Investigating Protein-Protein Interactions in the Chlorophyll Biosynthesis Pathway

Sarah Louise Hollingshead

A thesis submitted for the degree of Doctor of Philosophy

Department of Molecular Biology and Biotechnology

University of Sheffield

December 2013

Summary

The production of chemical energy from light energy is arguably the most important reaction known, which makes photosynthesis one of the most important processes on Earth. Nearly all life depends on the energy derived from light, and it is this process that supports Earths oxygenated atmosphere. Photosynthesis is initiated by the absorption of light, which is performed by specialised pigment molecules known as the chlorophylls. These are highly specialised pigment molecules that belong to the tetrapyrrole family, which also includes vitamin B12, sirohaem and haem. Chlorophyll is synthesised via a series of chemical reactions collectively known as chlorophyll biosynthesis.

Whilst the majority of enzymes required for chlorophyll biosynthesis are known and have, to some extent, been characterised, one enzyme, the magnesium protoporphyrin IX monomethylester cyclase (cyclase), remains an enigma. This enzyme catalyses the formation of the 5th isocyclic ring, altering the colour of the tetrapyrrole molecule from red to green. Despite being known for over 60 years, only one subunit of the cyclase has been identified. This is the catalytic or AcsF subunit, which contains a distinctive di-iron motif and is absolutely conserved across all known oxygenic photosynthetic organisms. One of the focuses of this work was to identify other subunits of the cyclase. This led to the discovery of Ycf54, a protein that is highly conserved among oxygenic photosynthetic organisms, which is essential for chlorophyll accumulation in *Synechocystis* sp. PCC6803. The latter half of this thesis focuses on investigating the structural and functional characteristics of *Synechocystis* Ycf54; a protein that had previously not been investigated. This work lead the identification of three residues in Ycf54 (D39, F40 and R82), which are required for protochlorophyllide synthesis and chlorophyll formation in *Synechocystis*, and the observation that these residues are all required for Ycf54 to interact with the catalytic (AcsF) subunit of the cyclase. Additionally, the crystal structure of wild type *Synechocystis* Ycf54 to a resolution of 1.2 Å, and the structures of two Ycf54 mutants (A9G and R82A), which have an *in vivo* phenotype were obtained using molecular replacement.

Furthermore this work presents the first investigations into the previously unknown protein *Slr0483* in *Synechocystis*. This protein contains the conserved C-terminal membrane anchoring CAAD (Cyanobacterial Aminoacyl-tRNA synthetases Appended Domain) domain and is found in all of the oxygenic photosynthetic organisms investigated. The experiments reported in this thesis show that Slr0483 is essential for photosystem stability and accumulation in *Synechocystis* and that this protein interacts with the chlorophyll biosynthesis enzymes protoporphyrin IX oxidase, protoporphyrin IX methyltransferase, the

cyclase subunit Sll1214 and the geranylgeranyl reductase ChlP, as well as the haem biosynthesis enzyme ferrochelatase. Leading to the hypothesis that Slr0483 may serve as a membrane anchor, localising the enzymes required for chlorophyll and haem biosynthesis, to the sites where these tetrapyrroles are required in the thylakoid membrane (i.e. next to the sites of photosystem and cytochrome assembly).

Acknowledgements

Firstly, I would like to thank my supervisor Professor Neil Hunter FRS for his continued support and advice throughout the four years I've been with the lab. Thanks in particular for not rejecting my rather late plea for a PhD, for jetting me off to several exotic conference locations and for lots of useful scientific discussions.

In relation to this work I'd like to thank Drs: Daniel Canniffe for introducing me to *Synechocystis*, lending me so many of his strains and teaching me how to use the HPLC, Phil Jackson for all his invaluable help, time and patience with the mass spectrometry, Roman Sobotka for sharing his knowledge on all things *Synechocystis* and Amanda Brindley for all her help and advice with protein purification.

A huge thank you to Sophie Bliss for all her help with the crystallography and introducing me to crystal trials, polishing cover slips, CCP4 and PyMOL (such fun!), Dave Armstrong for doing the NMR analysis on my mystery pigment (I couldn't have done that without you) and Jana Kopecna for her help with clear native gels.

Thanks to the whole of E12/E11 in its entirety for being a lovely bunch of people to work with and for making work enjoyable. In particular, I would like to thank Mo and Lizzy for keeping everything going. Jack, Dan, Paul, Lizzie and Gooch for all their great lab banter and the Dave's for generally being great chaps. Thanks for Katie for keeping me fuelled with lunch, Starbucks coffee and muffins.

Finally, thanks to all of the friends I've met in Sheffield over the last three years. Especially to Sophie, Abi, Simon, Katie and many bottles of sauvignon blank, for keeping me sane with our Wednesday drinking sessions and for helping me escape from the algae to go procure ice-creams and duck watch in the park.

Table of contents

Table of figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Table of tables

Table of abbreviations

AAA	ATPase's associated with a variety of activities	**LB**	Luria-Bertani
ALA	Δ-aminolaevulinic acid	**LDAO**	lauryldimethylamine-oxide
BOG	octyl-β-glucoside	**MgCH**	magnesium chelatase
BSA	bovine serum albumin	**MgPME**	magnesium protoporphyrin monomethylester
CAAD	cyanobacterial aminoacyl-tRNA synthetases appended domain	**NT**	amino terminal
Chlide	chlorophyllide	**OD**	optical density
CPOIII	coproporphyrinogen III oxidase	**p.s.i**	pounds per square inch
CT	carboxy-terminal	**PAGE**	polyacrylamide gel electrophoresis
Cyclase	magnesium protoporphyrin IX monomethylester cyclase	**PBS**	phosphate-buffered saline
β-DDM	n-dodecyl-β-D-maltoside	**Pchlide**	protochlorophyllide
dNTP	2'-deoxynucleotide-5'-triphosphate	**PCR**	polymerase chain reaction
DTT	Dithiothreitol	**POR**	light-dependent NADPH-protochlorophyllide oxidoreductase
DVR	divinyl reductase	**PS**	photosystem
FeCH	ferrochelatase	**PVDF**	polyvinylidene difluoride
FLAG	DYKDDDDK	**PPO**	protoporphyrinogen oxidase
GGPP	geranylgeranyl diphosphate	**QH_2O**	ultrapure water
GluRS	glutamate tRNA synthetase	**ROESY**	Rotating-frame Overhauser Effect SpectroscopY
GluTR	glutamyl-tRNA reductase	**SAM**	*S*-adenosyl-L-methionine
GSA	glutamate 1-semialdehyde	***Synechocystis***	*Synechocystis* sp. PCC 6803
GSAT	glutamate 1-semialdehyde aminotransferase	**TEAB**	tetraethylammonium bromide
GST	glutathione *S*-transferase	**TES**	2-[(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]ethanesulfonic acid, N-[Tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid
gun	genomes uncoupled	**TBS**	Tris-buffered saline
HPLC	high performance liquid chromatography	**Tris**	Tris(hydroxymethyl)methylamine
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid	**$tRNA^{Glu}$**	glutamate specific transfer ribonucleic acid
IPTG	isopropyl-β-D-thiogalacto-pyranoside	**YCF**	hypothetical chloroplast open reading frames

Table of gene names

Protein	Synechocystis gene ID	Arabidopsis thaliana gene ID
GUN4	sll0558	GUN4 At3g59400
ChlH	slr1055	GUN5 At5g13630
ChlI	slr1030	At4g18480
ChlD	slr1777	At1g08520
ChlM	slr0525	At4g25080
Sll1214	sll1214	CHL27 At3g56940
Sll1874	sll1874	CHL27 At3g56940
Ycf54	slr1780	LHAA At5g58250
DVR	slr1923	-
POR	slr0506	PORA At5g54190
		PORB At4g27440
		PORC At1g03630
ChlP	sll1091	CHLP At1g74470
ChlG	slr0056	CHLG At3g51820

1. Introduction

1.1 Photosynthesis

Photosynthesis is a biological mechanism by which light energy is harvested and converted to chemical energy. This process produces the molecular oxygen and reduced carbon sources essential for virtually all life on earth, the only exception being chemolithotrophic bacteria, which makes photosynthesis one of the most important biochemical processes on the planet. Photosynthesis is represented by the general equation:

$$2H_2A + CO_2 \xrightarrow{h\nu} (CH_2O) + H_2O + 2A \qquad \text{(Van Niel, 1962)}$$

where H_2A represents the source of hydrogen ions and electrons used to reduce carbon dioxide to carbohydrate (CH_2O) using the energy from light (*hv*), producing an oxidation product (A), and water in the process (Kiang *et al.*, 2007). Interestingly there are two types of photosynthetic reaction, classified as oxygenic or anoxygenic, which vary in the electron donor (H_2A) used and the oxidation product (A) produced.

Oxygenic photosynthesis is performed by plants, algae and cyanobacteria and can be represented by the equation:

$$12H_2O + 6CO_2 \xrightarrow{h\nu} (C_6H_{12}O_6) + 6H_2O + 6O_2$$

In this reaction water (H_2A) is cleaved yielding oxidised and reduced moieties. The reductant serves to reduce carbon dioxide to organic carbon and the oxidant gives rise to molecular oxygen; hence the name oxygenic photosynthesis.

In anoxygenic photosynthesis, a process performed by all other photosynthetic bacteria, a variety of simple molecules, excluding water, are used as the reductant. Examples of these include reduced sulphur compounds, molecular hydrogen and organic acids such as acetate or succinate; consequently the oxidation product is anoxygenic (Stanier *et al.*, 1959).

There are two stages of photosynthesis, one light-dependent and the other light-independent. In the light-dependent stage, visible radiation excites an electron in a pigment molecule, which in returning to a lower energy state, provides the energy needed for the synthesis of ATP and NAD(P)H. These products are then utilised in the light-independent

reaction to drive the assimilation of carbon dioxide into carbohydrate in a process known as the Calvin-Benson Cycle (Calvin and Benson, 1948).

Although represented by a deceptively simple equation, photosynthesis is a complex process requiring specialised membrane bound pigment protein complexes to facilitate the capture of solar energy and its channelling into carbon fixation. Energy capture is performed by pigment molecules mainly consisting of chlorophylls, bacteriochlorophylls, carotenoids and phycobilins. As the presence of these molecules on earth is so ubiquitous (chlorophylls and bacteriochlorophylls are among the most abundant pigment molecules on the planet) the collective reflection of green light from these pigments can be viewed from space (**Figure 1.1**)

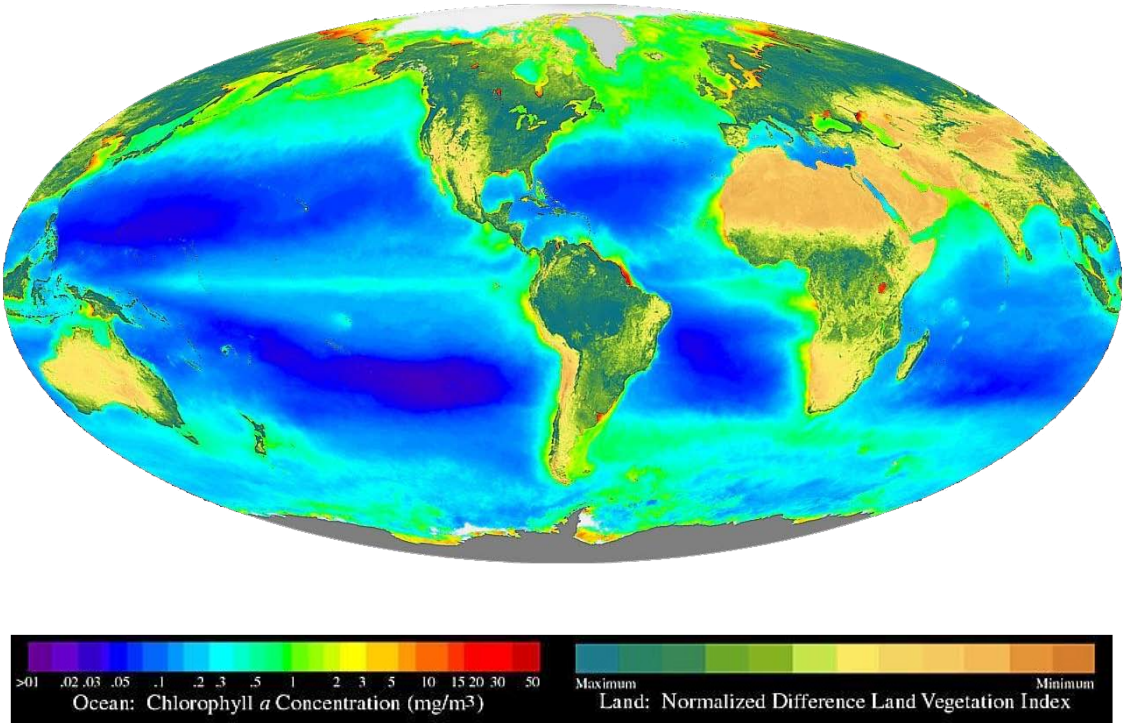

Figure 1.1 The global biosphere

Image of the global biosphere showing the average ocean phytoplankton chlorophyll concentration (data acquired between September 1997 and August 2000), combined with the Normalised Difference Vegetation Index over land. All data were derived from observations from the SeaWiFS sensor. The data used in this effort were acquired as part of the activities of NASA's Science Mission Directorate, and are archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Centre (DISC). The normalised difference land vegetation index is calculated from the scatter of solar radiation in the near-infrared region by photosynthetic vegetation using the equation: $Normalised\ difference\ land\ vegetation\ index = \frac{(a_{nir}-a_{vis})}{(a_{nir}+a_{vis})}$, where a_{nir} and a_{vis} represent surface reflectance's over ranges of wavelengths in the visible and near infrared regions (Asrar et al., 1984)

1.2 The origins of photosynthesis and photosynthetic bacteria

Although photosynthesis is generally regarded as an ancient process, probably evolving shortly after the origin of life (Olson and Blankenship, 2004), its origins will always be shrouded in mystery. Despite this, studies of fossils, phylogenetics and photosynthetic cell morphologies have enabled deductions to be made about the evolution of photosynthesis and its spread across several biological taxa.

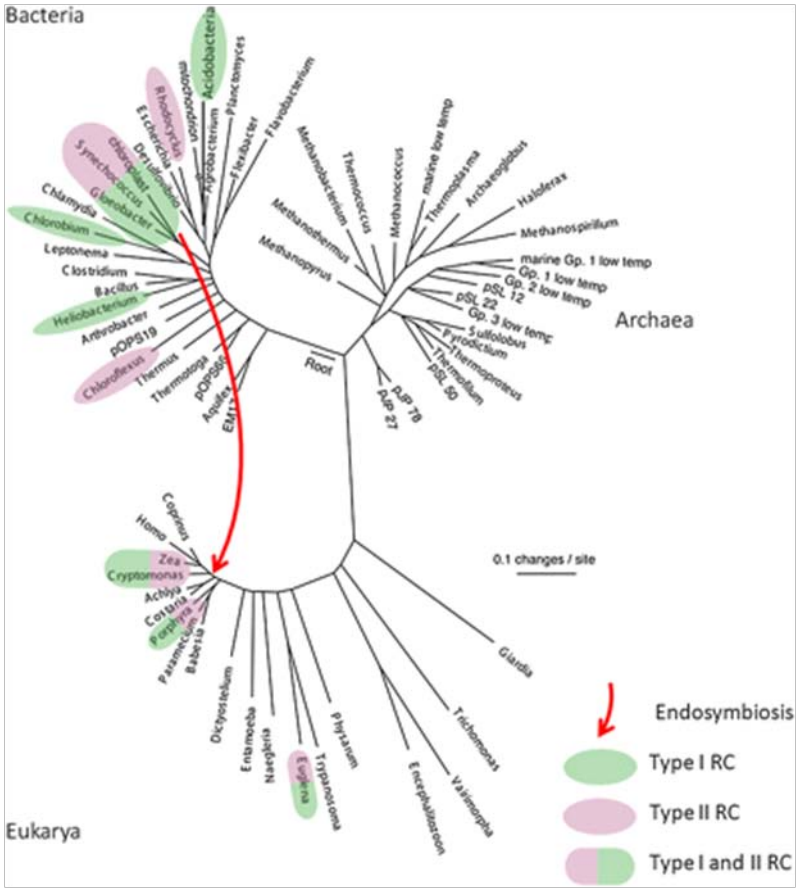

Figure 1.2 Hypothesised evolutionary relationships between the three domains of life based on small subunit rRNA phylogenetics

The taxa containing photosynthetic organisms are highlighted in green or purple, which denotes the reaction centre type; green represents reaction centre type I and purple represents reaction centre type II. The red arrow indicates the endosymbiosis event of the ancestral cyanobacterium (Pace, 1997; Blankenship, 2010).

Of the three domains of life, Eubacteria, Archaebacteria and Eukaryotes, chlorophyll-based photosynthesis is found only in subsets of the bacterial and eukaryotic domains, no chlorophyll-based photosynthetic Archaea has been discovered (**Figure 1.2**). The Archaea *Haleobacteria* are not classified as true photosynthetic organisms as they harness light

energy through bacteriorhodopsin using retinal, a carotenoid, and not a tetrapyrrole-based pigment (Oesterhelt and Stoeckenius, 1971).

A

Photosynthetic Bacteria					
Purple		Green		Cyanobacteria	Heliobacteria
Sulphur	Non-sulphur	Sulphur	Non-sulphur		
Thiospirillum	*Rhodobacter*	*Chlorobaculum*	*Chloroflexus*	*Synechocystis*	*Helicobacter*
Chromatium	*Rhodospirillum*	*Prosthecochloris*		*Synechococcus*	*Heliobacillus*
Thiocapsa	*Rhodopseudomonas*	*Pelodictyon*		*Cyanothece*	*Heliophilum*
Amoebacter	*Rhodocyclus*	*Chlorochromatium*		*Phormidium*	*Heliorestis*
Thiopedia	*Rhodomicrobium*	*Pelochromatium*		*Plectonema*	

B

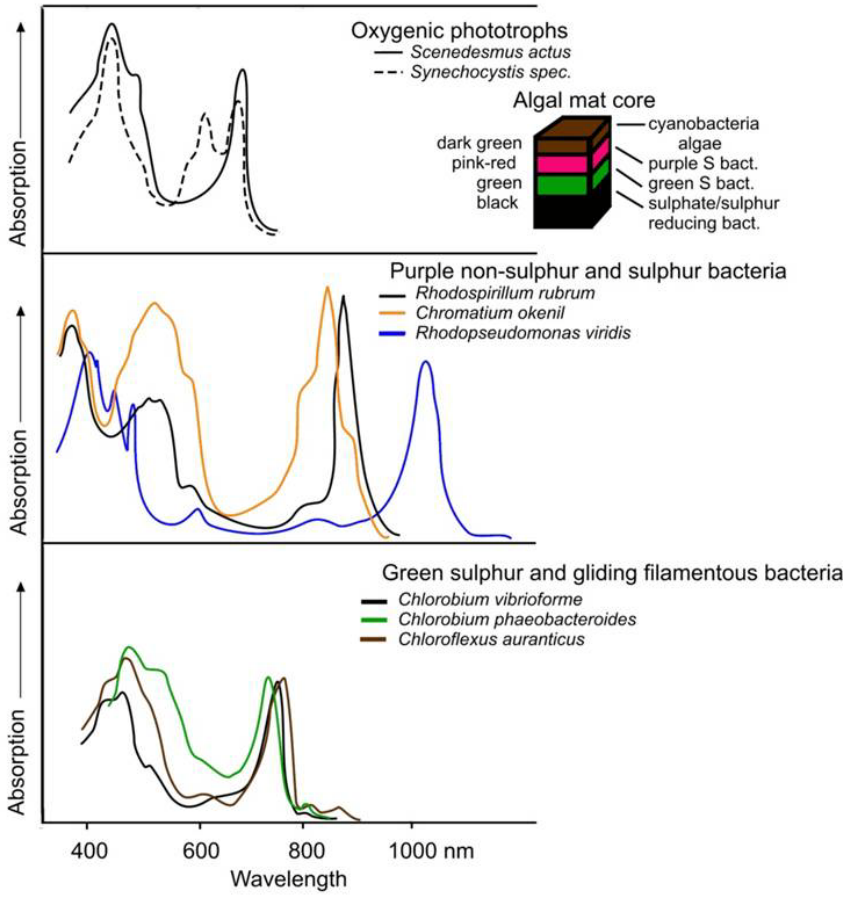

Figure 1.3 Taxonomy and absorption spectra of photosynthetic bacteria

Taxonomy of phototrophic bacteria (**A**). Absorption spectra of living photosynthetic organisms (**B**) (adapted from Garcia-Asua, 1999).

The ability to perform photosynthesis is found within six phyla of bacteria; these are the cyanobacteria, the proteobacteria or purple bacteria, the green sulphur bacteria, the firmicutes or heliobacteria, the green non-sulphur bacteria and the recently discovered

acidobacteria (Bryant *et al*., 2007; Raymond, 2008) (**Figure 1.3**). The extent of photosynthetic organisms within each phylum varies, with practically all cyanobacteria and green sulphur bacteria being phototrophic and the vast majority of species within the proteobacteria being non-phototrophs (Blankenship, 2010). A brief introduction to each of the photosynthetic phyla, with the exception of cyanobacteria, described in more detail later, is given below.

Proteobacteria (purple bacteria)

The proteobacteria are the most metabolically diverse phylum of organisms, which is reflected in their being named after Proteus, the Greek sea god, capable of assuming many identities. Organisms within this group have a remarkable range of metabolic capabilities, and include prokaryotes capable of growing aerobically, anaerobically, photoheterotrophically and photoautotrophically. Many organisms within this phylum also have the capability to utilise sulphur or phosphorus as an energy source, assimilate carbon and fix nitrogen. The photosynthetic members of the proteobacteria are known as the purple bacteria, owing to their purple pigmentation. These organisms are found littered among the alpha (*Rhodospirillum*, *Rhodobacter*), beta (*Rubrivivax gelatinosus*, *Rhodocyclus tenuis*) and gamma (*Halorhodospira halochloris*, *Ecotothiorhodospira*) proteobacteria (Hunter et al., 2009; Beatty, 2013).

Green sulphur bacteria

These organisms are anoxygenic phototrophs and typically inhabit freshwater lakes, sand flats and hyper saline waters such as the black sea. These bacteria are capable of growing to very high densities as their growth zones are limited to a narrow band in the water column that can provide both the reduced inorganic sulphur compounds and light intensities required for growth (Gloe et al., 1975; Overmann, 2001). They are differentiated from other types of photosynthetic bacteria, as organisms belonging to this phylum are non motile and contain *γ*-carotene as a major carotenoid component (Gibson et al., 1985). Both green and brown species make up the green-sulphur phylum; these are differentiated by their light-harvesting pigments. Green species contain bacteriochlorophyll *c* or *d*, and the carotenoids chlorobactene and hydroxy-chlorobactene. Whereas the brown species contain bacteriochlorophyll *e* and the carotenoids isorenieratene and β-isoernieratene (Gloe *et al*., 1975). The different pigment compositions are responsible for modifying light absorption, which is required to adapt the different organisms to different environments.

Firmicutes (heliobacteria)

First discovered in the 1980s, this group of strict anoxygenic phototrophic bacteria uniquely synthesise bacteriochlorophyll *g* and do not possess bacteriochlorophyll *a* (Gest and Favinger,

1983; Asao and Madigan, 2010). Requiring fairly high light intensities to grow, these organisms contain the simplest known photosynthetic complexes, which are located in the plasma membrane, as heliobacteria lack specialised photosynthetic membranes (Woese *et al*., 1985; Kimble-Long and Madigan, 2001). Heliobacteria may grow photoheterotrophically on a limited range of organic substrates or by fermentation of pyruvate in the dark, as they are unable to grow photoautotrophically by solar energy alone (Kimble *et al*., 1994; Ormerod *et al*., 1996).

Green non-sulphur bacteria (Chloroflexi)

Although referred to as "green non-sulphur" bacteria several members of this phylum utilise sulphide as a source of electrons for photosynthesis (Keppen *et al*., 2000; Klappenbach and Pierson, 2004). The first organism belonging to this phylum, *Chloroflexus aurantiacus*, was first isolated from a hot spring in 1974 (Pierson and Castenholz, 1974). Since then several other members of this family have been isolated. Organisms within this phylum characteristically form filaments, use chlorosomes as light harvesting antennae and display gliding motility.

Acidiobacteria

The most recent phototroph discovered is the bacterium *Candidatus* Chloracidobacterium thermophilum, which was isolated from a hot spring microbial mat in 2007 (Bryant *et al*., 2007). This newly discovered aerobic phototroph is distinctly unusual in that it contains chlorosomes and the Fenna-Matthews-Olsen (FMO) complex, previously seen only in anaerobic phototrophs. Although members of the *Acidobacter* phylum are notoriously difficult to isolate and grow under laboratory conditions, it has been possible to grow *Ca*. Cab. thermophilum *C. Thermophilum* in a mixed culture to study its FMO-chlorosome composition (Tsukatani *et al*., 2010; Wen *et al*., 2011).

1.2.1 The evolution of photosynthesis

The ability to photosynthesise is thought to have been passed onto eukaryotes through a single endosymbiosis event, where a non-photosynthetic protist engulfed and enslaved a cyanobacterial-like organism (Mereschkowsky, 1905; Mereschkowsky, 1910; Margulis, 1970; Price *et al*., 2012). Over time genes from the enslaved organism were transferred to the host nuclear DNA, lost through lack of selection pressure or substituted for a gene already present in the host, reducing it to the organelle now known as the plastid (Nugent and Palmer, 1988; Martin and Schnarrenberger, 1997; Delwiche, 1999). This organism then gave rise to the red, green and glaucophyte algae (Archibald, 2005). Plastids were then transferred horizontally across the eukaryotic lineages in secondary and tertiary endosymbiotic events to give rise to

the phyla Euglenids, Ciliates, Cryptomonads and Dinophysls to name a few examples (Cavalier-Smith, 1999; Yoon *et al*., 2004; Bhattacharya *et al*., 2004; Archibald, 2005) (**Figure 1.4**).

The evolution of photosynthesis within the photosynthetic bacteria remains unclear and is complicated by several factors including horizontal gene transfer, the components of the photosynthetic apparatus arising from various evolutionary pathways and the differing methods used to calculate evolutionary timelines. For example, phylogenetic studies that use non-photosynthetic genes as markers result in vastly different evolutionary pathways to studies that use photosynthetic genes as markers. A phylogenetic comparison of the non-photosynthetic genes, Hsp60 and Hsp70, (Gupta *et al*., 1999) suggests the Gram positive bacteria (of which heliobacteria are a subset) evolved first and subsequently diverged, along a linear pathway, into green non-sulphur bacteria, cyanobacteria, green sulphur bacteria and, lastly, purple bacteria. However, a comparison of the Mg-tetrapyrrole biosynthesis genes places the proteobacteria as the most ancient photosynthetic lineage, followed by the closely related green non-sulphur and green sulphur bacteria; emerging later are the heliobacteria, cyanobacteria and photosynthetic eukaryote lineages (Xiong *et al*., 2000). Consequently, it remains unresolved as to whether photosynthesis evolved once, subsequently being lost from many phyla and spreading to others via horizontal transfer events, or arose in multiple events across.

