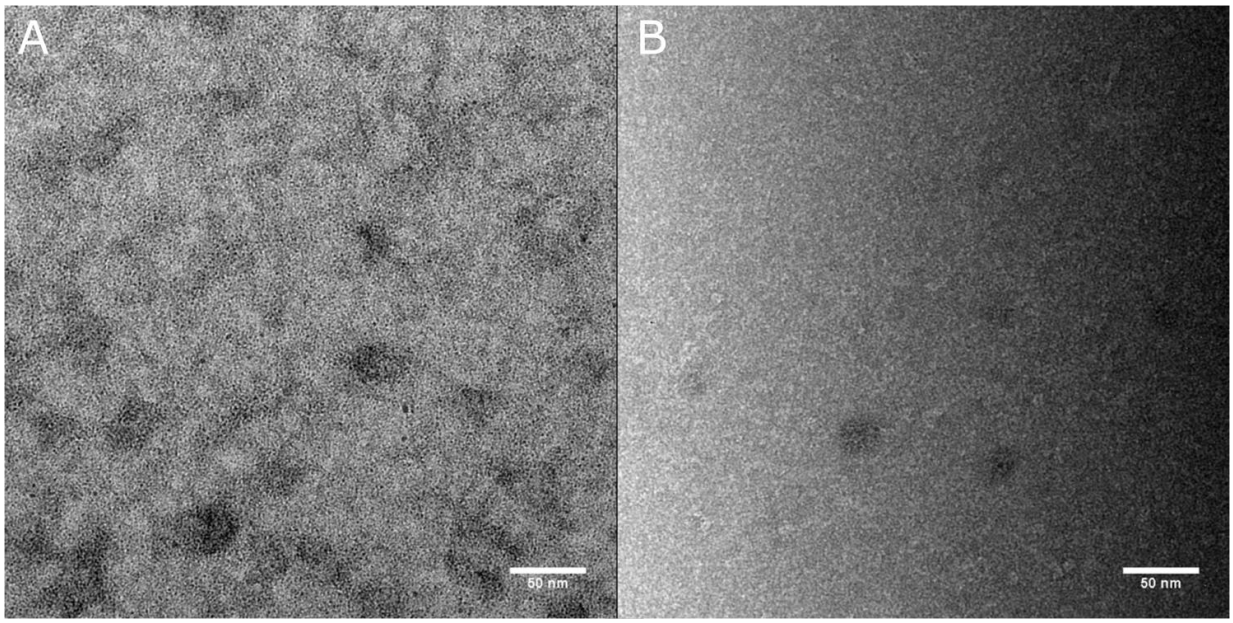


Figure 5.2. EM micrographs of MgCH complex purified from *Synechocystis* strain FLAG-ChlD Δ*chlD*. **A**, the micrograph of eluate with FLAG peptide present has high background noise. **B**, micrograph after desalting to remove FLAG peptide - there is some evidence of particles but these may also be staining artefacts.

The micrograph is less noisy than the initial image (Figure 5.2 **B**) and there may be visible particles but it is difficult to determine whether they are real or staining artefacts. The concentration of protein within the eluates is low, but MgCH is unstable when concentrating low volumes of sample. ChlD precipitates on membranes of spin concentrators so concentrating the sample was undesirable. It seems that significantly more cells would be needed to be grown for future work using this method.

An alternate strategy for purification from *Synechocystis* could be to construct a strain with a N-terminal His_6-tagged ChlD as it is possible that the presence of FLAG peptide is unsuitable for imaging using EM. Further processing steps like the buffer exchange could also disrupt the complex, especially if the complex is as transient as it appears to be.

5.2.1.3 Coexpression of subunits in *E. coli*

A vector that can be used by cells to coexpress all three genes of the *T. elongatus* MgCH had been constructed prior to this work by Dr Amanda Brindley. The pETcoco2 vector contains all the subunits and an N-terminal His_6 tag on ChlI. Transformation with this plasmid was conducted as in Section 2.5 (with 0.2 % (w/v) glucose supplemented LB agar plates), followed by cell growth in the same medium used to produce ChlD (Section 2.6) with the addition of 0.01 % arabinose to promote plasmid amplification. For induction, 400 μM IPTG was added to the flasks, which were then incubated at 20 ℃ overnight. Further steps were conducted in much the same way as for individual MgCH subunits.

All buffers used throughout purification of the coproduced MgCH subunits contained no NaCl, 13 mM MgCl and 3 mM ADP. Sonicated and clarified cell lysate was purified using a 2.5 ml Ni(II) Fast flow chelating Sepharose column as in Section 2.8.2.1 (wash 2 was omitted). The eluate was then subjected to size exclusion chromatography as in Section 2.8.2.3 (Figure 5.3 **A**). Protein eluted at lower elution volume than for any of the subunits individually, suggesting that a complex was stable enough to withstand the purification procedure. When separated by SDS PAGE, all three subunits were present in fractions L6-L10, with relatively low levels of contamination (Figure 5.3 **B**). Negatively stained grids were prepared for fractions L6-L10 as in Section 2.11.1 which were then imaged on a Philips CM100 microscope as before. Fractions L6-L9 were highly aggregated, but fraction L10 contained ring-like particles of *ca.* 140 Å diameter, as well as more irregular particles. The heterogeneity of this sample is similar to previous ones but it is more promising as there is evidence of complex formation, in a manner consistent with previous studies.

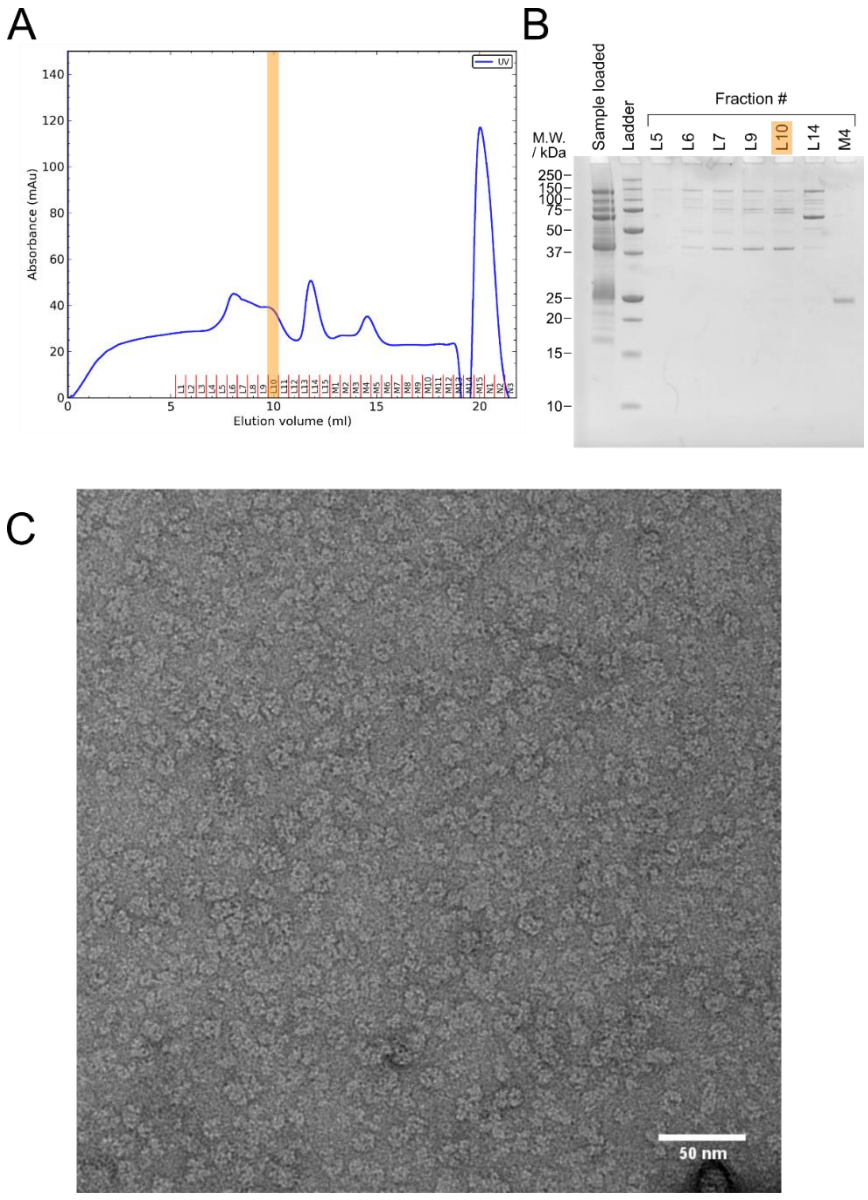


Figure 5.3. Purification and EM micrograph of co-expressed and produced MgCH complex. **A**, A_{280} trace from chromatogram of the size exclusion stage of the purification, orange shading highlighting fraction imaged in **C**. **B**, SDS PAGE of size exclusion eluates, labelled as in figure; molecular weights labelled on left in kDa. L6-L10 seem to contain all MgCH subunits (ChlH ~150 kDa; ChlD ~75 kDa; ChlI ~40 kDa). A ChlH-ChlD complex may be represented by fraction L14. **C**, EM micrograph of fraction L10 revealing both ring-like and more irregular particles. There still appears to be a degree of heterogeneity within the sample.

5.2.2 Using glutaraldehyde to stabilise the MgCH complex

Instability of the MgCH complex was proving to be an issue preventing the production of sample suitable for data collection using cryo-EM; to overcome such problems cross-linking reagents are sometimes used in the sample preparation of unstable multimeric complexes (Drulyte et al., 2018).

A commonly used method for this is the GraFix protocol developed by Stark (2010) where a protein complex is assembled and subjected to ultracentrifugation down a sucrose and glutaraldehyde gradient. A large complex migrates further and gradually encounters more cross-linking reagent to stabilise it whilst also separating it from smaller components. The glutaraldehyde concentration is suitably low, so that inter-complex cross-links are avoided.

This method requires significant optimisation for the complex being studied, so a more simple alternative cross-linking protocol was followed, as described in the study by Monroe et al. (2017). In this work, the authors simply added glutaraldehyde to a final concentration of 0.02 % (v/v) to the protein solution, which was then incubated for 30 min before quenching the cross-linking reaction with the addition of buffer including 1 M glycine.

MgCH subunits were purified separately as in Section 2.8.2 before being combined (40 μM ChlH, 20 μM ChlI, 10 μM ChlD) and 13 mM MgCl, 3 mM ADP added. The solution was incubated at 34 ℃ for 10 min to allow for complex assembly, before the addition of glutaraldehyde (0.02 % (v/v)) and a further incubation at 34 ℃ for 30 min. After this time, the same buffer supplemented with 1 M glycine was added 1 part MgCH complex solution: 1 part glycine solution, before loading onto a Sephracyl S300 column pre-equilibrated with size exclusion buffer (Section 2.8.2.3).

Protein eluted from 50 ml (Figure 5.4 **A**) and the early fractions were assessed by SDS PAGE using a 3-8 % Tris-Acetate precast gel (Section 2.10.2). Two relatively clean bands are present at molecular weights above the largest protein in the marker (250 kDa) that correspond to the broad first elution peak (fractions D2-E5). The subsequent peak appears at around 150 kDa so is most likely monomeric ChlH. Molecular weight standards (GE Healthcare) were applied to the column to allow for plotting of a calibration curve (Figure 5.4 **A**, inset).

Fractions E3-E9 were negatively stained as in Section 2.11.1 and imaged on a Philips CM100 microscope as before. Only fraction E3 contained particles with no obvious large aggregation (Figure 5.5); the particles were relatively large (*ca.* 20-30 nm) and homogenous and contained none of the ring-like structures observed previously. Fraction E8 and E9 contained particles that were a similar size to ChlH.

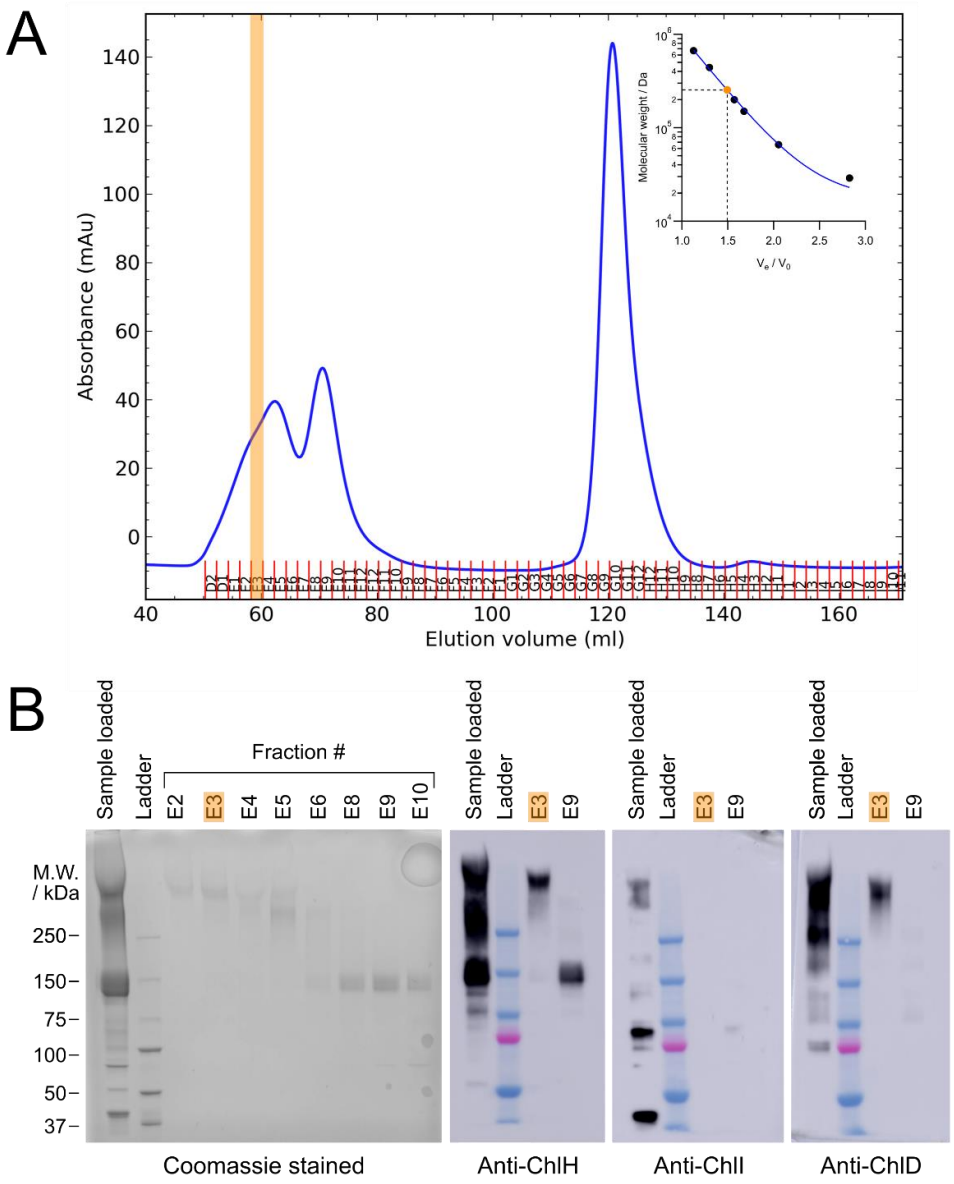


Figure 5.4. Purification of a cross-linked MgCH complex. **A**, chromatogram of the SEC performed using cross-linked MgCH. Orange, highlight of fraction E3, the only non-aggregated fraction from the first broad elution peak. Inset, calibration curve using molecular weight standards (black circles), with calculated results for fraction E3 (orange circle). **B**, SDS PAGE and immunoblot of the fractions from the first two elution peaks. Fraction E3 contains ChlH and ChlD, but not ChlI. However, in the sample that is loaded, ChlI is present in the same high molecular weight band purified within E3. Fraction E9 appears to be monomeric ChlH.

Using the retention time of fraction E3 and the calibration curve, it was calculated that the high band in the Tris-acetate gel was *ca.* 250 kDa (Figure 5.4 **A**, inset, orange marker). This is the same

size as the highest molecular weight standard run alongside the fractions but it perhaps migrated less through the gel due to the introduced cross-links preventing complete linearization by SDS.

Fraction E3 and E9 were immunoblotted (Section 2.10.4) for each of the MgCH subunits to assess their presence within the complex. Fraction E9 was confirmed to be ChlH, whilst fraction E3 contained ChlH and ChlD. Interestingly however, in the sample that is loaded onto the column, the high molecular weight band that is purified in fraction E3 appears to contain ChlI. When the anti-ChlI membrane was overexposed during imaging, a faint band could be detected in fraction E3 at the location of the high molecular weight band (data not shown). It is therefore unclear if this high molecular weight complex is ChlH-ChlD or the whole MgCH.

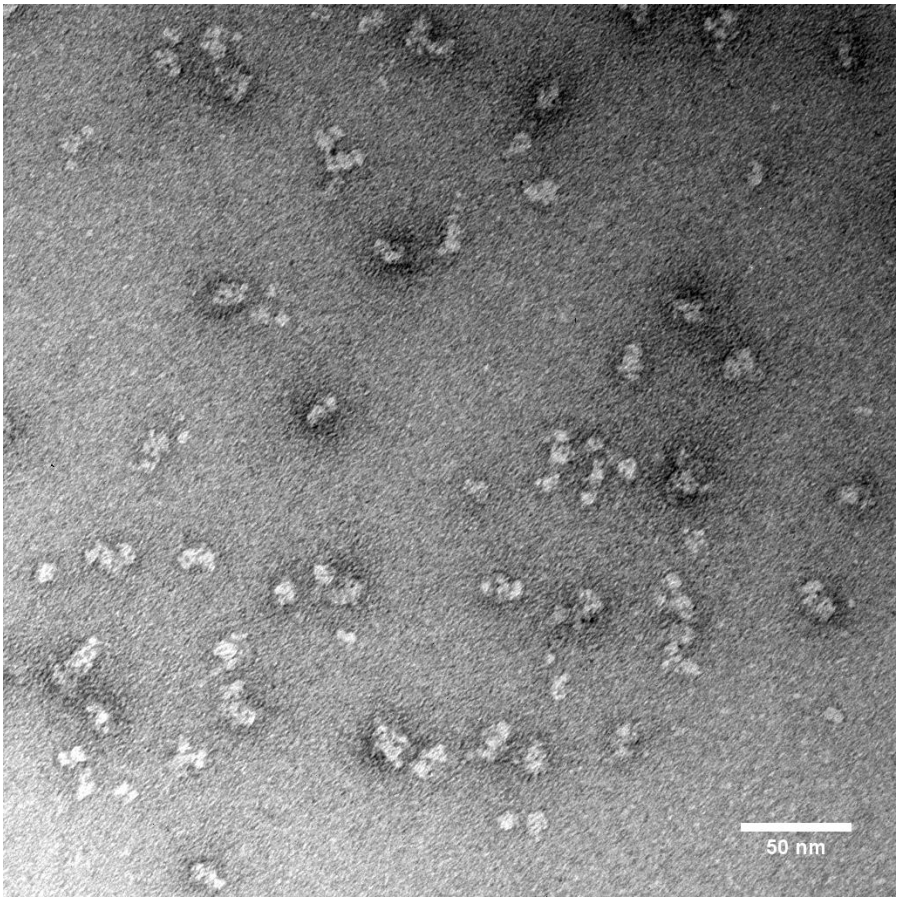


Figure 5.5. EM micrograph of negatively-stained purified cross-linked MgCH, fraction E3. Particles *ca.* 20-30 nm in length can be seen that are somewhat homogenous.

5.2.3 The ChlI-ChlD complex

An alternative method of purifying the MgCH complex was devised, using a tandem affinity purification procedure; it is known that ChlH and ChlI do not interact and are linked by ChlD (Farmer

et al. 2019; Chapter 3), so if two different tags are added to ChlI and ChlH then purifying using both of these tags should result in a whole MgCH complex. A StrepII tag on ChlH and a His_6 tag on ChlI were selected, so His_6*chlI* and *chlD* were cloned into the StrepII*chlH* construct used in Chapter 3 by Dr Amanda Brindley using the 'link and lock' method (McGoldrick et al., 2005). Within the construct, His_6*chlI* would be expressed first, followed by *chlD* and then StrepII*chlH*. This was transformed into *E. coli* strain Rosetta(DE3)pLysS and cells grown in the same conditions as would be used for WT ChlD production (Section 2.6). Cells were pelleted, sonicated, clarified and applied to a 2.5 ml Ni(II) Fast flow chelating Sepharose resin column in much the same way as previous MgCH complex purifications (i.e.: 13 mM MgCl, 3 mM ADP and no NaCl in the buffers) (Section 2.8.2.1). Eluates were separated by SDS PAGE as in Section 2.10.2 and it was revealed that ChlH was not produced by cells with this construct. However, the eluates did contain ChlI and ChlD of high purity; it was decided that these eluates would be used to try to isolate a ChlI-ChlD complex.

5.2.3.1 Purification and sample preparation

The eluates from the Ni(II) affinity column were further purified by size exclusion chromatography; eluates were pooled, concentrated to 0.5 ml and loaded onto a S200 Increase column pre-equilibrated with size exclusion buffer (Section 2.8.2.3) containing 13 mM MgCl, 3 mM ADP and no NaCl. The A_{280} trace (Figure 5.6 **A**) revealed several overlapping peaks eluting earlier than monomeric ChlI and ChlD, indicating complex formation. When separated by SDS PAGE using a 12 % Bis-Tris precast gel, Many of the fractions appeared to contain both ChlI and ChlD (Figure 5.6 **B**).

Fractions A5 to A8 containing purified ChlI-ChlD were diluted 20-fold and separately applied to a glow discharge treated carbon-coated 400 mesh Cu grid and stained with uranyl formate as in Section 2.11.1. The grid was screened on a Philips CM100 microscope with a 1K x 1K Gatan Multiscan 794 CCD. The micrographs revealed the presence of a large number of monodisperse particles of roughly similar sizes. Many ring-like structures were visible, as well as some particles predicted to be side-on rings (Figure 5.6 **C**). Due to time limitations preventing optimisation for cryo-EM, and in order to quickly screen for particle heterogeneity, the negatively stained grid corresponding to fraction A6 was chosen for data collection on as it was deemed to have the best particle distribution.