Perhaps the easiest origin to pinpoint should be that of the oxygen-producing cyanobacteria. These early organisms are thought to be responsible for the oxygenation of earth's atmosphere enabling the eventual evolution of aerobic life-forms and protective ozone layer. However, there is much controversy between the fossil and geological records. The geological evidence points towards molecular oxygen accumulating in the atmosphere ~ 2.7 billion years ago (Buick, 1992) , suggesting cyanobacteria will have evolved slightly before this time point, as the early oxygen produced will not have accumulated in the atmosphere, but would have been consumed in the oxidation of iron and sulphur compounds on the earth's surface (Blankenship, 1992; Xiong and Bauer, 2002). However, fossil evidence suggests cyanobacterial-like organisms were present much earlier, perhaps as early as 3.4 billion years ago (Schopf, 1993; Schopf *et al*., 2002). Therefore the question of when these organisms first evolved remains unanswered.

Today the cyanobacteria are widespread across the globe and can be found inhabiting terrestrial, freshwater and marine environments, as well as the more extreme habitats such as hot springs and saline lakes. They are a popular model organism for the study of oxygenic

photosynthesis, as they are possible progenitors of plant plastids, have relatively simple genetics in comparison to higher plants and are the only photosynthetic prokaryotes known to engage in oxygenic photosynthesis. Commonly known as "blue-green algae", cyanobacteria are Gram-negative photoautotrophs that have proven amenable to genetic manipulation, enabling studies of gene function and regulation. The work presented in this thesis was compiled through studies on the cyanobacterium *Synechocystis sp* PCC6803*,* from hereafter referred to as *Synechocystis*.

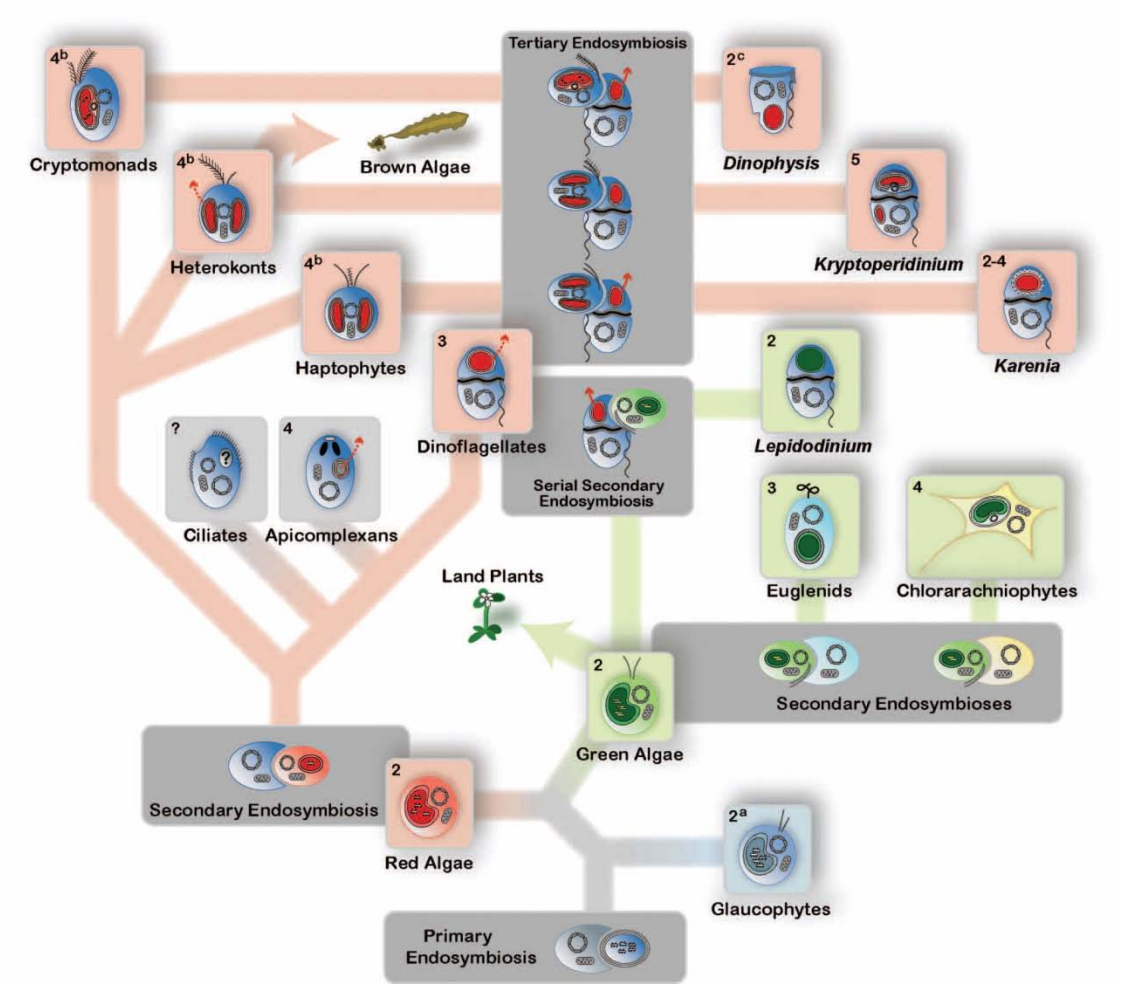

Figure 1.4 Hypothesised origin and spread of the plastid from primary endosymbiosis throughout the photosynthetic eukaryotes

This hypothesis is based on a consensus of morphological, molecular and biochemical data. Three photosynthetic lineages arise from the original endosymbiosis event, the red algae (eventually giving rise to brown algae), the green algae (eventually giving rise to the land plants) and the glaucophyptes. Both red and green algae are thought to have undergone secondary endosymbiosis events, as evidenced through phylogenetic studies and the number of membranes binding the arising plastid organelle. In the case of the Dinophysis, Kryptoperidinium, Kareina and Lepidodinium, the ancestral plastid has been replaced though a tertiary or serial secondary endosymbiosis event (Archibald, 2005).

1.3 *Synechocystis* as a model organism

Cyanobacteria have been popular model organisms since the 1970s. The cyanobacterium *Synechocystis* was first isolated from a freshwater lake and deposited in the Pasteur Culture Collection in 1968. However, it was not until the 1990s that this organism became a mainstream model organism for photosynthesis research. This can be attributed to the sequencing of the *Synechocystis* genome in 1996 (Kaneko *et al*., 1996), which was the first genome of a photosynthetic organism to be sequenced and only the third complete sequenced after those of *Haemophilus influenza* and *Mycoplasma genitalium*, as well as the generation of glucose tolerant *Synechocystis* strain that could propagate under photomixotrophic growth conditions (Rippka R et al., 1979; Williams, 1988) The completed sequence is 3,573,470 bp in length and contains an estimated 3,168 potential protein coding genes.

As discussed earlier, an ancient cyanobacterium is thought to be the progenitor of plant plastids, as members of this phylum are the only photosynthetic bacteria known to contain chlorophyll *a*. This homology is also reflected in plants and cyanobacteria using the same sets of proteins, with a few exceptions, to synthesise chlorophyll *a*, and assemble the photosynthetic apparatus (Vavilin and Vermaas, 2002).

Like many cyanobacteria, *Synechocystis* is naturally competent, capable of taking up DNA from the environment and incorporating it into its genome (Grigorieva and Shestakov, 1982). This natural phenomenon makes manipulation of the genome relatively straightforward. Constructs can be designed with flanking regions of DNA, homologous to the desired final location within the *Synechocystis* genome, enabling the insertion or deletion of genes using selection markers (Tsinoremas *et al*., 1994). Interestingly the *Synechocystis* genome is polyploidal with the most recent estimate suggesting there are ~60 copies of the genome in each cell (Griese *et al*., 2011), which is vastly higher than the 12 copies originally reported (Labarre *et al*., 1989). In terms of selection, this means any construct placed in the *Synechocystis* genome needs to be integrated into each genome copy to prevent the resulting mutant from returning to the wild type state.

1.4 Tetrapyrroles

Tetrapyrroles are a large group of organic energy-capturing molecules composed of four pyrroles (C_4H_4NH) (designated rings A-D) arranged in a cyclic (porphyrins) or linear (phycobilins and bilins) fashion. Each of these pyrrole subunits is connected, with the

exception of corrins, by single carbon methyne bridges. The carbon atoms that form the tetrapyrrole backbone are numbered from 1-20 starting from ring A (**Figure 1.5**). Chlorophylls form a structurally and functionally distinct group within the porphyrin family; characterised the presence of a fifth isocyclic ring, the "E" ring, and a phytol moiety esterified at the C17 position. In the majority of cases a magnesium ion is bound via co-ordinating bonds to the four central nitrogen atoms of the tetrapyrrole, the exception being a bacteriochlorophyll with a centrally co-ordinated zinc ion, found in the aerobic bacterium *Acidiphilium rubrum* (Matsuzawa *et al*., 2000). The pigment absorbs strongly in the red and blue-violet regions of the visible electromagnetic spectrum, reflecting green light to give photosynthetic organisms their characteristic colour (Zscheile *et al*., 1944). Chlorophylls are found in almost all photosynthetic organisms due to their ability for efficiently capturing solar radiation.

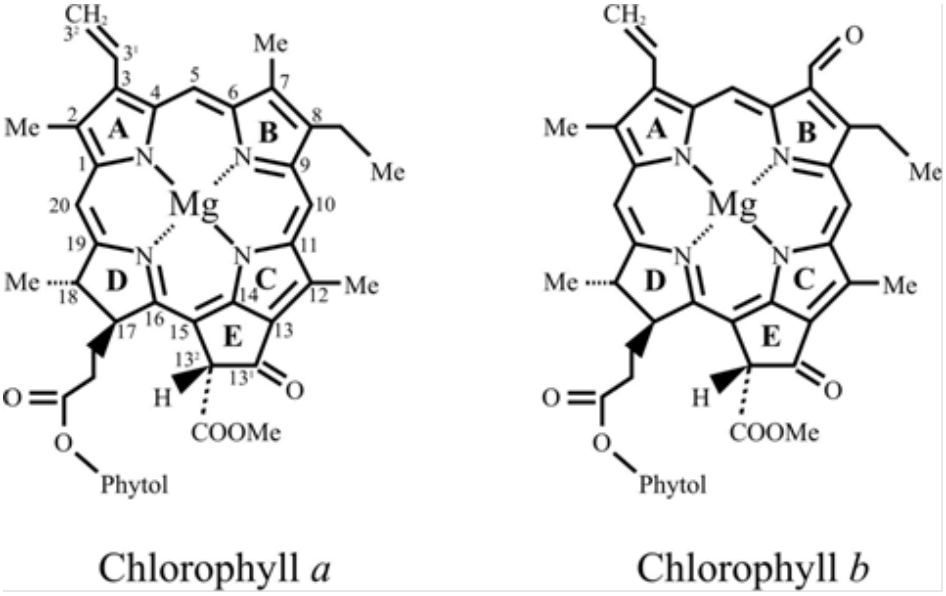

Figure 1.5 Structures of chlorophylls *a* and *b*

The carbon skeleton is numbered according to the IUPAC system. Chlorophylls *a* and *b* differ only at the C7 side group, where a methyl group has been modified to a carbonyl group.

There are five known types of chlorophyll: *a*, *b*, *c*, *d*, and the most recently discovered *f*. The most abundant of these is chlorophyll *a*, which plays a vital role in the reaction centres and core antennae of almost all oxygenic photosynthetic organisms including *Synechocystis*. *In situ* chlorophyll *a* has absorption maxima at 440 nm and 670 - 720 nm, which are located in two narrow bands at the edges of the visible spectrum. As a result photosynthetic organisms almost always contain other light-absorbing pigments to maximise their photoactive absorption range. Quite often these accessory pigments are other chlorophyll molecules, for example, chlorophyll *b*, which is found in all lineages of green eukaryotes, differs from chlorophyll *a* by the addition of a carbonyl-group at the C7 position (**Figure 1.5**). This

modification alters the *in situ* absorbance maxima to 460 nm and 630 - 680 nm, narrowing the green absorption gap and extending the photoactive wavelengths that can be harvested.

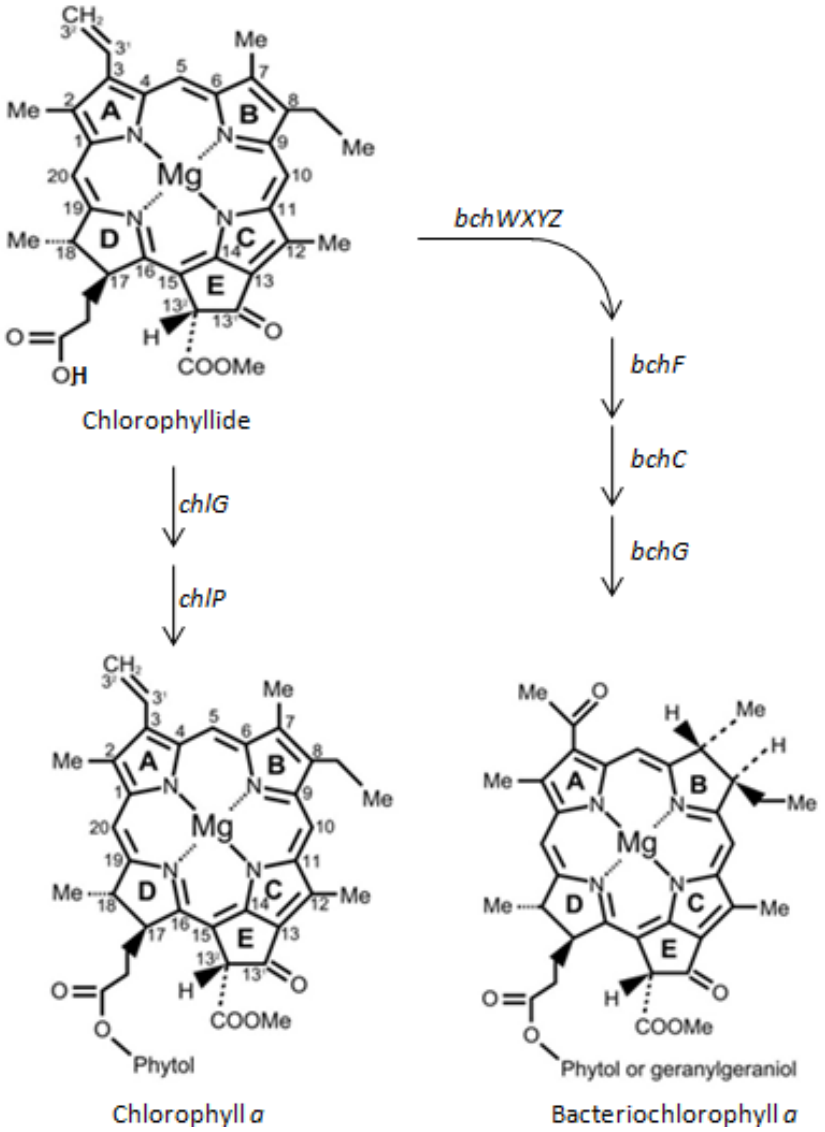

Figure 1.6 Bacteriochlorophyll *a* & chlorophyll *a* biosynthesis pathways

Both pathways share common intermediates up to chlorophyllide, at this point the tetrapyrrole ring in bacteriochlorophyll biosynthesis undergoes additional modifications. BchXYZ encode the chlorin reductase, which reduces the double bond between C_7 and C_8 (McGlynn and Hunter, 1993). BchF appears to hydrate the vinyl group to generate a hydroxyl intermediate, BchC further oxidises the hydroxyl group to a ketone, to form the C_3 oxo group of bacteriochlorophyllide (Wellington and Beatty, 1989; Burke *et al*., 1993a) and BchG is responsible for esterification of the geranylgeranyl or phytol tail.

The majority of photosynthetic bacteria, excluding the cyanobacteria, utilise bacteriochlorophyll, not chlorophyll as the major light harvesting pigment. Bacteriochlorophylls are biologically similar to chlorophylls, differing only in the number of

reduced pyrrole rings. Chlorophyll has only one reduced pyrrole ring, ring D, whereas bacteriochlorophyll has two reduced pyrrole rings, rings B and D. The 'tail' of bacteriochlorophyll may be phytol or geranylgeraniol, whilst the tail of chlorophyll only ever contains phytol (**Figure 1.6**). As *Synechocystis* only produces chlorophyll *a*, the biosynthesis of chlorophylls, and not bacteriochlorophylls, will be focused upon in this introduction.

In plants, algae and the majority of bacteria, tetrapyrroles, including chlorophylls, originate from a common biosynthetic pathway of which the earliest precursor is the 5-carbon compound δ-aminolaevulinic acid (ALA). In photosynthetic organisms, there are six enzymatic steps from the formation of ALA to the production of protoporphyrin IX, which represents the branch point between the haem and chlorophyll biosynthesis pathways. Here, the fate of the porphyrin molecule depends upon the metal ion chelated into the macrocycle, whereby the insertion of a magnesium ion represents the first committed step of chlorophyll biosynthesis and the insertion of an iron ion commits the porphyrin to haem biosynthesis.

Figure 1.7 and **Figure 1.8** depict the general chlorophyll biosynthetic pathway starting with the synthesis of ALA and ending with the synthesis of chlorophyll *b*. The pathway is generally divided into three sections; a) the formation of ALA, b) the condensation and modification of 8 ALA molecules to form protoporphyrin IX, c) The chelation of magnesium into protoporphyrin IX and the formation of chlorophyll *a*. A detailed description of the enzymes involved in chlorophyll biosynthesis is the focus of the next section of this chapter.

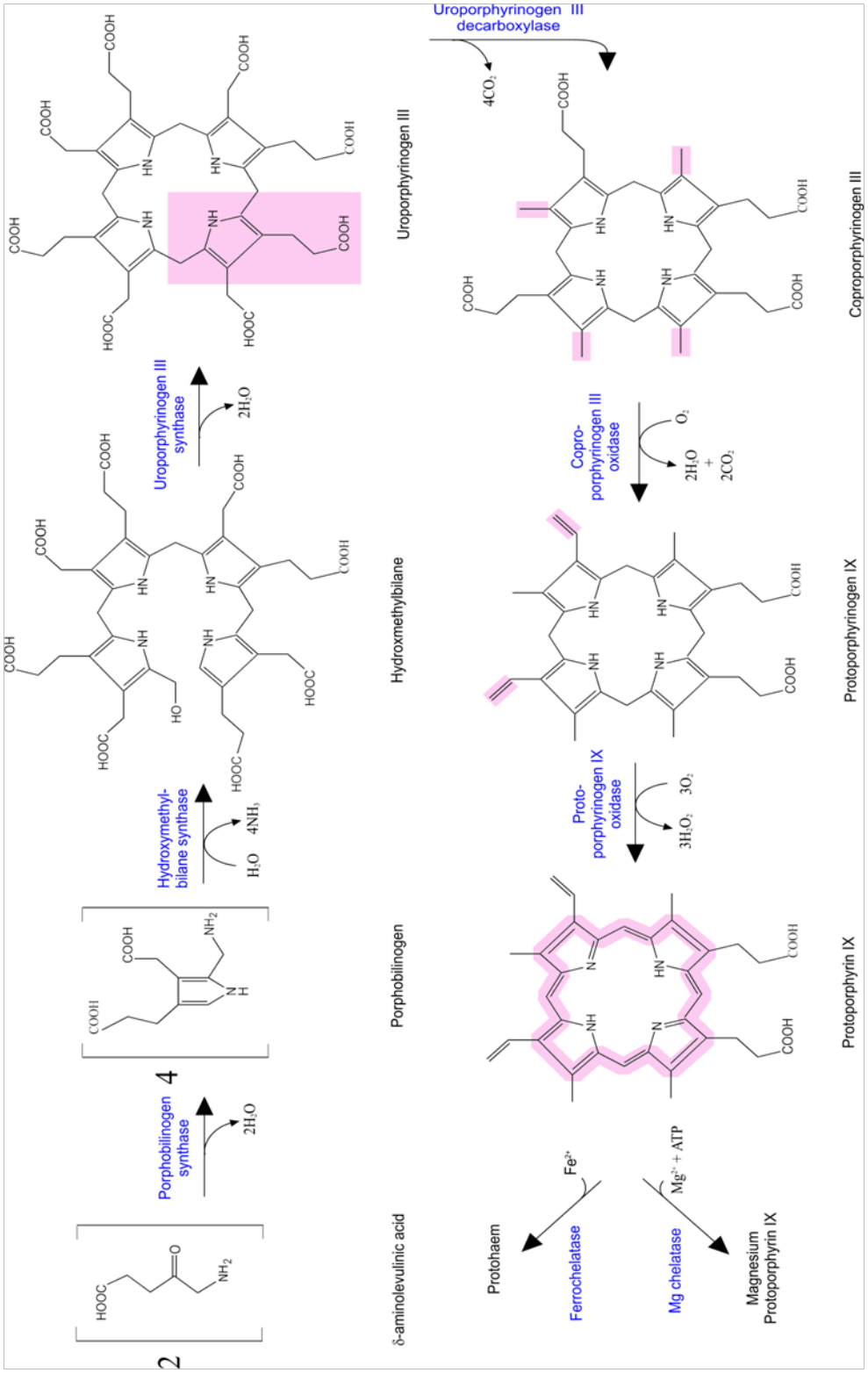

Figure 1.7 The chlorophyll biosynthetic pathway: δ–aminolaevulinic acid to protoporphyrin IX

The enzymes catalysing each step are shown in blue. Magenta shading indicates the groups modified at each stage.

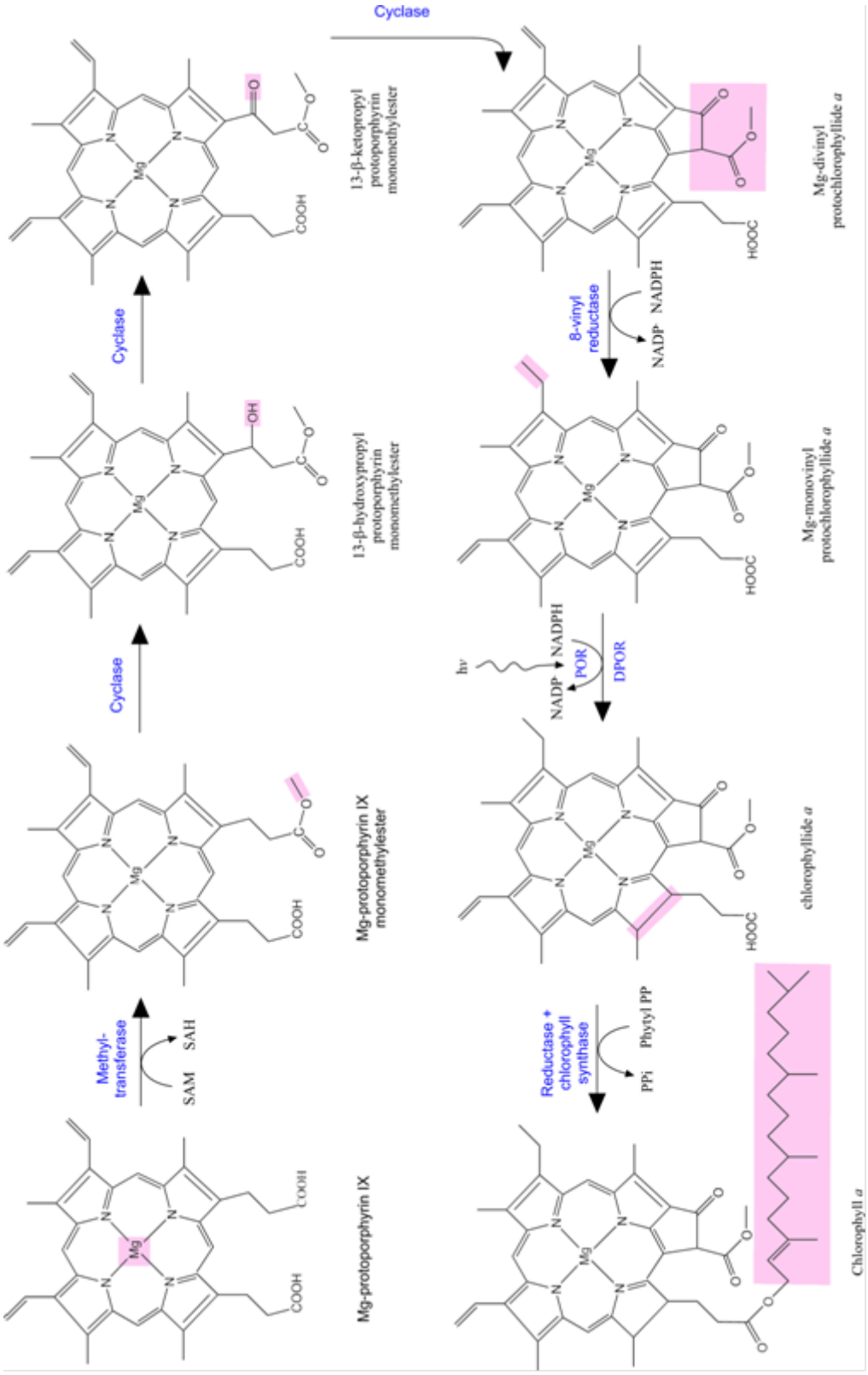

Figure 1.8 The chlorophyll biosynthetic pathway: Mg-protoporphyrin IX to chlorophyll *a*

The enzymes catalysing each step are shown in blue. Magenta shading indicates the groups modified at each stage.

1.4.1 Formation of δ-aminolaevulinic acid

ALA is the universal starting point for porphyrin biosynthesis, a precursor for many tetrapyrroles including the haems, vitamin B_{12}, siroheme and the bilins. There are two known pathways that lead to ALA formation, the C_4 or Shemin pathway and the C_5 pathway. In general, the Shemin pathway is used by non-plastid eukaryotes and α-proteobacteria, while the C_5 pathway is found in plants, algae and the majority of bacteria. There are a few exceptions to this rule, which are a small number of organisms that have the capability to utilise both pathways including: *Streptomyces nodosus*, *Euglena gracilis* (Mayer and Beale, 1992) and *Scenedesmus obliquus* (Petricek *et al*., 2006).

The Shemin pathway consists of a single reaction catalysed by the enzyme ALA synthase. Here, the enzyme condenses a molecule of succinyl-co-enzyme A and a molecule of glycine in a reaction that releases ALA and carbon dioxide as products and regenerates co-enzyme A (**Figure 1.9**). The final ALA molecule is composed of all five carbon atoms from the succinate molecule and a single carbon from glycine; the additional glycine carboxyl carbon is lost as carbon dioxide (Ano *et al*., 1999).