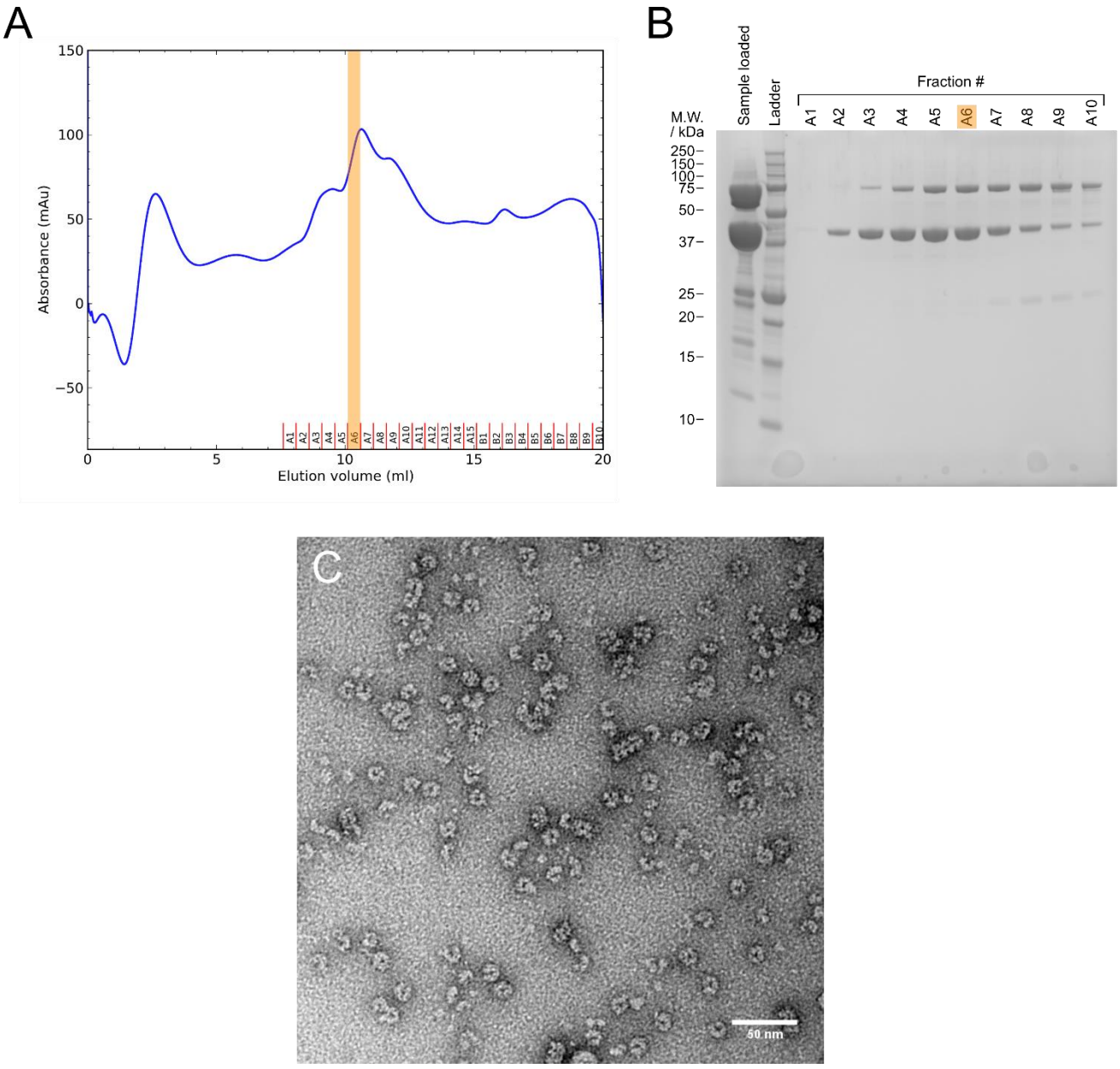


Figure 5.6. Purification and imaging of the ChlI-ChlD complex. **A**, chromatogram of the SEC performed, with the A_{280} trace displayed (blue). Fraction A6, which was chosen for data collection using EM, is highlighted in orange. **B**, SDS PAGE separation of the eluted fractions, revealing an absence of ChlH in the sample loaded. Within the fractions, both ChlI and ChlD were present to a high degree of purity. **C**, negatively stained micrograph of fraction A6 obtained from imaging on the CM100 microscope revealing the presence of ring-like structures.

5.2.3.2 Data collection and processing

The grid was clipped and loaded into a Technai Arctica cryo-microscope with a Falcon 3EC direct electron detector at room temperature (Section 2.11.2). The microscope was brought back to cryogenic temperatures and automated data collection was conducted using the EPU software (ThermoFisher). Imaging parameters can be found in Table 5.1.

Table 5.1. Imaging parameters for the automated data collection of the negatively stained ChlID grid.

Parameter	**Value**
Magnification	49000
Voltage (kV)	200
Electron exposure (e^-/ Å^2)	0.6 per frame (49.8 total on 83 frames)
Defocus range (µm)	-3 to -6 in steps of -0.5
Pixel size (Å)	2.86

Micrographs were dose-fractionated and corrected for beam-induced motion using MotionCor2; CTF estimation was done using CTFFIND4.1, and further processing conducted in RELION 3.0 (Rohou and Grigorieff, 2015; Zheng et al., 2017; Zivanov et al., 2018). From the corrected micrographs, *ca.* 1000 particles were picked manually and subjected to reference-free 2D classification. The resultant classes were used as templates for the autopicking algorithm within RELION. 89017 particles were picked (Figure 5.7) and extracted which were then aligned and classified as above.

The reference-free 2D classification of these particles revealed a large degree of heterogeneity in ring structures. The majority were hexameric rings, or incomplete hexameric rings, but there was also a wide range of larger complex ring structures that appear to be composed of the same monomer unit as the hexameric ring particles. The larger complexes could be heart-shaped, figure of eight-shaped, and also formed large oligomeric fibres. Whilst the presence of the larger complexes shows a propensity for ChlI and ChlD to form many different quaternary arrangements, 73 % of particles were classified into the 32 most populated 2D classes which are primarily hexameric rings (Figure 5.8).

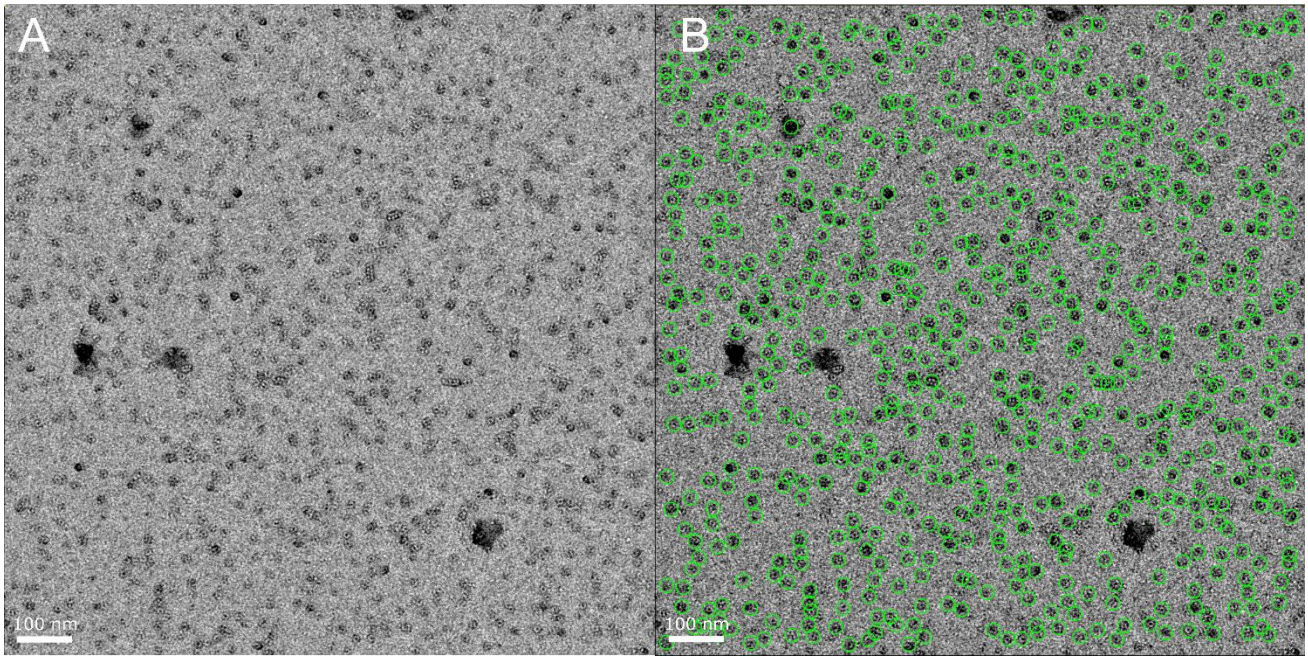


Figure 5.7. Example micrograph showing particle distribution (**A**) and the results of RELION's autopicking algorithm (**B**; green circles).

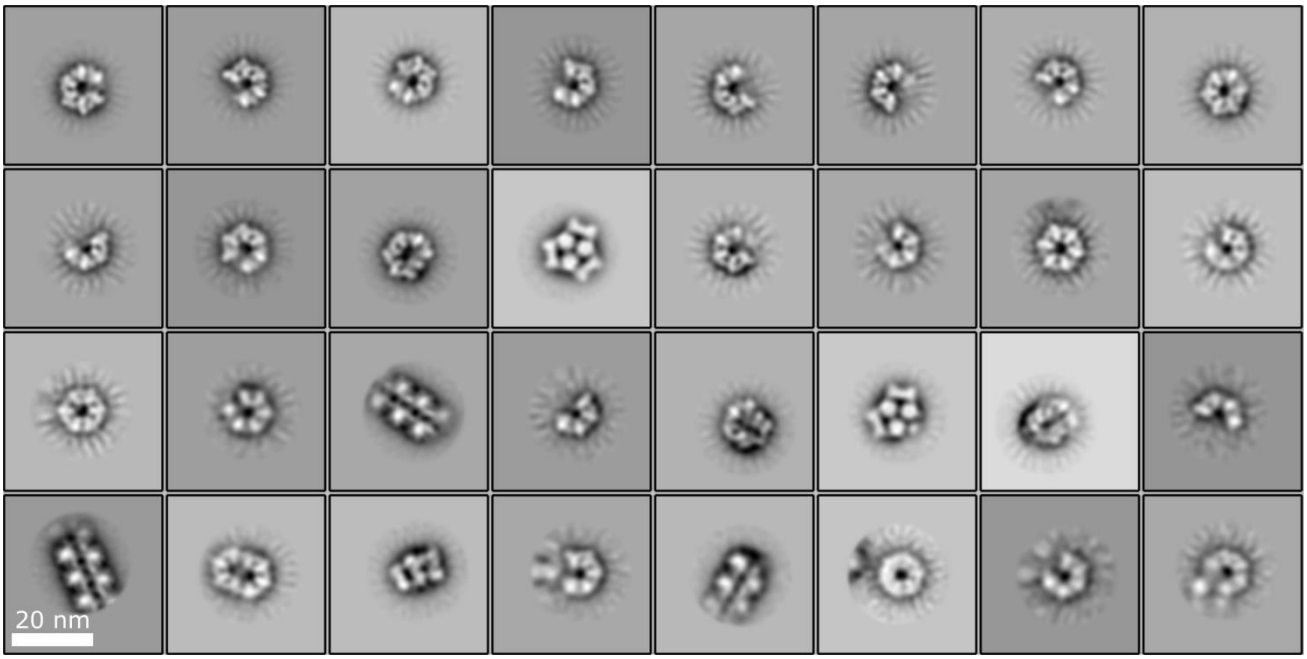


Figure 5.8. Top 32 most populated 2D classes from the reference-free 2D classification of the ChlID sample representing 65820 particles.

It appears that within the complete hexameric rings there is some degree of conformational heterogeneity; some show a more symmetrical ring, while others have small gap between two adjacent

monomers, perhaps related to the nucleotide binding state. Examples of the oligomeric fibrils are also present (e.g.: Figure 5.8, bottom-left class). There is also a potential side view (bottom, three from the left), which looks like it could comprise three stacked rings; the width of this class is lower than that of the rings, however. It is difficult to know what arrangement of ChlI and ChlD are contained within these classes, though the bottom-right class is interesting; it appears to be a hexameric ring, but with a protrusion extending from one of the monomers - perhaps an I_5D_1 heterohexamer.

Due to the large amounts of heterogeneity within this sample, and limited high-resolution information from the staining procedure, further processing was challenging and a reliable model not produced. The biological relevance of the wide-range of quaternary organisations adopted by ChlI and ChlD is hard to assess; the presence of other substrates and the other MgCH proteins may promote or suppress various states. One method to identify which monomeric components are ChlI or ChlD could be to label the molecule with nano-gold beads, as was done to identify the N-terminus of ChlH (Qian et al., 2012).

5.3 POR

5.3.1 Attempts to purify the POR ternary complex

Previous work has been conducted to isolate and characterise the POR ternary complex using various techniques, including EM (Armstrong, 2014). A protocol was developed to purify the ternary complex by mixing POR with $NADP^+$ and Pchlide and separating it by cation exchange chromatography in the same way that apo POR is purified. By using $NADP^+$, the complex can form but is not catalytically active when exposed to light. This methodology was followed in an attempt to repeat the previous work for use with cryo-EM.

5.3.1.1 Cation exchange chromatography

All sample preparation involving pigment was conducted under low light conditions. POR protein was produced and purified as in Sections 2.6 and 2.8.1, whilst pure Pchlide was produced according to Section 2.9. The dried Pchlide pigment was resuspended in POR cation exchange buffer (Section 2.8.1.3) supplemented with 50 mM β-OG by vortex mixing for 5 min and then clarified by centrifugation at 1700 *xg* for 5 min. The concentration of the solubilised Pchlide solution was calculated spectroscopically, using A_{630} and ε = 30.4 mM (in 80 % (v/v) acetone) and a 100 mM $NADP^+$ solution was prepared. 10 μM POR, 30 μM Pchlide and 1 mM $NADP^+$ were mixed in cation exchange buffer with 20 mM β-OG, below the critical micelle concentration (CMC),

thought to promote release of Pchlide to POR, before incubation for 10 min at 55 °C and clarification of any large aggregates by centrifugation at 1700 *xg* for 10 min. Cation exchange chromatography was then conducted as in Section 2.8.1.3.

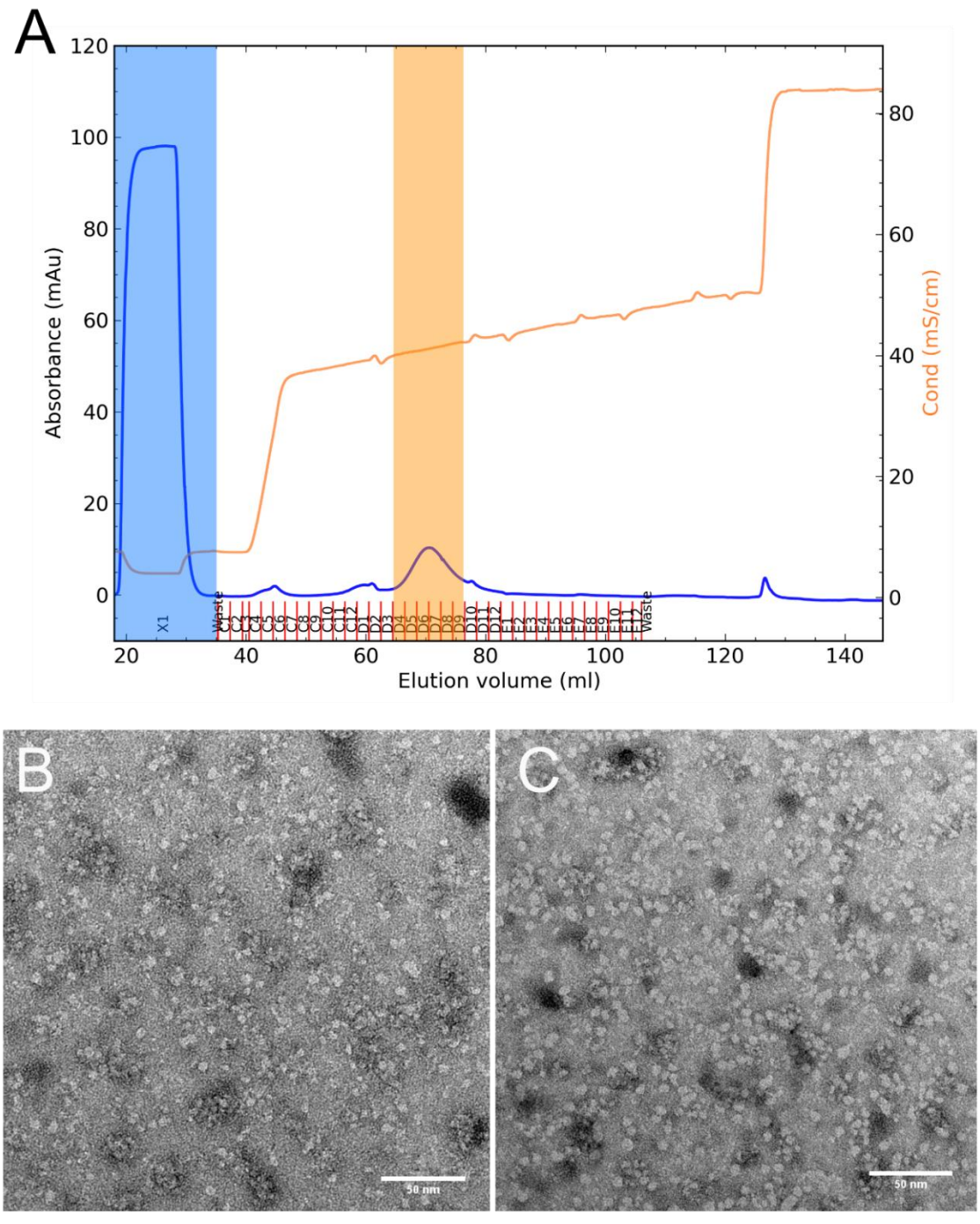


Figure 5.9. Purification of a POR ternary complex using cation exchange chromatography for analysis by EM. **A**, chromatogram of the purification, blue trace: A_{280}; orange trace: conductivity. Blue highlight: flow-through, imaged by EM in **B**; orange highlight: pooled fractions, imaged by EM in **C**. **B-C**, EM micrographs of fractions highlighted in **A**; the two samples look very similar and show no evidence of oligomerisation.

In the work by Dr David Armstrong (2014), the POR ternary complex was found in the flow-through of the column; it was suggested that the presence of β-OG around the ternary complex altered the charge of the complex, preventing it from binding to the column. In the purification performed in this work, protein was found in both the flow-through and the peak eluted during the salt gradient (Figure 5.9 **A**), though neither was green in colour. Both samples were negatively stained as in Section 2.11.1 and imaged on a Philips CM100 microscope; they appeared very similar, with monodisperse particles around 6 nm in diameter (Figure 5.9 **B**). A POR monomer is estimated to be *ca.* 4 nm in diameter so in the presence of detergent and heavy metal stain it is likely that these particles are monomeric POR. This result is in disagreement with the previous work where tetramers and pentamers were observed. Even if these samples contained both substrates, particles of that size are too small to study by cryo-EM.

It was decided to repeat the procedure but with a higher concentration of Pchlide and POR, to produce sample that could be more easily analysed spectroscopically. POR and substrates were produced as above and the sample was prepared as before; the two substrates (1 mM $NADP^+$, 30 μM Pchlide) were added to 25 μM POR in cation exchange buffer supplemented with 20 mM β-OG and incubated at 55 ℃ for 10 min before clarification. This mixture was then loaded onto a 5 ml HiTrap SP HP column attached to an ÄKTA FPLC system and purified as in Section 2.8.1.3.

Again, much of the protein does not bind to the column while some elutes during the salt gradient, though *ca.* 20 ml earlier than previously (Figure 5.10 **A**). The eluate of the salt gradient was green in colour, indicating the presence of Pchlide. This was checked spectroscopically and while the concentration was very low, the Soret band of Pchlide was clearly visible (Figure 5.10 **B**). The flow-through and independent fractions eluted during the salt gradient were applied to grids for negative staining and imaged on a Philips CM100 microscope. The micrograph of the flow-through (Figure 5.10 **C**) is very different from the previous micrograph of this fraction; the monomeric POR particles were absent and replaced by thin *ca.* 30 nm long particles. Within the micrograph of fraction D1 (Figure 5.10 **D**), lots of particles of a size similar to before are visible, but in addition there are some larger, roughly circular complexes *ca.* 30-40 nm in diameter, composed of what appear to be the monomeric particles. The heterogeneity of this sample makes it unsuitable for cryo-EM.

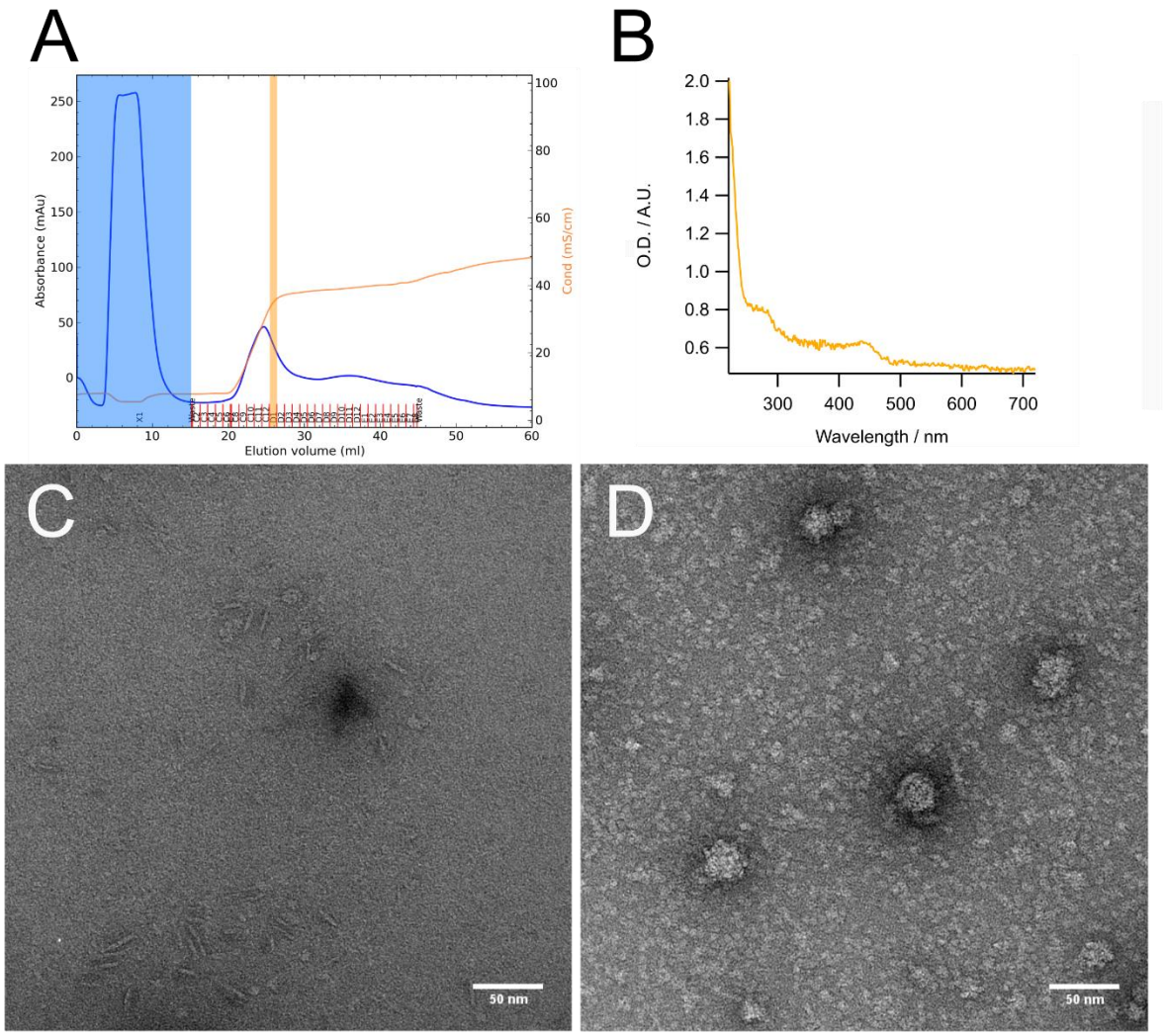


Figure 5.10. Second attempt at cation exchange chromatography to isolate the POR ternary complex for study by EM. **A**, chromatogram of the purification, traces coloured and fractions highlighted as previously. **B,** UV/Vis spectrum of fraction D1 (orange highlight, **A**) proving that Pchlide is in this fraction (broad peak at 440 nm). **C-D**, EM micrographs of negatively stained sample from the flow-through (**C**) and fraction D1 (**D**).

5.3.1.2 Size exclusion chromatography

While some progress was made to isolate the POR ternary complex, cation exchange chromatography does not remove heterogeneity of particle size. Therefore, further attempts at purification were conducted using size exclusion chromatography to try and isolate ternary complexes of specific sizes. It was also thought that it could be useful to isolate the active complex, as it could be that using $NADP^+$ was affecting the formation of the ternary complex.

Sample preparation and purification of an active POR ternary complex was conducted in the dark. Pure Pchlide was solubilised as above and mixed with NADPH and POR protein purified by immobilised nickel affinity chromatography and buffer exchanged into POR assay buffer (Section 2.13.1) using a PD10 desalting column (GE life sciences) according to the instructions provided by the manufacturer. Final concentrations were 5 µM POR, 26 µM Pchlide, 5 mM NADPH, and this solution was incubated at 50 °C for 10 min before being diluted 2-fold and clarified by centrifugation as above. The supernatant was then concentrated to 500 µl for loading onto a S200 Increase column pre-equilibrated with POR assay buffer. The resultant chromatogram (Figure 5.11 **A**) shows that the majority of the sample was too large for the column to separate and it eluted in the void volume. A smaller shoulder peak also eluted slightly later, as well as a peak eluting at a volume that corresponds to monomeric POR (as determined from a previous purification of apo-POR using the S200 Increase column).

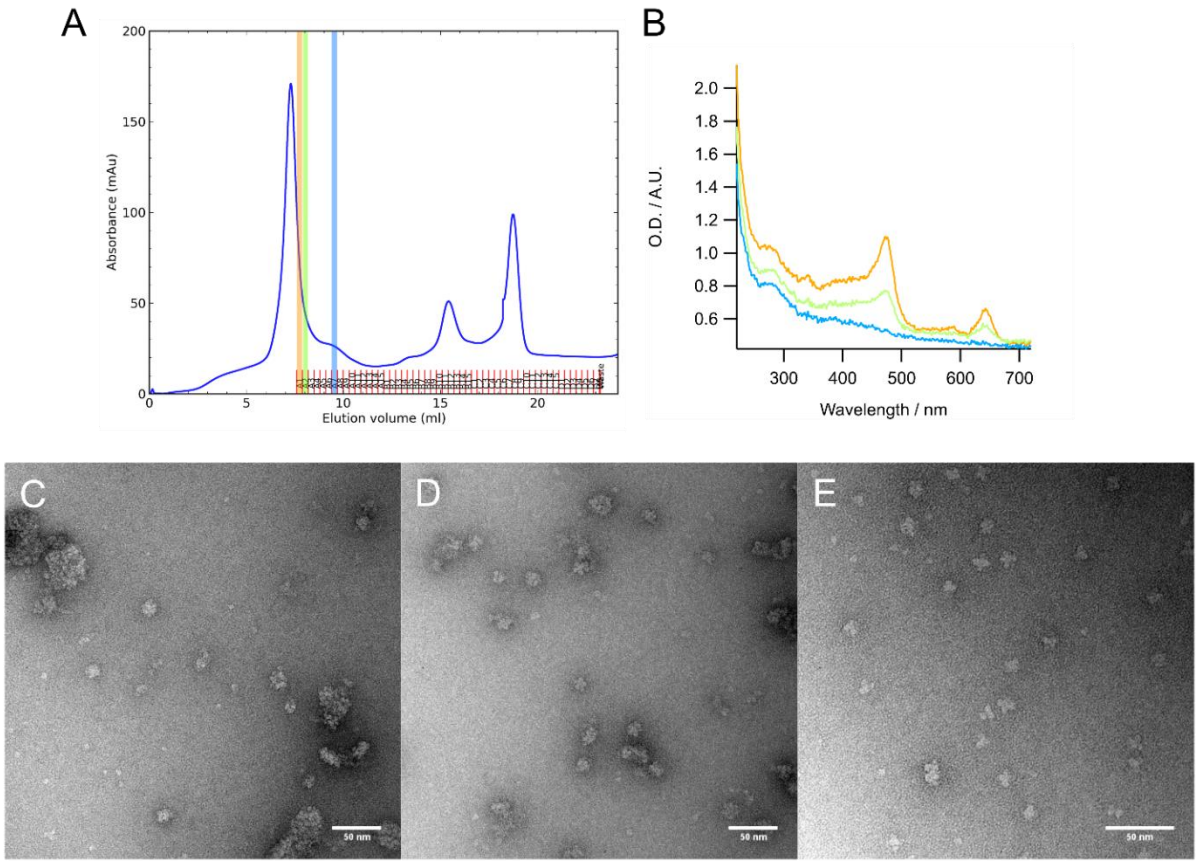


Figure 5.11. Purification of the active POR-ternary complex by size exclusion chromatography for study by EM. **A**, chromatogram from the size exclusion chromatography, blue trace corresponding to A_{280}; coloured highlights show fractions used for EM. **B**, UV/Vis spectra of the fractions highlighted in **A**, colours of the traces corresponding to highlights in **A**. **C-E**, EM micrographs of negatively stained sample; **C**: fraction A1 (orange); **D**: fraction A2 (green); **E**: fraction A7 (blue).

Three fractions from the earliest peaks were assessed spectroscopically to determine the presence of Pchlide, which was present in the first two fractions (Figure 5.11 **B**). These three fractions were then negatively stained (Section 2.11.1) and imaged on a Philips CM100 microscope (Figure 5.11 **C**-**E**). The three images are somewhat similar; they all show evidence for smaller POR complexes than have been observed previously, *ca.* 10-20 nm in size. It is possible to discern individual monomers in the complexes and it appears that the majority of these particles are tetrameric or pentameric, consistent with previous observations (Armstrong, 2014). It is interesting that there does not seem to be Pchlide present in fraction A7 (blue highlight) but that these complexes are still present; perhaps the presence of NADPH alone is enough to promote some quaternary organisation.

As much of the ternary complex eluted in the void volume of the S200 Increase column, a column that can separate larger complexes was required. POR ternary complex was produced in the same manner as previously and then loaded onto a Sephacryl S300 column pre-equilibrated with POR assay buffer (Section 2.13.1). The resultant chromatogram was vastly different to the previous result– the majority of the protein eluted as a monomer, with a broad base to the peak, suggesting that larger structures were eluting. These fractions were assessed for Pchlide as before and none was found. It is unclear what could have caused such a different result, it could be human error, slight differences in purification preparations or perhaps the different properties of the resin in the S300 column.

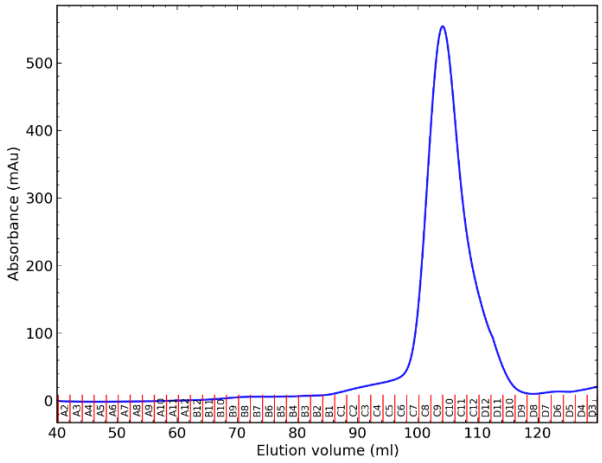


Figure 5.12. Chromatogram resulting from the size exclusion chromatography of POR ternary complex applied to a Sephracyl S300 column. The large major peak is most likely monomeric POR due to the high elution volume.

Finally, as time restraints were limiting further optimisation, cross-linking with glutaraldehyde in the same manner as was used for MgCH earlier in this chapter was used to try and stabilise the ternary complex. POR ternary complex was assembled and cross-linked as in Section 2.10.5 and then loaded onto a Sephacryl S300 column pre-equilibrated with POR assay buffer as before. The resultant chromatogram is similar to the previous one, though with a larger proportion of the protein present in the broad base of the peak (Figure 5.13 **A**). Two early fractions (highlighted in Figure 5.13 **A**) were selected for imaging by EM and were prepared and imaged as before (Figure 5.13 **B-C**). Again, no Pchlide was detected within the samples, but complex formation was observed in the micrographs. The particles were of a similar size to the smaller particles observed after the purification with the S200 Increase column (i.e.: *ca.* 10-20 nm), though individual monomer units are more difficult to distinguish. The fractions have a molecular weight of *ca.* 250 kDa calculated from the calibration curve generated earlier in this chapter (5.2.2); this is a stoichiometry of 6-7 POR monomers.

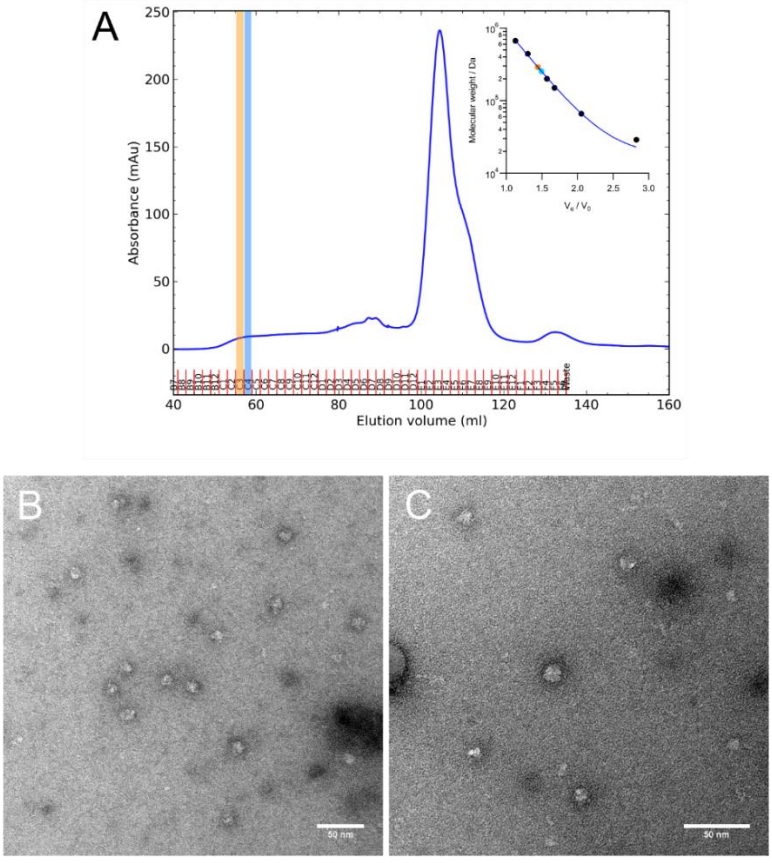


Figure 5.13. Purification of glutaraldehyde cross-linked POR ternary complex on a Sephacryl S300 size column. **A** resultant chromatogram, which is broadly similar to that in Figure 5.12. Highlighted are some high-molecular weight fractions that were imaged by EM in **A** and **B**. Inset, S300 calibration curve with fractions C3 and C4 plotted after calculating the molecular weight from the elution volume. **B-C**, EM micrographs of fractions C3 (**B**, orange) and C4 (**C**, blue).

5.4 Discussion

The quaternary organisation of MgCH and POR is key to their function within the cell and is currently poorly understood. Interactions between the subunits of MgCH have been characterised but the unknown overall architecture of this important protein machine hampers thorough understanding of the reaction mechanism; it is essential to know how the free energy of ATP hydrolysis in ChlI is transduced through ChlD to ChlH in order to drive chelation of magnesium and commit a photosynthetic cell to producing chlorophyll. The quaternary organisation of POR seems to be less integral to the reaction mechanism of catalysis, but it is clearly important *in vivo* during development of etioplasts. Currently there is no experimentally derived structure of POR bound to substrates, and the oligomerisation of the protein could hamper crystallisation attempts. Isolating one form of the POR ternary complex to be studied by cryo-EM would not only shed light on the oligomerisation mechanism of POR, but also on substrate binding and positioning, potentially aiding in understanding how this unique enzyme harnesses the light energy absorbed by Pchlide in order to drive catalysis.

In this chapter, progress was made towards purifying these complexes for study using EM, and while nothing suitable was produced for data collection using cryo-EM, further optimisations to the purification procedures could provide a suitable sample. The inherent heterogeneity of both of these complexes will be a major challenge to overcome; the AAA+ domains of ChlI and ChlD appear to form many different complexes and sub-complexes; the self-assembling nature and apparent wide range of oligomeric states of POR makes purifying a single species of POR ternary complexes from apo-POR challenging.

Multiple different methods were applied to try and form and stabilise the complexes. These included assembling the complex under assay-like conditions and then purifying the proteins; coexpression systems both in the native organism and in *E. coli*; and stabilisation through the use of cross-linking reagent. Many slight alterations in procedure resulted in slightly different results, though unfortunately nothing appropriate for taking further for single particle cryo-EM. Further optimisations and alterations to the protocols will allow for more development and hopefully the isolation of these complexes.

The use of glutaraldehyde to stabilise the complexes seemed successful to some extent, especially with MgCH, and with further optimisations to the protocol could allow for locking the dynamic MgCH into one state. This could also be expanded by using and optimising the GraFix protocol for use with these enzymes.

6 Investigating the reaction catalysed by POR using 2DES.

6.1 Introduction

The reaction catalysed by POR has been studied for many years, in part due to its biological importance in producing the important pigment chlorophyll, but also because of the unique mechanism of catalysis; it is one of only four discovered photoenzymes, proteins that catalyse reactions using the input of light energy (Björn, 2018). That the enzyme-substrate complex can be assembled in the dark and the reaction initiated with a light pulse has allowed for characterisation of the reaction mechanism to high temporal resolution, and extensive spectroscopic work has been conducted to try to understand this unique enzyme. With most enzymes, high-resolution temporal information is hard to acquire due to the often rate-limiting assembly of the enzyme-substrate complex or dissociation of the product from the enzyme. The spectral properties of the substrate, Pchlide, also aids with elucidating the reaction mechanism; distinct changes in both the fluorescence and absorption profile of the enzyme-substrate complex allows intermediates in the reaction to be monitored.

2-dimensional electronic spectroscopy (2DES), briefly reviewed in (Section 1.6.2), has revealed the possible relevance of electronic coherence to the photochemical steps of photosynthesis; it was shown that coherent states are sufficiently long-lived, at room temperature, to persist during the initial steps of electron transfer (Romero et al., 2014). Quantum coherence was originally observed for ultrafast steps in excitation energy transfer (Brixner et al., 2004, 2005), and then extended to the picosecond energy trapping steps of photosynthesis (Romero et al., 2014); thus it is timely to find out if quantum coherence also plays a role in the formation of product states in enzyme catalysis. Given the short timescales involved it was necessary to study an enzyme where the catalytic cycle can be triggered by femtosecond light pulses, and POR is an ideal model system. The ability to initiate the proton and hydride transfers catalysed by POR with a femtosecond laser presents a unique opportunity to perform the first 2D electronic spectroscopy on an enzyme and to see if the original coherence of the laser excitation is retained in the initial catalytic intermediates.

The work presented in this chapter was conducted with Dr Elisabet Romero (now at Institut Català d'Investigacló Química) at LaserLab Amsterdam, within the Vrije Universiteit after successful application for funding provided by LaserLab Europe. Dr Romero provided the expertise in conducting and analysing the 2DES experiments.

6.2 Sample preparation and experimental overview

In order to perform 2DES, large quantities of pure sample had to be produced. Each experiment required 2 ml of sample to an OD_{632} of *ca.* 0.3 in a 500 μm pathlength cuvette, with enzyme in excess so as to ensure that all Pchlide present was bound by POR. Sample would flow from a reservoir through the cuvette using a peristaltic pump, so that fresh POR ternary complex was exposed for each excitation event. POR was produced as in Sections 2.6 and 2.8.1, whilst Pchlide was prepared as in Section 2.9 (Figure 6.1).

2DES reveals coherent effects using three ultra-short and spectrally broad laser pulses each separated by controllable time delays (Figure 6.2). Fourier transform with respect to coherence time, τ (time between the first and second pulse) and rephasing time, t (time between the third and the emission of the photon echo signal) yields 2D spectra which correlates emission and excitation frequencies as a function of population time, T (time between the second and third pulse). Therefore, 2DES reveals the presence of populated states (signals on the diagonal of the 2D spectra) as well as correlations/couplings (coherences) between different spectral bands in the form of cross-peaks (signals below the diagonal in the 2D spectra) with high spectral and time resolution. The present study focuses on investigating the dynamics of populations and coherences by examining three different samples: Pchlide alone, solubilised in POR assay buffer (Section 2.13.1) (Pchlide); the non-productive POR ternary complex (POR_NADP_Pchlide); and the productive POR ternary complex (POR_NADPH_Pchlide).

In addition, the oscillation frequencies possibly present in the 2D traces (2D spectral point as a function of T) can be analysed to identify specific vibrational modes resonant with excitonic energy gaps. By comparing the oscillation frequencies between different positions in the 2D spectra, information could be obtained about the mechanisms that stabilise high energy states along the catalytic pathway in POR.