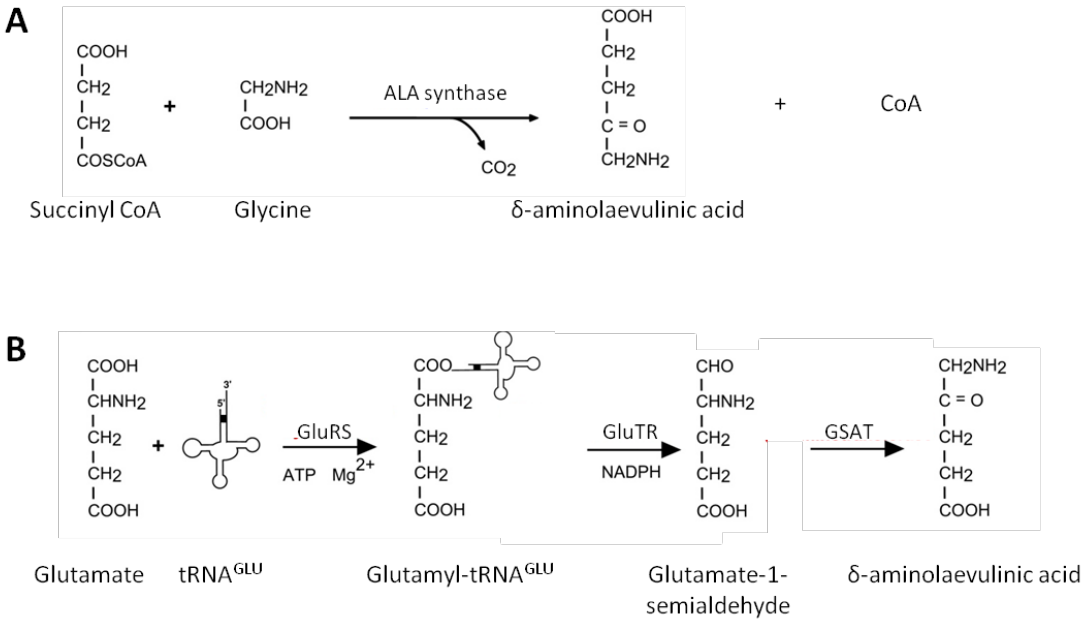

Figure 1.9 δ-aminolaevulinic acid formation by the Shemin and C_5 pathways

The Shemin pathway, showing the formation of ALA via the condensation of glycine and succinyl coenzyme A, catalysed by ALA synthase (**A**). The C_5 pathway, showing the formation of ALA via an activated $tRNA^{GLU}$ intermediate, catalysed by glutamyl tRNA synthetase, glutamyl tRNA reductase and glutamate-1-semialdehyde aminotransferase (**B**).

In the C_5 pathway the 5 carbon compounds glutamate, α-ketoglutarate or glutamine are the source of the carbon atoms in ALA (Beale and Castelfranco, 1974a; Beale and Castelfranco, 1974b). Unlike the Shemin pathway, the C_5 pathway consists of three different enzymes, glutamyl-tRNA synthetase (GluRS), glutamyl-tRNA reductase (GluTR) and glutamate-1-semialdehyde aminotransferase (GSAT), which work together to convert glutamate to ALA with the aid of a regenerated $tRNA^{GLU}$ (**Figure 1.9**). In this series of reactions, ALA formation begins with the ATP and Mg^{2+} dependent activation of glutamate, whereby the carboxyl group of glutamate is ligated to $tRNA^{GLU}$ by GluRS to produce glutamyl-$tRNA^{GLU}$. The carboxyl group is then reduced to a formyl group, producing glutamate-1-semialdehyde (GSA), in an NAD(P)H dependent reaction catalysed by GluTR that regenerates $tRNA^{GLU}$. Lastly, GSA is transaminated by GSAT by transfer of the amino group from C4 to C5 of GSA, culminating in the release of ALA (Jahn *et al*., 1992).

Glutamyl-tRNA synthase (GluRS)

Encoded by the *gltX* gene, GluRS is a member of the class I aminoacyl-tRNA synthetase family (Kern and Lapointe, 1979a; Kern and Lapointe, 1979b). This enzyme was first discovered when its activity was abolished in cell-free extracts pre-incubated with RNAse A; activity was returned upon addition of a ribonuclease inhibitor and tRNA (Huang *et al*., 1984; Schon *et al*., 1986; Huang and Wang, 1986). The reaction catalysed by GluRS, like all reactions catalysed by aminoacyl-tRNA synthetases, consists of a two-step process, in which glutamate is first converted to an aminoacyl adenylate in an activation step requiring Mg^{2+} and ATP and subsequently transferred to a tRNA to form $tRNA^{GLU}$ (Schimmel, 1979; Schimmel, 1987). It has been shown that *E. coli* GluRS is one of only three aminoacyl-tRNA synthetases that cannot carry out the activation of amino acid in the absence of tRNA (Schimmel, 1979; Schimmel, 1987). Interestingly, the GluRS is responsible for charging $tRNA^{GLU}$ for both protein and ALA synthesis, with only a small fraction of $tRNA^{GLU}$ being channelled into the ALA pathway (Beale *et al*., 1975). In the acidophilic bacterium *Acidithiobacillus ferrooxidans*, the activity of GluRS is regulated by intracellular levels of haem; increased levels of GluRS are observed under growth conditions with a high haem requirement and inhibition of GluRS by hemin is observed when intracellular haem is in excess (Levican *et al*., 2007).

The structure of *E. coli* GluRS, complexed with ATP and $tRNA^{GLU}$, has been resolved to a resolution of 2.5 Å (Perona *et al*., 1993) and the crystal structures of *Thermus thermophilus* GluRS (**Figure 1.10**) were solved for the un-complexed (Nureki *et al*., 1995) and $tRNA^{GLU}$ complexed enzyme (Sekine *et al*., 2001; Sekine *et al*., 2006). The GluRS structures show that the enzyme has five distinct domains, which are arranged in an elongated curve (**Figure 1.10 B**). Domain I hosts a Rossmann ATP-binding fold, with two characteristic ATP binding motifs

(HIGH/HVGG and KMSKS) and is conserved among all class-I tRNA synthetases. Domains II and III are located either side of the ATP-binding pocket of domain I and retain a conserved structure, which creates a groove to accommodate the tRNA acceptor arm (**Figure 1.10 A**). Domain III also contains three conserved α-helices, which form the basis for interactions with the tRNA anticodon stem. The two remaining domains, domains IV and V, are unique to GluRS and are required for recognition of the $tRNA^{GLU}$ anticodon. Mutational analysis of arginine R358 revealed this residue plays a critical role in GluRS recognition of the $tRNA^{GLU}$ anticodon (Sekine *et al*., 2006).

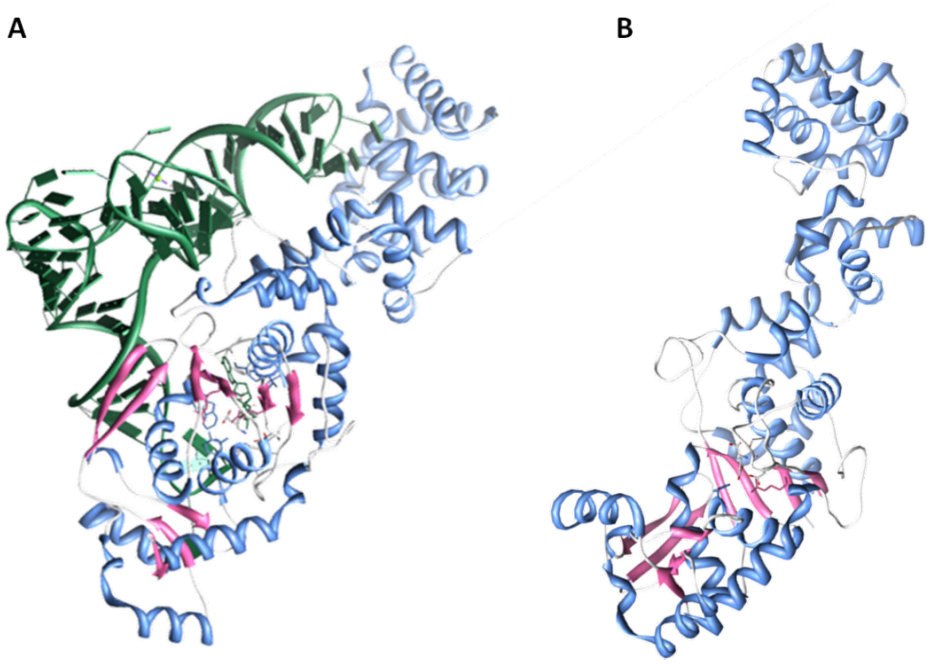

Figure 1.10 Crystal structure of *Thermus thermophilus* GluRS

The crystal structures of GluRS in a complex with $tRNA^{GLU}$, ATP and L-glutamate (**A**) and with L-glutamate (**B**) at resolutions of 2.2 Å and 1.98 Å, respectively (PBD 2DXI and PBD 2CUZ, Sekine *et al*., 2006).

Glutamyl-tRNA reductase (GluTR)

The second enzyme of the C_5 pathway, GluTR, is encoded by the *hemA* gene. Notorious as the most difficult enzyme of the C_5 pathway to study, GluTR is instable, present at very low abundance and produces a highly reactive intermediate, GSA (Beale, 2006). GluTR catalyses the NADPH-dependent reduction of the $tRNA^{GLU}$ bound glutamate to GSA. Before the crystal structure of GluTR was obtained, early investigations of GluTR resulted in much controversy regarding cofactor requirements, potential catalytic mechanisms and molecular weight. For example, gel filtration of *Bacillus subtilis* GluTR suggests the enzyme is an oligomer of 230,000 Da (Schroder *et al*., 1992), purified Barley GluTR is a pentamer of 270,000 Da (Bougri

and Grimm, 1996) and GluTR purified from *Chlamydomonas* is monomer of 130,000 Da (Chen *et al*., 1990).

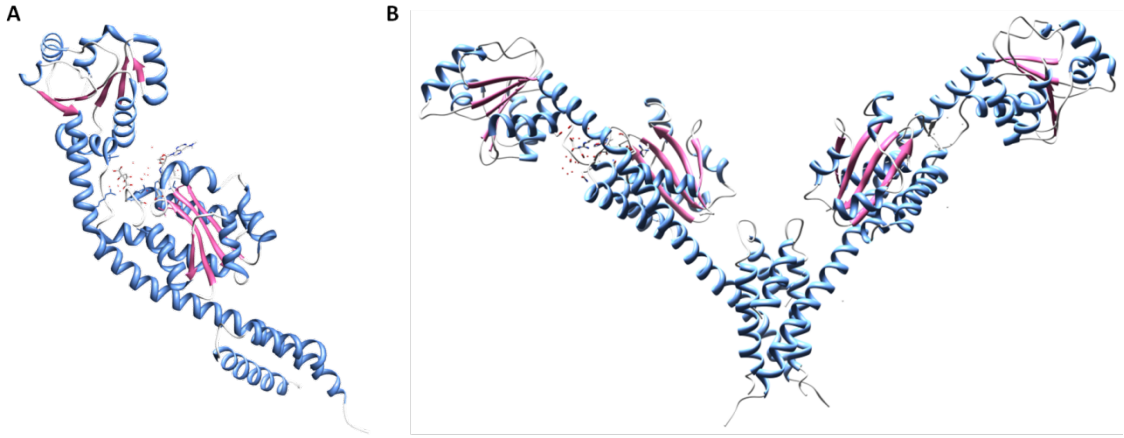

Figure 1.11 Crystal structure of *Methanopyrus kandlerii* GluTR

The crystal structure of GluTR in a complex with citric acid, glutamic acid and 6N-dimethyl-3-deoxyaminoadenosine, solved to a resolution of 1.95 Å (PBD 1GPJ, Moser *et al*., 2001b). GluTR monomer (**A**) and GluTR V-shaped diamer (**B**).

The crystal structure of GluTR from the hyperthermophilic archaeon *Methanopyrus kandlerii* was solved to a resolution of 1.9 Å (Moser *et al*., 1999; Moser *et al*., 2001a; Moser *et al*., 2001b). Although GluTR from *Methanopyrus kandlerii* was thought to be a tetramer, with a native molecular weight of 190,000 Da, the crystal structure revealed an extended V-shaped dimer, (**Figure 1.11 B**) which *in vivo* presumably forms a dimer of dimers. Each GluTR monomer consists of three distinct domains arranged along a α-helical spine (**Figure 1.11 A**). The first domain, located at the N-terminus, is the catalytic domain and is composed of two subdomains; a small βαββααβ section and a four helix bundle, comprised from three short α-helices and a section of the longer α-helical spine. Domains 1 and 2 are linked by the spinal α-helix and a short loop. The NADPH binding site is located in domain 2, which forms a classical NAD(P)H binding fold. Domain 3 is the dimerisation domain, where three helices from each monomer intertwine to form a six-helix bundle (**Figure 1.11 B**) (Moser *et al*., 2001b). Activity studies on this enzyme found only NADPH and no other co-factor was required for activity, however, the *Synechocystis* and *Chlorella vulgaris* GluTR's both require a divalent metal ion, Mg^{2+}, Mn^{2+} or Ca^{2+}, but not Zn^{2+}, for activity (Mayer *et al*., 1994).

Given the reactive nature of the substrate for GluTR, glutamyl-$tRNA^{GLU}$, it is not surprising a tight complex was found to form between GluTR and the preceding enzyme in the pathway, GluRS. The complex is formed only in the presence of glutamyl-$tRNA^{GLU}$ or free $tRNA^{GLU}$, ATP and glutamate. Upon the addition of NADPH and reduction of tRNA bound glutamate to GSA

the complex dissociates (Jahn, 1992), presumably so GluTR can form a complex with GSAT, the enzyme proceeding GluTR, to protect the highly reactive GSA intermediate (Luer *et al*., 2005; Layer *et al*., 2010). Immuno-precipitations, *in vivo* cross-linking reactions and sucrose gradient ultracentrifugation show GSAT and GluTR form a complex to enable channelling of GSA (Nogaj and Beale, 2005a).

Glutamate-1-semialdehyde aminotransferase (GSAT)

The third enzyme in the C_5 pathway, GSAT, is encoded by the *hemL* gene and represents a typical aminotransferase in structure (Hennig *et al*., 1997a). Although ALA can be formed spontaneously at physiological pH in the presence of high concentrations of GSA, GSAT is predominantly responsible for the transamination of GSA to ALA. This is achieved by interchanging the amino and aldehyde groups at the C1 and C2 positions (Gough *et al*., 1989). Unusually, GSAT is able to directly exchange the amino and aldehyde groups without the need for an additional acceptor or donor molecule; a unique feature among aminotransferases (Smith *et al*., 1992). The only cofactor required by GSAT for function is pyridoxal-phosphate, which was deduced by the observed inhibition of GSAT by the pyridoxal-phosphate inhibitors amino-oxyacetate and gabaculine (Soper and Manning, 1982; Hoober *et al*., 1988). Purified GSAT proteins from a range of species have been shown to have a molecular mass of 80 – 100 kDa and many of the plant and bacterial proteins have high similarity to the aspartate aminotransferase family (Avissar and Beale, 1989; Elliott *et al*., 1990; Mehta and Christen, 1994).

X-ray crystallography was used to solve the crystal structures of GSAT from *Synechococcus* (**Figure 1.12 A**) (Hennig *et al*., 1997a), *Thermosynechococcus elongatus* (**Figure 1.12 B**) (Schulze et al., 2006) and *Bacillus subtilis* (**Figure 1.12 C**) (Ge *et al*., 2010), to resolutions of 2.4 Å, 2.9 Å and 2.3 Å respectively). All three proteins crystallised as dimers and contain three domains (**Figure 1.12**). The structures also show GSAT is structurally related to the extensive pyridoxal-phosphate-dependent protein family; other members include: aminotransferases, decarboxylases, racemases, synthetases and mutases.

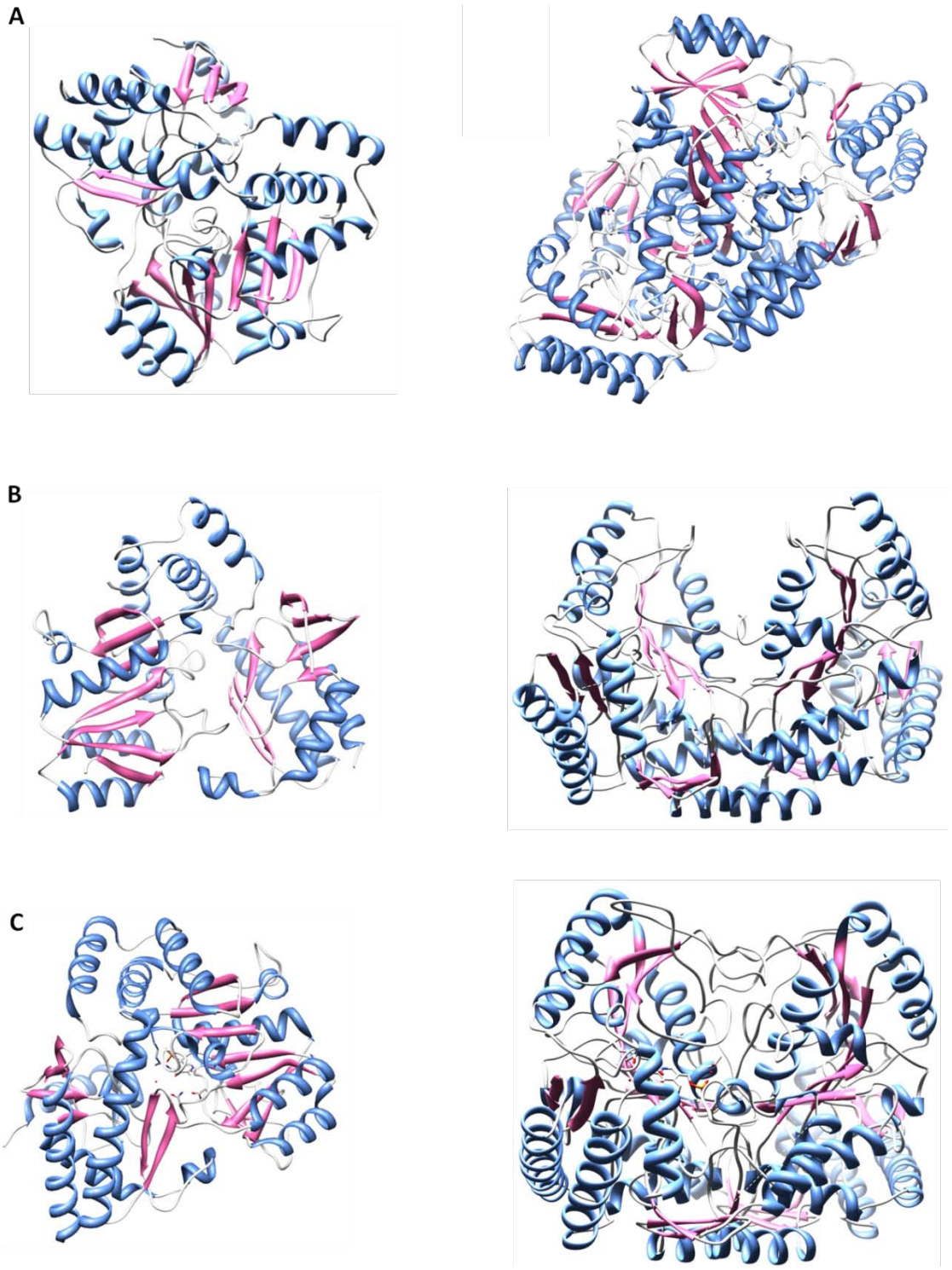

Figure 1.12 Crystal structures of GSAT from *Synechococcus*, *Thermosynechococcus elongatus* and *Bacillus subtilis*

Synechococcus GSAT asymmetric unit and homodimer (**A**), in a complex with pyridoxal-phosphate and PMP at a resolution of 2.40 Å (PDB 2GSA, Hennig *et al*., 1997). *Thermosynechococcus elongatus* GSAT asymmetric unit and homodimer (**B**) solved to a resolution of 2.85 Å (PDB 2CFB, Schulze *et al*., 2006). *Bacillus subtilis* GSAT asymmetric unit and homodimer (**C**), in a complex with pyridoxalphosphate at a resolution of 2.3 Å (PDB 3BS8, Ge *et al*., 2010).

The *Synechococcus* GSAT (**Figure 1.12 A**) is an ellipsoidally shaped dimeric protein, whose monomers interact along a large convoluted interface. Each monomer can be divided into three domains, an N-terminal domain, the cofactor binding domain, and the C-terminal domain. The N-terminal domain is ~70 residues in length and comprised of an α-helix and a three-stranded anti-parallel β-sheet. The catalytic, co-factor binding domain is composed of a central seven-stranded β-sheet, which has six parallel strands and one anti-parallel strand and the C-terminal domain is made up of a three-stranded β-sheet, coated on the outer surface by four α-helices. The *Bacillus* GSAT is structurally very similar, differing slightly by the presence of three additional short segments of α-helix in the N-terminal domain and in domain 3, where *Synechococcus* GSAT has a four-stranded anti-parallel β-sheet and *B. subtilis* GSAT has a two-stranded antiparallel β-sheet. Both crystal structures show the two active sites are located at the dimer interface and, at this fixed snapshot in time, they show one monomer is in a closed active position, whilst the other is in an open conformation. It is thought the dimer oscillates between one monomer being closed in the active site position, whilst the other monomer is open, allowing for substrate to enter or leave the active site (Hennig *et al*., 1997b).

1.4.2 Condensation of δ-aminolaevulinic acid to porphobilinogen

ALA dehydratase (ALAD), also known as porphobilinogen synthase, catalyses the asymmetric condensation of two ALA molecules to form porphobilinogen in the first step common to all known tetrapyrrole biosynthetic pathways. Porphobilinogen, a monopyrrole, is the first pyrrole molecule in the tetrapyrrole pathway and is defined by a five-membered ring with the formula C_4R_4NH, where R may be hydrogen or a functional group. It is porphobilinogen that forms the building block for the formation of the remainder of the tetrapyrrole (Jordan, 1991).

Primary sequence analysis of ALAD, which is encoded by the gene *hemB*, reveals a high level of homology across a core sequence of 330 residues in plants, animals and bacteria (Jaffe et al., 1995; Jaffe, 1999). The ALAD monomer is highly conserved between organisms and is ~ 35 kDa in size. The protein usually functions as a homooctamer with the exception of *Rhodobacter capsulatus* ALAD, which functions as a hexamer (Bollivar *et al*., 2004).

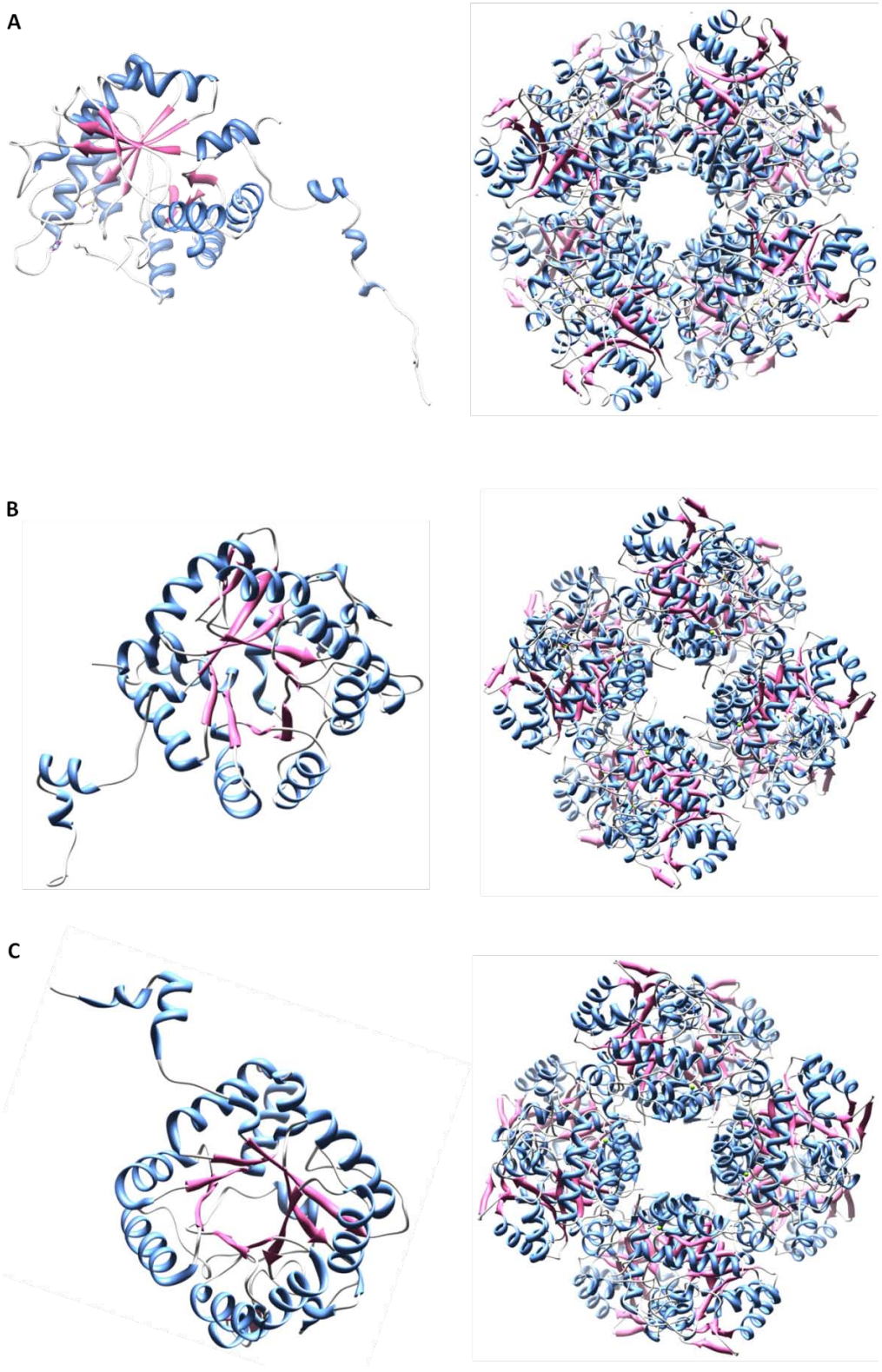

Figure 1.13 Crystal structures of ALAD enzymes

Saccharomyces cerevisiae ALAD monomer and homooctamer (**A**) solved to a resolution of 2.30 Å (PDB 1AW5, Erskine et al., 1998). *Pseudomonas aeruginosa* ALAD monomer (with a co-ordinated magnesium and sulphate ion) and homooctamer (**B**) at a resolution of 1.67 Å (PDB 1B4K, Frankenberg et al., 1999). *Chlorobium virbrioforme* ALAD monomer and homooctamer (**C**) at a resolution of 2.60 Å (PDB 1W1Z, Coates *et al*., 2004).

All ALADs have a requirement for divalent metal ions and usually fall into one of two categories; Mg^{2+} ion dependent or Zn^{2+} ion dependent. Human and yeast ALAD contain two Zn^{2+} ions, one of which is co-ordinated by highly conserved cysteine residues and plays a role in catalysis. The *E. coli* ALAD contains the active site Zn^{2+} ion and an additional Mg^{2+} ion, which stimulates activity (Mitchell and Jaffe, 1993). The plant ALADs contain a highly conserved Asp metal binding domain that co-ordinates an essential Mg^{2+} ion (Senior *et al*., 1996).

The crystal structures of ALADs have been solved for many organisms including, *Saccharomyces cerevisiae* (**Figure 1.13 A**) (Erskine *et al*., 1997), *Pseudomonas aeruginosa* (**Figure 1.13 B**) (Frankenberg *et al*., 1999), *E. coli* (Erskine *et al*., 1999) and *Chlorobium virbrioforme* (**Figure 1.13 C**) (Rhie *et al*., 1996). These reveal the ALAD structure is highly conserved and contains eight active sites, of which only four are functional in mammals. In general, the active site is located at the centre of a triose-phosphate isomerase-like αβ barrel with a mobile section of the protein serving as a lid, gating solvent access (Erskine *et al*., 1997; Jaffe, 2004). The ALAD monomers all contain a αβ-barrel like fold located in the major domain of the monomer. The ALAD monomers all stack into a concentric double ring to form a remarkably similar higher octameric structure.