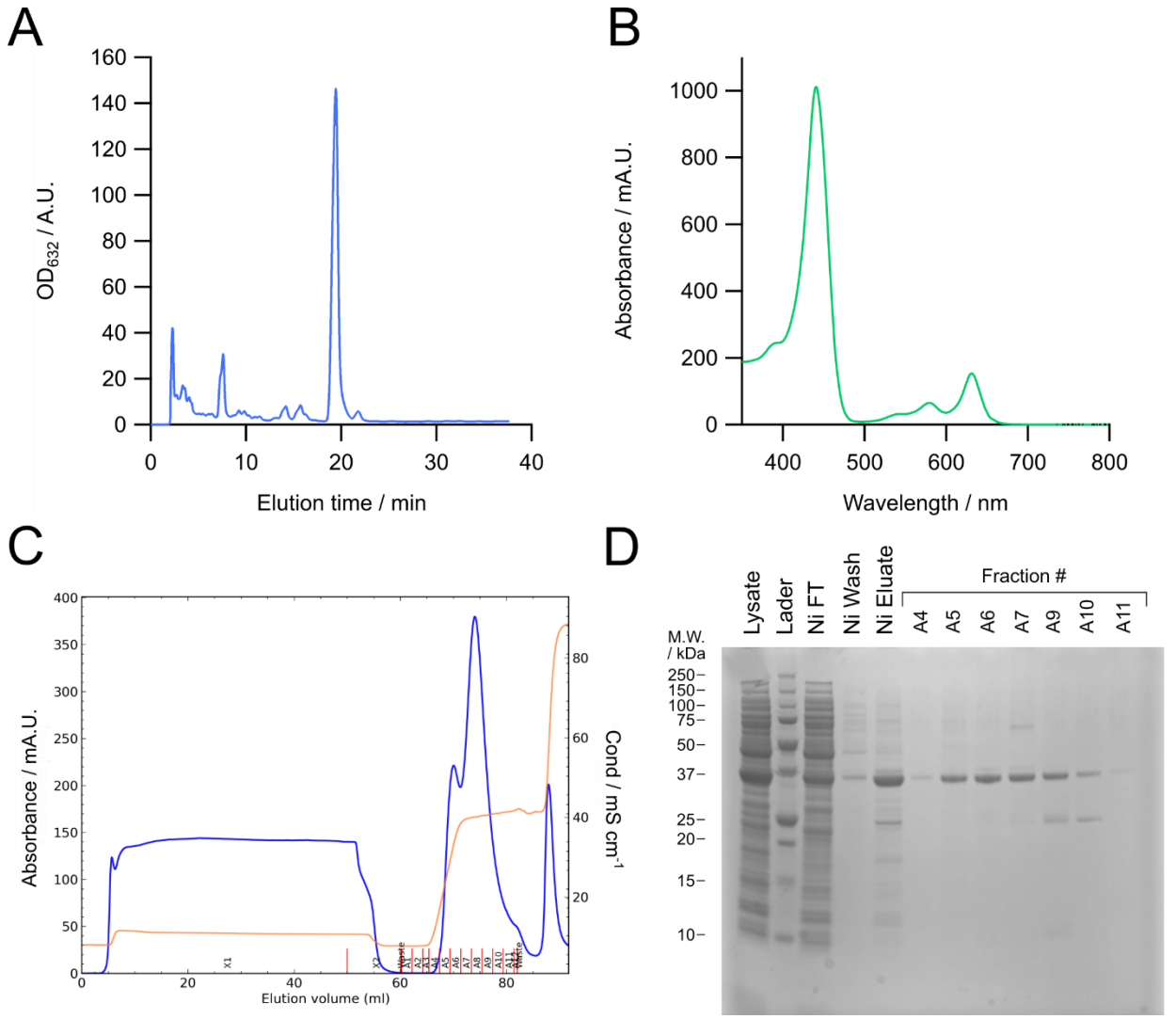


Figure 6.1. Representative purification of Pchlide and POR. **A**, HPLC trace of pigment from *Rba. Sphaeroides* strain Δ*bchJ* grown, extracted and purified as detailed in Section 2.9; pure Pchlide pigment present in the large peak at 20 min elution time. **B**, Spectra of 20 min eluate, showing purity of the Pchlide pigment. **C**, purification of POR on a Hitrap SP HP cation exchange column as detailed in Section 2.8.1.3; OD_{280} trace, blue; conductivity, orange. **D**, SDS PAGE of the POR purification, including immobilised nickel affinity chromatography step; Ni FT: Flowthrough from Ni-TED column; fractions from the SP HP column in **C**.

The absorption profile of each sample was collected and compared against the laser profile of the 2DES setup (Figure 6.3). The laser profile utilised in the 2DES experiments is the blue-most range that can be obtained with the home-built non-collinear optical parametric amplifier (NOPA) which uses the second harmonic (515 nm) of the laser fundamental (1033 nm) to amplify a portion of a white light continuum (Manzoni and Cerullo, 2016). While the laser light does not cover the blue side of the Q_Y absorption band of Pchlide, it is sufficient to excite the red side of this band to allow investigation of the light-dependent reduction catalysed by POR.

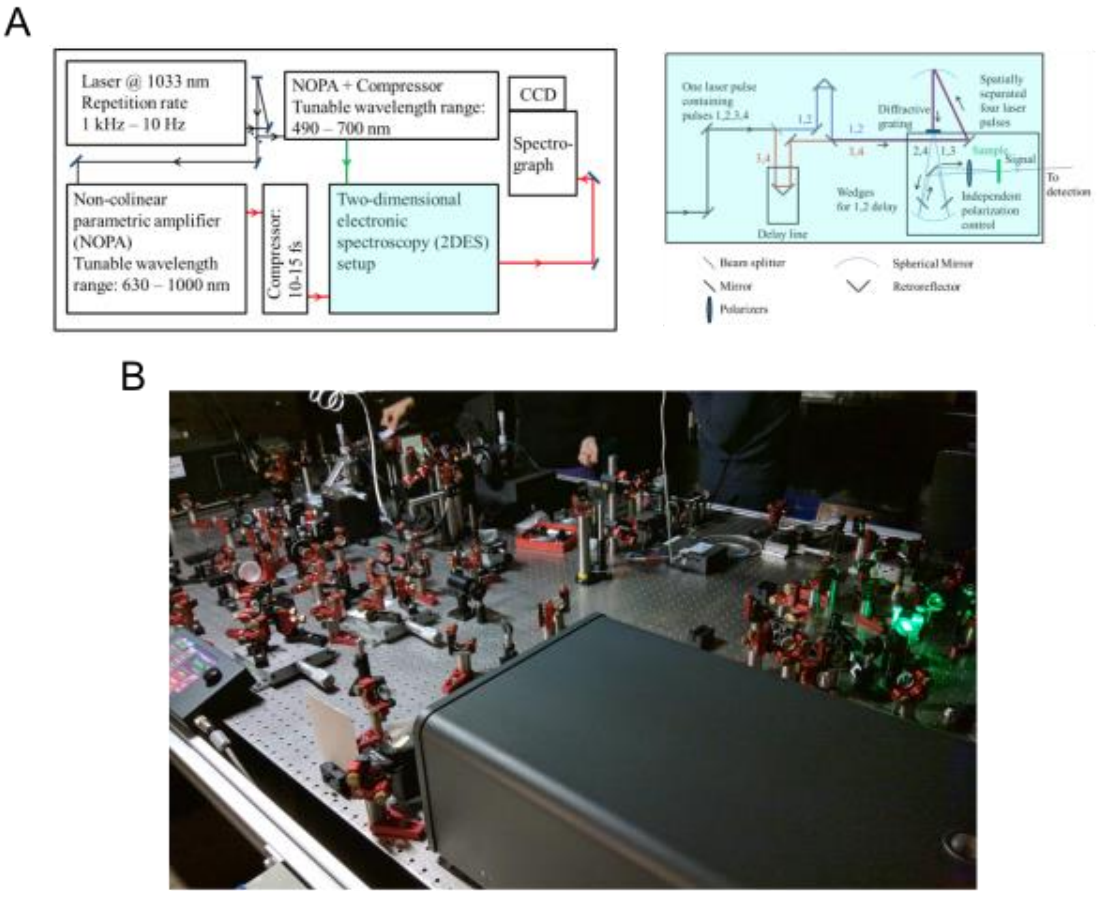


Figure 6.2. Schematic representation of the 2DES setup. **A**, all of the optical systems used within the setup with a more detailed schematic of the optics used in the 2DES specific part of the system (pale blue). **B**, photograph of the whole 2DES setup.

Also of note is that these experiments were all conducted at room temperature; the POR protein used is from *T. elongatus*, a thermophilic cyanobacterium, and as such assays would normally be conducted at 55 °C. However, performing the 2DES at the lower temperature was deemed acceptable, as the experiments had to run over a long period of time (~10 hours) so heating to 55 °C would accelerate the reaction, potentially using up all substrate as well as reducing the stability of POR. In addition, the light-dependent reactions occur even at cryogenic temperatures (Heyes and Hunter, 2004), so conducting the experiments at the lower temperature should not affect the investigated reaction.

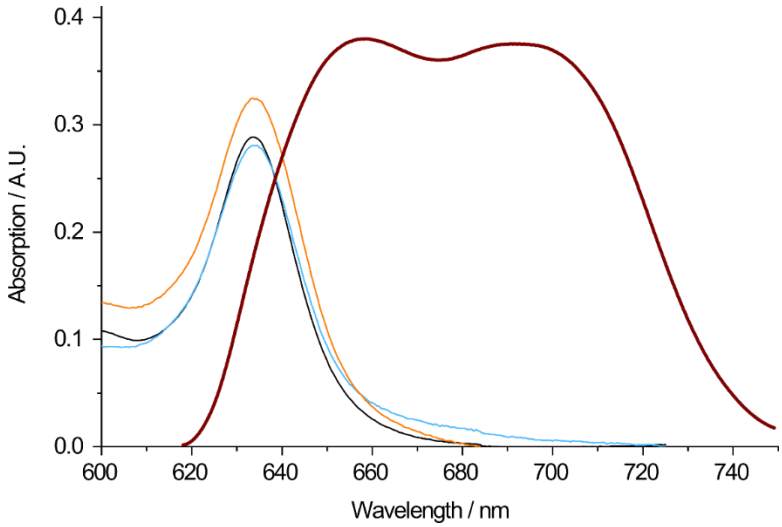


Figure 6.3. Room temperature absorption spectra of laser profile used in the 2DES (brown) compared to each of the samples investigated. Pchlide alone (black); POR_NADP_Pchlide (cyan); and POR_NADPH_Pchlide (orange). Data not normalised.

6.3 Pchlide alone

All of the following experiments were conducted in the dark. Dry, pure Pchlide was solubilised as in Section 2.13.1 and diluted to the appropriate concentration for the 2DES experiment. The solution was added to the cuvette for use within the 2DES system. Room temperature 2DES at a range of coherence and population times and was performed as described in Section 2.17. Data collected at -24 fs to 100 fs population time suffered from the so-called "coherent artefact", a signal appearing due to coherent interactions of the three laser pulses used in the experiment around time zero, and is typical for time-resolved experiments. 2D data in this range were not suitable for analysis; this was the case for all samples.

The absolute total 2D spectra of Pchlide alone (Figure 6.4) reveal a maximum signal on the diagonal at 645 nm; this is higher than the 634 nm maximum when conducting standard 1D spectroscopy but it is the result of exciting Pchlide on the red side of its absorption (Figure 6.3). It is a convolution of the laser profile and the absorption spectrum of Pchlide, which applies to all further experiments also. The spectral tail that extends from λ_τ = 645 nm further to the red along the λ_t axis is attributed to the vibrational band of Pchlide. From the 2D spectra, but more clearly from the spectral evolution of the diagonal peak (Figure 6.5), it can be seen that amplitude increases as a function of *T*. Both the spectral shape and evolution are consistent with chlorin chromophores in solution; they have been observed before for chlorophyll *a* (Bricker et al., 2015) and pheophorbide in methanol (personal communication, Dr Romero). The single peak on the diagonal and absence

of any off-diagonal peaks mean that there is no coupling occurring between any of the pigment molecules.

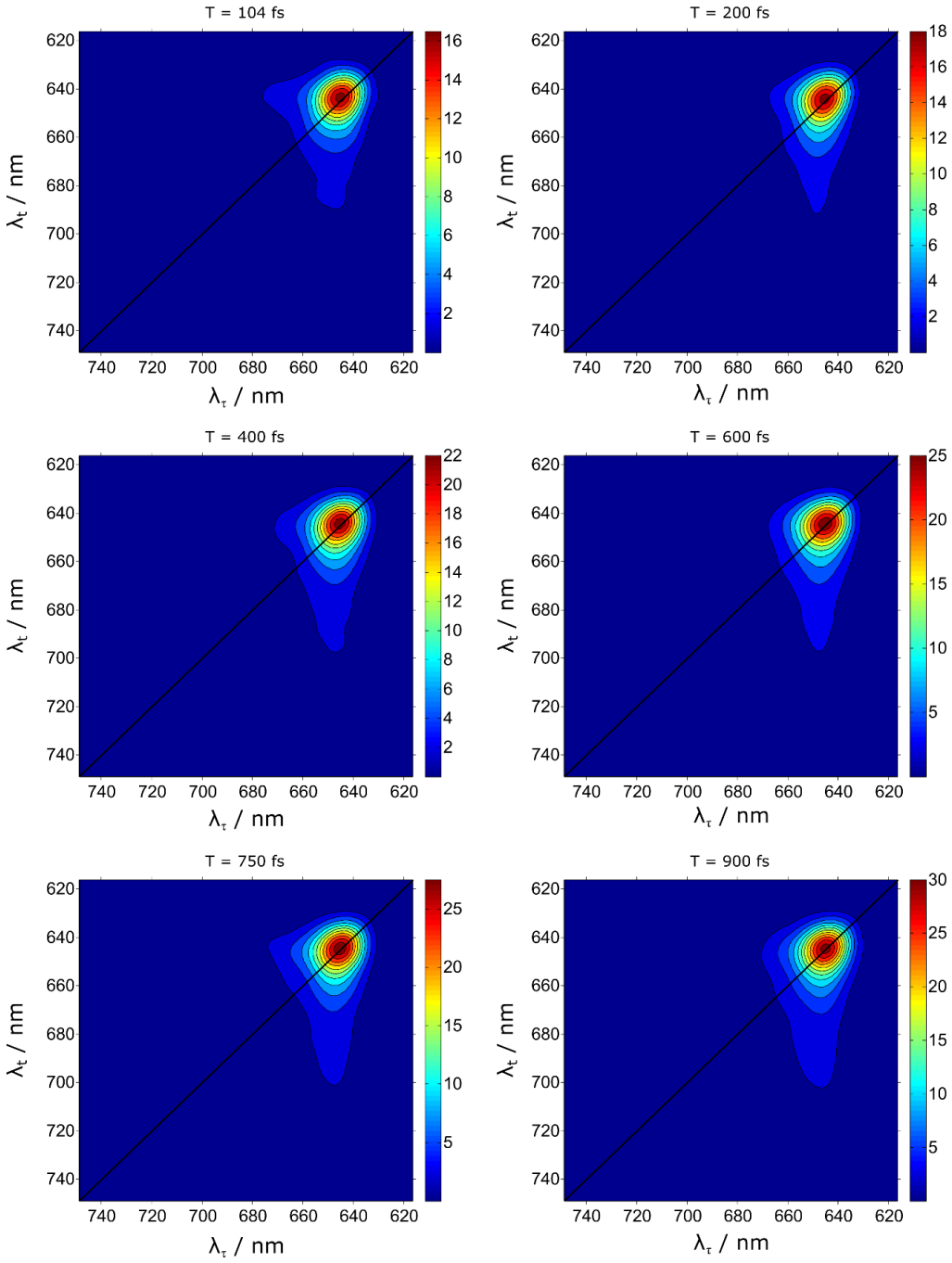


Figure 6.4. Absolute total 2D spectra of the Pchlide alone sample at several representative population times (*T*). Amplitude values (A.U.) represented as heat map (scale on the right of each 2D spectra).

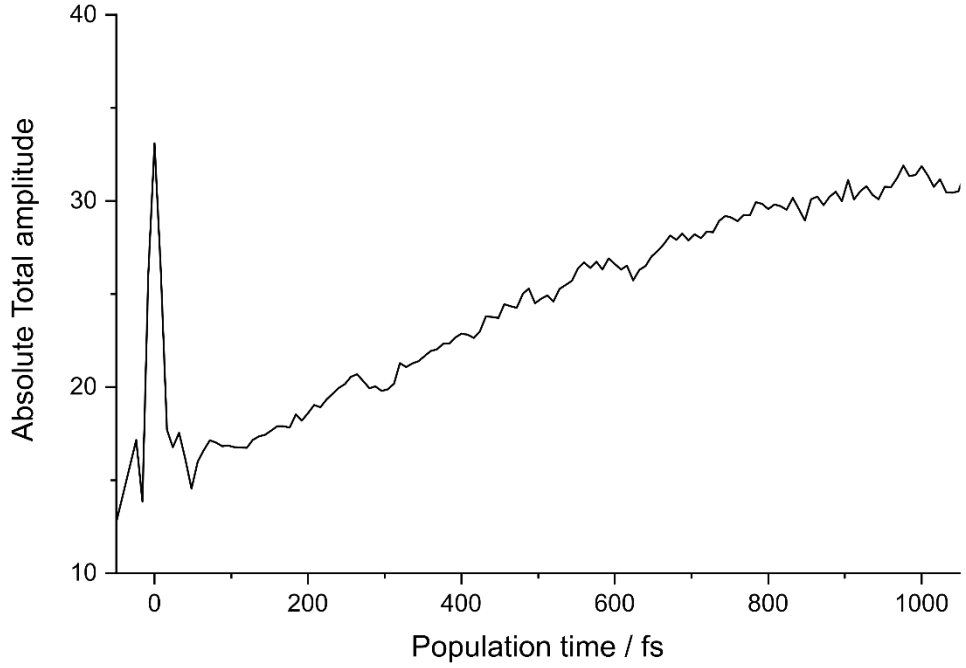


Figure 6.5. 2D spectral evolution of the maximal amplitude point of the 2D spectrum for the Pchlide alone sample, (λ_τ, λ_t) = (645, 645) nm. Absolute total amplitude increases as a function of *T*, consistent with previous results (Bricker et al. (2015); personal communication, Dr Romero).

6.4 The non-productive POR ternary complex

Dry pure Pchlide was solubilised in the same was for the previous sample and mixed with POR and $NADP^+$, to a final concentration of 20 μM Pchlide (OD_{632} = 0.3 in 500 μm pathlength cuvette; ε = 30.4 mM^{-1} cm^{-1}), 100 μM POR, and 250 μM $NADP^+$ such that 98 % of Pchlide was bound to POR, and 99.99 % of POR was bound to $NADP^+$ (using previously published K_d values (McFarlane et al., 2005)). The sample was heated to 55 ℃ for 10 min to promote substrate binding and clarified by centrifugation at 1500 x *g* for 10 min at 4 ℃, the supernatant was then loaded into the 2DES apparatus. It is of note that no green pellet formed during the centrifugation step, indicating complete binding by POR. When 2DES was performed, an unexpected problem occurred for data acquisition further than 1000 fs *T*: a white spot (most likely protein precipitate/aggregate) formed on the cuvette internal surface, exactly on the position where the laser was hitting the cuvette. Therefore, the 2D data for $T > 1$ ps have been rejected as this white precipitate was likely bleached, insoluble protein, which scattered excitation light and attenuated the light incident on soluble protein-pigment complexes. It is not clear why this occurred, but it may be due to highly-excited Pchlide degrading or generating radicals, contributing to protein denaturation.

The 2D spectra (Figure 6.6) of this sample are broadly similar to the Pchlide alone. However, the maximal peak of Pchlide absorption is shifted to the red along the λ_t axis; (λ_τ, λ_t) = (645, 650) nm, a shift typical of Pchlide bound to POR (Heyes et al., 2002). In addition, the vibrational band of Pchlide is more pronounced in this sample. The spectral evolution of the POR_NADP_Pchlide sample differs more from Pchlide alone, however, showing an exponential decay in amplitude as a function of *T*, which has also been observed for the bacteriochlorin pair in LH1 (Ferretti et al., 2014), potentially indicating a dark state of pigment formed within the protein environment that promotes faster excited state decay. This decay of amplitude within a protein environment has also been observed for pheophorbide bound to maquette protein (personal communication, Dr Romero).

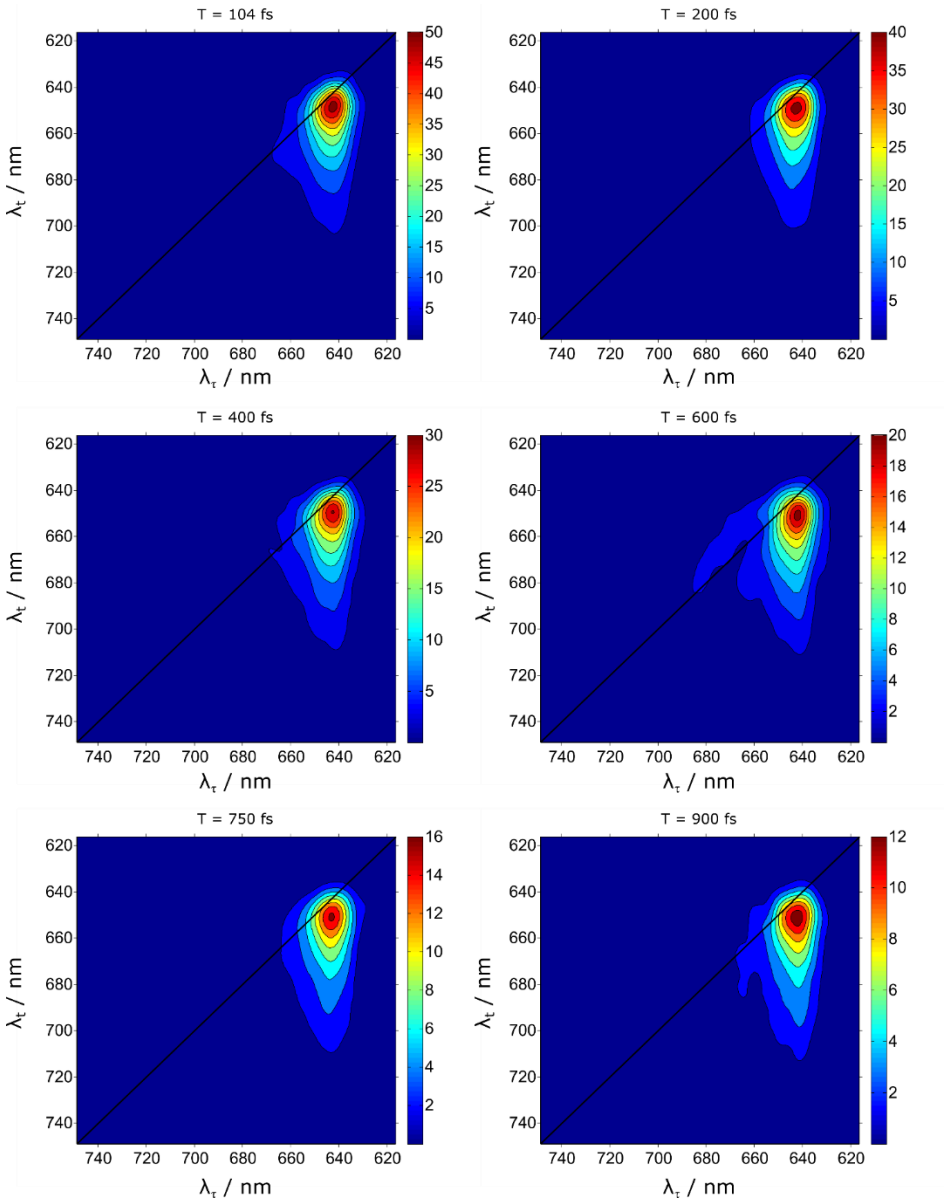


Figure 6.6. Absolute total 2D spectra of the non-productive POR ternary complex (POR_NADP_Pchlide) at several representative population times (*T*). Amplitude values represented as heat map, as in Figure 6.4.

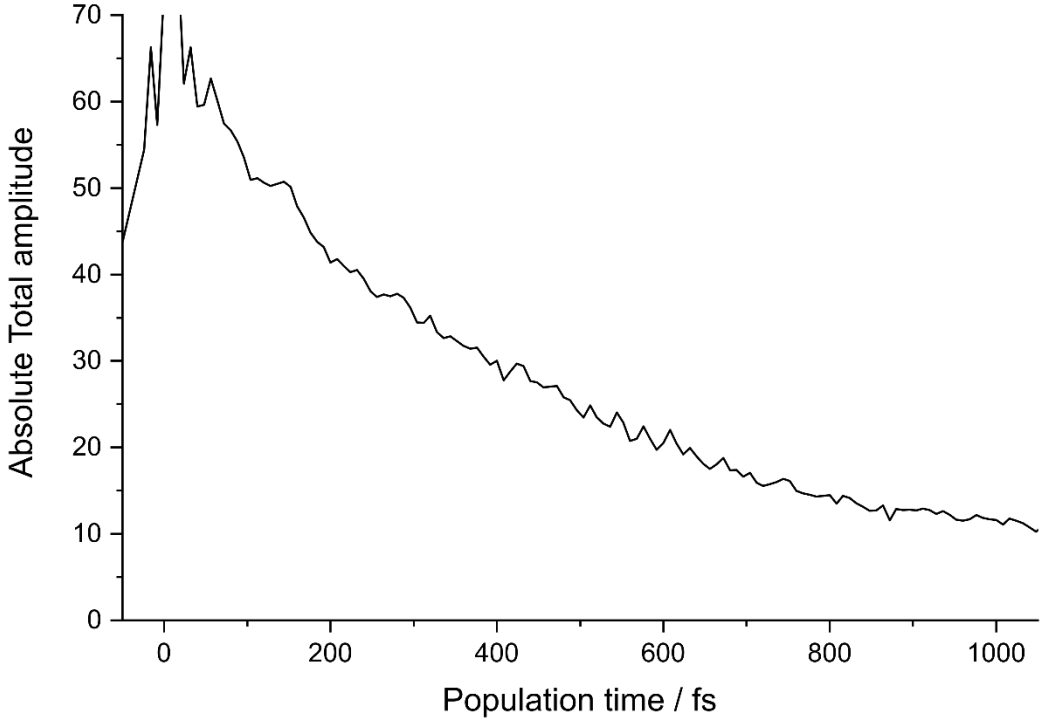


Figure 6.7. 2D spectral evolution of the maximal amplitude point (λ_τ, λ_t) = (645, 650) nm showing the exponential decay of the absolute total amplitude as a function of *T*.

6.5 The active POR ternary complex

The active POR ternary complex was assembled in the same manner as the inactive complex (Section 6.4) but with NADPH instead of $NADP^+$ which was then loaded into the 2DES setup and subjected to 2DES. The 2D spectra are substantially different from previous samples (Figure 6.8); Chlide formation as a diagonal peak at 675 nm was observed very early in the experiment and an off-diagonal cross-peak between Pchlide and Chlide (λ_τ, λ_t) = (645, 675) nm was present. The presence of this cross-peak points to several possibilities: i) Coherence/ coupling between the reactant Pchlide and Chlide product; ii) the formation of an intermediate state (potentially a Pchlide triplet state) prior to formation of Chlide; or iii) the formation of an intermediate state (perhaps a resonant form of Pchlide and Chlide) which occurs during the conversion of Pchlide to Chlide. With regard to (i) above it is unlikely that there is coupling between Pchlide on one monomer of POR and the Chlide product on another monomer within an oligomer, as the predicted distance between active sites is too far for coupling to be observed.

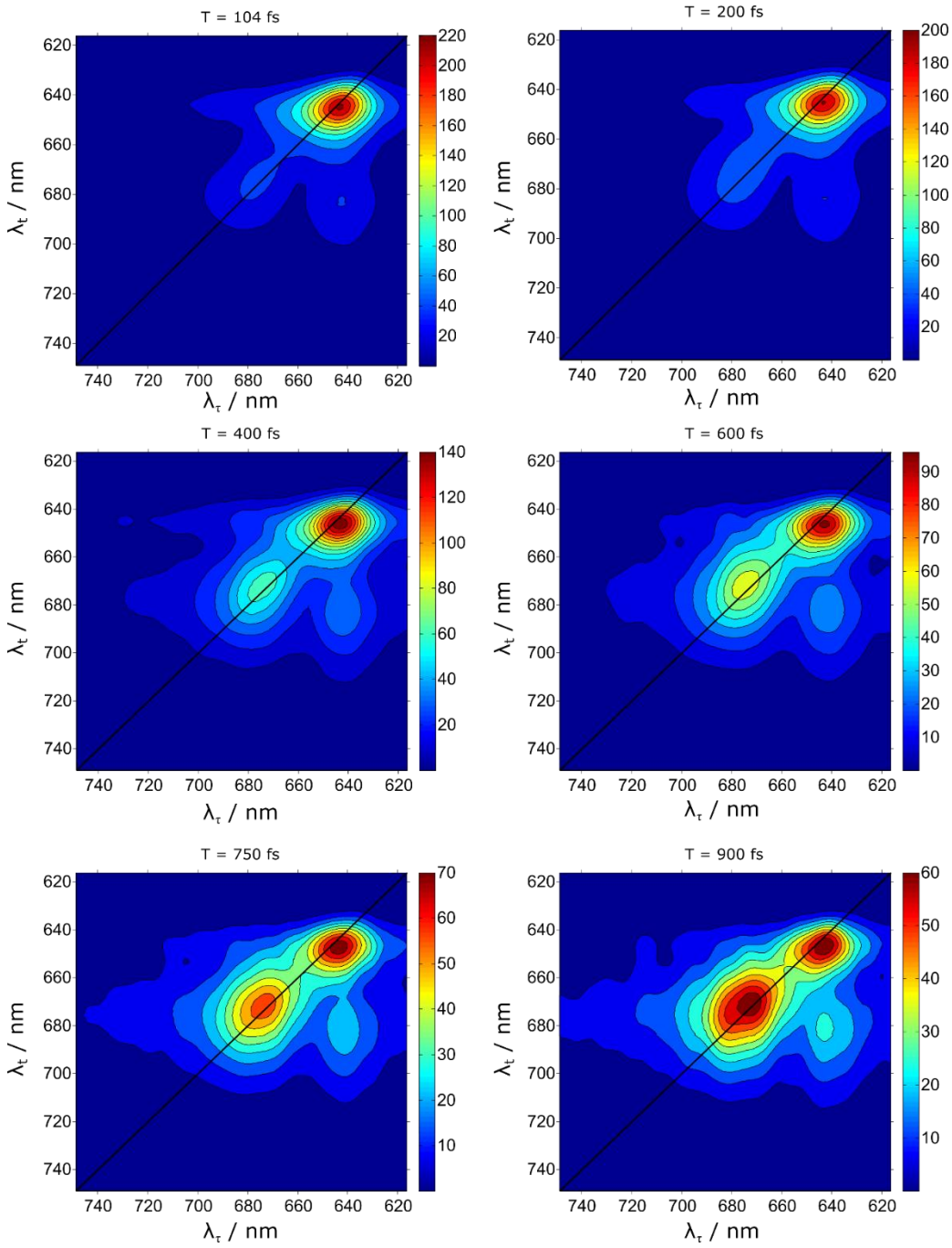


Figure 6.8. The absolute total 2D spectra of the productive POR ternary complex (POR_NADPH_Pchlide) at representative population times, revealing the presence of a cross-peak between Pchlide and Chlide, as well as changes in the amplitudes of the three prominent peaks as a function of *T*.

To distinguish between these hypotheses, the dynamics of the 2D spectra were investigated at three points, the Pchlide diagonal at 645 nm; the Chlide diagonal at 675 nm and the cross-peak at (λ_τ, λ_t) = (645, 675) nm (Figure 6.9). The dynamics of all three points show three distinct time courses: 100 – 330 fs, 330 – 800 fs, and 800 – 1000 fs. Within the first time period the Pchlide 2D trace

decays while the Chlide 2D trace grows, and the cross-peak remains constant; during the second the Pchlide signal decays, the Chlide signal remains almost constant and the cross-peak amplitude decays; and during the third the Pchlide amplitude decays at a much slower rate than during the first time period, the Chlide 2D signal increases at the same rate that Pchlide decays, and the cross-peak remains nearly constant.

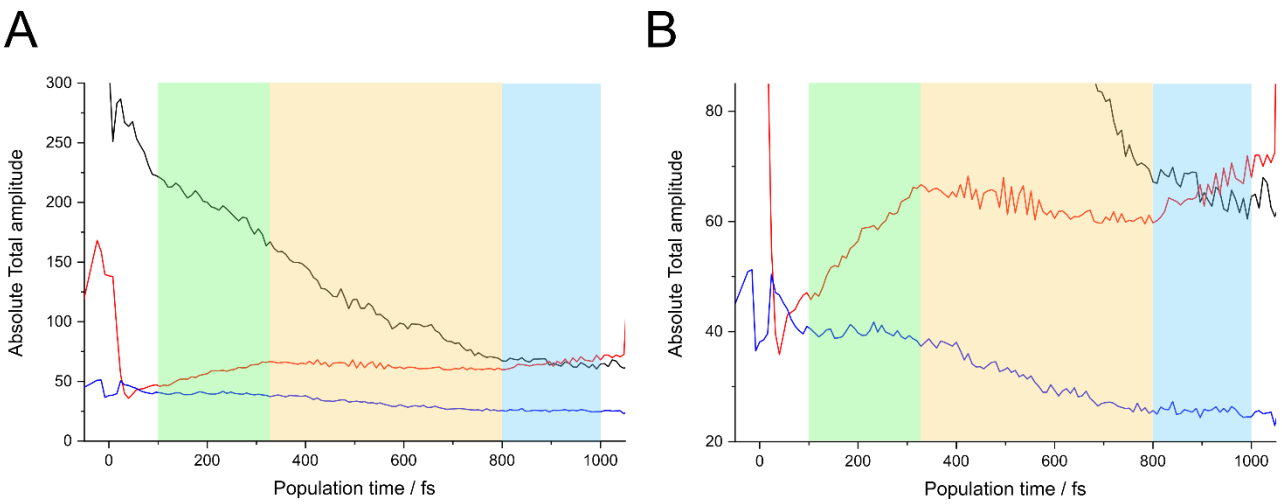


Figure 6.9. 2D spectral evolution of the three peaks described within the text. Black, the Pchlide diagonal peak; red, the Chlide diagonal peak; blue, the cross-peak of Pchlide and Chlide. **A** and **B** differ only in the scale of the absolute amplitude for a clearer visualisation of the three distinct time periods mentioned within the text (coloured shading).

The measured dynamics indicate that the growth of the Chlide signal, which lies on the diagonal at 675 nm, is accompanied by the decay of the Pchlide signal on the diagonal at 645 nm; it is noteworthy that during the transfer of amplitude from the 645 to the 675 nm band the amplitude of the off-diagonal cross-peak between Pchlide and Chlide (λ_τ, λ_t) = (645, 675) nm remains constant, relative to the amplitudes of the separate Pchlide and Chlide signals. This is an interesting observation, but it does not help to distinguish between options (ii) and (iii) proposed above to explain the data. It appears likely that this off-diagonal cross-peak denotes the formation of an intermediate state in the conversion of Pchlide to Chlide. Most likely, the complex dynamics observed are a combination of different processes, presumably the formation of Chlide together with the decay of the Pchlide excited state.

Within the decays, the presence of periodic oscillations as a function of *T*, so-called "quantum beats", can be assessed to identify vibrational modes resonant with excitonic energy gaps; these oscillations are not clear from the decay data in Figure 6.9 and attempts to generate 2D frequency maps (which only display spectral features that are coherent at that frequency) resulted in spectra

that were not dissimilar from the original 2D spectra (at low frequencies), or very noisy (at high frequencies). It was therefore concluded that no clear quantum beats are present within this dataset.

The ultrafast formation of a signal corresponding to Chlide after the first 100 fs following excitation is inconsistent with the much slower μs formation of Chlide previously observed (Scrutton et al., 2012). This formation is also much faster than the reported formation of the triplet state of Pchlide (Brandariz-De-Pedro et al., 2017). Therefore, the third proposed hypothesis is favoured: the formation of a resonant form between Pchlide and Chlide (not quite Pchlide, and not quite Chlide) as a precursor of the final Chlide product. While these observations are not conclusive, if the hypothesis is correct then the presented work demonstrates a previously unknown intermediate state in the formation of Chlide on an ultrafast time scale (several hundreds of fs), much faster than previously observed.

6.6 Discussion

The light-driven reduction of Pchlide catalysed by POR is a unique and complex reaction that occurs on an ultrafast time-scale; the requirement of light for initiation does, however, present interesting opportunities of studying enzymatic catalysis on these ultrafast time-scales. The reaction has been extensively studied using a variety of spectroscopic methods, both time-resolved and static, but it is still unclear exactly how POR performs its catalysis, exemplified by multiple studies suggesting different mechanisms of action (Archipowa et al., 2018; Heyes et al., 2014; Sytina et al., 2008). With the work presented here, a novel intermediate produced on the time-scale of several hundred femtoseconds after excitation may also exist.

Whilst this work showed no evidence of quantum beats, preventing insight into potential stabilising mechanisms of POR, the presence of a cross-peak when the active complex is excited and catalysis promoted is interesting. It is difficult to assess exactly what has caused the cross-peak, but the spectral evolution of the three key spectral features is certainly linked, suggesting that it is catalytically relevant. More work would be required to confirm these results; conducting the same experiment at 77 K could reveal oscillations in the spectral evolution that are hidden at room temperature. However, there are difficulties with this that do not apply to the photosynthetic complexes that have previously been studied; the non-reversible reduction of Pchlide to Chlide adds another dynamic that must be accounted for, whereas photosynthetic pigments “reset” after excitation. Indeed, this dynamic has made the current work challenging to process and interpret, though

the procedure used here should have ensured that fresh unexposed POR ternary complex was exposed to each pulse of excitation energy. In an ideal situation, an active POR ternary complex would flow from a reservoir through the cuvette for excitation and then either be disposed of or somehow recycled. Producing that much protein and substrate would be very time consuming, however.

The presented work is the first attempt at applying the challenging and complex technique of 2DES to investigate enzyme catalysis. It is fascinating to observe the formation of Chlide on a previously undiscovered ultrafast time-scale and with further work may reveal that more quantum phenomena are involved in biological systems.

7 Conclusions and future work

Magnesium chelatase and protochlorophyllide oxidoreductase are important enzymes, and consequently they have been studied for several decades. The work presented in this thesis adds more biochemical and structural information about these complexes, both of which still lack a complete structural characterisation owing to multiple factors, touched upon in the previous chapters, including dynamic formation and dissolution; multiple substrate binding events; and the ability to form multiple non-productive oligomeric states.

The discovery of ChlD as the mediator between ATPase and chelation activity in MgCH opens further avenues of research to understand the specific mechanism behind this transduction of energy to promote chelation. An analogy with dynein, the motor protein that transports cargo along microtubules, is tempting to make; changes in conformation within the AAA+ ring with nucleotide binding and hydrolysis are translated across long distances through a mediating protein to promote certain structural states (Schmidt et al., 2012). Structural characterisation of the proline-rich acidic linker domain of ChlD may shed some light on the mechanism.

The application of XLMS to the ChlH-ChlD complex could be expanded to other members of the MgCH machine such as ChlH and Gun4, or ChlI and ChlD, to aid characterisation of the active complex. Should a preparation of the whole MgCH suitable for analysis using cryo-EM be achieved, utilising XLMS alongside the EM work would potentially allow for easier reconstruction of the complex, and act as an additional validation method, as has been used recently (Schmidt and Urlaub, 2017). Even without EM work, using the XLMS protocol and carefully considered kinetic titrations (much like the ones used in Chapter 3) could result in a structural analysis of the whole, active MgCH complex.

Substrate binding within the chelating subunit, ChlH, was explored extensively in Chapter 4, including preliminary structural characterisation of ChlH in a substrate bound form. Further binding studies using tryptophan fluorescence of the key mutants investigated would also shed more light on the mechanism of binding and chelation. It could also be of interest to test whether the presence of the other subunits affect substrate binding. In addition, future work could determine the structure of ChlH in complex with the substrate and product, to aid in developing a more comprehensive understanding of the reaction mechanism. Furthermore, information about the conformational state of ChlH under more native conditions could be revealed, which has thus far been inconsistent with

the crystal structure; even apo-ChlH appears to be in a less extended conformation than is observed in the crystal structures. Alternative methods of producing cryo-EM grids suitable for data collection will have to be attempted in order to generate high-resolution density maps and an atomic model of the ChlH in a more native state. This could be achieved using a novel freezing method whereby a piezoelectric system is used to spray a fine mist of protein solution onto grids for plunge freezing, removing the physically harsh and inconsistent blotting step from the process (Dandey et al., 2018; Noble et al., 2018).

The presence of an interesting 2-D class resulting from the negative stain data of the ChlI-ChlD complex may provide evidence for the productive ChlID complex utilised to drive chelation (Section 5.2.3.2, Figure 5.8). The hexameric ring structure with extra density protruding from one of the subunits could be a $ChlI_5$-$ChlD_1$ complex which could interact with ChlH via the protruding density which is presumably the integrin I-like domain of ChlD. Preliminary native mass-spectrometry data conducted within the lab group has also suggested a stoichiometry of $ChlI_5$:$ChlD_1$:$ChlH_{1/2}$ (data not shown). Previous kinetic analysis conducted by Adams (2012) also provided evidence of a $ChlI_5$-$ChlD_1$ complex; when ChlD is titrated into a constant [ChlI] as well as all other components of the MgCH reaction, maximal activity is observed at a 5:1 ratio of [ChlI]:[ChlD] (Adams, 2012).

Leading on from this, and with other novel experimental evidence presented within this thesis, a new mechanism of the MgCH enzyme can be proposed (Figure 7.1). ChlH binds protoporphyrin IX in the determined active site (potentially with the assistance of Gun4; Chapter 4) and magnesium is bound, delivered vectorally through the tunnel of residues determined in Section 4.2.4. ChlI binds $MgATP^{2-}$ and magnesium, whilst ChlD binds $MgATP^{2-}$ or MgADP (determining which will require more work) and a $ChlI_5D_1$ complex forms with interactions mediated through the AAA+ domains (Adams et al., 2016b). This complex can interact with ChlH via the integrin I-like domain of ChlD (Chapter 3; Farmer et al., 2019) and a conformational change, induced by ATP hydrolysis in ChlI is transferred through ChlD to ChlH to promote chelation in ChlH. The product, Mg-PPIX is then released or, more probably in native systems, retrieved by Gun4 and passed onto ChlM, which is most likely closely associated, if not interacting directly with ChlH. It is unclear whether the $ChlI_5$-$ChlD_1$ complex would disassemble after supplying the driving force to ChlH, or whether it would remain intact, bind new substrates, and move onto a new ChlH molecule. The latter scenario could perhaps be supported by the kinetic evidence from Chapter 3 suggesting that

one ChlID complex can interact with and power ten ChlH subunits under saturating conditions (Section 3.3.2).

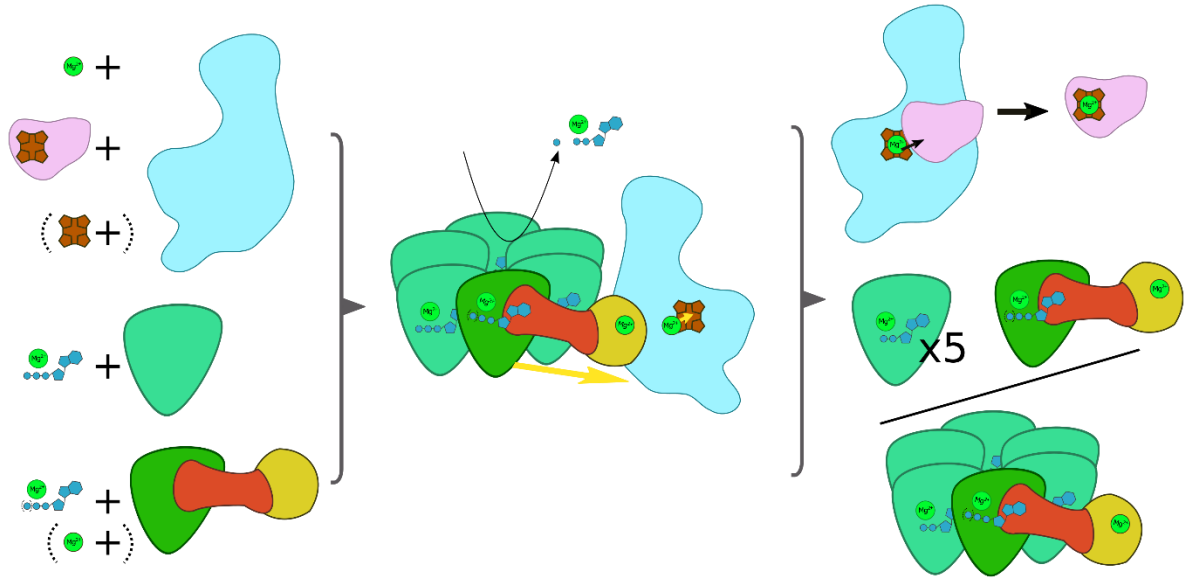


Figure 7.1. The proposed reaction scheme of MgCH. Colour scheme as in Figure 1.18. Briefly, ChlH binds its substrates, P_{IX} and Mg^{2+} and associates with a $ChlI_5$-$ChlD_1$ complex that has itself assembled upon binding of nucleotide and Mg^{2+}. ATP hydrolysis in ChlI is detected by and transferred through ChlD to ChlH (yellow arrow) to promote chelation. The ChlID complex disassociates from ChlH to either break apart itself or move onto a new ChlH to power whilst MgP_{IX} is released from ChlH with or without the assistance of Gun4. See text for detailed description of mechanism.