1.4.3 Porphobilinogen deaminase and uroporphyrinogen III synthase

The formation of the first stable cyclic tetrapyrrole, Uroporphyrinogen III, requires two enzymes; porphobilinogen deaminase (PBGD) and uroporphyrinogen III synthase (UROS), which were first described in the late fifties (Bogorad, 1958a; Bogorad, 1958b; Bogorad, 1958c).

PBGD, also known as hydroxymethylbilane synthase, catalyses the stepwise condensation of four porphobilinogen molecules to form linear hydroxymethylbilane, in a reaction where each amino group is lost as ammonia (Jordan PM, 1991). The product is highly unstable and has a half life of 4 minutes at pH 8 and 37°C. In the absence of UROS the linear tetrapyrrole can spontaneously cyclise, giving rise to the biologically inactive isomer uroporphyrinogen I (Jordan and Seehra, 1979). Following synthesis hydroxymethylbilane is released into solution and picked up by UROS. This enzyme functions as both a cyclase and isomerase catalysing an unusual reaction where the D pyrrole ring inverts, via a spiro-pyrrolenine intermediate, interchanging the C_α atoms (Stark *et al*., 1993). In organisms where UROS is defective or absent, hydroxymethylbilane spontaneously converts to uroporphyrinogen I, which is then oxidised to Uroporphyrin I. This oxidation product can be found in the urine of patients suffering from congenital porphyria (Rimington and MILES, 1951; Louie et al., 1996).

Porphobilinogen deaminase (PBGD)

PBGD is encoded by the *hemC* gene, which, when translated forms a monomeric enzyme with a molecular mass ranging between 34 and 44 kDa, with optimal enzyme activity at pH 8 (Jordan, 1991). All PBGBs have a covalently tethered dipyrromethane cofactor, attached via a thioester link to an active site cysteine residue, which is absolutely conserved across many divergent species (Jordan *et al*., 1988; Warren and Jordan, 1988a; Witty *et al*., 1993). This cofactor is assembled by the apoenzyme from two molecules of porphobilinogen. Once formed, the protein-cofactor complex is permanent and highly important for stabilising the PBGD structure (Scott *et al*., 1989). The dipyrromethane acts as a scaffold, upon which four porphobilinogen molecules sequentially attach, functioning as a reaction primer to enable the stepwise construction of linear hydroxymethylbilane. After the sequential addition of four porphobilinogen molecules, a "hexapyrrole" is formed, composed of four hydroxymethylbilane rings and two dipyrromethane rings. This hexapyrrole is then cleaved in a hydration reaction to yield hydroxymethylbilane and regenerate the dipyrromethane co-factor (**Figure 1.14**) (Warren and Jordan, 1988b; Shoolingin-Jordan *et al*., 2003a). The dipyrromethane also appears to limit the total number of ligated porphobilinogen molecules to four, as the addition of the fourth molecule stimulates hydrolytic cleavage and the release of HMB (Battersby *et al*., 1983).

The structure of PBGB from *E. coli* (**Figure 1.15**) has been solved by X-ray crystallography to resolutions of 1.9 Å and 1.76 Å (Louie *et al*., 1992; Louie *et al*., 1996). The structures reveal that PBGB is a flexible three-domain polymerase, with a single catalytic site. The three domains are of similar size and domains 1 and 2 have a similar αβ topology, namely, a doubly wound five-stranded β-sheet composed of four parallel strands and one anti-parallel strands, with α-helices packed against each face of the sheet. Domain 3 consists of a three-stranded anti-parallel β-sheet, covered on one face by three α-helices. Although the dipyrromethane cofactor is covalently linked to C242 in domain 3, it was found to form an extensive network of salt-bridges and hydrogen bonds in the cleft between domains 1 and 2 (Louie *et al*., 1992; Louie *et al*., 1996).

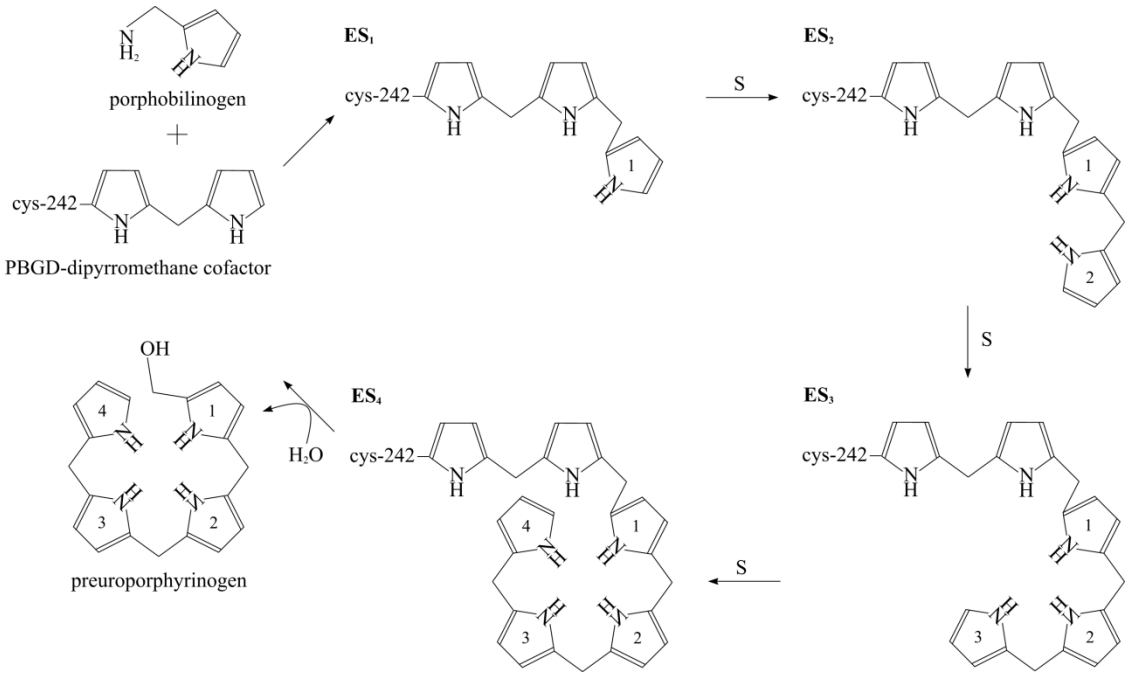

Figure 1.14 The catalytic cycle of porphobilinogen deaminase (PBGD)

The elongation of the enzyme-bound dipyrromethane cofactor by the addition of four successive porphobilinogen molecules. The enzyme cycles between four substrate-bound states, ES_1, ES_2, ES_3 and ES_4, until a "hexapyrrole" is formed. The "hexapyrrole" is then cleaved by hydrolysis at the junction between pyrroles 2 and 3 to yield hydroxymethylbilane and regenerate dipyrromethane (Shoolingin-Jordan *et al.*, 2003b).

Figure 1.15 Crystal structure of *E. coli* porphobilinogen deaminase

The crystal structure of *E. coli* PBGD solved to a resolution of 1.76 Å (PBD 1PDA, Louie *et al.*, 1992). Highlighted in orange is the highly conserved C242, through which the dipyrromethane cofactor is ligated.

Uroporphyrinogen III synthase (UROS)

UROS is encoded by the *hemD* gene (Rimington and MILES, 1951; Hansson *et al*., 1991) and is responsible for producing uroporphyrinogen III, the first cyclic tetrapyrrole in the pathway. Early studies on the enzyme purified from rat liver suggested a folate derived cofactor was required for activity (Kohashi *et al*., 1984). However, this was disputed when later studies on UROS from *Eulena gracilis* showed this was not the case (Hart and Battersby, 1985). Rather, these studies showed that UROS is a monomeric protein with a molecular mass ranging between 31 and 39 kDa.

The crystal structure for human UROS (**Figure 1.16**) was solved to a resolution of 1.85 Å. This protein was revealed to have two domains arranged in an unusual dumbbell shape, where the two domains are connected by a two-stranded antiparallel β-sheet. The two domains have similar αβ topology, in which a parallel β-sheet is surrounded by α-helices. Domain 1 is structurally related to the flavodoxin-like fold family, whilst domain 2 has a DNA glycosylase-like fold. The large open cleft between the two domains is home to the active site and is lined with ten conserved residues. Of these only three were found to diminish activity when mutated: Arg65Ala, Thr103Ala and Tyr168Phe (Mathews *et al*., 2001).

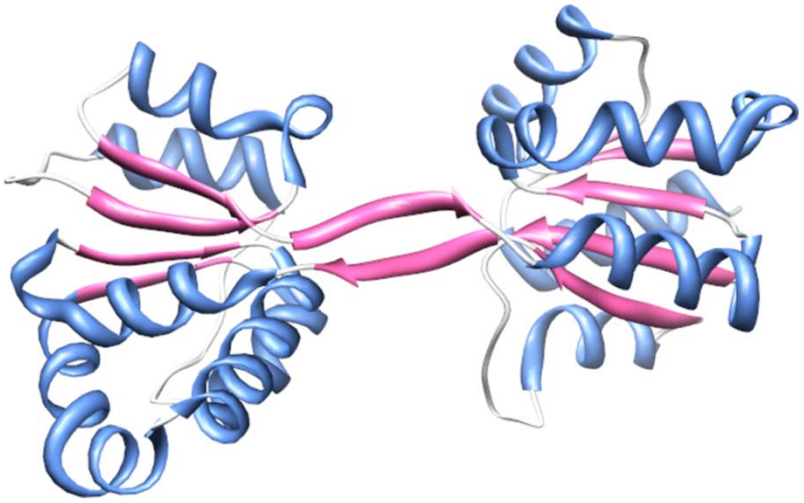

Figure 1.16 Crystal structure of human uroporphyrinogen III synthase

The crystal structure of the UROS monomer, from humans, solved to a resolution of 1.84 Å (PDB 1JR2, Matthews *et al*., 2001).

The formation of uroporphyrinogen III represents a branch point in the tetrapyrrole biosynthetic pathway, where uroporphyrinogen III can be directed down one of two routes; C-methylation by uroporphyrinogen III methylase produces sirohydrochlorin and initiates the synthesis of sirohaem, coenzyme F_{430} or vitamin B_{12}, while decarboxylation by

uroporphyrinogen III decarboxylase (UROD) leads to coporphyrinogen III formation and the synthesis of haems and chlorophylls.

1.4.4 Uroporphyrinogen III to coproporphyrinogen III

UROD, encoded by the *haem* gene, catalyses the sequential decarboxylation of the four acetate side chains of uroporphyrinogen III to give rise to coporphyrinogen III. The reaction yields four methyl groups at the corresponding positions in the pyrrole rings. Under physiological substrate conditions mechanistic studies show the decarboxylation proceeds in an ordered clockwise manner beginning with the acetate side chain of D ring, followed by decarboxylation of the acetate side chains on rings A, B and C (Phillips *et al*., 2009). When uroporphyrinogen III is present in excess, decarboxylation was observed to occur in a random fashion (Luo and Lim, 1993). UROD can also decarboxylate uroporphyrinogen I, but the product, coproporphyrinogen I, cannot be further metabolised. Unusually, UROD was found not to require any prosthetic group or cofactor to function (Layer *et al*., 2010).

Sequence analysis of UROD genes reveals that the primary sequence is highly conserved and has an apparent monomer molecular mass of ~40 kDa (Romeo *et al*., 1986; Whitby *et al*., 1998). The crystal structures have been solved for UROD from humans, 1.6 Å and 1.7 Å, (Whitby *et al*., 1998; Phillips *et al*., 2003), *Nicotiana tabacum*, 2.3 Å, (Martins *et al*., 2001) and *Bacillus subtilis*, 2.3 Å (Fan *et al*., 2007).

Human UROD crystallises as a homodimer, where each monomer is comprised of only a single domain. This domain forms a distorted αβ-barrel, in which a central eight-stranded parallel β-barrel is encased in a swirl of eight α-helices. An additional nine α-helices and a β-hairpin are located on the loops between the alternating α-helices and β-strands that form the central barrel (**Figure 1.17 A**) (Whitby *et al*., 1998).

The active site of UROD is located in a deep cleft formed by loops L1-L4 and L8. It is thought uroporphyrinogen III inserts into this cleft, which shields the porphyrin from the solvent. Many of the 37 conserved residues are located around the active site cleft and the majority of these residues have a role in forming a stabilising hydrophobic core (Whitby *et al*., 1998). This structure hints towards two invariant residues, Asp86 and Tyr164, which point into the cleft, having a role in catalytic activity. This was verified by the subsequent structure of human UROD, which revealed Asp86 to form a hydrogen-bond with the substrate. When this residue was substituted with an Asn, near-complete inhibition of wild-type activity was observed (Phillips *et al*., 2003). Tyr164 was not found to be critical for catalysis, but required

for wild type levels of function, as the mutant Tyr164Phe retained 25 – 30 % of wild type activity (Phillips *et al*., 2003).

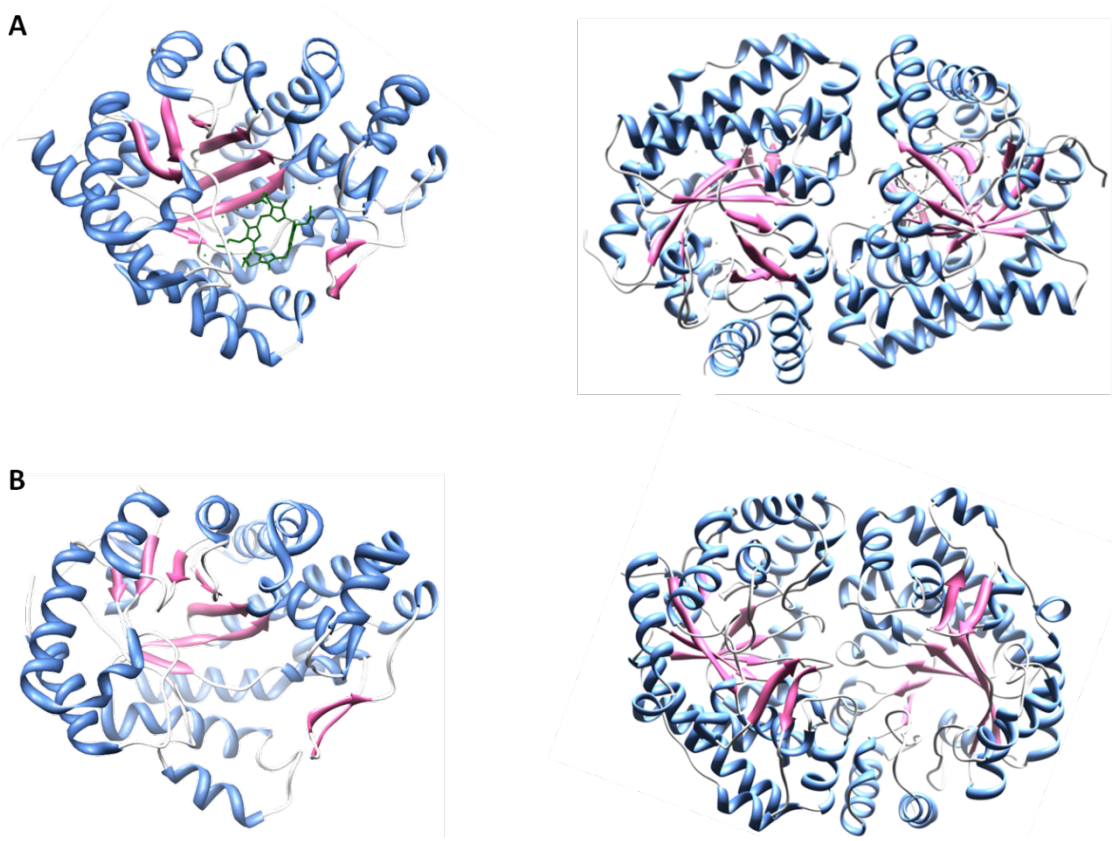

Figure 1.17 Crystal structures of human and tobacco uroporphyrinogen III decarboxylase

Crystal structure of human UROD (**A**), monomer, with co-ordinated coporphyrinogen III, and dimer, solved to a resolution of 1.75 Å (PDB 1R3Y, Phillips *et al*., 2003). Crystal structure of tobacco UROD (**B**) solved to a resolution of 2.30 Å (PDB 1J93, Martins *et al*., 2001).

In humans, a deficiency in UROD is responsible for one of the most common forms of porphyria, cutanea tarda, which is characterised by the blistering of skin in areas exposed to high levels of sunlight. This phenomenon is a result of uroporphyrinogen III accumulation, which when exposed to light generates reactive-oxygen-species (ROS) (Kushner *et al*., 1976). I plants, accumulation of uroporphyrinogen III in plants, leads to the formation of necrotic lesions in leaves (Mock and Grimm, 1997; Hu *et al*., 1998; Mock *et al*., 1998), thought to result from the release of uroporphyrinogen III into the cytoplasm, where it becomes oxidised (Shalygo *et al*., 1998).

1.4.5 Coproporphyrinogen III to protoporphyrinogen IX

Coporphyrinogen III is transformed to protoporphyrinogen IX by the enzyme coproporphyrinogen III oxidase (CPOIII). The conversion requires the oxidative

decarboxylation of the propionate side chains on rings A and B to yield two vinyl groups, forming the divinyl compound protoporphyrinogen IX.

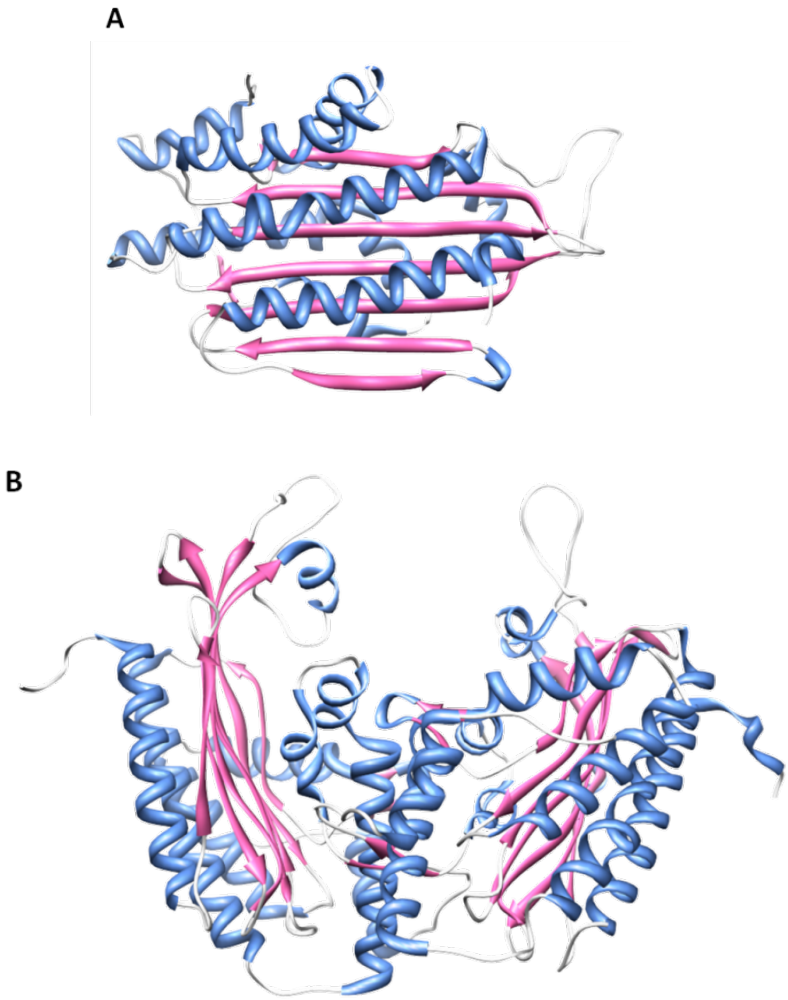

Figure 1.18 Crystal structure of oxygen-dependent coproporphyrinogen III oxidase

The crystal structure of *Saccharomyces cerevisiae* oxygen-dependent CPOIII solved to a resolution of 2.40 Å, monomer (**A**), homodimer (**B**) (PBD 1TLB, Phillips *et al.*, 2004).

There are two distinct types of CPOIII found in nature, one for the oxygen-dependent reaction (*HemF*) and the other for the oxygen-independent reaction (*HemN*). The oxygen-dependent enzyme is predominately found in eukaryotes and is dependent on molecular oxygen as the terminal electron acceptor. The oxygen-independent enzyme, found in bacteria, belongs to the radical SAM (S-adenosyl methionine) group of proteins and does not use oxygen as the terminal electron acceptor. The two forms of CPOIII share no sequence similarity and are thought to have evolved independently under different oxygen environments (Heinemann *et al*., 2008). *Synechocystis*, along with many other bacteria, has been shown to have both types of CPOIII; it is thought the two enzymes operate at different

environmental oxygen levels enabling a stable supply of tetrapyrroles in environments with fluctuating oxygen (Goto *et al*., 2010).

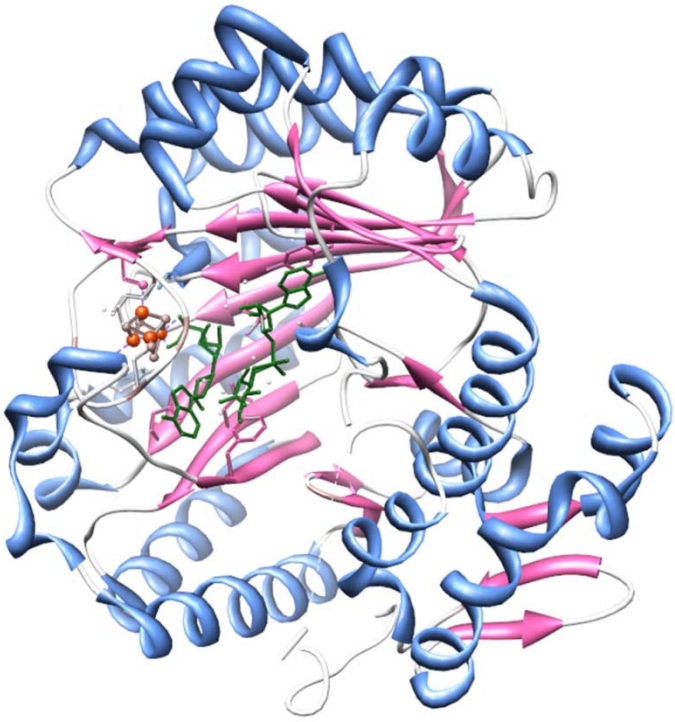

Figure 1.19 Crystal structure of oxygen-independent coporphyrinogen III oxidase

The crystal structure of *E. coli* CPOIII solved to a resolution of 2.07 Å. Highlighted in green are the two SAM cofactors, in orange are the four Fe ions located in an Fe-S cluster with four S atoms (light brown). The FeS cluster is ligated via three conserved cysteine residues (PDB 1OLT, Layer *et al*., 2003).

In mammals, the oxygen-dependent HemF CPOIII is mostly associated with the inner side of the outer mitochondrial membrane (Grandchamp *et al*., 1978) and is targeted to the mitochondria by a targeting sequence 110 amino acids long (Delfau-Larue *et al*., 1994). The structures of HemF from yeast (Phillips *et al*., 2004) and humans (Lee *et al*., 2005) have been solved and reveal that HemF is a single domain homodimeric protein (**Figure 1.18 B**). Interestingly, HemF adopts an unusual secondary structure consisting of a relatively flat seven-stranded anti-parallel β-sheet covered on both sides by α-helices (**Figure 1.18 A**) (Phillips *et al*., 2004).

In plants, CPOIII is found exclusively in the plastid, as demonstrated by immune-detection and analyses of CPOIII activity in peas and soya beans (Smith *et al*., 1993). Reduction of CPOIII by anti-sense RNA expression in transgenic tobacco leaves, leads to coporphyrinogen III accumulation and light intensity-dependent leaf necrosis (Kruse *et al*., 1995). Although the anti-photosensitisation processes were activated following knockdown of CPOIII activity, these were not sufficient to prevent reactive oxygen species-induced cell death (Kruse *et al*.,

1995). Reactive oxygen species induced cell death is also observed in the *Arabidopsis lin2* (lesion initiation) mutant, which encodes a defective CPOIII gene (Ishikawa *et al*., 2001).

The oxygen-independent CPOIII is an iron sulphur protein and contains an unusual oxygen-labile 4Fe-4S cluster (**Figure 1.19**). This cluster is co-ordinated though three conserved cysteine residues located in the characteristic CxxxCxxC motif (Sofia *et al*., 2001). One of the four FeS clusters is ligated to the s-adenosyl-L-methionine (SAM) cofactor that initiates radical based catalysis (Layer *et al*., 2003). **Figure 1.19** shows the crystal structure of HemN, which was the first member of the radical-SAM family to have its structure determined by X-ray crystallography. The radical SAM group of proteins all share the property of using an iron-sulphur cluster to reductively cleave SAM to produce a radical, which can go on to perform a wide variety of chemical transformations. The structure shows HemN is a monomeric protein, consisting of two distinct domains; a large N-terminal domain and a smaller C-terminal domain. The N-terminal domain consists of an unusual twelve-stranded curved β-sheet, surrounded on the outer surface by α-helices; it is home to the iron-sulphur cluster and the binding sites for two SAM molecules (Layer *et al*., 2003).

1.4.6 Protoporphyrinogen IX to protoporphyrin IX

The last enzyme common to both the haem and chlorophyll biosynthesis pathways is protoporphyrinogen IX oxidase (PPO), which catalyses the flavin-dependent conversion of protoporphyrinogen IX to protoporphyrin IX in a reaction that oxidises the macrocycle by the withdrawal of six electrons. The oxygen-dependent reaction consumes three molecules of oxygen and generates three molecules of peroxide.

In plants, PPO is encoded by two different genes, PPOI and PPOII. These two isoforms share few similarities in the same species, but the sequence identity between the same isoform from different species can be as high as 70% (Lermontova *et al*., 1997; Che *et al*., 2000; Watanabe *et al*., 2001). PPO activity is detected in both the mitochondria and chloroplast, making it the first enzyme in tetrapyrrole biosynthesis to be found in two subcellular locations (Jacobs and Jacobs, 1987; Smith *et al*., 1993). As a target for the phthalimide-type and diphenylether-type herbicides, PPO is of considerable commercial interest (Matringe *et al*., 1989a; Camadro *et al*., 1991). Herbicide inhibition of PPO leads to the accumulation of protoporphyrin IX in the plastids, which, upon leakage into the cytosol, is oxidised to protoporphyrin IX by unspecific oxidases attached to the cells plasma membrane. These excessive levels of protoporphyrin IX go on to generate lethal amounts of reactive oxygen species, causing membrane peroxidation and, ultimately, cell death (Matringe and Scalla, 1988; Matringe *et al*., 1989b; Jacobs and Jacobs, 1993).