Much biochemical work has been conducted in previous studies and within this thesis to characterise POR and MgCH, but completing investigations into their quaternary organisation as started in Chapter 5 could be key to developing a more complete understanding of the reaction mechanism of these enzymes, especially with regard to MgCH. Promising progress in this direction was made, and given more time and effort, it is certainly possible that cryo-EM of the intact MgCH complex will be developed to obtain an atomic model.

Future work on purifying the whole MgCH enzyme could include developing the tandem affinity protocol initiated in Chapter 5, enabling a strict selection procedure for the intact chelation machine and further optimisations of the cross-linking protocol to enable more effective stabilisation. Using the full set of substrates may be necessary in order to promote full complex formation, as may the addition of the Gun4 protein. Gun4 is often overlooked when conducting *in vitro* work as it is not required for activity but *in vivo* this protein is essential for chlorophyll production so its

addition may increase the stability of the MgCH complex. Other proteins in the chlorophyll biosynthetic pathway could also form a complex with MgCH *in vivo*, with ChlH known to promote activity in ChlM (Shepherd et al., 2005). Indeed, a method not dissimilar to that used to study the ChlH-ChlD interaction in Chapter 3 could be conducted to investigate ChlH-Gun4 and ChlH-ChlM complexes. It is probable that *in vivo* the various enzymes of the chlorophyll biosynthesis pathway form a metabalon to enable efficient and safe production of the highly reactive chlorins.

For POR, optimisation of the size exclusion chromatography protocol could be a promising avenue to pursue as there seems to be some consistency in the formation of the smaller 5-7mers of POR. A complex of this size would be perfect for studying using cryo-EM, the technique that is perhaps required to solve the substrate bound structure of POR. Utilising cryo-electron tomography to analyse the *in vivo* POR complexes involved in photosynthetic membrane maturation could also be an interesting route to explore. Recent advances in cryo-electron tomography could mean that the structures of PLB in their native organisms could be resolved at a higher resolution than previously possible, which would enable further understanding of the role they play in the development of etioplasts.

POR is an extremely interesting enzyme to study due to the unique light-initiation of catalysis, which allows access to ultra-high resolution temporal information about the reaction mechanism. Complex spectroscopic work continues to be conducted on POR, including within Chapter 6 of this thesis. Careful consideration would have to be given to determine how to approach future work using 2DES on POR. Completing the work started in Chapter 6 would result in a very novel study, as the first use of 2DES on an enzyme, and an interesting observation of chemistry happening within an enzyme system on a truly ultra-fast timescale, should the current hypothesis be accurate. Clearly, complex physical phenomena are occurring in this enzyme and the chemistry that it catalyses. An intermediate on the timescale observed is currently very difficult to explain and reconcile with previous data.

Ultimately, there are many avenues of further research that are possible and will develop understanding of these two biologically interesting and mechanistically unique enzymes.

References

Adams, N.B.P. (2012). The roles of ChlI and ChlD, the AAA+ subunits of Magnesium Chelatase. University of Sheffield.

Adams, N.B.P., and Reid, J.D. (2012). Nonequilibrium isotope exchange reveals a catalytically significant enzyme-phosphate complex in the ATP hydrolysis pathway of the AAA+ ATPase magnesium chelatase. Biochemistry *51*, 2029–2031.

Adams, N.B.P., and Reid, J.D. (2013). The allosteric role of the AAA+ domain of ChlD protein from the magnesium chelatase of Synechocystis species PCC 6803. J. Biol. Chem. *288*, 28727–28732.

Adams, N.B.P., Marklew, C.J., Qian, P., Brindley, A. a, Davison, P. a, Bullough, P. a, and Hunter, C.N. (2014). Structural and functional consequences of removing the N-terminal domain from the magnesium chelatase ChlH subunit of Thermosynechococcus elongatus. Biochem. J. *464*, 315–322.

Adams, N.B.P., Brindley, A.A., Hunter, C.N., and Reid, J.D. (2016a). The catalytic power of magnesium chelatase: a benchmark for the AAA + ATPases. FEBS Lett. *590*, 1687–1693.

Adams, N.B.P., Vasilev, C., Brindley, A.A., and Hunter, C.N. (2016b). Nanomechanical and thermophoretic analyses of the nucleotide-dependent interactions between the AAA+ subunits of magnesium chelatase. J. Am. Chem. Soc. *138*, 6591–6597.

Adhikari, N.D., Orler, R., Chory, J., Froehlich, J.E., and Larkin, R.M. (2009). Prophyrins promote the association of GENOMES UNCOUPLED 4 and a Mg-chelatase subunit with chloroplast membranes. J. Biol. Chem. *284*, 24783–24796.

Al-Karadaghi, S., Hansson, M., Nikonov, S., Jönsson, B., and Hederstedt, L. (1997). Crystal structure of ferrochelatase: The terminal enzyme in heme biosynthesis. Structure *5*, 1501–1510.

Al-Karadaghi, S., Franco, R., Hansson, M., Shelnutt, J. a, Isaya, G., and Ferreira, G.C. (2006). Chelatases: distort to select? Trends Biochem. Sci. *31*, 135–142.

Archipowa, N., Kutta, R.J., Heyes, D.J., and Scrutton, N.S. (2018). Stepwise hydride transfer in a biological system: insights into the reaction mechanism of the light-dependent protochlorophyllide oxidoreductase. Angew. Chemie Int. Ed. *57*, 2682–2686.

Armstrong, D.R. (2014). Structural and functional studies of the light-dependent protochlorophyllide oxidoreductase enzyme. University of Sheffield.

Axelsson, E., Lundqvist, J., Sawicki, A., Nilsson, S., Schröder, I., Al-Karadaghi, S., Willows, R.D., and Hansson, M. (2006). Recessiveness and dominance in barley mutants deficient in Mg-chelatase subunit D, an AAA protein involved in chlorophyll biosynthesis. Plant Cell *18*, 3606–3616.

Baaske, P., Wienken, C.J., Reineck, P., Duhr, S., and Braun, D. (2010). Optical thermophoresis for quantifying the buffer dependence of aptamer binding. Angew. Chemie - Int. Ed. *49*, 2238–2241.

Baker, M.E. (1994). Protochlorophyllide reductase is homologous to human carbonyl reductase and pig 20 beta-hydroxysteroid dehydrogenase. Biochem. J. *300 (Pt 2*, 605–607.

Beale, S.I. (1990). Biosynthesis of the tetrapyrrole pigment precursor, aminolevulinic acid, from glutamate. Plant Physiol. *93*, 1273–1279.

Beale, S.I. (1999). Enzymes of chlorophyll biosynthesis. Photosynth. Res. *60*, 43–73.

Beale, S.I., Gough, S.P., and Granick, S. (1975). The biosynthesis of delta-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley. Ann. Clin. Res. *72*, 2719–2723.

Begley, T.P., and Young, H. (1989). Protochlorophyllide reductase. 1. Determination of the regiochemistry and the stereochemistry of the reduction of protochlorophyllide to chlorophyllide. J. Am. Chem. Soc. *111*,

3095–3096.

Benach, J., Atrian, S., Gonzàlez-Duarte, R., and Ladenstein, R. (1999). The catalytic reaction and inhibition mechanism of Drosophila alcohol dehydrogenase: Observation of an enzyme-bound NAD-ketone adduct at 1.4 Å resolution by X-ray crystallography. J. Mol. Biol. *289*, 335–355.

Björn, L.O. (2018). Photoenzymes and related topics: an update. Photochem. Photobiol. 459–465.

Blankenship, R.E. (2014). Molecular mechanisms of photosynthesis (Wliey Blackwell).

Blankenship, R.E., Tiede, D.M., Barber, J., Brudvig, G.W., Fleming, G., Ghirardi, M., Gunner, M.R., Junge, W., Kramer, D.M., Melis, A., et al. (2011). Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. Science *332*, 805–809.

Block, M.A., Tewari, A.K., Albrieux, C., Maréchal, E., and Joyard, J. (2002). The plant S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase is located in both envelope and thylakoid chloroplast membranes. Eur. J. Biochem. *269*, 240–248.

Boddi, B., Lindsten, A., Ryberg, M., and Sundqvist, C. (1989). On the aggregational states of protochlorophyllide and its protein complexes in wheat etioplasts. 135–143.

Bogorad, L., and Granick, S. (1953). The enzymatic synthesis of porphyrins from porphobilinogen. Proc. Natl. Acad. Sci. *39*, 1176–1188.

Bollivar, D.W., Clauson, C., Lighthall, R., Forbes, S., Kokona, B., Fairman, R., Kundrat, L., and Jaffe, E.K. (2004). Rhodobacter capsulatus porphobilinogen synthase, a high activity metal ion independent hexamer. BMC Biochem. *5*, 1–12.

Böttcher, B., Wynne, S.A., and Crowther, R.A. (1997). Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. Nature *386*, 88–91.

Boynton, T.O., Daugherty, L.E., Dailey, T.A., and Dailey, H.A. (2009). Identification of Escherichia coli HemG as a novel, menadione-dependent flavodoxin with protoporphyrinogen oxidase activity. Biochemistry *48*, 6705–6711.

Brandariz-De-Pedro, G., Heyes, D.J., Hardman, S.J.O., Shanmugam, M., Jones, A.R., Weber, S., Nohr, D., Scrutton, N.S., and Fielding, A.J. (2017). Direct evidence of an excited-state triplet species upon photoactivation of the chlorophyll precursor protochlorophyllide. J. Phys. Chem. Lett. *8*, 1219–1223.

Bricker, W.P., Shenai, P.M., Ghosh, A., Liu, Z., Enriquez, M.G.M., Lambrev, P.H., Tan, H.S., Lo, C.S., Tretiak, S., Fernandez-Alberti, S., et al. (2015). Non-radiative relaxation of photoexcited chlorophylls: Theoretical and experimental study. Sci. Rep. *5*, 1–16.

Brindley, A.A., Raux, E., Leech, H.K., Schubert, H.L., and Warren, M.J. (2003). A story of chelatase evolution: Identification and characterization of a small 13-15-kDa “ancestral” cobaltochelatase (CbiXs) in the archaea. J. Biol. Chem. *278*, 22388–22395.

Brindley, A.A., Adams, N.B.P., Hunter, C.N., and Reid, J.D. (2015). Five glutamic acid residues in the C-terminal domain of the ChlD subunit play a major role in conferring Mg^{2+} cooperativity on magnesium chelatase. Biochemistry acs.biochem.5b01080.

Brixner, T., Mančal, T., Stiopkin, I. V., and Fleming, G.R. (2004). Phase-stabilized two-dimensional electronic spectroscopy. J. Chem. Phys. *121*, 4221–4236.

Brixner, T., Stenger, J., Vaswani, H.M., Cho, M., Blankenship, R.E., and Fleming, G.R. (2005). Two-dimensional spectroscopy of electronic couplings in photosynthesis. Nature *434*, 625–628.

Calvin, M., and Bassham, J.A. (1955). The photosynthetic cycle (Berkeley, CA).

Camadros, J., Ibrahamp, N.G., and Leverell, R.D. (1984). Kinetic studies of human liver ferrochelatase. J. Biol. Chem. *259*, 5678–5682.

Canniffe, D.P., Jackson, P.J., Hollingshead, S., Dickman, M.J., and Hunter, C.N. (2013). Identification of

an 8-vinyl reductase involved in bacteriochlorophyll biosynthesis in Rhodobacter sphaeroides and evidence for the existence of a third distinct class of the enzyme. Biochem. J. *450*, 397–405.

Canniffe, D.P., Chidgey, J.W., and Hunter, C.N. (2014). Elucidation of the preferred routes of C8-Vinyl reduction in chlorophyll and bacteriochlorophyll biosynthesis. Biochem. J. *462*, 433–440.

Castelfranco, P.A., Weinstein, J.D., Schwarcz, S., Pardo, A.D., and Wezelman, B.E. (1979). The Mg insertion step in chlorophyll biosynthesis. Arch. Biochem. Biophys. *192*, 592–598.

Chen, G.E., Hitchcock, A., Jackson, P.J., Chaudhuri, R.R., Dickman, M.J., Hunter, N., and Canniffe, D.P. (2016). Two unrelated 8-vinyl reductases ensure production of mature chlorophylls in Acaryochloris marina. J. Bacteriol. *198*, 1393–1400.

Chen, G.E., Canniffe, D.P., and Hunter, C.N. (2017). Three classes of oxygen-dependent cyclase involved in chlorophyll and bacteriochlorophyll biosynthesis. Proc. Natl. Acad. Sci. *114*, 6280–6285.

Chen, G.E., Canniffe, D.P., Barnett, S.F.H., Hollingshead, S., Brindley, A.A., Vasilev, C., Bryant, D.A., and Neil Hunter, C. (2018). Complete enzyme set for chlorophyll biosynthesis in Escherichia coli. Sci. Adv. *4*.

Chen, M., Schliep, M., Willows, R.D., Cai, Z.-L., Neilan, B.A., and Scheer, H. (2010). A red-shifted chlorophyll. Science (80-.). *329*, 1318–1319.

Chen, X., Wang, X., Feng, J., Chen, Y., Fang, Y., Zhao, S., Zhao, A., Zhang, M., and Liu, L. (2014). Structural insights into the catalytic mechanism of synechocystis magnesium protoporphyrin IX O-methyltransferase (ChlM). J. Biol. Chem. *289*, 25690–25698.

Chen, X., Pu, H., Fang, Y., Wang, X., Zhao, S., Lin, Y., Zhang, M., Dai, H.-E., Gong, W., and Liu, L. (2015a). Crystal structure of the catalytic subunit of magnesium chelatase. Nat. Plants *1*, 15125.

Chen, X., Pu, H., Wang, X., Long, W., Lin, R., and Liu, L. (2015b). Crystal Structures of GUN4 in Complex with Porphyrins. Mol. Plant *8*, 1125–1127.

Cheng, Y.C., and Silbey, R.J. (2006). Coherence in the B800 ring of purple bacteria LH2. Phys. Rev. Lett. *96*, 1–4.

Chidgey, J.W., Linhartová, M., Komenda, J., Jackson, P.J., Dickman, M.J., Canniffe, D.P., Koník, P., Pilný, J., Hunter, C.N., and Sobotka, R. (2014). A cyanobacterial chlorophyll synthase-HliD complex associates with the Ycf39 protein and the YidC/Alb3 insertase. Plant Cell *26*, 1267–1279.

Chidgey, J.W., Jackson, P.J., Dickman, M.J., and Hunter, C.N. (2017). PufQ regulates porphyrin flux at the haem/bacteriochlorophyll branchpoint of tetrapyrrole biosynthesis via interactions with ferrochelatase. Mol. Microbiol. *106*, 961–975.

Cutler, S.R., Rodriguez, P.L., Finkelstein, R.R., and Abrams, S.R. (2010). Abscisic acid: emergence of a core signaling network.

Dailey, H.A., and Dailey, T.A. (1996). Protoporphyrinogen oxidase of Myxococcus xanthus. *271*, 8714–8718.

Dailey, H. a, Dailey, T. a, Wu, C.K., Medlock, a E., Wang, K.F., Rose, J.P., and Wang, B.C. (2000). Ferrochelatase at the millennium: structures, mechanisms and [2Fe-2S] clusters. Cell. Mol. Life Sci. *57*, 1909–1926.

Dailey, H.A., Gerdes, S., Dailey, T.A., Burch, J.S., and Phillips, J.D. (2015). Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. Proc. Natl. Acad. Sci. U. S. A. *112*, 2210–2215.

Dailey, H.A., Dailey, T.A., Gerdes, S., Jahn, D., Jahn, M., O'Brian, M.R., and Warren, M.J. (2017). Prokaryotic heme biosynthesis: multiple pathways to a common essential product. Microbiol. Mol. Biol. Rev. *81*, 1–62.

Dandey, V.P., Wei, H., Zhang, Z., Tan, Y.Z., Acharya, P., Eng, E.T., Rice, W.J., Kahn, P.A., Potter, C.S.,

and Carragher, B. (2018). Spotiton: New features and applications. J. Struct. Biol. *202*, 161–169.

Davison, P. a, and Hunter, C.N. (2010). Abolition of magnesium chelatase activity by the gun5 mutation and reversal by Gun4. FEBS Lett. *585*, 183–186.

Davison, P.A., Schubert, H.L., Reid, J.D., Iorg, C.D., Heroux, A., Hill, C.P., and Hunter, C.N. (2005). Structural and biochemical characterization of Gun4 suggests a mechanism for its role in chlorophyll biosynthesis. Biochemistry *44*, 7603–7612.

Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J., and Blanche, F. (1992). Assay, purification, and characterization of cobaltochelatase, a unique complex enzyme catalyzing cobalt insertion in hydrogenobyrinic acid a,c-diamide during coenzyme B12 biosynthesis in Pseudomonas denitrificans. J. Bacteriol. *174*, 7445–7451.

Dietzek, B., Kiefer, W., Hermann, G., Popp, J., and Schmitt, M. (2006). Solvent effects on the excited-state processes of protochlorophyllide: a femtosecond time-resolved absorption study. J. Phys. Chem. B *110*, 4399–4406.

Dietzek, B., Tschierlei, S., Hermann, G., Yartsev, A., Pascher, T., Sundström, V., Schmitt, M., and Popp, J. (2009). Protochlorophyllide a: a comprehensive photophysical picture. Chemphyschem *10*, 144–150.

Drulyte, I., Johnson, R.M., Hesketh, E.L., Hurdiss, D.L., Scarff, C.A., Porav, S.A., Ranson, N.A., Muench, S.P., and Thompson, R.F. (2018). Approaches to altering particle distributions in cryo-electron microscopy sample preparation. Acta Crystallogr. Sect. D Struct. Biol. *74*, 1–12.

Duhr, S., and Braun, D. (2006). Why molecules move along a temperature gradient. Proc. Natl. Acad. Sci. U. S. A. *103*, 19678–19682.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu. Rev. Biophys. Biomol. Struct. *35*, 93–114.

Farmer, D.A., Brindley, A.A., Hitchcock, A., Jackson, P.J., Johnson, B., Dickman, M.J., Hunter, C.N., Reid, J.D., and Adams, N.B.P. (2019). The ChlD subunit links the motor and porphyrin binding subunits of magnesium chelatase. Biochem. J. *476*, 1875–1887.

Fernandez-Leiro, R., and Scheres, S.H.W. (2017). A pipeline approach to single-particle processing in RELION. Acta Crystallogr. Sect. D Struct. Biol. *73*, 496–502.

Ferretti, M., Novoderezhkin, V.I., Romero, E., Augulis, R., Pandit, A., Zigmantas, D., and Grondelle, R. Van (2014). The nature of coherences in the B820 bacteriochlorophyll dimer revealed by two-dimensional electronic spectroscopy. Phys. Chem. Chem. Phys. *16*, 9930–9939.

Filling, C., Berndt, K.D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., Ladenstein, R., Jörnvall, H., and Oppermann, U. (2002). Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. J. Biol. Chem. *277*, 25677–25684.

Fodje, M.N., Hansson, A., Hansson, M., Olsen, J.G., Gough, S., Willows, R.D., and Al-Karadaghi, S. (2001). Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase. J. Mol. Biol. *311*, 111–122.

Fortian, A., Castao, D., Gonzalez, E., Lan, A., Falcon-Perez, J.M., and Millet, O. (2011). Structural, thermodynamic, and mechanistical studies in uroporphyrinogen III synthase: Molecular basis of congenital erythropoietic porphyria. In Advances in Protein Chemistry and Structural Biology, pp. 43–74.

De Franceschi, N., Hamidi, H., Alanko, J., Sahgal, P., and Ivaska, J. (2015). Integrin traffic - the update. J. Cell Sci. *128*, 839–852.

Fuesler, T.P., Hanamoto, C.M., and Castelfranco, P.A. (1982). Separation of Mg-protoporphyrin IX and Mg-protoporphyrin IX monomethyl ester synthesized de novo by developing cucumber etioplasts. Plant Physiol. *69*, 421–423.

Gabruk, M., Nowakowska, Z., Skupien-Rabian, B., Kędracka-Krok, S., Mysliwa-Kurdziel, B., and Kruk,

J. (2016). Insight into the oligomeric structure of PORA from A. thaliana. Biochim. Biophys. Acta - Proteins Proteomics *1864*, 1757–1764.

Garbarino, J.E., and Gibbons, I. (2002). Expression and genomic analysis of midasin, a novel and highly conserved AAA protein distantly related to dynein. BMC Genomics *3*, 18.

Gaudreault, F., and Najmanovich, R.J. (2015). FlexAID: revisiting docking on non-native-complex structures. J. Chem. Inf. Model. *55*, 1323–1336.

Gholami, S., Nenov, A., Rivalta, I., Bocola, M., Bordbar, A.K., Schwaneberg, U., Davari, M.D., and Garavelli, M. (2018). Theoretical model of the protochlorophyllide oxidoreductase from a hierarchy of protocols. J. Phys. Chem. B *122*, 7668–7681.

Ghosh, D., Sawicki, M., Pletnev, V., Erman, M., Ohno, S., Nakajin, S., and Duax, W.L. (2001). Porcine carbonyl reductase structural basis for a functional monomer in short chain dehydrogenase/reductases. J. Biol. Chem. *276*, 18457–18463.

Gibson, K.D., Neuberger, A., and Scott, J.J. (1954). The enzymatic conversion of aminolaevulic acid to porphobilinogen. Biochem. J. *58*, xli–xlii.

Gibson, L.C., Willows, R.D., Kannangara, C.G., von Wettstein, D., and Hunter, C.N. (1995). Magnesium-protoporphyrin chelatase of Rhodobacter sphaeroides: reconstitution of activity by combining the products of the bchH, -I, and -D genes expressed in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. *92*, 1941–1944.

Ginsberg, N.S., Cheng, Y.-C., and Fleming, G.R. (2009). Two-dimensional electronic spectroscopy of molecular aggregates. Acc. Chem. Res. *42*, 1352–1363.

Glaeser, R.M. (2016). Specimen Behavior in the Electron Beam. In Methods in Enzymology, (Elsevier Inc.), pp. 19–50.