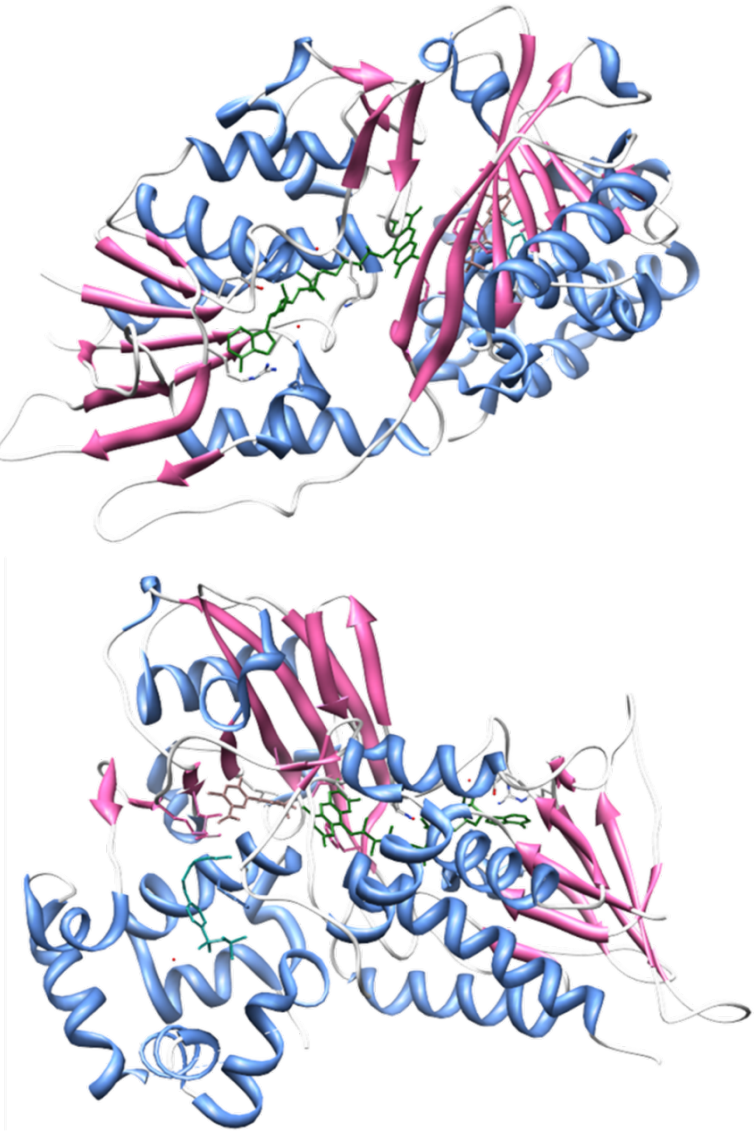

Figure 1.20 Crystal structure of protoporphyrinogen IX oxidase

The crystal structure of PPO from tobacco solved to a resolution of 2.9 Å (PDB 1SEZ, Koch *et al*., 2004). The monomeric asymmetric unit is viewed from opposing angles; highlighted in green are the non-covalently bound flavin cofactor, in light brown is the phenyl-pyrazol inhibitor (a class of herbicide) and in green the Triton-X-100 molecule that inserts into the binding pocket.

The crystal structure of mitochondrial PPO from tobacco has been solved to a resolution of 2.9 Å (**Figure 1.20**). This PPO is a dimeric yellow protein with a monomer molecular weight of 55 kDa that crystallised with a non-covalently bound FAD cofactor (Koch *et al*., 2004). The protein is composed of three domains; FAD-binding, substrate-binding and membrane-binding. The FAD binding domain is composed of a nine-stranded anti-parallel β-sheet coated on one side by α-helices and a β-hairpin. This domain shows significant structural and sequence homology to other flavoenzymes and displays a *p*-hydroxybenzoate-hydroxylase-fold-like topology. Although PPO is a tightly bound membrane protein, it contains no transmembrane helices. Rather, it binds monotopically to the inner mitochondrial membrane

with two conserved membrane anchor sequences in a manner similar to that observed for squalene-hopene cyclase (Wendt *et al*., 1999; Koch *et al*., 2004). The membrane binding domain is composed solely of α-helices numbering eight in total; two of these helices are involved in dimeric assembly of PPO. The active site is located between the FAD and substrate-binding domains; several conserved residues are located at the interface of this pocket one of which, Asn67 (that corresponds to the highly conserved Arg59 in mammals), is responsible for the human inherited disorder variegate porphyria when mutated to Trp (Maneli *et al*., 2003). The substrate binding domain is composed of a seven-stranded β-sheet, with two α-helices located on the outer face and co-crystallises with a Triton-X-100 molecule inserted into the binding pocket (**Figure 1.20**) (Koch et al., 2004).

In bacteria there are two forms of PPO, an oxygen-dependent (*hemY*) protein and an oxygen-independent (*hemG*) protein. Unlike CPOIII, only a single form of PPO is found in a given cell. Recently a third PPO gene, *hemJ*, was found in *Synechocystis*, which lacks both the *hemY* and *hemG* homologues. Bioinformatic analyses of this protein indicate it may contain transmembrane helices and be flavin-independent (Kato *et al*., 2010; Boynton *et al*., 2011).

1.4.7 The branch-point of haem and chlorophyll biosynthesis

Protoporphyrin IX lies at the branch point of the chlorophyll and haem biosynthesis pathways. Here, the fate of the porphyrin is dependent on the metal ion chelated into the macrocycle. Insertion of a ferrous ion (Fe^{2+}) by ferrochelatase (FeCH) commits the porphyrin to the haem branch, while insertion of a magnesium ion (Mg^{2+}) by magnesium chelatase (MgCH) directs the porphyrin towards the chlorophyll biosynthetic pathway.

Although the two enzymes perform a relatively similar role, namely insertion of a metal ion into the protoporphyrin IX macrocycle, they are unrelated in structure and have opposing energy requirements. FeCH is a single subunit enzyme that catalyses the energetically favourable insertion of Fe^{2+} into protoporphyrin IX without the requirement of ATP, whereas MgCH is a multi-subunit protein that catalyses the energetically unfavourable insertion of Mg^{2+} into protoporphyrin IX. The removal of the hydration shell surrounding the Mg^{2+} ion is thought to be the reason for the absolute requirement of ATP (Reid and Hunter, 2004).

1.4.7.1 Ferrochelatase

As described earlier, FeCH catalyses the insertion of a ferrous ion into the centre of the macrocycle in a reaction that generates protohaem. FeCHs are encoded by *hemH* and exist in a variety of different forms with a momomeric molecular mass of approximately 36–40 kDa (Dailey *et al*., 2000). There are examples of FeCHs existing as soluble monomers, membrane-associated monomers and membrane-associated homodimers. Iron sulphur clusters [2Fe-2S]

have been found in a limited number of prokaryotic FeCHs and all animal FeCHs, but have not yet been reported in any FeCHs isolated from plants. As well as iron, FeCHs are capable of chelating a variety of divalent metal ions *in vitro*, including copper, zinc, cobalt, nickel, mercury, cadmium, lead and manganese. However, not all of these porphyrin-ion compounds can leave the active site and so they effectively block FeCH activity (Dailey *et al*., 2000; Medlock *et al*., 2009). FeCH is also able to utilise protoporphyrin IX molecules with a variety of side groups on rings A and B (Dailey *et al*., 2000).

The location of FeCH within the cell is organism dependent; in mammalian and yeast cells the enzyme is associated with the mitochondrial membrane (Dailey and Karr, 1987; Karr and Dailey, 1988a), in plant cells, which have two FeCHs (type-I and type-II), the enzyme is localised in the mitochondria (type I) and chloroplast (type I and type II) (Chow *et al*., 1997; 1998; Masuda *et al*., 2000; Suzuki *et al*., 2000; 2002; 2003) and in prokaryotes FeCH is found either in the cytoplasm or associated with the plasma membrane (Hansson and Hederstedt, 1994).

Eukaryotic FeCHs are nuclear encoded and synthesised in the cytoplasm as a pre-protein with a mitochondria or plastid targeting sequence (Karr and Dailey, 1988b). Initial breakthroughs into the study of eukaryotic FeCHs arrived with the cloning of *Saccharomyces cerevisiae* FeCH (Labbe-Bois, 1990), which was closely followed by recombinant expression of human (Taketani *et al*., 1992) and murine FeCHs (Taketani *et al*., 1990; Brenner and Frasier, 1991). In non-plant eukaryotes a single FeCH is produced, which, following its translocation to the inner membrane of the mitochondria is proteolytically cleaved into its mature form. Most eukaryotic FeCHs are membrane associated, although they do not contain a transmembrane domain.

FeCHs from algae, cyanobacteria and plant type II FeCHs all possess a distinctive hydrophobic C-terminal extension, with a conserved chlorophyll *a/b* (CAB) binding domain, connected to the FeCH catalytic core by a proline-rich linker (Suzuki *et al*., 2002; Sobotka *et al*., 2011). The cyanobacterium *Synechocystis* contains such a FeCH; initially it was thought the C-terminal extension, which forms a putative transmembrane segment, was required for membrane localisation. However, in a strain of *Synechocystis* (ΔH324) that contains a truncated enzyme lacking the transmembrane segment, FeCH was still found to associate with the membrane, but the enzyme was momomeric. Therefore it appears the C-terminal CAB-domain extension is required for dimer formation. Additional analysis of truncated FeCH mutants revealed the proline rich linker is essential for catalytic activity (Sobotka *et al*., 2008a; 2011).

The crystal structure of human FeCH has been solved to a resolution of 2.0 Å (**Figure 1.21**). This protein exists as a homodimer with a molecular mass of ~86 kDa; the mature protein is composed of residues 65-423, as residues 1-64 comprise the mitochondrial targeting tag, which is proteolytically removed upon FeCH arrival in the organelle. Each monomer contains an N-terminal and a C-terminal domain; the N-terminal domain (residues 65-247) forms part of the active site pocket (residues 80-130) and the C-terminal domain (residues 248-423) co-ordinates the [2Fe-2S] cluster (residues 390-423) and helps stabilise the homodimer (Wu *et al*., 2001). The two domains are folded into similar αβ domains, where a four-stranded parallel β-sheet is flanked by α-helices in a βαβ motif, and are connected by a hinge region formed by two α-helices in a similar fold to those found in other periplasmic binding proteins (Wu *et al*., 2001). Although the human FeCH is a homodimer, the *Bacillus* FeCH is a monomer. On closer examination of the crystal structures, it is apparent the human FeCH dimer is stabilised via interactions between the neighbouring C-terminal extensions, which are absent in the *Bacillus* FeCH.

The active site pocket of human FeCH is comprised of two hydrophobic lips and the lower lip is lined with several highly conserved residues (His263, Glu343, His341, Phe337, Tyr123, Arg164 and Tyr165) arranged in hydrophobic ridges. Three cholate detergent molecules were found inserted into the active site of the crystal structure, interacting with the aforementioned hydrophobic residues. The N-terminal region around the active site pocket has also been proposed as the region required for membrane association, as yeast containing mutant FeCH lacking this region was found to accumulate in the mitochondrial matrix (Gora *et al*., 1996; 1999). Given the hydrophobic nature and poor solubility of the porphyrin substrate, it is not surprising that the membrane binding domain and entrance to the active site pocket are in close proximity, allowing the active site to be orientated towards the membrane. This positioning allows for the substrate and haem product to enter and leave the active site via the membrane, thus eliminating any problems arising from aggregation of these poorly soluble molecules in the mitochondrial matrix. The active site also contains a Co^{+2} ion coordinated by residues His231 and Asp383, which lies on the opposite side of the active site to the hydrophobic entrance.

The observation that the porphyrin macrocycle of N-methylmesoporphyrin is distorted when bound to the active site in *Bacillus* FeCH, led to the prediction that this distortion enables the FeCH mediated insertion of Fe^{2+} into the macrocycle (Al-Karadaghi *et al*., 1997). However, this hypothesis was contradicted when the structure of human FeCH bound to protoporphyrin IX showed only a slight distortion in the porphyrin, which was found rotated by 100° and positioned 4.5 Å deeper into the active site pocket, relative to N-methylmesoporphyrin in

Bacillus FeCH (Al-Karadaghi *et al*., 1997; Wu *et al*., 2001; Medlock *et al*., 2007). It was later showed that the FeCH reaction proceeded in at least two steps; a rapid substrate-binding step, followed by a slower, irreversible, metal chelation step (Hoggins *et al*., 2007). In the rapid substrate-binding step, the metal ion binding is fast, whereby the metal ion initially binds at a remote site and moves to the active site, which is concurrently accompanied by the binding of the porphyrin molecule. In the slower irreversible step, the Fe^{2+} ion is inserted into the porphyrin maccrocycle and the metalloporphyrin is released from the active site (Hoggins et al., 2007).

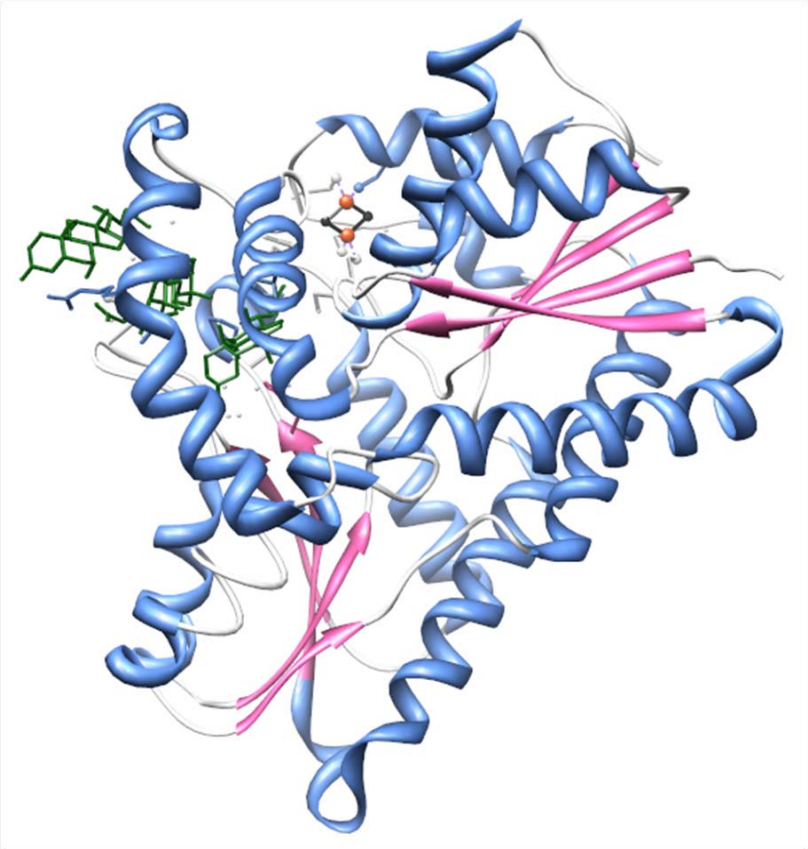

Figure 1.21 Crystal structure of ferrochelatase

The crystal structure of human FeCH monomer solved to a resolution of 2.00 Å (PDB 1HRK, Wu *et al*., 2001). Highlighted in green are the three cholate molecules that crystallise in the active site, in black (S) and orange (Fe) is the 2Fe-2S cluster, ligated to four conserved cysteine residues.

The [2Fe-2S] cluster is located in an extension of the C-terminal domain not found in the *Bacillus subtilis* FeCH, which does not contain an iron-sulphur cluster (Al-Karadaghi *et al*., 1997). Four highly conserved cysteine residues (Cys196, Cys403, Cys406 and Cys411) are involved in ligating the [2Fe-2S] cluster and are located in the binding motif $Cx_{206}Cx_2Cx_4C$, where x represents any amino acid residue. An unusual feature of this [2Fe-2S] cluster is the absence of additional hydrogen bonding to the iron or sulphur atoms and the limited

hydrogen bonding between the C-terminal extension and the core of the molecule (Wu *et al*., 2001). However, this region appears to be essential for FeCH activity as a chimeric yeast FeCH, where this 40 residue C-terminal extension had been exchanged for the corresponding region in mouse FeCH, exhibited no enzyme activity (Gora *et al*., 1999).

1.4.7.2 Magnesium chelatase

The second enzyme at the haem-chlorophyll branch point is MgCH, which is responsible for the chelation of a divalent magnesium ion into the centre of the porphyrin macrocycle: the first committed step of chlorophyll biosynthesis. Unlike FeCH, MgCH is a multi-subunit protein. The three genes required for MgCH composition are encoded within the loci *bchI, bchD,* and *bchH* or *chlI, chlD* and *chlH*, depending on whether the host organism ultimately synthesises bacteriochlorophyll (*bch*) or chlorophyll (*chl*).

MgCH activity was first observed *in vivo* using *Rhodobacter sphaeroides* extracts grown anaerobically in the light, or under low oxygen conditions in the dark. These experiments were carried out in the presence of the metal chelators EDTA or its related chelators EGTA, N-(2-hydroxyethyl)-ethylenediamine-NN'N'-triacetate and trans-1,2-diaminocyclohexanetetra-acetate, to prevent the spontaneous insertion of divalent zinc into the macrocycle (Gorchein, 1972). A few years later MgCH activity was demonstrated in intact isolated cucumber chloroplasts, where the Mg^{2+} chelation, observed in etioplasts or developing chloroplasts, was found to be abolished in plastids heat treated to 100°C (Smith and Rebeiz, 1977). Subsequently, MgCH was demonstrated to have an absolute requirement for ATP and MgCH activity was blocked by AMP (Pardo *et al*., 1980). In MgCH assays using a cell free system based on pea extracts, a six minute lag time was observed between ATP addition and MgCH activity, indicating a requirement for ATP to activate the enzyme prior to chelation (Walker and Weinstein, 1991a). As FeCH is relatively promiscuous in the metal ions it can insert into protoporphyrin IX, the specificity of MgCH for inserting a divalent magnesium ion was also investigated; it was found no other metal ions could serve as a substrate or inhibit MgCH activity (Walker and Weinstein, 1991b).

The three *bch* genes (*bchIDH*) required for MgCH activity were first identified when disruptions in any of these three loci in *Rhodobacter sphaeroides* or *Rhodobacter capsulatus*, resulted in abolition of Mg-protoporphyrin IX production and accumulation of the MgCH substrate, protoporphyrin IX (Coomber *et al*., 1990; Bollivar *et al*., 1994c; Suzuki *et al*., 1997; Naylor *et al*., 1999). The corresponding *chl* genes (*chlDH*) were later identified in *Synechocystis* by the same method (Jensen *et al*., 1996a). Higher plant MgCH genes were similarly localised in barley (*Hordeum vulgare*), when isolated etioplasts with lesions in either of the three loci, *Xantha-f, Xantha-g* or *Xantha-h*, were found to accumulate protoporphyrin

IX when fed with the chlorophyll precursor ALA (Jensen *et al*., 1996c; Hansson *et al*., 1999). These studies provided the first evidence that three different protein subunits were required for MgCH activity and that this enzyme is highly conserved between chlorophyll and bacteriochlorophyll producing organisms.

MgCH activity was first reconstructed *in vitro* when individual extracts from *E. coli* containing over-expressed BchH, BchI or BchD from *Rhodobacter sphaeroides* were found to produce Mg-protoporphyrin IX when combined (Gibson *et al*., 1995). The subsequent expression of the *Synechocystis* ChlHID homologues *in vitro* also demonstrated MgCH activity (Jensen *et al*., 1996a). Later experiments showed stoichiometries of $BchH_{36}$:$BchI_4$:$BchD_1$ and $ChlH_4$:$ChlI_2$:$ChlD_1$ were optimum for *in vitro* steady state assays (Gibson *et al*., 1999; Jensen *et al*., 1999). MgCH activity has also been studied on recombinant proteins from tobacco (Papenbrock *et al*., 1997), maize (Sawers *et al*., 2006), rice (Zhou *et al*., 2012) and the thermophilic cyanobacterium *Thermosynechococcus elongatus* (Qian et al., 2012a).

The MgCH subunits, ChlH/BchH, ChlI/BchI and ChlD/BchD, have predicted molecular masses of 120-155 kDa, 37-47 kDa and 60-87 kDa, respectively. Homology analyses, using bioinformatics, of each Chl/Bch pair from various species, show varying degrees of pairwise identity. For example, there is 38% sequence identity between BchH from *Rhodobacter sphaeroides* and *Xantha-f* from *Barley* and 86% identity between the ChlHs from *Arabidopsis* and *Antirrhinum*. Similarly 50-90% identity exists between the I subunits and 28-58% between the D subunits from various species. Interestingly, the N-terminus of the I and D subunits share 40% sequence identity, suggesting that the two proteins evolved from a common ancestor protein following a gene duplication event (Jensen *et al*., 1996b). Similar homology analyses showed the H subunit shares significant sequence similarities with CobN, the largest subunit of aerobic cobaltochelatase, which is a large heterotrimeric protein involved in vitamin B_{12} biosynthesis (Walker and Willows, 1997).Although the catalytic mechanism of MgCH remains unknown, much has been deduced about the function and role of each subunit.

The MgCH I and D subunits have been shown to form a complex, whose formation is dependent on the presence of ATP and Mg^{2+} (Gibson *et al*., 1999). Several studies indicate that I forms homo-oligomeric structures upon addition of ATP and Mg^{2+}, however some controversy remains as to their exact stochiometry. Jensen *et al* (1998) used gel filtration studies to show dimeric *Synechocystis* ChlI formed high molecular aggregates corresponding to hexamers or octamers and that pre-incubation of I with a non-hydrolysable form of ATP removed the lag-phase for the formation of the *Synechocystis* ID complex. However, single-

particle analyses of ChlI from the same species, showed a ring-like structure with seven-fold rotational symmetry, suggesting ChlI may exist as a heptamer (Reid *et al*., 2003). Single particle images showed BchI to form hexameric rings, onto which the authors modelled the BchI monomer (Fodje *et al*., 2001), whose crystal structure (*Rhodobacter capsulatus* BchI) has been solved to a resolution of 2.1 Å. The formation of hexameric rings by BchI was later confirmed with additional analyses of single particle EM images (Willows *et al*., 2004). These data suggest that the higher organisation of the I subunit is variable and dependent on whether the host organism ultimately produces chlorophyll or bacteriochlorophyll, although the exact stochiometry of ChlI is not fully elucidated.

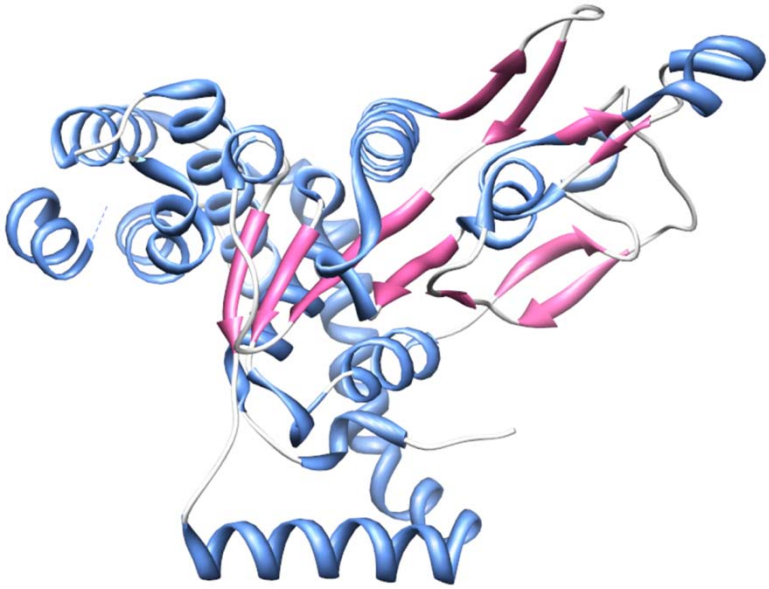

Figure 1.22 Crystal structure of BchI from *Rhodobacter capsulatus*

The crystal structure of BchI from Rhodobacter capsulatus solved to a resolution of 2.1 Å (PDB 1G8P, Fodje *et al*., 2001).

Functional and sequence analyses have identified the I subunit as having ATPase activity and a AAA^+ (ATPases Associated with various cellular Activities) domain (Jensen *et al*., 1999; Neuwald *et al*., 1999). The amino acids involved in nucleotide binding in AAA^+ proteins are highly conserved and sequence analysis of I reveals a Walker A (GxxxxGK), a Walker B (D/Exx) and a SRH (Second Region of Homology) motif. These conserved motifs are also found in the D subunit, along with a C-terminal integrin I domain (Fodje *et al*., 2001). The Walker A motif is required for binding the adenine and ribose rings of the nucleotide, whilst the residues in the Walker B motif contribute towards the binding of triphosphate groups. The SRH motif, which is also found in the H subunit, is thought to be involved in detecting nucleotide binding and hydrolysis (Guenther *et al*., 1997). The integrin I domain found in D binds to certain amino acid sequences such as LDV, RGE and RGD. The LDV motif has been identified in ChlH and

BchH, RGE is found in BchI and there is a RGD in ChlI, indicating many possibilities for H-D, I-D interactions.

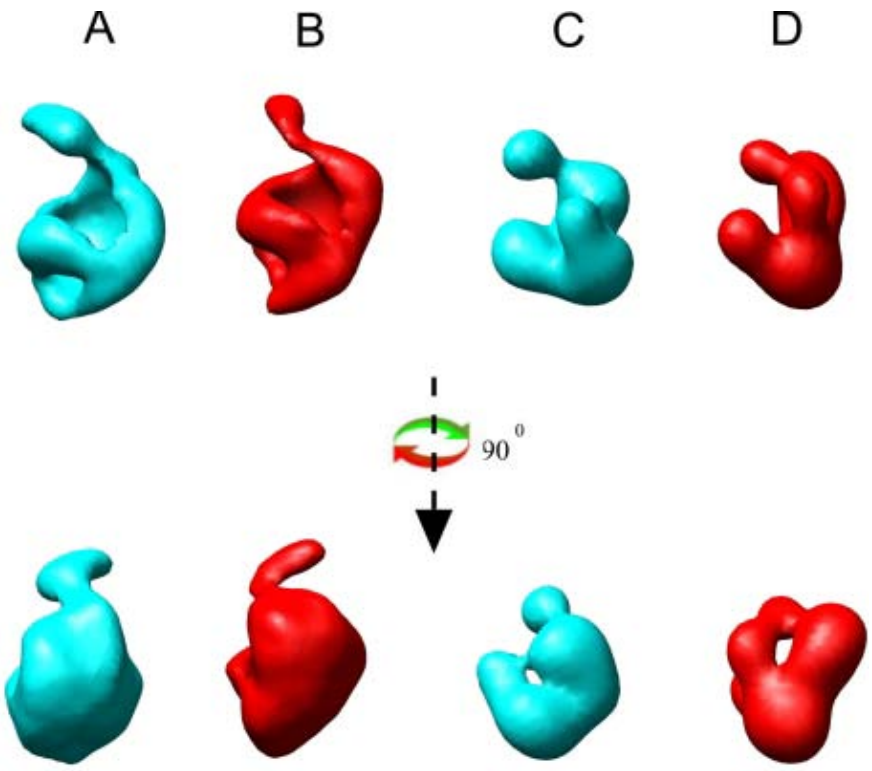

Figure 1.23 Single particle reconstruction of ChlH and BchH

3D models of single particle reconstituted ChlH (A&B) and BchL (C&D). Models were calculated for apo-ChlH/BchL (cyan) and porphyrin complexed ChlH/BchH (red),(Qian et al., 2012b).

The crystal structure of BchI (**Figure 1.22**) has three main domains, an N-terminal domain, AAA^+ domain and a C-terminal domain (Fodje *et al*., 2001). The Walker A and Walker B nucleotide binding motifs reside in the N-terminal domain, which comprises of a twisted five-stranded parallel β-sheet, flanked by α-helices to form a classic αβα Rossmann fold. The AAA^+ domain is located on top of the N-terminal domain and is composed of five loosely associated β-strands and two short α-helical segments. The main feature of this domain is the large sections of loop regions that may be involved in interactions with other chelatase subunits (Fodje *et al*., 2001). The C-terminal domain is composed of four α-helices, connected in an up-and-down manner characteristic of four helix bundles (Fodje *et al*., 2001).