Glaeser, R.M., Han, B.G., Csencsits, R., Killilea, A., Pulk, A., and Cate, J.H.D. (2016). Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens. Biophys. J. *110*, 749–755.

Gorchein, A. (1972). Magnesium protoporphyrin chelatase activity in Rhodopseudomonas spheroides . Studies with whole cells. Biochem. J. *127*, 97–106.

Gorchein, A. (1973). Control of magnesium–protoporphyrin chelatase activity in Rhodopseudomonas spheroides. Biochem. J. *134*, 833–845.

Gough, S.P., Petersen, B.O., and Duus, J.Ø. (2000). Anaerobic chlorophyll isocyclic ring formation in Rhodobacter capsulatus requires a cobalamin cofactor. Proc. Natl. Acad. Sci. *64*, 6908–6913.

Gräfe, S., Saluz, H.P., Grimm, B., and Hänel, F. (1999). Mg-chelatase of tobacco: the role of the subunit CHL D in the chelation step of protoporphyrin IX. Proc. Natl. Acad. Sci. U. S. A. *96*, 1941–1946.

Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. Elife *4*, 1–19.

Gribun, A., Cheung, K.L.Y., Huen, J., Ortega, J., and Houry, W.A. (2008). Yeast Rvb1 and Rvb2 are ATP-dependent DNA helicases that form a heterohexameric complex. J. Mol. Biol. *376*, 1320–1333.

Griffiths, W.T. (1975). Characterization of the terminal stages of chlorophyll (ide) synthesis in etioplast membrane preparations. Biochem. J. *152*, 623–655.

Griffiths, W.T. (1978). Reconstitution of chlorophyllide formation by isolated etioplast membranes. Biochem. J. *174*, 681–692.

Griffiths, W.T. (1980). Substrate-specificity studies on protochlorophyllide reductase in barley (Hordeum vulgare) etioplast membranes. Biochem. J. *186*, 267–278.

Grigorieff, N. (2016). Frealign: an exploratory tool for single-particle cryo-EM. Methods Enzymol. *579*, 191–226.

Guo, R., Luo, M., and Weinstein, J.D. (1998). Magnesium chelatase from developing pea leaves. Plant

Physiol. *116*, 605–615.

Gupta, A.J., Duhr, S., and Baaske, P. (2018). Microscale thermophoresis (MST). In Encyclopedia of Biophysics, (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 1–5.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. *6*, 519–529.

Hansson, a, Willows, R.D., Roberts, T.H., and Hansson, M. (2002). Three semidominant barley mutants with single amino acid substitutions in the smallest magnesium chelatase subunit form defective AAA+ hexamers. Proc. Natl. Acad. Sci. U. S. A. *99*, 13944–13949.

Heldt, O., Lawrence, A.D., Lindenmeyer, M., Deery, E., Heathcote, P., Rigby, S.E., and Warren, M.J. (2005). Aerobic synthesis of vitamin B12: Ring contraction and cobalt chelation. Biochem. Soc. Trans. *33*, 815–819.

Henderson, R., and Unwin, P.N.T. (1975). Three-dimensional model of purple membrane obtained by electron microscopy. Nature *257*, 28–32.

Hennig, M., Grimm, B., Contestabile, R., John, R.A., and Jansonius, J.N. (1997). Crystal structure of glutamate-1-semialdehyde aminomutase: An 2-dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity. Proc. Natl. Acad. Sci. *94*, 4866–4871.

Heyes, D.J., and Hunter, C.N. (2002). Site-directed mutagenesis of Tyr-189 and Lys-193 in NADPH: protochlorophyllide oxidoreductase from Synechocystis. Biochem. Soc. Trans. *30*, 601–604.

Heyes, D.J., and Hunter, C.N. (2004). Identification and characterization of the product release steps within the catalytic cycle of protochlorophyllide oxidoreductase. Biochemistry *43*, 8265–8271.

Heyes, D.J., and Hunter, C.N. (2005). Making light work of enzyme catalysis: protochlorophyllide oxidoreductase. Trends Biochem. Sci. *30*, 642–649.

Heyes, D.J., and Scrutton, N.S. (2009). Conformational changes in the catalytic cycle of protochlorophyllide oxidoreductase: what lessons can be learnt from dihydrofolate reductase? Biochem. Soc. Trans. *37*, 354–357.

Heyes, D.J., Martin, G.E.M., Reid, R.J., Hunter, C.N., and Wilks, H.M. (2000). NADPH:protochlorophyllide oxidoreductase from Synechocystis: Overexpression, purification and preliminary characterisation. FEBS Lett. *483*, 47–51.

Heyes, D.J., Ruban, A. V, Wilks, H.M., and Hunter, C.N. (2002). Enzymology below 200 K: the kinetics and thermodynamics of the photochemistry catalyzed by protochlorophyllide oxidoreductase. Proc. Natl. Acad. Sci. U. S. A. *99*, 11145–11150.

Heyes, D.J., Hunter, C.N., van Stokkum, I.H.M., van Grondelle, R., Groot, M.L., Grondelle, R. Van, and Groot, M.L. (2003). Ultrafast enzymatic reaction dynamics in. Nat. Struct. Biol. *10*, 388–389.

Heyes, D.J., Heathcote, P., Rigby, S.E.J., Palacios, M. a., van Grondelle, R., and Hunter, C.N. (2006). The First Catalytic Step of the Light-driven Enzyme Protochlorophyllide Oxidoreductase Proceeds via a Charge Transfer Complex. J. Biol. Chem. *281*, 26847–26853.

Heyes, D.J., Sakuma, M., and Scrutton, N.S. (2007). Laser excitation studies of the product release steps in the catalytic cycle of the light-driven enzyme, protochlorophyllide oxidoreductase. J. Biol. Chem. *282*, 32015–32020.

Heyes, D.J., Menon, B.R.K., Sakuma, M., Scrutton, N.S., and Oxidoreductase, P. (2008a). Conformational events during ternary enzyme-substrate complex formation are rate limiting in the catalytic cycle of the light-driven enzyme protochlorophyllide oxidoreductase. Biochemistry *47*, 10991–10998.

Heyes, D.J., Sakuma, M., de Visser, S.P., and Scrutton, N.S. (2008b). Nuclear quantum tunneling in the light-activated enzyme protochlorophyllide oxidoreductase. J. Biol. Chem. *284*, 3762–3767.

Heyes, D.J., Sakuma, M., and Scrutton, N.S. (2009). Solvent-slaved protein motions accompany proton but

not hydride tunneling in light-activated protochlorophyllide oxidoreductase. Angew. Chemie Int. Ed. *48*, 3850–3853.

Heyes, D.J., Levy, C., Sakuma, M., Robertson, D.L., and Scrutton, N.S. (2011). A twin-track approach has optimized proton and hydride transfer by dynamically coupled tunneling during the evolution of protochlorophyllide oxidoreductase. J. Biol. Chem. *286*, 11849–11854.

Heyes, D.J., Hardman, S.J.O., Hedison, T.M., Hoeven, R., Greetham, G.M., Towrie, M., and Scrutton, N.S. (2014). Excited-state charge separation in the photochemical mechanism of the light-driven enzyme protochlorophyllide oxidoreductase. Angew. Chem. Int. Ed. Engl. *54*, 1512–1515.

Hobbs, C., Reid, J.D., and Shepherd, M. (2017). The coproporphyrin ferrochelatase of Staphylococcus aureus : mechanistic insights into a regulatory iron-binding site. Biochem. J. *474*, 3513–3522.

Hoffert, M.J. (2002). Advanced technology paths to climate stability: energy for a greenhouse planet. Science (80-.). *298*, 981–987.

Hoggins, M., Dailey, H.A., Hunter, C.N., and Reid, J.D. (2007). Direct measurement of metal ion chelation in the active site of human ferrochelatase. Biochemistry *46*, 8121–8127.

Hollingshead, S.L. (2013). Investigating protein-protein interactions in the chlorophyll biosynthesis pathway. University of Sheffield.

Hollingshead, S., Kopečná, J., Armstrong, D.R., Bučinská, L., Jackson, P.J., Chen, G.E., Dickman, M.J., Williamson, M.P., Sobotka, R., and Hunter, C.N. (2016). Synthesis of chlorophyll-binding proteins in a fully segregated Δycf54 strain of the cyanobacterium Synechocystis PCC 6803. Front. Plant Sci. *7*, 1–15.

Humphries, J.D., Byron, A., and Humphries, M.J. (2006). Integrin ligands at a glance. J. Cell Sci. *119*, 3901–3903.

Islam, M.R., Aikawa, S., Midorikawa, T., Kashino, Y., Satoh, K., and Koike, H. (2008). slr1923 of Synechocystis sp. PCC6803 is essential for conversion of 3,8-divinyl(proto)chlorophyll(ide) to 3-monovinyl(proto)chlorophyll(ide). Plant Physiol. *148*, 1068–1081.

Jackson, A.H., Sancovich, H.A., Ferramola, A.M., Evans, N., Games, D.E., Matlin, S.A., Elder, G.H., and Smith, S.G. (1976). Macrocyclic intermediates in the biosynthesis of porphyrins. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *273*, 191–206.

Jaffe, E.K. (1995). Porphobilinogen synthase, the first source of heme's asymmetry. J. Bioenerg. Biomembr. *27*, 169–179.

Jaffe, E.K. (2003). An unusual phylogenetic variation in the metal ion binding sites of porphobilinogen synthase. Chem. Biol. *10*, 25–34.

Jaffe, E.K. (2016). The remarkable character of porphobilinogen synthase. Acc. Chem. Res. *49*, 2509–2517.

Jensen, P.E., Gibson, L.C., Henningsen, K.W., and Hunter, C.N. (1996). Expression of the chlI, chlD, and chlH genes from the Cyanobacterium synechocystis PCC6803 in Escherichia coli and demonstration that the three cognate proteins are required for magnesium-protoporphyrin chelatase activity. J. Biol. Chem. *271*, 16662–16667.

Jensen, P.E., Gibson, L.C., and Hunter, C.N. (1998). Determinants of catalytic activity with the use of purified I, D and H subunits of the magnesium protoporphyrin IX chelatase from Synechocystis PCC6803. Biochem. J. *334*, 335–344.

Jensen, P.E., Gibson, L., and Hunter, C.N. (1999). ATPase activity associated with the magnesium-protoporphyrin IX chelatase enzyme of Synechocystis PCC6803: evidence for ATP hydrolysis during Mg2+ insertion, and the MgATP-dependent interaction of the ChlI and ChlD subunits. Biochem. J. *339*, 127.

Jensen, P.E., Reid, J.D., and Hunter, C.N. (2000). Modification of cysteine residues in the ChlI and ChlH subunits of magnesium chelatase results in enzyme inactivation. Biochem. J. *352*, 435.

Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H.M., Duhr, S., Baaske, P., and Breitsprecher, D. (2014). Microscale thermophoresis: interaction analysis and beyond. J. Mol. Struct. *1077*, 101–113.

Ji, X., Mo, T., Liu, W.Q., Ding, W., Deng, Z., and Zhang, Q. (2019). Revisiting the mechanism of the anaerobic coproporphyrinogen III oxidase HemN. Angew. Chemie - Int. Ed. *58*, 6235–6238.

Jordan, P.M. (1991). Chapter 1 The biosynthesis of 5-aminolaevulinic acid and its transformation into uroporphyrinogen III. In New Comprehensive Biochemistry, pp. 1–66.

Kang, Z., Wang, Y., Gu, P., Wang, Q., and Qi, Q. (2011). Engineering Escherichia coli for efficient production of 5-aminolevulinic acid from glucose. Metab. Eng. *13*, 492–498.

Karger, G.A., Reid, J.D., and Hunter, C.N. (2001). Characterization of the binding of deuteroporphyrin IX to the magnesium chelatase H subunit and spectroscopic properties of the complex. Biochemistry *40*, 9291–9299.

Karlberg, T., Lecerof, D., Gora, M., Silvegren, G., Labbe-Bois, R., Hansson, M., and Al-Karadaghi, S. (2002). Metal binding to Saccharomyces cerevisiae ferrochelatase. Biochemistry *41*, 13499–13506.

Kaschner, M., Loeschcke, A., Krause, J., Minh, B.Q., Heck, A., Endres, S., Svensson, V., Wirtz, A., von Haeseler, A., Jaeger, K.E., et al. (2014). Discovery of the first light-dependent protochlorophyllide oxidoreductase in anoxygenic phototrophic bacteria. Mol. Microbiol. *93*, 1066–1078.

Kato, K., Tanaka, R., Sano, S., Tanaka, A., and Hosaka, H. (2010). Identification of a gene essential for protoporphyrinogen IX oxidase activity in the cyanobacterium Synechocystis sp. PCC6803. Proc. Natl. Acad. Sci. *107*, 16649–16654.

Kavanagh, K.L., Jörnvall, H., Persson, B., and Oppermann, U. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell. Mol. Life Sci. *65*, 3895–3906.

Kern, D., Potier, S., Boulanger, Y., and Lapointe, J. (1979). The Monomeric Glutamyl-tRNA of Escherichia. *254*, 518–524.

Klement, H., Helfrich, M., Oster, U., Schoch, S., and Rüdiger, W. (1999). Pigment-free NADPH:protochlorophyllide oxidoreductase from Avena sativa L. Purification and substrate specificity. Eur. J. Biochem. *265*, 862–874.

Kopečná, J., De Vaca, I.C., Adams, N.B.P., Davison, P.A., Brindley, A.A., Hunter, C.N., Guallar, V., and Sobotka, R. (2015). Porphyrin binding to Gun4 protein, facilitated by a flexible loop, controls metabolite flow through the chlorophyll biosynthetic pathway. J. Biol. Chem. *290*, 28477–28488.

Kühlbrandt, W., Wang, D.N., and Fujiyoshi, Y. (1994). Atomic model of plant light-harvesting complex by electron crystallography. Nature *367*, 614–621.

Lambert, N., Chen, Y.N., Cheng, Y.C., Li, C.M., Chen, G.Y., and Nori, F. (2013). Quantum biology. Nat. Phys. *9*, 10–18.

Larkin, R.M., Alonso, J.M., Ecker, J.R., and Chory, J. (2003). GUN4, a regulator of chlorophyll synthesis and intracellular signaling. Science *299*, 902–906.

Lebedev, N., and Timko, M.P. (1999). Protochlorophyllide oxidoreductase B-catalyzed protochlorophyllide photoreduction in vitro: insight into the mechanism of chlorophyll formation in light-adapted plants. Proc. Natl. Acad. Sci. U. S. A. *96*, 9954–9959.

Lee, J.O., Bankston, L.A., and Robert C Liddington, M.A.A. and (1995). Two conformations of the integrin A-domain (I-domain): a pathway for activation? Structure *3*, 1333–1340.

Lesk, A.M. (1995). NAD-binding domains of dehydrogenases. Curr. Opin. Struct. Biol. *5*, 775–783.

Levican, G., Katz, A., de Armas, M., Nunez, H., and Orellana, O. (2007). Regulation of a glutamyl-tRNA synthetase by the heme status. Proc. Natl. Acad. Sci. *104*, 3135–3140.

Li, X., Zhou, N., Chen, W., Zhu, B., Wang, X., Xu, B., Wang, J., Liu, H., and Cheng, L. (2017). Near-atomic resolution structure determination of a cypovirus capsid and polymerase complex using cryo-EM at 200 kV. J. Mol. Biol. *429*, 79–87.

Lobo, S.A.L., Scott, A., Videira, M.A.M., Winpenny, D., Gardner, M., Palmer, M.J., Schroeder, S., Lawrence, A.D., Parkinson, T., Warren, M.J., et al. (2015). Staphylococcus aureus haem biosynthesis: Characterisation of the enzymes involved in final steps of the pathway. Mol. Microbiol. *97*, 472–487.

Louie, G. V, Brownlie, P.D., Lambert, R., Cooper, J.B., Blundell, T.L., Wood, S.P., Malashkevich, V.N., Hadener, A., Warren, M.J., and Shoolingin-jordan, P.M. (1996). The three-dimensional Structure of Escherichia coli porphobilinogen deaminase at 1.76-A resolution. Proteins-Structure Funct. Genet. *25*, 48–78.

Ludtke, S.J. (2016). Single-particle refinement and variability analysis in EMAN2.1. Methods Enzymol. *579*, 159–189.

Lundqvist, J., Elmlund, D., Heldt, D., Deery, E., Söderberg, C. a G., Hansson, M., Warren, M., and Al-Karadaghi, S. (2009). The AAA(+) motor complex of subunits CobS and CobT of cobaltochelatase visualized by single particle electron microscopy. J. Struct. Biol. *167*, 227–234.

Lundqvist, J., Elmlund, H., Wulff, R.P., Berglund, L., Elmlund, D., Emanuelsson, C., Hebert, H., Willows, R.D., Hansson, M., Lindahl, M., et al. (2010). ATP-induced conformational dynamics in the AAA+ motor unit of magnesium chelatase. Structure *18*, 354–365.

Lundqvist, J., Braumann, I., Kurowska, M., Müller, A.H., and Hansson, M. (2013). Catalytic turnover triggers exchange of subunits of the magnesium chelatase AAA+ motor unit. J. Biol. Chem. *288*, 24012–24019.

Luo, B.-H., Carman, C. V, and Springer, T. a (2007). Structural basis of integrin regulation and signaling. Annu. Rev. Immunol. *25*, 619–647.

Luo, S., Luo, T., Liu, Y., Li, Z., Fan, S., and Wu, C. (2018). N-terminus plus linker domain of Mg-chelatase D subunit is essential for Mg-chelatase activity in Oryza sativa. Biochem. Biophys. Res. Commun. *497*, 749–755.

Manzoni, C., and Cerullo, G. (2016). Design criteria for ultrafast optical parametric amplifiers. J. Opt. (United Kingdom) *18*, 103501.

Martin, G.E., Timko, M.P., and Wilks, H.M. (1997). Purification and kinetic analysis of pea (Pisum sativum L.) NADPH:protochlorophyllide oxidoreductase expressed as a fusion with maltose-binding protein in Escherichia coli. Biochem. J. *325 (Pt 1*, 139–145.

Masgrau, L., Basran, J., Hothi, P., Sutcliffe, M.J., and Scrutton, N.S. (2004). Hydrogen tunneling in quinoproteins. Arch. Biochem. Biophys. *428*, 41–51.

Masgrau, L., Roujeinikova, A., Johannissen, L.O., Hothi, P., Basran, J., Ranaghan, K.E., Mulholland, a J., Sutcliffe, M.J., Scrutton, N.S., and Leys, D. (2006). Atomic description of an enzyme reaction dominated by proton tunneling. Science (80-.). *312*, 237–241.

Masgrau, L., Ranaghan, K.E., Scrutton, N.S., Mulholland, A.J., and Sutcliffe, M.J. (2007). Tunneling and classical paths for proton transfer in an enzyme reaction dominated by tunneling: oxidation of tryptamine by aromatic amine dehydrogenase. J. Phys. Chem. B *111*, 3032–3047.

Masuda, T. (2008). Recent overview of the Mg branch of the tetrapyrrole biosynthesis leading to chlorophylls. Photosynth. Res. *96*, 121–143.

Mathews, M.A.A., Schubert, H.L., Whitby, F.G., Alexander, K.J., Schadick, K., Bergonia, H.A., Phillips, J.D., and Hill, C.P. (2001). Crystal structure of human uroporphyrinogen III synthase. EMBO J. *20*, 5832–5839.

Mattes, T.A., Deery, E., Warren, M.J., and Escalante-Semerena, J.C. (2017). Cobalamin biosynthesis and insertion. In Encyclopedia of Inorganic and Bioinorganic Chemistry, pp. 1–24.

McFarlane, M.J., Hunter, C.N., and Heyes, D.J. (2005). Kinetic characterisation of the light-driven protochlorophyllide oxidoreductase (POR) from Thermosynechococcus elongatus. Photochem. Photobiol. Sci. *4*, 1055–1059.

McGoldrick, H.M., Roessner, C.A., Raux, E., Lawrence, A.D., McLean, K.J., Munro, A.W., Santabarbara, S., Rigby, S.E.J., Heathcote, P., Scott, A.I., et al. (2005). Identification and characterization of a novel vitamin B12 (cobalamin) biosynthetic enzyme (CobZ) from Rhodobacter capsulatus, containing flavin, heme, and Fe-S cofactors. J. Biol. Chem. *280*, 1086–1094.

McMullan, G., Faruqi, A.R., and Henderson, R. (2016). Direct Electron Detectors. In Methods in Enzymology, (Elsevier Inc.), pp. 1–17.

Medlock, A., Swartz, L., Dailey, T.A., Dailey, H.A., and Lanzilotta, W.N. (2007). Substrate interactions with human ferrochelatase. Proc. Natl. Acad. Sci. U. S. A. *104*, 1789–1793.

Medlock, A.E., Carter, M., Dailey, T. a, Dailey, H. a, and Lanzilotta, W.N. (2009). Product release rather than chelation determines metal specificity for ferrochelatase. J. Mol. Biol. *393*, 308–319.

Menon, B.R.K., Waltho, J.P., Scrutton, N.S., and Heyes, D.J. (2009). Cryogenic and laser photoexcitation studies identify multiple roles for active site residues in the light-driven enzyme protochlorophyllide oxidoreductase. J. Biol. Chem. *284*, 18160–18166.

Menon, B.R.K., Davison, P. a., Hunter, C.N., Scrutton, N.S., and Heyes, D.J. (2010). Mutagenesis Alters the Catalytic Mechanism of the Light-driven Enzyme Protochlorophyllide Oxidoreductase. J. Biol. Chem. *285*, 2113–2119.

Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M.I., Pragani, R., Boxer, M.B., Earl, L.A., Milne, J.L.S., et al. (2016). Breaking cryo-EM resolution barriers to facilitate drug discovery. Cell *165*, 1698–1707.

Meyerson, J.R., Rao, P., Kumar, J., Chittori, S., Banerjee, S., Pierson, J., Mayer, M.L., and Subramaniam, S. (2014). Self-assembled monolayers improve protein distribution on holey carbon cryo-EM supports. Sci. Rep. *4*, 7084.