The first clue as the role of the unusually large H subunit was the observation that *E. coli* cells over-expressing BchH from *Rhodobacter sphaeroides* were tinted red. Spectra of these cells revealed the pigment to be protoporphyrin IX and it was found incubation of ChlH with protoporphyrin IX significantly reduced the lag time for Mg-protoporphyrin IX formation (Gibson *et al*., 1995; Jensen *et al*., 1998). Subsequent deuteroprotoporphyrin IX binding

studies using fluorescence quenching and co-elution on a gel filtration column confirmed this observation (Karger *et al*., 2001), with protoporphyrin IX-BchH non-covalently binding in a 1:1 ratio (Willows *et al*., 1996); it should be noted the binding of protoporphyrin IX to ChlH/BchH is observed by a red shift in the porphyrins fluorescence excitation and emission spectra (Karger *et al*., 2001).

Given the size of the H subunit (~120-155 kDa) and the importance of regulating the flux of substrate down the chlorophyll biosynthesis pathway, it was unsurprising that H was found to be associated with other functions within the tetrapyrrole biosynthesis pathway. Investigations into receptors for the plant hormone abscisic acid (ABA) identified the C-terminal half of ChlH from *Arabidopsis* as an ABA binding protein (Shen et al., 2006a; Wu et al., 2009a). These authors demonstrated that ABA-mediated signalling was a positive regulator for seed germination, post-germination growth and stomatal movement. They also demonstrated ChlH was expressed in all plant tissues, not just green tissues, and suggested ChlH could be an ABA receptor, mediating ABA signalling at the whole plant level (Shen et al., 2006a; Wu et al., 2009a). However, a recent study on the barley homologue Xantha-F showed that ABA had no effect on MgCH activity and could not bind to the Barley ChlH/Xantha-F subunit (Muller and Hansson, 2009). It has also been postulated that ChlH has a role in sugar catabolism, via regulation of the anti-sigma factor, SigE (Osanai *et al*., 2005; Osanai *et al*., 2009). The H subunit has also been shown to increase the activity of the next enzyme in the pathway, Mg-protoporphyrin-IX-methyltransferase (methyltransferase) (Alawady *et al*., 2005; Shepherd *et al*., 2005a). It is not known whether Mg-protoporphyrin IX, the product from the MgCH reaction, dissociates from the H subunit. Therefore, it could be that the H subunit facilitates substrate channelling by acting as a carrier that transfers Mg-protoporphyrin IX to the methyltrasferase.

The structure of ChlH from the thermophilic cyanobacterium *Thermosynechococcus elongatus* was recently elucidated using single particle reconstruction and small-angle X-ray scattering (Qian et al., 2012a). Three dimensional structures of apo-ChlH and porphyrin-ChlH to a resolution of 30 Å were constructed from negatively strained proteins using single particle reconstruction. These showed that ChlH forms a cage-like conformation, with a globular N-terminal domain. It is thought the porphyrin molecule enters into this "cage", where it is non-covalently bound, ready for Mg^{2+} chelation. Although structures for both the apo- and protoporphyrin IX- ChlH were solved, at this resolution no significant alterations in conformation could be observed upon porphyrin binding (Qian et al., 2012a). Previous to this study, a single-particle EM reconstruction of negatively stained BchH from *Rhodobacter capsulatus*, revealed that BchH has a three-lobed structure (Sirijovski *et al*., 2008). Here, the

structures of apo- and protoporphyrin IX BchH, reconstructed to a resolution of 25 Å, revealed an alteration in conformation brought about by porphyrin binding, suggesting that both the N- and C-terminal domains interact with the substrate.

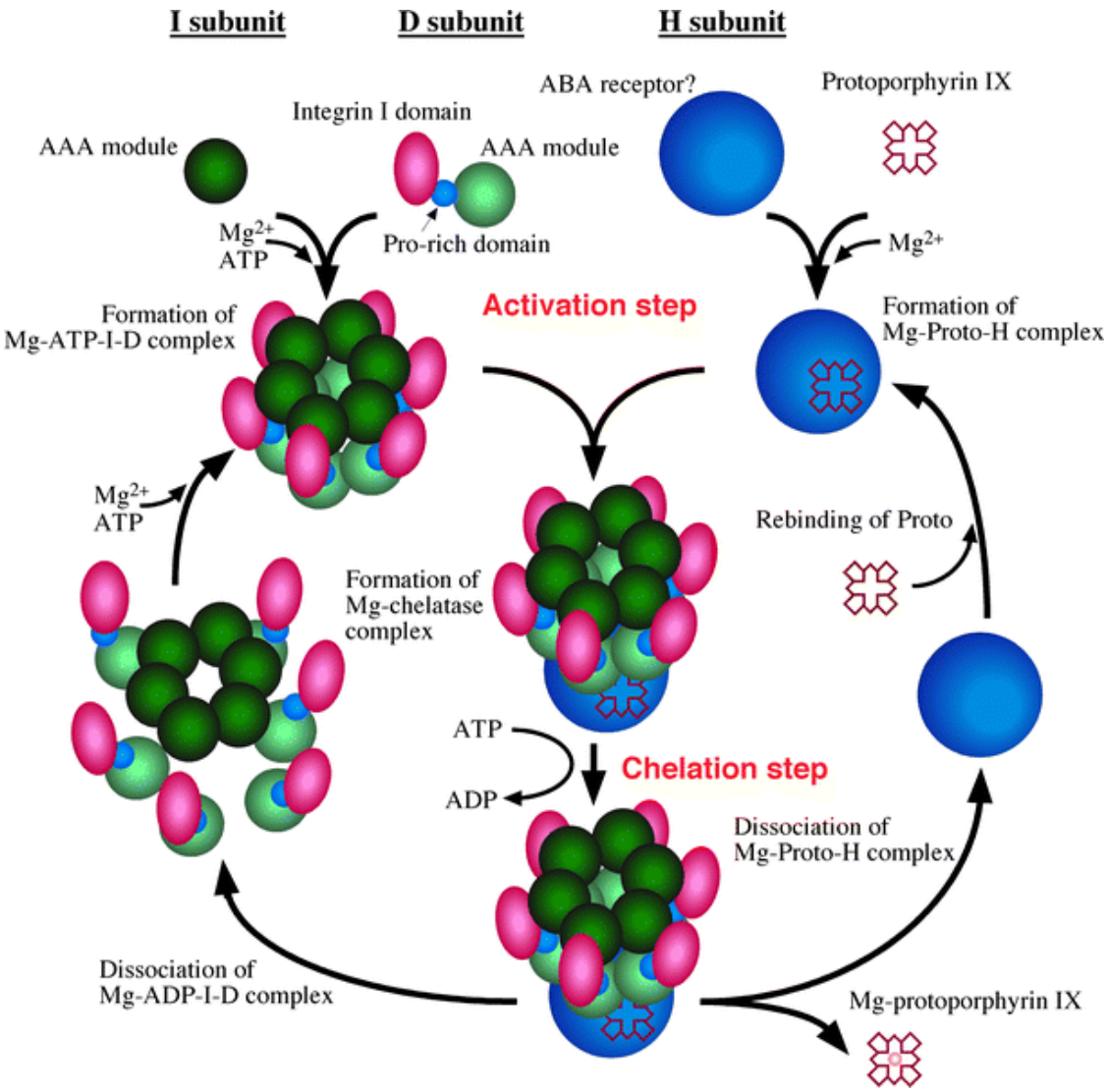

Figure 1.24 Model of the catalytic cycle of Mg-chelatase

Activation step: six I subunits are assembled into a hexameric ring in an ATP and Mg^{2+} dependent process. Six D subunits form a hexameric ring independently of ATP. Two hexameric rings come together to form a two tiered ATP-Mg-I-D complex. H subunit binds to protoporphyrin IX and Mg^{2+}. Chelation step: H subunit binds to ATP-Mg-I-D complex triggering ATP hydrolysis. After the formation of Mg-protoporphyrin IX the complex dissociates (Masuda, 2008).

Although the exact mechanism and stochiometry of the chelatase has not yet been fully elucidated, a mechanistic model of the MgCH reaction has been proposed based on the information known about substrate binding and subunit interactions (Warren *et al.*, 2009). The H subunit, bound to protoporphyrin IX, is thought to interact with a Mg^{2+} ATP charged I-D complex, to form a short-lived Mg^{2+}-ATP-protoporphyrin IX-H-I-D super-complex. Chelation of

Mg^{2+} into protoporphyrin IX is thought to be driven by ATP hydrolysis, resulting in the formation of Mg^{2+}ADP and Mg-protoporphyrin IX. The super-complex is thought to dissociate to Mg^{2+}ADP-I-D and Mg-protoporphyrin IX-H, whereupon fresh Mg^{2+}ATP can enter the I-D complex to recharge the enzyme for another reaction cycle (**Figure 1.24**).

The genome of *Arabidopsis* contains single genes encoding the D and H subunits, whereas the I subunit is encoded by two homologous genes, *CHLI1* and *CHLI2*. *Arabidopsis ΔCHLI1/ ΔCHL2* strains displayed an albino phenotype, and *ΔCHLI1* displayed a pale green phenotype whereas *ΔCHLI2* had no discernible effects (Apchelimov *et al*., 2007). It appears that while CHLI2 could be used to synthesise some Mg-protoporphyrin IX it could not wholly complement *CHLI1*; interestingly when *CHLI2* was put under the control of the promoter for *CHLI1*, wild type pigmentation returned, demonstrating that the only difference between the two genes is expression and not an inability to function as part of the MgCH complex (Huang and Li, 2009). The green sulphur bacterium *Chlorobaculum tepidum* has three paralogues of the *bchH* gene; *bchH, bchS* and *bchT*. Strains in which these genes have been individually inactivated all display a wild-type phenotype, as do the double mutant strains *bchH/bchT* and *bchS/bchT*; it was not possible to create a *bchS/bchH* double mutant. These data suggest that *Chlorobaculum tepidum* BchT alone cannot support the minimal bacteriochlorophyll requirements for normal growth (Johnson and Schmidt-Dannert, 2008a; Gomez Maqueo *et al*., 2009).

In higher plants, plastids differentiate into organelles that are consistent with the function of the tissue in which they are located (Kirk and Tilney-Basset, 1978; Whittaker and Danks, 1978). For example, plastids located in the mesophyll cells of the leaf develop into chloroplasts, whereas plastids located in the root cortical cells differentiate into amyloplasts (starch containing plastids). In dark-grown plants the plastids located in the mesophyll cells develop into etioplasts, identified by their characteristic array of pro-lamellar bodies, and upon exposure to light these etioplasts develop into chloroplasts (Susek *et al*., 1993). Thus it was hypothesised that environmental and intrinsic signals could combine to determine the fate of the undifferentiated plastid. Such signals would require transmissions from the nucleus to the plastid, and retrograde, feedback, signalling from the plastid to the nucleus. It was hypothesised that genes required for retrograde signalling from the plastid to the nucleus could be elucidated by identifying mutants in which the plastid and nuclear genomes had become uncoupled. One such screen identified the *Arabidopsis GUN* (genomes uncoupled) loci as components of plastid-nucleus signal transduction (Susek *et al*., 1993; Mochizuki *et al*., 2001), these mutants were no longer able to repress the transcription of nuclear genes encoding proteins required for photosynthesis or chlorophyll biosynthesis

upon blockage of chloroplast development. The gun1-*5* mutant was found to have de-repressed expression of *Lhcb* (encodes a light-harvesting chlorophyll *a/b* binding protein), a gene that is normally repressed in dark grown mutants and is expressed in response to chloroplast development, in the presence of Norflurazon (an herbicide that stops biosynthesis of photoprotective carotenoids). Four of these mutants, gun2, gun 3, gun 4 and gun 5 were found to affect certain plastid enzymes involved in the biosynthesis of tetrapyrroles. gun 2 was shown to encode haem oxygenase, gun3 encodes phytochromobilin synthase, gun 4 encodes a stimulator of MgCH and gun5 encodes the ChlH subunit of MgCH (Mochizuki et al., 2001; Larkin et al., 2003a).

A single point mutation, A990V, in the gun5/*ChlH* gene was identified as the propagator of the gun5 phenotype (Mochizuki *et al*., 2001). This point mutation is located within a region highly conserved among H genes leads to a 25% decrease in chlorophyll levels in comparison to wild type; furthermore MgCH assays with the *Arabidopsis* gun5 mutant show comparable MgCH activity when compared to wild type (Strand *et al*., 2003). These results show the signalling phenotype of gun5 mutant is independent of gun5MgCH activity. Moreover, comparison of the gun5 mutant phenotype with the phenotype of another MgCH subunit mutant (ChlI), showed ChlI mutants did not exhibit a GUN phenotype and therefore only ChlH could play a role in retrograde signalling (Mochizuki *et al*., 2001).

1.4.7.3 Gun4

The genomes uncoupled screen also identified gun4 as an important factor for chlorophyll accumulation in *Arabidopsis*. Mutants with a lesion at the gun4 locus were found to have reduced levels of the chlorophyll precursor Mg-protoporphyrin IX and chlorophyll; *Arabidopsis* plants with this lesion exhibited a pale phenotype (Vinti *et al*., 2000; Mochizuki *et al*., 2001).

The gene encoding gun4 was mapped to a 99-kb region on chromosome 3 using a series of BAC clones until one was identified that could rescue the gun4 phenotype. One open reading frame on the complementing BAC encoded a previously uncharacterised protein, highly conserved in all oxygenic photosynthetic organisms, with a putative N-terminal chloroplast signal sequence. A single missense mutation (Leu88Phe) was responsible for the observed gun4 phenotype, this mutation is thought to affect gun4stability, as gun4mutants express wild type levels of gun4 mRNA, but have significantly reduced GUN4 protein (Larkin et al., 2003a). Further analysis of GUN4, identified the protein as co-precipitating with ChlH (GUN5), a subunit of the MgCH, when purified from thylakoid membranes, and capable of binding protoporphyrin IX, the substrate of MgCH. When GUN4 was assayed for effects on MgCH activity, it was found that although not essential for activity *in vivo* or *in vitro*, *Synechocystis*

GUN4 (sll0558) could stimulate MgCH activity. Interestingly, porphyrin binding studies with *Synechocystis* GUN4 found the protein bound the product of MgCH, Mg-protoporphyrin IX, more strongly than the substrate (Larkin et al., 2003a). This demonstrates GUN4 interaction with the H subunit may play a role in chlorophyll biosynthesis by sequestering Mg-protoporphyrin IX or delivering Mg-protoporphyrin IX to the next enzyme in the pathway, ChlM.

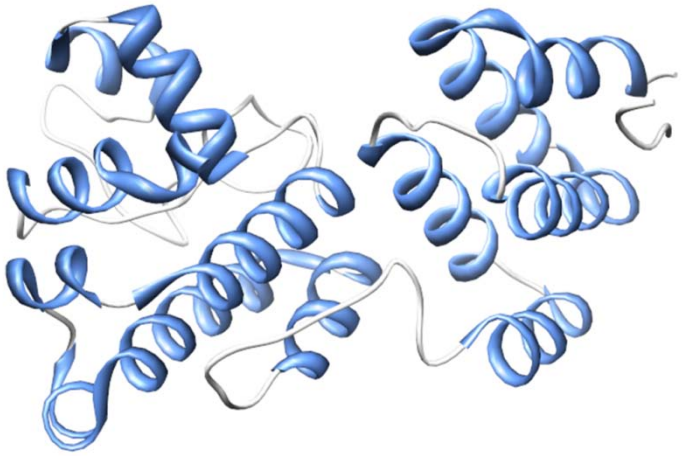

Figure 1.25 Crystal structure of GUN4

The crystal structure of GUN4 from *Thermosynechococcus elongatus* solved to a resolution of 1.5 Å, (PDB 1Z3X, Davison *et al*., 2005).

Detailed kinetic analyses of the effects of *Synechocystis* and *Thermosynechococcus elongatus* GUN4 proteins on the activity of *Synechocystis* MgCH activity were carried out (Davison et al., 2005a). These assays demonstrated that GUN4 is able to alter the threshold of Mg^{2+} required for MgCH activity. In the absence of GUN4 MgCH is essentially inactive at concentrations of Mg^{2+} below 2 mM, and addition of GUN4 essentially restores MgCH activity to the levels observed under non-Mg^{2+} limiting conditions (Davison et al., 2005a). Stimulation of MgCH by GUN4 at these concentrations of Mg^{2+} is physiologically relevant, as chloroplast concentrations of Mg^{2+} can be as low as 0.5 mM in the dark, increasing to ~6 mM in daylight (Reid and Hunter, 2004). Interestingly, GUN4 was found to restore MgCH activity to the levels observed in wild type with no GUN4 present, in *Synechocystis* ChlH mutants carrying the *GUN5* or *cch* mutations (Davison and Hunter, 2011). These mutant ChlH proteins are both able to bind the substrate and assemble into the MgCH complex with the ChlI/ChlD subunits *in vitro*, but have no MgCH activity. The partial restoration of MgCH activity by GUN4 provides an explanation for the continued chlorophyll biosynthesis observed in the *Arabidopsis GUN5* mutant (Davison and Hunter, 2011). Further evidence for the role of GUN4 in stimulating MgCH activity comes from ALA feeding experiments in *Arabidopsis*, which

demonstrate that porphyrin bound forms of GUN4 associated more stably with the thylakoid membranes, possibly promoting the formation of the MgCH by recruiting ChlH to the membrane (Adhikari *et al*., 2009; Adhikari *et al*., 2011).

Overexpression of *GUN4* in *Arabidopsis* leads to the general inactivation of the enzymes in the chlorophyll biosynthesis pathway (Peter and Grimm, 2009), a response that does not correlate with the observed mRNA levels, suggesting that GUN4 may operate post-translationally on regulating chlorophyll biosynthesis. These authors also showed that GUN4 deficiency prevented ALA synthesis and chlorophyll accumulation during photoperiodic growth. The abolition of GUN4 in *Synechocystis* leads to hampered MgCH activity and decreased levels of chlorophyll binding proteins, in particular the photosystem II antenna protein CP47 (Sobotka *et al*., 2008b). Similar investigations in *Chlamydomonas reinhardtii* found that disruption of the *GUN4* gene drastically diminished accumulation of light-harvesting complex proteins and reduced chlorophyll levels to half those observed in wild type (Formighieri *et al*., 2012). The authors also identified 803 genes with altered expression levels in Δgun4 compared to wild type. These transcripts included genes encoding proteins for the light harvesting complexes, tetrapyrrole biosynthesis, signal transduction, transcription and chromatin remodelling (Formighieri *et al*., 2012).

The structures of *Synechocystis* and *Thermosynechococcus elongatus* GUN4 have been solved to a resolution of 1.78 Å and 1.5 Å respectively (Verdecia et al., 2005; Davison et al., 2005a). Both structures reveal that GUN4 is a highly helical two domain protein, with a novel porphyrin binding fold that bears no resemblance to known porphyrin-binding proteins. The two helical domains are linked by a single loop. The N-terminal domain is structurally similar to a TPR (tetratricopeptide repeat) domain; proteins with this motif often function in protein-protein interactions, although the function of this in GUN4 is currently unknown. The C-terminal domain has a cluster of highly conserved surface exposed residues, which form a small hydrophobic cleft, or “greasy palm”. Mutational analyses and porphyrin binding studies have shown these are required for porphyrin binding (Verdecia et al., 2005; Davison et al., 2005a).

1.4.8 *S*-adenosyl-L-methionine Mg-protoporphyrin IX methyltransferase

The enzyme *S*-adenosyl-L-methionine magnesium protoporphyrin IX methyltransferase, catalyses the esterification of the carboxylic acid on the C13 side chain with the methyl group of *S*-adenosyl-L-methionine to form magnesium protoporphyrin IX monomethyl ester (MgPME). The methyltransferase is a recombinantly soluble, single subunit protein that is dependent upon SAM as a cofactor and methyl donor group. The reaction is thought to be

essential, as the methylation product of the methyltransferase prevents the spontaneous decarboxylation of the propionate group during the formation of the isocyclic ring in the next step of chlorophyll biosynthesis (Beale, 1999). This enzyme is designated ChlM in chlorophyll-producing organisms and BchM in bacteriochlorophyll-producing organisms.

Initial studies of the methyltransferase, demonstrated that this protein is membrane localised in *Rhodobacter sphaeroides* (Tait and Gibson, 1961; Gibson *et al*., 1963) and in chloroplast preparations (Radmer and Bogorad, 1967). Although not an integral membrane protein, the methyltransferase has a very hydrophobic region that could be involved in anchoring the protein to the membrane. *Arabidopsis* and spinach methyltransferases have been shown to have a dual localisation and are present in both the chloroplast envelope and thylakoid membranes (Block *et al*., 2002). The significance of the dual location is unknown, but it may be that both the substrate and product play a role in plastidic signalling (Susek *et al*., 1993; Kropat *et al*., 2000; Pontier *et al*., 2007).

The first evidence for methyltransferase activity came from the recombinant expression of *Rhodobacter capsulatus* BchM in *E. coli*. In this experiment, an *E. coli* cell free extract containing over-expressed BchM was demonstrated to incorporate ^{14}C-methyl-labelled SAM onto the C13 propionate group of Mg-protoporphyrin IX (Bollivar *et al*., 1994a). Similarly, Gibson and Hunter (1994) observed methyltransferase activity with BchM from *Rhodobacter sphaeroides* recombinantly over-expressed in *E. coli*. It is interesting to note that although *Rhodobacter sphaeroides* and *Rhodobacter capsulatus* are closely related organisms there is a 5 kDa difference in the BchM molecular masses; the *Rhodobacter capsulatus* enzyme has a molecular mass of 22.3 kDa, and the *Rhodobacter sphaeroides* enzyme has a molecular mass of 27.5 kDa.

The cyanobacterial ChlM homologue was first identified by functional complementation of a *Rhodobacter capsulatus* BchM mutant using a cosmid library of the *Synechocystis* genome (Smith *et al*., 1996). Expression of this gene, driven from the strong *puc* operon promoter, was able restore chlorophyll biosynthesis levels to similar levels observed in wild type. Surprisingly, the deduced primary sequence of ChlM displays only 29% sequence identity to the BchM primary sequence, although both enzymes contain the essential putative SAM-binding motif (Gibson and Hunter, 1994; Bollivar *et al*., 1994a; Smith *et al*., 1996). Secondary structure threading techniques were used to assign the methyltransferase to the class of small molecule methyltransferases, all of which contain a common seven-stranded β-sheet and a SAM binding fold, enabling the methyltransferase to be compared with better characterised enzymes, such as glycine-N-methyltransferase and M.*HhaI* (Shepherd and

Hunter, 2004). Although there is not yet a crystal structure of the methyltransferase, the structure of the closely related methyltransferase BchU has been solved to a resolution of 2.27 Å. BchU, a dimeric protein involved in bacteriochlorophyll *c* biosynthesis, is a class I methyltransferase (Harada *et al*., 2005; Wada *et al*., 2006). The C-terminal domain of BchU exhibits classic fold of methyltransferases, a seven-stranded β-sheet, which, in this case, is flanked by six α-helices to form an open αβα sandwich.

There is evidence for coupling of enzyme activity between the methyltransferase and the H subunit of MgCH, the preceding enzyme in the chlorophyll biosynthesis pathway: Addition of *Rhodobacter capsulatus* BchH to an *E. coli* cell free extract containing over-expressed *Rhodobacter capsulatus* BchM, results in a seven fold increase in BchM activity (Hinchigeri *et al*., 1997), purified recombinant *Synechocystis* ChlH has a stimulatory effect on purified *Synechocystis* ChlM activity (Shepherd *et al*., 2005b) and a yeast two-hybrid system is positive for an interaction between Tobacco ChlM and ChlH (Alawady *et al*., 2005). Additionally, tobacco plants deficient in ChlM have been shown to have reduced MgCH activity, suggesting a feedback loop or coupling occurs between these two enzymes (Alawady and Grimm, 2005). Interestingly, the coupling of BchM from *Chlorobaculum tepidum* with each of its three BchH paralogues (BchH, BchT, BchS) gives wildly different results on BchM activity measured in stopped assays. Addition of BchH to BchM assays resulted in a 30% reduction in reaction velocity, whereas addition of BchS or BchT resulted in BchM assay velocity increasing by 60% and 30% respectively (Johnson and Schmidt-Dannert, 2008b). Although the exact mechanism by which H and BchM interact is unknown, the results provide further evidence for H and BchM functioning together to regulate the flux of porphyrins down the chlorophyll biosynthesis pathway and for the formation of a substrate channelling super-complex between the chlorophyll biosynthesis enzymes.

Kinetic analysis of the methyltransferase has proved difficult in the past due to the similar spectral properties displayed by the substrate and product of the reaction. Heterologous expression of *Synechocystis* ChlM in *E. coli* has enabled characterisation of the kinetic properties of this enzyme. The kinetic mechanism of ChlM was determined using the water soluble MgP analogue magnesium deuteroporphyrin IX (d-MgP) (Shepherd *et al*., 2003). The steady state kinetics revealed that the reaction proceeds via a ternary complex and product inhibition studies using *S*-adenosyl-L-homocysteine (SAH) indicated that substrate binding was not ordered, instead occurring via a random binding mechanism, whereby SAH may bind productively to either free enzyme or a ChlM:d-MgP complex (Shepherd *et al*., 2003). Kinetic analyses performed in the presence of the MgCH subunit ChlH revealed that ChlH dramatically accelerated the formation and breakdown of an intermediate in the catalytic

cycle of ChlM, thus suggesting that ChlH is directly involved in the reaction mechanism (Shepherd and Hunter, 2004; Shepherd *et al*., 2005c). A method for overcoming the problem of spectral similarity of the substrate and product of this reaction was recently described (McLean and Hunter, 2009). These authors showed that by coupling the ChlM reaction to SAH nucleosidase and adenine deaminase, SAH produced during the reaction could be converted to adenine and hypoxanthine, enabling the accumulation of hypoxanthine to be monitored by a decrease in absorbance at 265 nm (McLean and Hunter, 2009).

1.4.9 Mg-protoporphyrin IX monomethylester cyclase

Mg-protoporphyrin IX monomethylester cyclase (cyclase) catalyses the conversion of MgPME to 3,8-divinylprotochlorophyllide (Pchlide), in a reaction accompanied by a red to green colour change (**Figure 1.26**). Although studied in some detail, the enzyme(s) responsible for the cyclisation reaction remains elusive and the least understood in the chlorophyll biosynthesis pathway. Despite attempts over the past 60 years the active enzyme has never been completely biochemically purified or successfully expressed in a recombinant system.