Miyazawa, A., Fujiyoshi, Y., and Unwin, N. (2003). Structure and gating mechanism of the acetylcholine receptor pore. Nature *423*, 949–955.

Mobius, K., Arias-Cartin, R., Breckau, D., Hannig, A.-L., Riedmann, K., Biedendieck, R., Schroder, S., Becher, D., Magalon, A., Moser, J., et al. (2010). Heme biosynthesis is coupled to electron transport chains for energy generation. Proc. Natl. Acad. Sci. *107*, 10436–10441.

Mochizuki, N., Brusslan, J. a, Larkin, R., Nagatani, a, and Chory, J. (2001). Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc. Natl. Acad. Sci. U. S. A. *98*, 2053–2058.

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. Elife *6*, 1–22.

Moser, J., Schubert, W.D., Beier, V., Bringemeier, I., Jahn, D., and Heinz, D.W. (2001). V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. EMBO J. *20*, 6583–6590.

Muraki, N., Nomata, J., Ebata, K., Mizoguchi, T., Shiba, T., Tamiaki, H., Kurisu, G., and Fujita, Y. (2010). X-ray crystal structure of the light-independent protochlorophyllide reductase. Nature *465*, 110–114.

Murata, K., and Wolf, M. (2018). Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. Biochim. Biophys. Acta - Gen. Subj. *1862*, 324–334.

Nagata, N., Tanaka, R., Satoh, S., and Tanaka, A. (2005). Identification of a vinyl reductase gene for chlorophyll synthesis in Arabidopsis thaliana and implications for the evolution of prochlorococcus species. Plant Cell *17*, 233–240.

Neuwald, a F., Aravind, L., Spouge, J.L., and Koonin, E. V (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. *9*, 27–43.

van Niel, C.. (1962). The present status of the comparative study of photosynthesis. Ann. Rev. Plant Physiol. *13*, 1–26.

Noble, A.J., Wei, H., Dandey, V.P., Zhang, Z., Tan, Y.Z., Potter, C.S., and Carragher, B. (2018). Reducing effects of particle adsorption to the air–water interface in cryo-EM. Nat. Methods *15*, 793–795.

Nogaj, L.A., and Beale, S.I. (2005). Physical and kinetic interactions between glutamyl-tRNA reductase and glutamate-1-semialdehyde aminotransferase of Chlamydomonas reinhardtii. J. Biol. Chem. *280*, 24301–24307.

Nogales, E., Wolf, S.G., and Downing, K.H. (1998). Structure of the αβ tubulin dimer by electron crystallography. Nature *391*, 199–203.

Nomata, J., Kondo, T., Mizoguchi, T., Tamiaki, H., Itoh, S., and Fujita, Y. (2014). Dark-operative protochlorophyllide oxidoreductase generates substrate radicals by an iron-sulphur cluster in bacteriochlorophyll biosynthesis. Sci. Rep. *4*, 1–8.

Nomata, J., Terauchi, K., and Fujita, Y. (2016). Stoichiometry of ATP hydrolysis and chlorophyllide formation of dark-operative protochlorophyllide oxidoreductase from Rhodobacter capsulatus. Biochem. Biophys. Res. Commun. *470*, 704–709.

Osanai, T., Imashimizu, M., Seki, A., Sato, S., Tabata, S., Imamura, S., Asayama, M., Ikeuchi, M., and Tanaka, K. (2009). ChlH, the H subunit of the Mg-chelatase, is an anti-sigma factor for SigE in Synechocystis sp. PCC 6803. Proc. Natl. Acad. Sci. U. S. A. *106*, 6860–6865.

Oster, U., C. E., B., and W. Rüdiger (1997). Characterization of chlorophyll a and bacteriochlorophyll a synthases by heterologous expression in Escherichia coli. J. Biol. Chem. *272*, 9671–9676.

Palovcak, E., Wang, F., Zheng, S.Q., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A., and Cheng, Y. (2018). A simple and robust procedure for preparing graphene-oxide cryo-EM grids. J. Struct. Biol. *204*, 80–84.

Pandey, A., Andersen, J.S., and Mann, M. (2000). Use of mass spectrometry to study signaling pathways. Sci. Signal. *2000*, pl1–pl1.

Papenbrock, J., Gräfe, S., Kruse, E., Hänel, F., and Grimm, B. (1997). Mg-chelatase of tobacco: identification of a Chl D cDNA sequence encoding a third subunit, analysis of the interaction of the three subunits with the yeast two-hybrid system, and reconstitution of the enzyme activity by co-expression of recombinant CHL D,. Plant J. *12*, 981–990.

Perona, J.J., Rould, M.A., and Steitz, T.A. (1993). Structural basis for transfer RNA aminoacylation by Escherichia coli glutaminyl-tRNA synthetase. Biochemistry *32*, 8758–8771.

Petersen, B.L., Jensen, P.E., Gibson, L.C.D., Stummann, B.M., Hunter, C.N., and Henningsen, K.W. (1998). Reconstitution of an active magnesium chelatase enzyme complex from the bchI, -D, and -H gene products of the green sulfur bacterium Chlorobium vibrioforme expressed in Escherichia coli. J. Bacteriol. *180*, 699–704.

Phillips, J.D., Whitby, F.G., Warby, C.A., Labbe, P., Yang, C., Pflugrath, J.W., Ferrara, J.D., Robinson, H., Kushner, J.P., and Hill, C.P. (2004). Crystal structure of the oxygen-dependant coproporphyrinogen oxidase (Hem13p) of Saccharomyces cerevisiae. J. Biol. Chem. *279*, 38960–38968.

Phillips, J.D., Warby, C.A., Whitby, F.G., Kushner, J.P., and Hill, C.P. (2009). Substrate shuttling between active sites of uroporphyrinogen decarboxylase is not required to generate coproporphyrinogen. J. Mol. Biol. *389*, 306–314.

Plow, E.F., Haas, T.A., Zhang, L., Loftus, J., and Smith, J.W. (2000). Ligand binding to integrins. J. Biol. Chem. *275*, 21785–21788.

Proctor, M.S., Chidgey, J.W., Shukla, M.K., Jackson, P.J., Sobotka, R., Hunter, C.N., and Hitchcock, A. (2018). Plant and algal chlorophyll synthases function in Synechocystis and interact with the YidC/Alb3 membrane insertase. FEBS Lett. *592*, 3062–3073.

Qian, P., Marklew, C.J., Viney, J., Davison, P.A., Brindley, A.A., Söderberg, C., Al-Karadaghi, S., Bullough, P.A., Grossmann, J.G., and Hunter, C.N. (2012). Structure of the cyanobacterial magnesium chelatase H subunit determined by single particle reconstruction and small-angle X-ray scattering. J. Biol. Chem. *287*, 4946–4956.

Reid, J.D., and Hunter, C.N. (2002). Current understanding of the function of magnesium chelatase. Biochem. Soc. Trans. *30*, 643–645.

Reid, J.D., and Hunter, C.N. (2004). Magnesium-dependent ATPase activity and cooperativity of magnesium chelatase from Synechocystis sp. PCC6803. J. Biol. Chem. *279*, 26893.

Reid, J.D., Siebert, C.A., Bullough, P.A., and Hunter, C.N. (2003). The ATPase activity of the ChlI subunit of magnesium chelatase and formation of a heptameric AAA+ ring. Biochemistry *42*, 6912–6920.

Reinbothe, C., Lebedev, N., and Reinbothe, S. (1999). A protochlorophyllide light-harvesting complex involved in de-etiolation of higher plants. Nature *397*, 80–84.

Richter, A.S., Hochheuser, C., Fufezan, C., Heinze, L., Kuhnert, F., and Grimm, B. (2016). Phosphorylation of GENOMES UNCOUPLED 4 alters stimulation of Mg chelatase activity in angiosperms. Plant Physiol. *172*, 1578–1595.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyano- bacteria. J. Gen. Microbiol. *11*, 1–61.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Romão, C. V., Ladakis, D., Lobo, S.A.L., Carrondo, M.A., Brindley, A.A., Deery, E., Matias, P.M., Pickersgill, R.W., Saraiva, L.M., and Warren, M.J. (2011). Evolution in a family of chelatases facilitated by the introduction of active site asymmetry and protein oligomerization. Proc. Natl. Acad. Sci. U. S. A. *108*, 97–102.

Romero, E., Augulis, R., Novoderezhkin, V.I., Ferretti, M., Thieme, J., Zigmantas, D., and van Grondelle, R. (2014). Quantum coherence in photosynthesis for efficient solar-energy conversion. Nat. Phys. *10*, 676–682.

Romero, E., Novoderezhkin, V.I., and Van Grondelle, R. (2017). Quantum design of photosynthesis for bio-inspired solar-energy conversion. Nature *543*, 355–365.

Romes, E.M., Sobhany, M., and Stanley, R.E. (2016). The crystal structure of the ubiquitin-like domain of ribosome assembly factor Ytm1 and characterization of its interaction with the AAA-ATPase midasin. J. Biol. Chem. *291*, 882–893.

De Rosier, D.J., and Klug, A. (1968). Reconstruction of three dimensional structures from electron micrographs. Nature *217*, 130–134.

Rüdiger, W., Benz, J., and Guthoff, C. (1980). Detection and partial characterization of activity of chlorophyll synthetase in etioplast membranes. Eur. J. Biochem. *109*, 193–200.

Russo, C.J., and Passmore, L.A. (2016). Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. J. Struct. Biol. *193*, 33–44.

Sawicki, A., Zhou, S., Kwiatkowski, K., Luo, M., and Willows, R.D. (2017). 1-N-histidine phosphorylation of ChlD by the AAA+ ChlI2 stimulates magnesium chelatase activity in chlorophyll synthesis. Biochem. J. BCJ20161094.

Scheer, H. (1991). Structure and occurence of chlorophylls. In Chlorophylls, pp. 3–23.

Scheres, S.H.W. (2015). Semi-automated selection of cryo-EM particles in RELION-1.3. J. Struct. Biol.

189, 114–122.

Scheres, S.H.W. (2016). Processing of structurally heterogeneous cryo-EM data in RELION (Elsevier Inc.).

Schimmel, P. (1987). Aminoacyl transfer RNA synthetases: general scheme of structure function relationships in the polypeptides and recognition of transfer-RNAs. Annu. Rev. Biochem. *56*, 125–158.

Schlau-Cohen, G.S., Dawlaty, J.M., and Fleming, G.R. (2012). Ultrafast multidimensional spectroscopy: Principles and applications to photosynthetic systems. IEEE J. Sel. Top. Quantum Electron. *18*, 283–295.

Schmidt, C., and Urlaub, H. (2017). Combining cryo-electron microscopy (cryo-EM) and cross-linking mass spectrometry (CX-MS) for structural elucidation of large protein assemblies. Curr. Opin. Struct. Biol. *46*, 157–168.

Schmidt, H., Gleave, E.S., and Carter, A.P. (2012). Insights into dynein motor domain function from a 3.3-Å crystal structure. Nat. Struct. Mol. Biol. *19*, 492–497.

Schoefs, B., and Franck, F. (2004). Protochlorophyllide reduction: mechanisms and evolution¶. Photochem. Photobiol. *78*, 543.

Schubert, H.L., Raux, E., Wilson, K.S., and Warren, M.J. (1999). Common chelatase design in the branched tetrapyrrole pathways of heme and anaerobic cobalamin synthesis. Biochemistry *38*, 10660–10669.

Schubert, H.L., Phillips, J.D., Heroux, A., and Hill, C.P. (2008). Structure and mechanistic implications of a uroporphyrinogen III. Biochemistry *47*, 8648–8655.

Scrutton, N.S., Groot, M.L., and Heyes, D.J. (2012). Excited state dynamics and catalytic mechanism of the light-driven enzyme protochlorophyllide oxidoreductase. Phys. Chem. Chem. Phys. *14*, 8818–8824.

Seidel, S. a I., Dijkman, P.M., Lea, W. a., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J.S., Srinivasan, P., Baaske, P., Simeonov, A., et al. (2013). Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. Methods *59*, 301–315.

Sekine, S. ichi, Shichiri, M., Bernier, S., Chênevert, R., Lapointe, J., and Yokoyama, S. (2006). Structural bases of transfer RNA-dependent amino acid recognition and activation by glutamyl-tRNA synthetase. Structure *14*, 1791–1799.

Shafqat, J., El-Ahmad, M., Danielsson, O., Martinez, M.C., Persson, B., Pares, X., and Jornvall, H. (1996). Pea formaldehyde-active class III alcohol dehydrogenase: common derivation of the plant and animal forms but not of the corresponding ethanol-active forms (classes I and P). Proc. Natl. Acad. Sci. *93*, 5595–5599.

Shen, M., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. Protein Sci. *15*, 2507–2524.

Shen, Y., and Ryde, U. (2005). Reaction mechanism of porphyrin metallation studied by theoretical methods. Chemistry *11*, 1549–1564.

Shen, Y.-Y., Wang, X.-F., Wu, F.-Q., Du, S.-Y., Cao, Z., Shang, Y., Wang, X.-L., Peng, C.-C., Yu, X.-C., Zhu, S.-Y., et al. (2006). The Mg-chelatase H subunit is an abscisic acid receptor. Nature *443*, 823–826.

Shepherd, M., and Hunter, C.N. (2004). Transient kinetics of the reaction catalysed by magnesium protoporphyrin IX methyltransferase. Biochem. J. *382*, 1009–1013.

Shepherd, M., Mclean, S., and Hunter, C.N. (2005). Kinetic basis for linking the first two enzymes of chlorophyll biosynthesis. FEBS J. *272*, 4532–4539.

Sirijovski, N., Lundqvist, J., Rosenbäck, M., Elmlund, H., Al-Karadaghi, S., Willows, R.D., and Hansson, M. (2008). Substrate-binding model of the chlorophyll biosynthetic magnesium chelatase BchH subunit. J. Biol. Chem. *283*, 11652–11660.

Skotnicová, P., Sobotka, R., Shepherd, M., Hájek, J., Hrouzek, P., and Tichý, M. (2018). The cyanobacterial protoporphyrinogen oxidase HemJ is a new b-type heme protein functionally coupled with coproporphyrinogen III oxidase. J. Biol. Chem. *293*, 12394–12404.

Smith, B.B., and Rebeiz, C.A. (1977). Chloroplast biogenesis: Detection of Mg-protoporphyrin chelatase in vitro. Arch. Biochem. Biophys. *180*, 178–185.

Smith, K.M., and Goff, D.A. (1986). Syntheses of some proposed biosynthetic precursors to the isocyclic ring in chlorophyll a. J. Org. Chem. *51*, 657–666.

Snider, J., Thibault, G., and Houry, W.A. (2008). The AAA+ superfamily of functionally diverse proteins. Genome Biol. *9*, 216.1-216.8.

Snijder, J., Borst, A.J., Dosey, A., Walls, A.C., Burrell, A., Reddy, V.S., Kollman, J.M., and Veesler, D. (2017). Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids. J. Struct. Biol. *198*, 38–42.

Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. *33*, 244–248.

Song, Y., Pu, H., Jiang, T., Zhang, L., and Ouyang, M. (2016). Crystal structure of glutamate-1-semialdehyde-2,1-aminomutase from Arabidopsis thaliana . Acta Crystallogr. Sect. F Struct. Biol. Commun. *72*, 448–456.

Stark, H. (2010). GraFix: Stabilization of fragile macromolecular complexes for single particle Cryo-EM (Elsevier Masson SAS).

Stetefeld, J., Jenny, M., and Burkhard, P. (2006). Intersubunit signaling in glutamate-1-semialdehyde-aminomutase. Proc. Natl. Acad. Sci. *103*, 13688–13693.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. *41*, 207–234.

Sytina, O. a., Heyes, D.J., Hunter, C.N., Alexandre, M.T., van Stokkum, I.H.M., van Grondelle, R., and Groot, M.L. (2008). Conformational changes in an ultrafast light-driven enzyme determine catalytic activity. Nature *456*, 1001–1004.

Sytina, O.A., van Stokkum, I.H.M., Heyes, D.J., Hunter, C.N., van Grondelle, R., and Groot, M.L. (2010). Protochlorophyllide excited-state dynamics in organic solvents studied by time-resolved visible and mid-infrared spectroscopy. J. Phys. Chem. B *114*, 4335–4344.

Tanaka, R., and Tanaka, A. (2007). Tetrapyrrole biosynthesis in higher plants. Annu. Rev. Plant Biol *58*, 321–346.

Tarahi Tabrizi, S., Sawicki, A., Zhou, S., Luo, M., and Willows, R.D. (2016). GUN4-protoporphyrin IX is a singlet oxygen generator with consequences for plastid retrograde signalling. J. Biol. Chem. *2*, jbc.C116.719989.

Townley, H.E., Griffiths, W.T., and Nugent, J.P. (1998). A reappraisal of the mechanism of the photoenzyme protochlorophyllide reductase based on studies with the heterologously expressed protein. FEBS Lett. *422*, 19–22.

Townley, H.E., Sessions, R.B., Clarke, A.R., Dafforn, T.R., and Trevor Griffiths, W. (2001). Protochlorophyllide oxidoreductase: A homology model examined by site-directed mutagenesis. Proteins Struct. Funct. Genet. *44*, 329–335.

Trott, O., and Olson, A.J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. NA-NA.

Troup, B., Jahn, M., Hungerer, C., and Jahn, D. (1994). Isolation of the hemF operon containing the gene for the Escherichia coli aerobic coproporphyrinogen III oxidase by in vivo complementation of a yeast HEM13 mutant. J. Bacteriol. *176*, 673–680.

Troup, B., Hungerer, C., and Jahn, D. (1995). Cloning and characterization of the Escherichia coli hemN gene encoding the oxygen-independent coproporphyrinogen III oxidase. J. Bacteriol. *177*, 3326–3331.

Valera, V., Fung, M., Wessler, A.N., and Richards, W.R. (1987). Synthesis of 4R- and 4S-tritium labeled

NADPH for the determination of the coenzyme stereospecificity of NADPH: protochlorophyllide oxidoreductase. Biochem. Biophys. Res. Commun. *148*, 515–520.

Verdecia, M.A., Larkin, R.M., Ferrer, J.-L., Riek, R., Chory, J., and Noel, J.P. (2005). Structure of the Mg-chelatase cofactor GUN4 reveals a novel hand-shaped fold for porphyrin binding. PLoS Biol. *3*, e151.

Viney, J., Davison, P.A., Hunter, C.N., and Reid, J.D. (2007). Direct measurement of metal-ion chelation in the active site of the AAA+ ATPase magnesium chelatase. Biochemistry *46*, 12788–12794.

Wade, R.H. (1992). A brief look at imaging and contrast transfer. Ultramicroscopy *46*, 145–156.

Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., et al. (2018). SPHIRE-crYOLO: A fast and accurate fully automated particle picker for cryo-EM. BioRxiv 356584.

Walker, C.J., and Weinstein, J.D. (1991). Further characterization of the magnesium chelatase in isolated developing cucumber chloroplasts: substrate specificity, regulation, intactness, and ATP requirements. Plant Physiol. *95*, 1189–1196.

Walker, C.J., and Weinstein, J.D. (1994). The magnesium-insertion step of chlorophyll biosynthesis is a two-stage reaction. Biochem. J. *299*, 277–284.

Wang, P., Wan, C., Xu, Z., Wang, P., Wang, W., Sun, C., Ma, X., Xiao, Y., Zhu, J., Gao, X., et al. (2013). One divinyl reductase reduces the 8-vinyl groups in various intermediates of chlorophyll biosynthesis in a given higher plant species, but the isozyme differs between species. Plant Physiol. *161*, 521–534.

Webb, B., and Sali, A. (2002). Comparative protein structure modeling ssing MODELLER. Curr. Protoc. Bioinforma. 1–37.

Wendler, P., Ciniawsky, S., Kock, M., and Kube, S. (2012). Structure and function of the AAA+ nucleotide binding pocket. Biochim. Biophys. Acta - Mol. Cell Res. *1823*, 2–14.

Wilks, H.M., and Timko, M.P. (1995). A light-dependent complementation system for analysis of NADPH:protochlorophyllide oxidoreductase: identification and mutagenesis of two conserved residues that are essential for enzyme activity. Proc. Natl. Acad. Sci. U. S. A. *92*, 724–728.

Willows, R.D., Gibson, L.C.D., Kanangara, C.G., Hunter, C.N., and Wettstein, D. (1996). Three separate proteins constitute the magnesium chelatase of Rhodobacter sphaeroides. Eur. J. Biochem. *235*, 438–443.

Willows, R.D., Hansson, A., Birch, D., Al-Karadaghi, S., and Hansson, M. (2004). EM single particle analysis of the ATP-dependent BchI complex of magnesium chelatase: An AAA+ hexamer. J. Struct. Biol. *146*, 227–233.

Wu, C.K., Dailey, H.A., Rose, J.P., Burden, A., Sellers, V.M., and Wang, B.C. (2001). The 2.0 Å structure of human ferrochelatase, the terminal enzyme of heme biosynthesis. Nat. Struct. Biol. *8*, 156–160.

Yamazaki, S., Nomata, J., and Fujita, Y. (2006). Differential Operation of Dual Protochlorophyllide Reductases for Chlorophyll Biosynthesis in Response to Environmental Oxygen Levels in the Cyanobacterium Leptolyngbya boryana. Plant Physiol. *142*, 911–922.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhang, J., Weng, H., Ding, W., and Kang, Z. (2017). N-terminal engineering of glutamyl-tRNA reductase with positive charge arginine to increase 5-aminolevulinic acid biosynthesis. Bioengineered *8*, 424–427.

Zhao, G.-J., and Han, K.-L. (2008). Site-specific solvation of the photoexcited protochlorophyllide a in methanol: formation of the hydrogen-bonded intermediate state induced by hydrogen-bond strengthening. Biophys. J. *94*, 38–46.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zigmantas, D., Read, E.L., Mancal, T., Brixner, T., Gardiner, A.T., Cogdell, R.J., and Fleming, G.R. (2006). Two-dimensional electronic spectroscopy of the B800-B820 light-harvesting complex. Proc. Natl. Acad. Sci. *103*, 12672–12677.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*, 1–22.