In 1950, Granick laid out the first hypothesis of how the cyclisation reaction was likely to proceed based on the β-oxidation of fatty acids. He suggested the isocyclic ring would be formed by β-oxidation of the Mg-PME 6-methylpropionate side chain to form methyl-β-oxopropionate, progressing through the intermediates 6-acrylate, 6-β-hydroxypropionate and 6-β-ketopropionate (Granik, 1950). The isolation of two of these intermediates, 6-β-hydroxypropionate and 6-β-ketopropionate, from an organelle-free system further supported his hypothesis (Wong *et al*., 1985). Subsequently, chemically synthesised 6-acrylate proved to be an inactive substrate in organisms that form chlorophyll aerobically (Walker *et al*., 1988). Experiments with O^{18} labelled molecular oxygen, showed the direct incorporation of atmospheric O_2 into the carbonyl group of the isocyclic ring (Walker *et al*., 1989), suggesting the chlorophyll producing cyclase is a mono-oxygenase. In contrast, the bacteriochlorophyll producing *Rhodobacter sphaeroides* was found to incorporate oxygen donated from water, indicating the involvement of a hydratase (Porra *et al*., 1996), giving the first evidence for at least two cyclase proteins; one oxygen-dependent and one oxygen-independent.

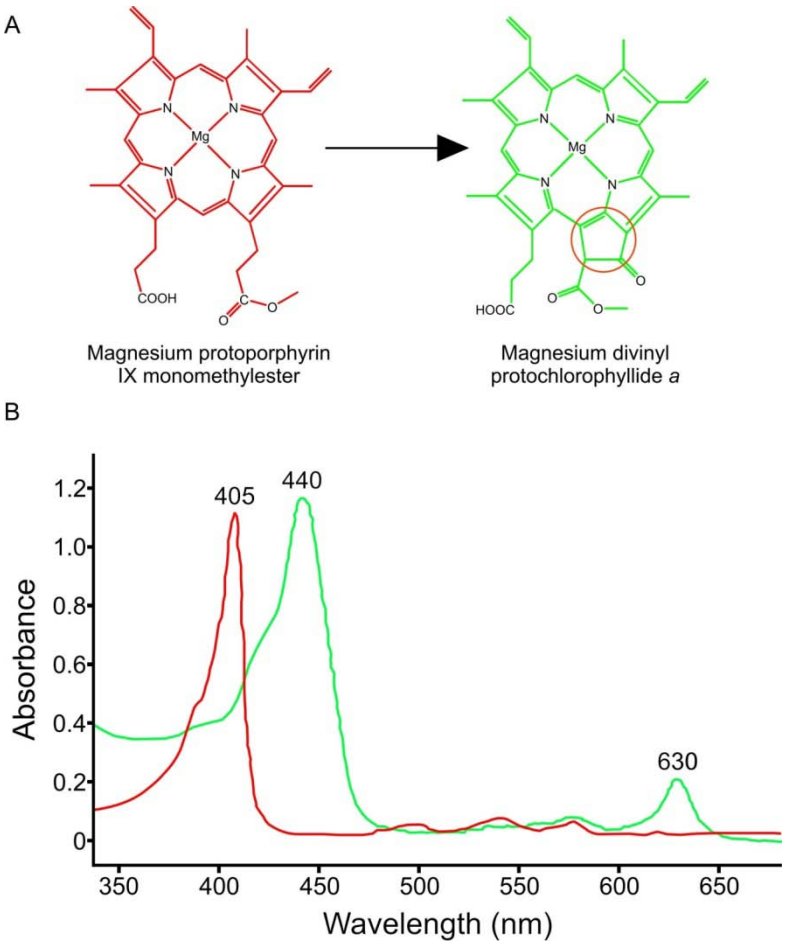

Figure 1.26 The cyclase reaction

The formation of the fifth ring, circled on divinyl-Pchlide, is responsible for the red to green colour change observed with this reaction (**A**). Absorbance spectra for the substrate Mg-PME (red) and product Pchlide (green) (**B**).

Experiments using lysates of whole cells from several photosynthetic organisms suggest the oxygen-dependent cyclase requires two cofactors, iron and NADPH, and is composed of a number of different subunits (Wong and Castelfranco, 1984a; Wong and Castelfranco, 1985; Walker et al., 1988; Walker et al., 1991; Whyte and Castelfranco, 1993; Bollivar and Beale, 1996a; Rzeznicka et al., 2005). The first evidence of multiple subunits was from cyclase assays using whole cell lysate fractions from cucumber cotyledon etioplasts, where cyclase activity could only be observed in the presence of the combined membrane and soluble fractions (Wong and Castelfranco, 1984a). Later, a partial biochemical purification of the cucumber soluble protein suggested that it had a molecular mass of over 30 kDa (Walker *et al*., 1991). Similar studies, using extracts from the green algae *Chlamydomonas reindhardtii* and cyanobacterium *Synechocystis*, also demonstrated the absolute requirement of both a membrane-bound and soluble subunit for cyclase activity (Bollivar and Beale, 1996a). Studies using barley provided additional confirmation of this distribution, although genetic studies

have also identified an additional membrane subunit encoded within the mysterious *viridis-k* locus (Gough, 1972; Rzeznicka *et al*., 2005). In all cases any attempts to further purify the membrane subunit from the crude membrane extracts resulted in abolition of cyclase activity, further highlighting the extreme difficulty of solving the mystery of the cyclase.

Early evidence the cyclase requires iron as a co-factor was the observation that MgPME (the substrate of the cyclase) accumulated in plants deficient in iron (Spiller *et al*., 1982). Subsequently, iron chelators were found to inhibit cyclase activity when added to assays using, intact wheat leaf chloroplasts (Nasrulhaq-Boyce et al., 1987a), lysates of cucumber cotyledon chloroplasts (Walker *et al*., 1991), the membrane fraction of *Synechocystis* and lysates of *Chlamydomonas reinharttii* chloroplasts (Bollivar and Beale, 1996a). Further probing into the type of iron required for cyclase activity, revealed the cyclase to co-ordinate a non-haem, Fe^{2+} ion, located within the membrane bound subunit, as deduced from the discovery that hydrophilic Fe^{2+} chelators and all Fe^{3+} chelators had little effect on cyclase activity (Bollivar and Beale, 1996a).

The first gene associated with anaerobic cyclase activity is *bchE*, which was indentified in mutants of *Rhodobacter sphaeroides* and *Rhodobacter capsulatus* that accumulated MgPME (Hunter and Coomber, 1988; Yang and Bauer, 1990; Gough *et al*., 2000). Later, the aerobic cyclase was found to require the unrelated gene *acsF* (Pinta *et al*., 2002), of which homologues have been found in all the oxygenic photosynthetic organisms investigated. These include the genes *Crd1* and *Cth1* in *Chlamydomonas reinhardtii* (Moseley *et al*., 2000; Moseley *et al*., 2002), *sll1214* and *sll1874* in *Synechocystis* (Minamizaki *et al*., 2008a; Peter *et al*., 2009), *Chl27* in *Arabidopsis* (Tottey et al., 2003a) and *Xantha-I* in barley (Rzeznicka *et al*., 2005). Further bioinformatics analyses of the distribution of the anaerobic and aerobic cyclase genes revealed that the *bchE* and *acsF* genes are found together in a wide range of photosynthetic organisms. These include the filamentous anoxygenic phototrophs: *Chloroflexus aurantiacus*, *Roseiflexus* sp. and *Roseiflexus castenholzii*, and the purple bacteria: *Rhodobacter sphaeroides, Bradyrhizobium* sp, *Rhodopseudomonas palustris*, *Methylobacterium* sp., *Rubrivivax gelatinosus* and *Roseobacter denitrificans* and the cyanobacterium *Synechocystis*. With the exception of *Synechocystis*, all of the bacteria mentioned above are capable of growing under aerobic and anaerobic conditions, suggesting the aerobic and anaerobic cyclases are differentially expressed to cope with changing environmental oxygen conditions.

A recent study into the distribution of the AcsF and BchE genes throughout the *Proteobacteria* revealed that the different functional groups of the phototrophic

Proteobacteria employed different forms of the cyclase (Boldareva-Nuianzina *et al*., 2013). The strictly anaerobic purple sulphur bacteria were found to contain only the *bchE* gene, whereas the purple non-sulphur bacteria analysed were found to contain both *acsF* and *bchE* and all of the aerobic anoxygenic phototrophs studied contained the *acsF* gene – almost half of these also contained the *bchE* gene. Further phylogenetic analyses into the distribution of *bchE* showed that unlike *acsF* its distribution is not conserved within the photosynthetic gene cluster. Rather, *bchE* is found in different places of the genome, suggesting the gene was transferred several times during evolution. This is further supported by the variance of the *bchE* gene sequences found within the *Gammaproteobacteria (Boldareva-Nuianzina et al., 2013)*. Interestingly, the authors also noted *bchE* was present concurrently with *bchJ* – a gene perhaps falsely attributed divinyl reductase activity, and whose exact function also remains unknown (Canniffe *et al*., 2013).

Following discovery of the *AcsF* gene in *Rubrivivax gelatinosa*, the *Crd1* (Copper response defect 1) and *Crt1* (Copper target homolog 1) genes in *Chlamydomonas reinhardtii* were discovered via an independent investigation into the effects of copper depletion on protein expression. Moseley *et al*. (2000 & 2002), showed that *Crd1* expression was significantly up-regulated under copper limiting or oxygen limiting conditions and was essential for the maintenance of PSI and its associated light-harvesting complexes. Upon probing the localisation of *Crd1*, it was found to be solely localised in the plastid membrane. *Crt1* was identified in the expressed sequence tag database as a paralog of *Crd1* and although found to have a similar function, it was found to be differentially expressed, regulated by a mechanism of metalloregulation (Moseley *et al*., 2002). The *Arabidopsis* CHL27 gene was identified by its homology to *Crd1* and was found to be required for cyclase activity. CHL27-antisense plants fed with ALA accumulated MgPME and were blocked at the cyclase step (Pinta et al., 2002; Tottey et al., 2003a). CHL27 knock down mutants were also found to have severe growth retardation, defects in chloroplast development and damaged PSII reaction centres (Bang *et al*., 2008). Micro-array analyses of these mutants revealed widespread repression of genes involved in photosynthesis, including those for light-harvesting complex I & II, PSI and PSII (Bang *et al*., 2008). As with *Crd1*, localisation experiments showed the *chl27* gene product was in the chloroplast, associated with both the inner-envelope and thylakoid membranes (Tottey et al., 2003a). Interestingly, *Arabidopsis* contains only one copy of the CHL27 gene, whereas the photosynthetic organisms *Chlamydomonas reinhardtii* and *Synechocystis* have two isoforms; one reason for this could be the requirement of micro-organisms to adapt rapidly to changing environmental conditions.

Synechocystis contains three *bchE* homologues (*slr0905, sll1242* and *slr0309*) and two *acsF* homologues (*sll1214* and *sll1874*). Of the two *acsF* homologues, *sll1214* is thought to be the sole cyclase operating under aerobic conditions, with *sll1874* a cyclase that operates under micro-oxic conditions. Experimental evidence demonstrated that the Δ*sl1214* mutant accumulated MgPME and could only grow under micro-oxic conditions, whereas the Δ*sll1874* mutant was able to grow aerobically, but not under micro-oxic conditions (Minamizaki *et al*., 2008a; Peter *et al*., 2009). Knockout mutants in each of the three *bchE* homologues were able to propagate under anaerobic conditions analogously to wild type cells, suggesting these genes do not confer anaerobic cyclase activity in *Synechocystis* (Minamizaki *et al*., 2008b). Although one subunit of the cyclase has been successfully identified, this subunit on its own does not confer cyclase activity, demonstrated by assays on *E. coli* whole cell lysates in which *sll1214* and *sll1874* had been over-expressed (Hollingshead S, unpublished data). This is concurrent with the evidence gathered from whole cell cyclase assays which all demonstrate the absolute requirement of both a soluble and insoluble subunit (Walker et al., 1988; Walker et al., 1989; Walker et al., 1991; Bollivar and Beale, 1996a; Rzeznicka et al., 2005). Thus, there remains at least one cyclase subunit, located within the soluble fraction, to be identified.

The expression of the AcsF homologue in the green filamentous micro-organism *Chloroflexus aurantiacus* has also been investigated (Tang *et al*., 2009); like *Synechocystis* this organism also has genes encoding the aerobic and anaerobic cyclase. The authors were surprised to see AcsF expressed under anaerobic growth conditions, when it was previously thought the AcsF gene was only expressed under aerobic conditions (Pinta *et al*., 2002; Minamizaki *et al*., 2008b). Pairwise comparison of the *Chloroflexus aurantiacus* AcsF gene to those of other organisms revealed a 50 % or lower identity, much lower than the 70 % found between genes from most other organisms. This led Tang *et al*. (2009) to propose that the *Chloroflexus aurantiacus* AcsF may have additional roles to the cyclisation reaction, which still occur under anaerobic growth conditions. These hypotheses include: a role in the biosynthesis of different bacteriochlorophylls, a role as an iron transport protein or a role as a site for electron transfer under anaerobic growth conditions (Tang *et al*., 2009).

Cobalamin has been identified as a required co-factor of the anaerobic cyclase (Gough *et al*., 2000). Mg-protoporphyrins, particularly MgPME, were found to accumulate in the B_{12}-requiring *Rhodobacter capsulatus* mutants, which carry the *bluE* and *bluB* genes. These were converted to Pchlide upon the addition of exogenous cobalamin, which restored anaerobic cyclase activity. This observation has led to the proposal of a mechanism for the anaerobic reaction; adenosylcobalamin is thought to form an adenosyl radical, which in turn leads to the formation of the 13^1-radical of MgPME. Withdrawal of an electron from the 13^1-radical

would give rise to a 13^1-cation, which may be attacked by a hydroxyl ion to give the 13^1-hydroxy intermediate of MgPME; finally the withdrawal of three hydrogen atoms leads to the eventual cyclisation and formation of Pchlide. *BchE* has also been identified as containing an iron sulphur cluster, which may be involved in the initial electron transfer steps of the mechanism, as iron-deficient *Rhodobacter capsulatus* was found to accumulate MgPME (Cooper, 1963). Interestingly, *bchE* also contains the cysteine motif (CXXX-CXXC) essential for the formation of iron-sulphur clusters in many anaerobic enzymes (Johnson, 1998). This was later confirmed with UV-visible spectral analysis of the heterologously expressed *Rhodobacter gelatinosus* BchE protein (Ouchane *et al*., 2004).

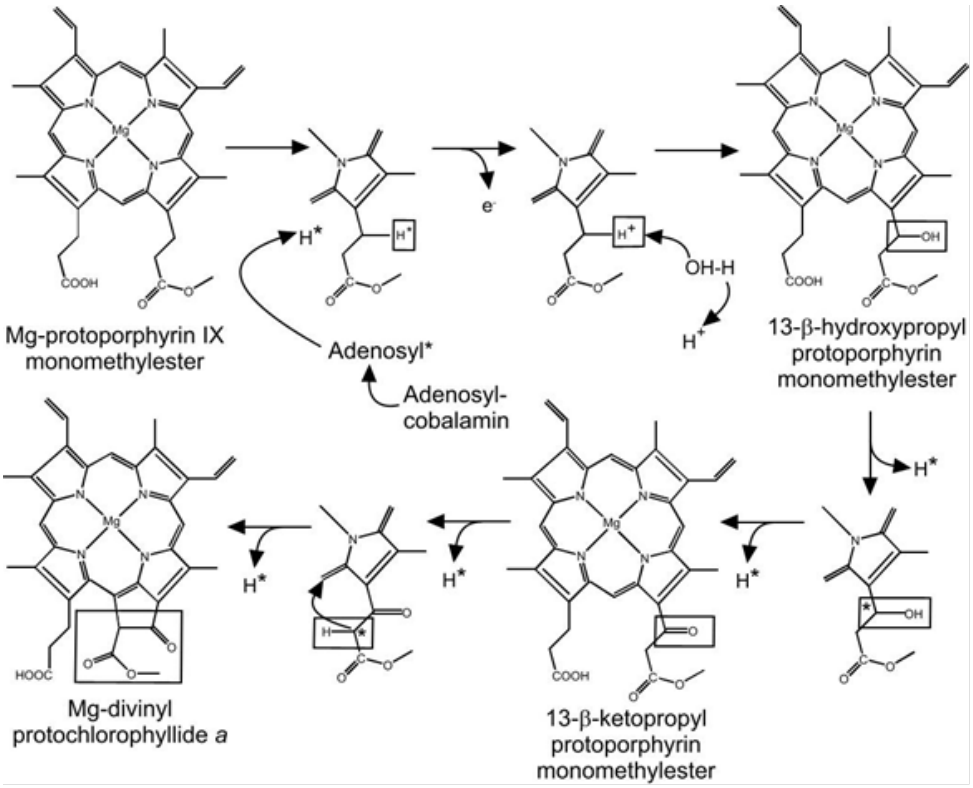

Figure 1.27 The proposed reaction mechanism for the anaerobic cyclase

The reaction proceeds via the β-hydroxy and β-keto intermediates, while the monovinyl species can also act as a substrate for the enzyme. Radicals are denoted by an asterisk (Gough et al., 2000).

It has been suggested that the chloroplast localised NADPH thioredoxin reductase and 2-cys peroxidase have a role in promoting cyclase. An *Arabidopsis* mutant, in which the NADPH thioredoxin reductase gene, *ntrc*, was disrupted, accumulated the chlorophyll precursor pigments protoporphyrin IX, Mg-protoporphyrin IX and MgPME. The hypothesis for the perturbation at the cyclase step, and the accumulation of pigments preceding this step, is that the reactive oxygen species generated during the cyclase reaction would then go on to destroy the proteins responsible for Pchlide catalysis. Stenbaek *et al*. (2008) also

demonstrated that the addition of NADPH thioredoxin reductase and 2-cys peroxidase to the cyclase reaction resulted in a stimulation of aerobic cyclase activity.

1.4.10 Protochlorophyllide reductase

After the formation of the isocyclic ring, Pchlide may be reduced at the 8-vinyl group of ring D by the divinyl reductase (described in section 1.4.11) followed by reduction of the C17-C18 double bound catalysed by POR, or Pchlide may be directly reduced by POR. The addition of a hydride across the C17-C18 double bond results in the formation of chlorophyllide (chlide); this reaction may be catalysed by two different Pchlide reductases. The first is a light dependent enzyme (POR), which is one of only two enzymes in biology whose catalytic activity is known to be activated by light. The second (DPOR) can undertake catalysis in the dark. Phylogenetic analyses of the two POR genes show that primitive anoxygenic photosynthetic bacteria contain only dark POR and angiosperms (flowering plants) contain only light POR. In contrast, cyanobacteria, green algae, mosses and most gymnosperms possess both light and dark POR. The lack of dark POR in angiosperms makes these organisms dependent on light for chlorophyll biosynthesis.

One theory for the distribution of light and dark POR enzymes across photosynthetic organisms is based on the extreme oxygen sensitivity of dark POR. Dark POR is thought to have evolved first in the ancient anoxygenic photosynthetic organisms. However, dark POR has a half life that rapidly declines upon exposure to oxygen (Nomata et al., 2006); dark POR of anoxygenic and oxygenic photosynthetic organisms have a very similar half life in the presence of oxygen, suggesting that dark POR was not able to evolve to adopt greater oxygen tolerance upon the advent of oxygenic photosynthesis (Yamazaki *et al*., 2006; Yamamoto *et al*., 2009). Instead, bacteria capable of oxygenic photosynthesis are thought to have adapted to the increasing oxygen content of the atmosphere by acquiring the oxygen insensitive light POR (Reinbothe *et al*., 2010).

1.4.10.1 The “light” POR

The unique, light activated reduction of Pchlide to Chlide by POR is perhaps the most interesting reaction in chlorophyll biosynthesis. Unlike the dark POR, POR is encoded within the nuclear DNA as a single polypeptide, which gives rise to a protein of ~37 kDa. The first evidence of POR activity was from experiments performed on material from dark grown barley angiosperms (Griffiths, 1975a; Griffiths, 1975b; Griffiths, 1978). This dark-grown etiolated material accumulates Pchlide and POR at high levels, which with the addition of NADPH form large aggregated ternary complexes, known as “prolamellar bodies” (Griffiths, 1978). On illumination these prolamellar bodies rapidly dissociate as the hydride molecule is

transferred from NADPH to Pchlide, forming Chlide (Heyes *et al*., 2006). Although several spectral forms of Pchlide have been identified in these prolamellar bodies, they have been attributed to the pigment interactions with membranes or proteins shifting the spectra, rather than different species of Pchlide (Boddi *et al*., 2003). The differing spectral properties of Pchlide and Chlide allow UV-VIS spectroscopic assays to be performed with ease. This, coupled with the very nature of POR, its absolute requirement of light for catalysis, has allowed the investigation of the very fast initial phase of the reaction – the transfer of hydride from NADPH to Pchlide - which is completed on a sub-picosecond timescale; impossible to analyse for most enzyme catalysed reactions.

Extensive studies have been made of the POR reaction using a combination of site-directed mutagenesis, substrate analogues and various forms of spectroscopy. This has enabled a detailed mechanism for the reaction of POR to be compiled. Experiments with the 4R and 4S ^{3}H-labelled isomers of NADPH show that the *pro-S* hydride of the nicotinamide ring is the proton transferred to the C17 position of Pchlide (Begley and Young, 1989). Two highly conserved residues, which correspond to Tyr189 and Lys193 in *Synechocystis* POR, are essential for POR activity. During POR catalysis, after the transfer of the hydride from NADPH to C17 of Pchlide, the proton at the C18 position is thought to be added from Tyr189. It is thought Lys193 is necessary to lower the pKa of tyrosine, thus facilitating deprotonation (Wilks and Timko, 1995). Analyses of the POR reaction at low temperatures (120K) led to the discovery that the catalytic mechanism consists of an initial light-driven step, which is followed by several dark reactions (Heyes *et al*., 2002; Heyes *et al*., 2003; Heyes and Hunter, 2004).

The rate limiting step of POR is the initial formation of the ternary complex, i.e. the binding of NADPH to POR. Recent studies have shown that it is essential POR is pre-bound to NADPH; an intricate process consisting of three distinguishable kinetic phases. These phases are characterised by structural changes to the protein, which allow for the correct docking of NADPH (Heyes *et al*., 2008). These data are one explanation for the formation of large prolamellar bodies of ternary complex, as the formation of large quantities of protein already bound to substrate will rapidly reduce the rate limiting step and allow for the organism to produce Chlide and then chlorophyll as soon as the light cycle begins. It has been suggested that phosphorylation of POR is required for the aggregation of POR into these prolamellar bodies, although the exact mechanism of this phosphorylation remains unclear (Wiktorsson *et al*., 1996; Kovacheva *et al*., 2000).

Prokaryotic chlorophyll-producing organisms containing the dark POR gene contain only a single copy of the POR-encoding gene (Rowe and Griffiths, 1995). However, three different isoforms of the POR gene have been identified in higher plants, all of which are differentially regulated depending upon light patterns and development stage (Holtorf *et al*., 1995; Armstrong *et al*., 1995; Oosawa *et al*., 2000). These different isoforms are probably required at different stages in the greening process to enable the plant to regulate its requirement for chlorophylls. Two of these isoforms, PORA and PORB, are thought to form a high mass light-harvesting complex within the prolamellar bodies of dark grown material. Here, they are found in a ratio of 5 to 1 for PORA to PORB; this stochiometry is thought to optimise this step of chlorophyll biosynthesis, so Chlide formation can occur rapidly, even under low light intensities (Reinboth *et al*., 1999). However, it should be noted that the existence of these complexes remains controversial and the matter of ongoing debate (Masuda and Takamiya, 2004).

No structural data for POR is currently available, as the enzyme is thought to be composed of several mobile loops, rendering crystallisation impossible. Although work is ongoing to ascertain the structure by NMR, the molecular mass of light POR at 37 kDa is at the very upper limit of this technique, making progress slow and costly. However, bioinformatics analysis of the PORs primary sequence reveals this protein is a member of the short-chain dehydrogenase/reductase superfamily of enzymes. Structures of several members of this family have been solved, and several groups have used these to construct homology models of light POR from *Synechocystis* (Townley *et al*., 2001), pea (Dahlin *et al*., 1999) and barley (Buhr *et al*., 2008). All three models suggest POR has a central seven-stranded parallel β-sheet, which is predicted to be surrounded by nine α-helices. A 33 amino acid loop, connecting two of these β-strands, was initially suggested to be required for Pchlide binding, however this was later shown to be involved in the formation of prolamellar bodies (Reinbothe and Reinbothe, 1996; Townley *et al*., 2001; Reinbothe *et al*., 2010). POR also has a highly conserved Rossmann binding motif (GXXXGXG), which is located at the N-terminus and the site of NADPH binding.

1.4.10.2 The “dark” POR

The dark POR is encoded within three genes, *chl/bchB*, *chl/bchL*, and *chl/bchN*. These were first identified in *Rhodobacter capsulatus*, when mutations in any of these three loci resulted in the loss of bacteriochlorophyll formation and the accumulation of Pchlide (Zsebo and Hearst, 1984; Coomber *et al*., 1990; Yang and Bauer, 1990; Burke *et al*., 1993b). The corresponding genes in chlorophyll-producing eukaryotes were found encoded within the chloroplast genome (Kohchi *et al*., 1988). The first biochemical characterisation of the dark

POR came from over-expression and purification of all three *Rhodobacter capsulatus* subunits from *Rhodobacter capsulatus* (Fujita and Bauer, 2000). Two of the three genes, *bchL* and *bchN* were expressed as S-tag fusion proteins, whilst the third gene, *bchB*, was untagged. BchN was found to co-purify with BchB, indicating the formation of a tight complex. Fujita and Bauer (2000) went on to demonstrate dark POR activity was absolutely dependent on the presence of all three subunits, as well as ATP and reduced dithionite.

Although all three dark POR genes are encoded on the chloroplast genome, classical genetic analyses have identified algal mutants that exhibit a *yellow-in-the-dark* (*y*) phenotype. Mutations in any of these *y* loci result in the loss of chlorophyll production and the accumulation of a small amount of Pchlide; their yellow phenotype is caused by carotenoid accumulation (Armstrong, 1997). Each of these mutants is capable of synthesising Childe in the light, when the light POR comes into play, but production of Chlide in the dark is prohibited. *Chlamydomonas reinhardtii* has seven of these nuclear encoded loci, named *y-1* to *y-10*, and mutations in any of these loci results in abolition of Chlide production (Sager, 1955; Ford and Wang, 1980a; Ford and Wang, 1980b). Analyses of wild type *Chlamydomonas reinhardtii* cells showed that although ChlN and ChlB were present in similar concentrations under light and dark grown wild type cells, ChlL was only present in cells grown in the dark or at light intensities below 15 µmol m^{-2} sec^{-1} (Cahoon and Timko, 2000). These authors found that ChlL expression was negatively photoregulated by the redox potential within the chloroplast, and that accumulation of ChlL was dependent upon expression of the nuclear encoded *y* genes. The existence of *y* genes is common to the green algae, as similar *y* genes have also been identified in *Chlorella* sp. and *Scenedesmus* sp. (Bogorad, 1976; Schulz and Senger, 1993).

Bioinformatics analyses show the primary sequences of all three subunits have significant similarities to the *nifH* (35 % identity to *chlL*), *nifD* (19% identity to *chlN*) and *nifK* (19% identity to *chlB*) subunits of nitrogenase, an enzyme that catalyses the reduction of dinitrogen to ammonia (Fujita *et al*., 1993). Nitrogenase is an oxygen-sensitive enzyme that catalyses the ATP-dependent reduction of atmospheric N_2 to ammonia. It is composed of a Fe protein (a NifH dimer), which contains a 4Fe-4S cluster in each subunit, and a MoFe protein (a NifD-NifK heterodimer), which has a 8Fe-7S cluster and a Mo-7Fe-9S-homocitrate cluster (Einsle et al., 2002). The Fe protein is an ATP-dependent electron donor for the MoFe protein and the MoFe protein provides the active site for dinitrogen reduction (Peters *et al*., 1995; Igarashi and Seefeldt, 2003). In a similar manner to the nitrogenase, dark POR is composed of an L-protein (a BchL dimer) and a catalytic component, the NB-protein (a BchN-BchB heterotetramer) (Fujita and Bauer, 2000; Fujita and Bauer, 2003).

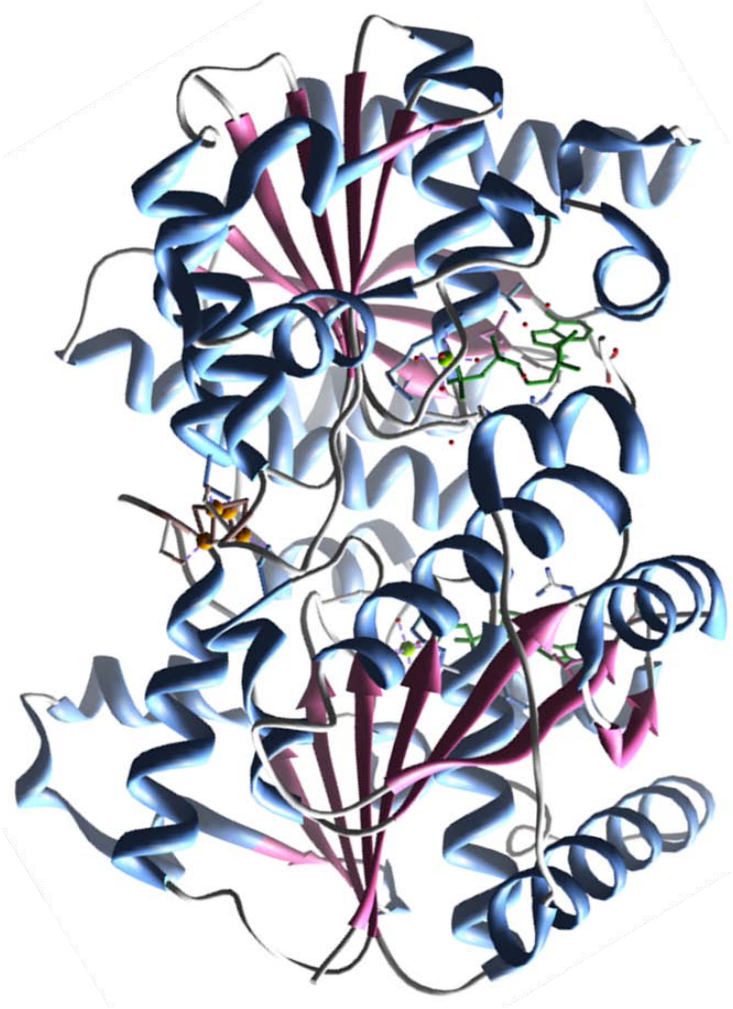

Figure 1.28 Crystal structure of the L-protein of *Rhodobacter sphaeroides* dark POR

The crystal structure of *Rhodobacter sphaeroides* L-protein subunit of dark POR solved to a resolution of 1.63 Å. The homodimer in a complex with two Mg-ADP molecules, highlighted in green, in orange are the four Fe ions located in a Fe-S cluster with four S atoms (light brown). The FeS cluster is ligated via four conserved cysteine residues, Cys126 and Cys160, two from each subunit (PDB 3FWY, Sarma *et al*., 2008).

Indeed, the recently solved crystal structures of the dark POR L-protein and NB-protein from *Rhodobacter sphaeroides* and *Rhodobacter capsulatus* respectively, confirm that the sequence identity of the dark POR proteins to the nitrogenase proteins is reflected in the crystal structure (Sarma *et al*., 2008; Muraki *et al*., 2010). The crystal structure of the L-protein, bound to MgADP, was solved to a resolution of 1.6 Å (Sarma *et al*., 2008). Recombinant L-protein was obtained by over-expression in an *Azotobacter vinelandii* system and purified using a polyhistidine tag. Analyses of the iron content showed each mole of BchL to contain 3.6 moles of Fe, consistent with the presence of a 4Fe-4S cluster.

Comparison of the BchL structure to that of the nitrogenase Fe subunit showed that these two proteins share an overall structural similarity including the subunit bridging 4Fe-4S

cluster and the nucleotide binding site (**Figure 1.28**). However, despite this apparent high structural similarity, BchL was unable to replace the nitrogenase Fe subunit in a nitrogenase substrate reduction assay (Sarma *et al*., 2008).

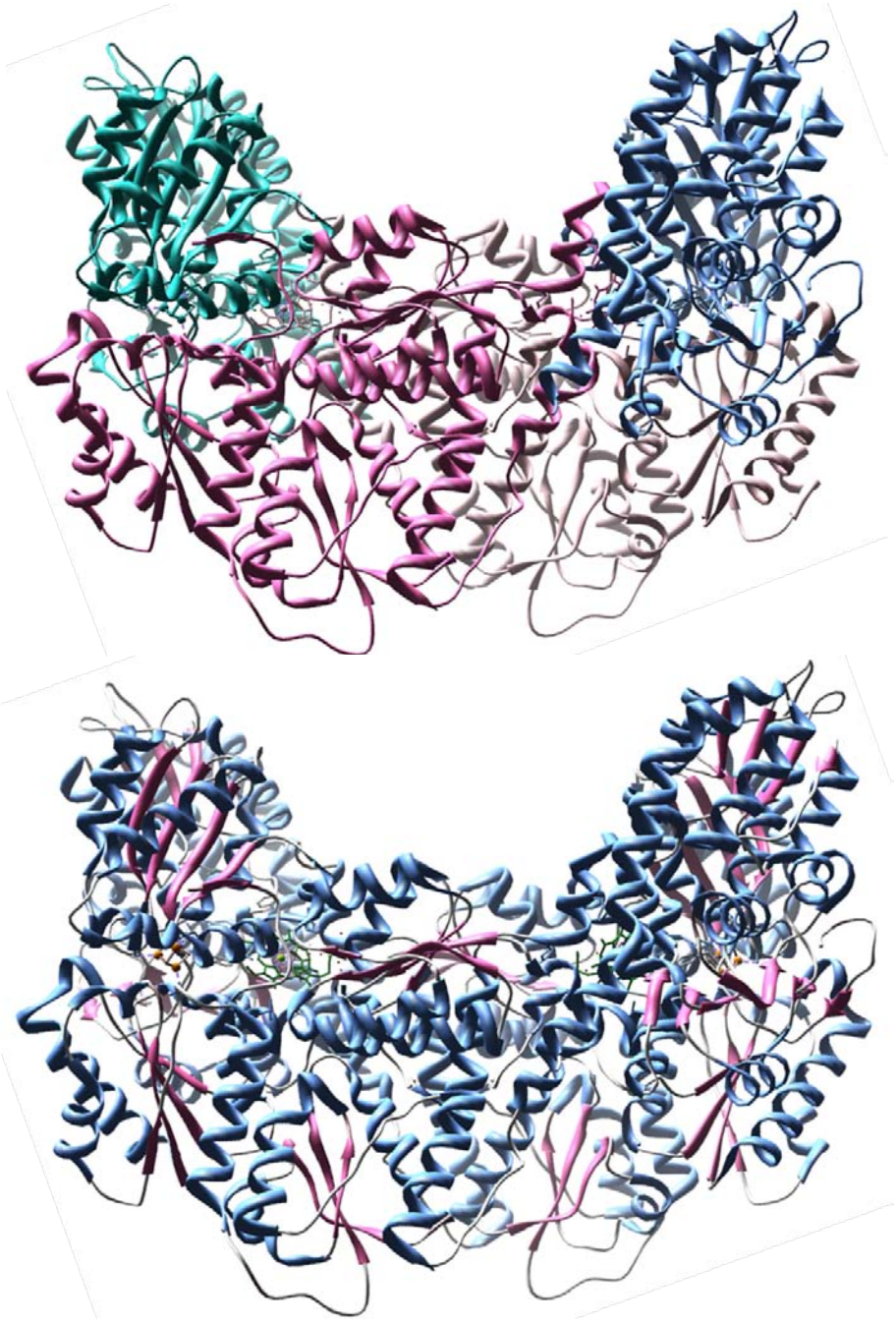

Figure 1.29 Crystal structure of NB-protein from dark POR

The crystal structure of *Rhodobacter capsulatus* NB-protein subunit of dark POR solved to a resolution of 2.8 Å. Top, highlighted in cyan and light blue are the two BchN subunits and highlighted in pink and light pink are the two BchB subunits. Bottom, the hetero-tetramer in a complex with two Pchlide molecules (green), in orange are the four Fe ions located in an Fe-S cluster with four S atoms (light brown). The FeS cluster is co-ordinated via cysteine residues Cys26, Cys51 and Cys112 from BchN and Asp36 from BchB (PDB 3AEK, Muraki *et al*., 2010).

The crystal structures for both the Pchlide-bound (**Figure 1.29**), and Pchlide-unbound NB proteins were solved to resolutions of 2.3 Å and 2.8 Å respectively. As with the L-protein, the NB-protein was found to have high structural similarity to the nitrogenase MoFe-protein, with the spatial arrangements of the iron sulphur cluster and Pchlide of the NB-protein being almost identical to the iron-sulphur cluster and FeMo-cofactor of the FeMo-protein (Muraki *et al*., 2010). The only notable difference between the NB-protein and the FeMo-protein is the unique C-terminal region possessed by BchB that is 100 residues in length. This region is conserved among all BchB/ChlB proteins and probably plays an important role in dark POR function; however, due to the disordered nature of this region, it is rendered "invisible" in the crystal structure (Muraki *et al*., 2010; Reinbothe *et al*., 2010).

Each BchN-BchB hetero-dimer was co-ordinated with one 4Fe-4S cluster and one Pchlide molecule; interestingly the iron-sulphur cluster co-ordinates with residues from both proteins; three Cys residues from BchN (Cys26, Cys51 and Cys112) and one Asp residue from BchB (Asp36). The Pchlide molecule was located within a hydrophobic substrate binding cavity. Prior to determination of the NB-protein crystal structure, it was thought Cys95 in the BchB protein also played a part in the iron-sulphur cluster co-ordination, as a BchB protein in which this residue is mutated is unable to form an active complex with BchN (Fujita and Bauer, 2003; Nomata *et al*., 2008; Brocker *et al*., 2008). However, this residue was found located 6 Å away from the closest Fe ion, suggesting instead that Cys95 plays a crucial role in stabilising the BchB-BchN dimer (Muraki *et al*., 2010). The authors also found that the BchB Asp36 was not required for assembly of the protein-iron sulphur complex, but that it was absolutely required for catalytic activity. A significant structural change was found to occur upon Pchlide binding, in the form of the partial unwinding of the α-helix located next to the catalytic region of the BchN-BchB dimer. This goes on to form a lid structure, which closes to form a hydrophobic pigment binding cavity (Muraki et al., 2010).

Naturally, the reaction mechanism employed by dark POR is fundamentally different to that of light POR. The mechanism of double bond reduction as inferred from the structural data is thought to begin with the transfer of a single electron from the 4Fe-4S cluster of the L-protein to the 4Fe-4S cluster of the NB-protein. This electron is then transferred from the NB-protein 4Fe-4S cluster to the conjugated electron cloud of Pchlide. Simultaneously two protons, one from BcbB-Asp274 and one from the propionate side chain of Pchlide, is transferred to C17 and C18, resulting in the production of a cation radical. The reaction is completed upon the transfer of a second electron to Pchlide from the NB-protein 4Fe-4S cluster and the formation of a single bond between C17 and C18 (Muraki *et al*., 2010; Reinbothe *et al*., 2010).

Ferredoxin then reduces the oxidised L-protein, regenerating it ready for the next reaction cycle (Nomata *et al*., 2005).

Very recently an *in silico* study has proposed the L-protein of dark POR and POR contain a tyrosine-phenylalanine-tyrosine (TFT) motif, the first such similarity reported between the two enzymes (Gabruk *et al*., 2012). This sequence is thought to be specific for the POR reaction, as a similar motif is not present in the NifH subunit of the nitrogenase, or any short chain dehydrogenases. The exact function of this TFT motif remains unknown due to a lack of experimental data, but existing studies using POR and L-protein mutants show this region is highly important for wild type levels of POR activity (Dahlin *et al*., 1999; Sarma *et al*., 2008).

1.4.11 Divinyl reductase

Chlorophylls may be classified into two groups depending on the number of vinyl side chains they possess; 3-8-divinyl (DV) or 3-monovinyl (MV), with the DV molecule containing two vinyl groups at positions 3 and 8 on the macrocycle and the MV molecule having a vinyl group at position 3 and an ethyl group at position 8. Conversion of the 8-vinyl group to an ethyl group is catalysed by the enzyme known as the divinyl reductase (DVR). Although the 8-vinyl reduction is an essential step in chlorophyll biosynthesis (the only exception being the marine plankton *Prochlorococcus sp*, which contain DV-chlorophylls), uncertainty still remains as to the exact stage in the pathway the reaction occurs. Various MV and DV chlorophyll precursors have been detected in plants and algae, which cannot be accounted for by the stepwise operation of a linear biosynthetic pathway (Rebeiz *et al*., 1994), thus leading to the hypothesis that several DVR enzymes are responsible for reducing different DV intermediates (Tripathy and Rebeiz, 1986; Tripathy and Rebeiz, 1988a; Whyte and Griffiths, 1993). One theory is that MV and DV pools of different intermediates represent separate routes for chlorophyll biosynthesis, which can be altered by physiological and environmental factors, such as age and light conditions (Rebeiz *et al*., 1983; Shioi and Takamiya, 1992). An alternative theory is that a single 8-vinyl reductase exists with broad substrate specificity; therefore the different MV/DV intermediates observed would simply be an effect of variation in the rate of 8-vinyl reduction on different substrates.

Several groups have reported the ability of photosynthetic organisms to reduce the 8-vinyl group to an 8-ethyl group at different stages of the chlorophyll biosynthetic pathway. It was suggested the substrate for the DVR reaction in *R. sphaeroides* was DV-Pchlide, as this precursor accumulates in cells treated with 8-hydroxyquinoline, an inhibitor of bacteriochlorophyll biosynthesis (Jones, 1963a; Jones, 1963b). In 1974, Ellsworth and Hsing demonstrated that DV-MgPME could be converted to MV-MgPME in crude homogenates of

etiolated wheat plastids. Tripathy and Rebeiz (1988) then showed that exogenous DV-Pchlide could be partially converted to MV-Pchlide by barley plastids. Later, Parham and Rebeiz (1992 & 1995) showed DV-Chlide could be converted to MV-Chlide in cucumber etioplast membranes isolated from etiolated cotyledons, and by extracts from barley and maize leaves. Meanwhile, Whyte and Griffiths (1993) found that etioplasts isolated from cucumber and wheat tissue would convert an exogenous DV-Pchlide pool to a predominantly MV-Pchlide pool in the dark. All the evidence supports a multi-branched pathway, but without the isolation of an active enzyme the authors could not confirm the presence of a single, broad specificity, DVR enzyme or a family of closely related DVR enzymes with different substrate specificities.

A *bchJ* mutant was isolated from *Rhodobacter capsulatus* that was able to synthesise bacteriochlorophyll, but which accumulated and excreted DV-Pchlide, which lead to the conclusion that *bchJ* encoded a DVR (Bollivar *et al*., 1994c). However, in a later study, it was reported that mutants in any of the dark-POR genes (*bchB, bchL* or *bchN*) accumulated a mixture of DV-Pchlide and MV-Pchlide, with a *bchJ/bchL* double mutant accumulating Pchlide with a significantly greater DV to MV ratio. It was also reported that the MgPME pool in *Rhodobacter capsulatus* normally contained a mixture of DV and MV molecules, the ratio of which was not affected in the *bchJ* mutant. This led the authors to conclude that a second DVR may be present, which can convert earlier intermediates to a MV form (Suzuki and Bauer, 1995). Subsequently, the *bchJ* homologue in *Chlorobaculum tepidum* was disrupted; the resulting mutant produced detectable amounts of MV-bacteriochlorophylls, but secreted DV-Pchlide into the culture medium (Chew and Bryant, 2007). An identical result was observed in *Rhodobacter sphaeroides* after disruption of the *bchJ* gene (Canniffe *et al*., 2013), suggesting *bchJ* is not a DVR, but that it may play an important role in substrate channelling or regulation of (bacterio)chlorophyll biosynthesis.

In 2005, two groups independently showed that the AT5G18660 gene of *Arabidopsis* encodes a DVR (Nakanishi et al., 2005; Nagata et al., 2005a). When expressed in *E. coli*, AT5G18660 catalyses the conversion of DV-Chlide to MV-Chlide, in a reaction that requires NADPH (Nagata et al., 2005a). Phylogenetic analysis reveals genes homologous to the *A. thaliana* DVR gene across higher plants, green algae, photosynthetic bacteria and *Synechococcus sp*, but not in other cyanobacterial lineages or red algae. One such homologue of AT5G18660, *bciA* (*CT1063*), was identified in *Chlorobaculum tepidum;* disruption of this gene led to accumulation of DV-chlorophylls. The recombinant BciA protein expressed in *E. coli* was found to reduce DV-Pchlide to MV-Pchlide *in vitro* (Nagata et al., 2005b). Open reading

frame, rsp_3070, was found to encode an orthologue of the *bciA* gene in *Rhodobacter sphaeroides* (Chew and Bryant, 2007).

A DVR gene in *Synechocystis*, open reading frame *slr1923*, was identified, which was unrelated to *bciA* or b*chJ* through a bioinformatics screen. Subsequent mutation of this gene resulted in mutants that were unable to synthesise MV-chlorophylls or grow under high light intensities (Islam *et al*., 2008; Ito *et al*., 2008). In *Synechocystis* strains where the *slr1923* gene was replaced by either the *Rhodobacter sphaeroides bchJ* or *bciA* (*rps_3070*) genes, the *bciA* strain was found to have WT levels of MV-chlorophyll, whereas the *bchJ* strain did not (Canniffe *et al*., 2013). This provides further evidence that cyanobacterial lineages (excluding *Synechococcus*) and red algae contain a class of DVR unrelated to the enzyme in plants and non-cyanobacteria, and that BchJ is not a functioning DVR.

Very recently it was reported that one DVR enzyme is responsible for reducing the 8-vinyl groups in various chlorophyll biosynthesis groups, with DVR proteins from different species having diverse and differing substrate preferences, despite being homologues (Wang *et al*., 2013). Enzyme assays using four recombinant DVR proteins from rice, maize, Arabidopsis and cucumber were performed using five DV substrates: chlorophyll, Chlide, Pchlide, MgPME and Mg-Protoporphyrin IX. The rice and maize DVRs were found to act on all five DV substrates, whereas the cucumber and *Arabidopsis* proteins could only act on three of the substrates (DV-Chlide, DV-Pchlide and DV-MgPME). This study led to the conclusion that a single DVR with broad substrate specificity is responsible for reducing the 8-vinyl groups of chlorophyll biosynthetic intermediates in higher plants (Wang *et al*., 2013).

1.4.12 Geranylgeranyl reductase and chlorophyll synthase

In the final stages of chlorophyll *a* biosynthesis Chlide is esterifyed with activated geranylgeranyl, a C_{20} isoprenoid alcohol, which constitutes 30% of the total molecular weight of the chlorophyll molecule. Chlorophyll synthase (ChlG) is unique in that it links the isoprenoid and tetrapyrrole metabolism pathways, through the esterification of the tetrapyrrole Chlide to the isoprenoid geranylgeranyl (see **Figure 1.31** for a summary of the isoprenoid pathway). The Chlide- geranylgeranyl molecule is then sequentially reduced to phytol by the enzyme geranylgeranyl reductase, also known as ChlP.

Chlorophyll synthase, a single enzyme located in the thylakoid membrane, is able to esterify Chlide with either geranylgeranyldiphosphate or phytoldiphosphate; the substrate used is organism dependent. *A. thaliana* and *Synechocystis* can use geranylgeranyldiphosphate as a substrate, but have a preference for phytoldiphosphate, whereas *Rhodobacter capsulatus*

and *Rhodobacter sphaeroides* require geranylgeranyldiphosphate as a substrate (Rudiger *et al*., 1980; Soll *et al*., 1983; Oster and Rüdiger, 1997; Oster *et al*., 1997a; Addlesee *et al*., 2000). The tetrapyrrole moiety used is also organism-dependent. Chlorophyll synthase from chlorophyll-producing organisms may use Chlide *a* or Chlide *b* as a substrate, but not bacterio-Chlide, whilst the bacteriochlorophyll synthase from *Rhodobacter capsulatus* can use bacterio-Chlide as a substrate, but not Chlide (Benz and Rüdiger, 1981; Oster and Rüdiger, 1997; Oster *et al*., 1997a). Heterologous expression in *E. coli* of the *chlG* gene from *Synechocystis* or the *bchG* genes from *Rhodobacter sphaeroides* and *Rhodobacter capsulatus* was shown to produce an active (bacterio)chlorophyll synthase able to esterify (bacterio)Chlide with geranylgeranyldiphosphate *in vitro* (Oster *et al*., 1997b; Addlesee *et al*., 2000).

Initially it was not known if phytol was directly attached to Chlide or if a precursor was attached, then modified to phytol. In 1974, Liljenberg found low levels of Chlide-geranylgeranyl in developing horse chestnut leaves, a year later Ogawa and colleagues reported a short-lived intermediate between Chlide and chlorophyll; evidence which suggests a precursor is attached to Chlide and then later modified to phytol (Liljenberg, 1974; Ogawa *et al*., 1975). This hypothesis was further supported by Bollivar when he identified a *R. capsulatus* mutant disrupted in the gene *bchP* which accumulated bacteriochlorophyll esterifyed with geranylgeraniol, not phytol (Bollivar *et al*., 1994b). The electron transfer from the light-harvesting complex to the reaction centre was not affected in this mutant. However the mutant's growth rate was significantly retarded, possibly due to reduced stability of the pigment-protein complexes. *Rhodobacter sphaeroides* with a mutation in the *bchP* gene exhibits a significantly reduced level of light-harvesting LH2 complex and a red-shifted B850 absorbance maximum. However, all normal functions could be restored when the *bchP* mutant was complemented with its homologous counterpart, *chlP*, from *Synechocystis* (Addlesee *et al*., 1996). Recently Shpilyov and colleagues (2005) inactivated the *Synechocystis chlP* gene. As with *Rhodobacter capsulatus* the mutant was found to accumulate excessive levels of chlorophyll-geranylgeranyl.

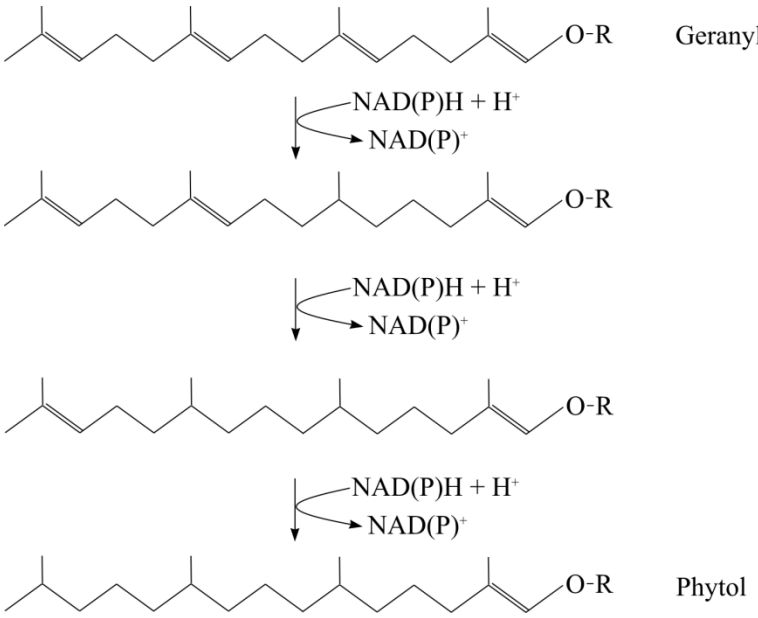

Figure 1.30 Terminal hydrogenation steps of chlorophyll biosynthesis

Sequential reduction of geranylgeranyl to phytol as performed by ChlP/BchP. R = chlorophyllide or pyrophosphate.

Surprisingly an increase in the total abundance of phycobilisomes was observed, perhaps up-regulated to counteract the decreasing levels of other light harvesting pigments (Shpilyov *et al*., 2005). The *ΔchlP* strain of *Synechocystis* was incapable of growing photoautotrophically and could only be cultivated with the aid of glucose as an external carbon source. The authors proposed the increased rigidity of the phytol moiety caused instability within the pigment-protein complexes, leading to photo-oxidative stress and rapid degradation of the photosystems (Shpilyov *et al*., 2005; Shpilyov *et al*., 2012) .

Shortly after esterification the geranylgeranyl-reductase catalyses three successive reductions across the C=C double bonds at positions C_6, C_{10} and C_{14} (**Figure 1.30**). The reaction requires NADH/NADPH and ATP, although in the absence of ATP, NADPH alone is able to hydrogenate Chlide-geranylgeranyl (Castelfranco and Beale, 1983). Schoch and colleagues observed etiolated oat seedlings grown under anaerobic conditions accumulated Chlide-geranylgeranyl, which suggests oxygen may be essential for hydrogenation and a possible explanation for the presence of bacteriochlorophyll-geranylgeranyl in certain species of anaerobic bacteria (Schoch *et al*., 1980). The *A. thaliana* geranylgeranyl-reductase is a multifunctional enzyme, capable of catalysing the stepwise reduction of geranylgeranyl-diphosphate into phytol-diphosphate as well as the reduction of Chlide-geranylgeranyl into Pchlide-geranylgeranyl (Keller *et al*., 1998). It is not yet known if the *Synechocystis* geranylgeranyl-reductase reduces geranylgeranyl to phytol before or after esterification to

Chlide. Moreover, no *in vitro* kinetic analyses have been reported for the geranylgeranyl-reductase, thus little is known about the enzyme's mechanism.

1.5 Aims

This project seeks to investigate the interactions between the known chlorophyll biosynthesis enzymes. To achieve this, a number of *Synechocystis* strains have been generated that contain a known chlorophyll biosynthesis enzyme with an N-terminal FLAG-tag. These strains, in conjunction with C-terminal FLAG-tagged chlorophyll biosynthesis enzymes that will be generated in this work, will form the initial basis for FLAG-pulldown affinity experiments.

Ultimately it is hoped these FLAG-pulldown experiments will further our understanding of:

- whether the chlorophyll biosynthesis enzymes form a membrane associated super complex
- identify the unknown components of the third committed enzyme in the chlorophyll biosynthesis pathway, Mg-protoporphyrin IX monomethylester cyclase, an enzyme which has remained an enigma for over 60 years
- identify as yet unknown proteins that may be involved in the efficient channelling of chlorophyll to the sites of photosynthesis in the thylakoid membranes
- identify as yet unknown proteins that are involved in inserting chlorophyll into the photosynthetic complexes and chlorophyll binding proteins

Additionally, this project aims to raise antibodies against several chlorophyll biosynthesis enzymes to which there are no antibodies currently available. This will enable:

- comprehensive investigations into the locations of the chlorophyll biosynthesis enzymes
- analysis of chlorophyll biosynthesis enzyme accumulation in photosynthetic mutants

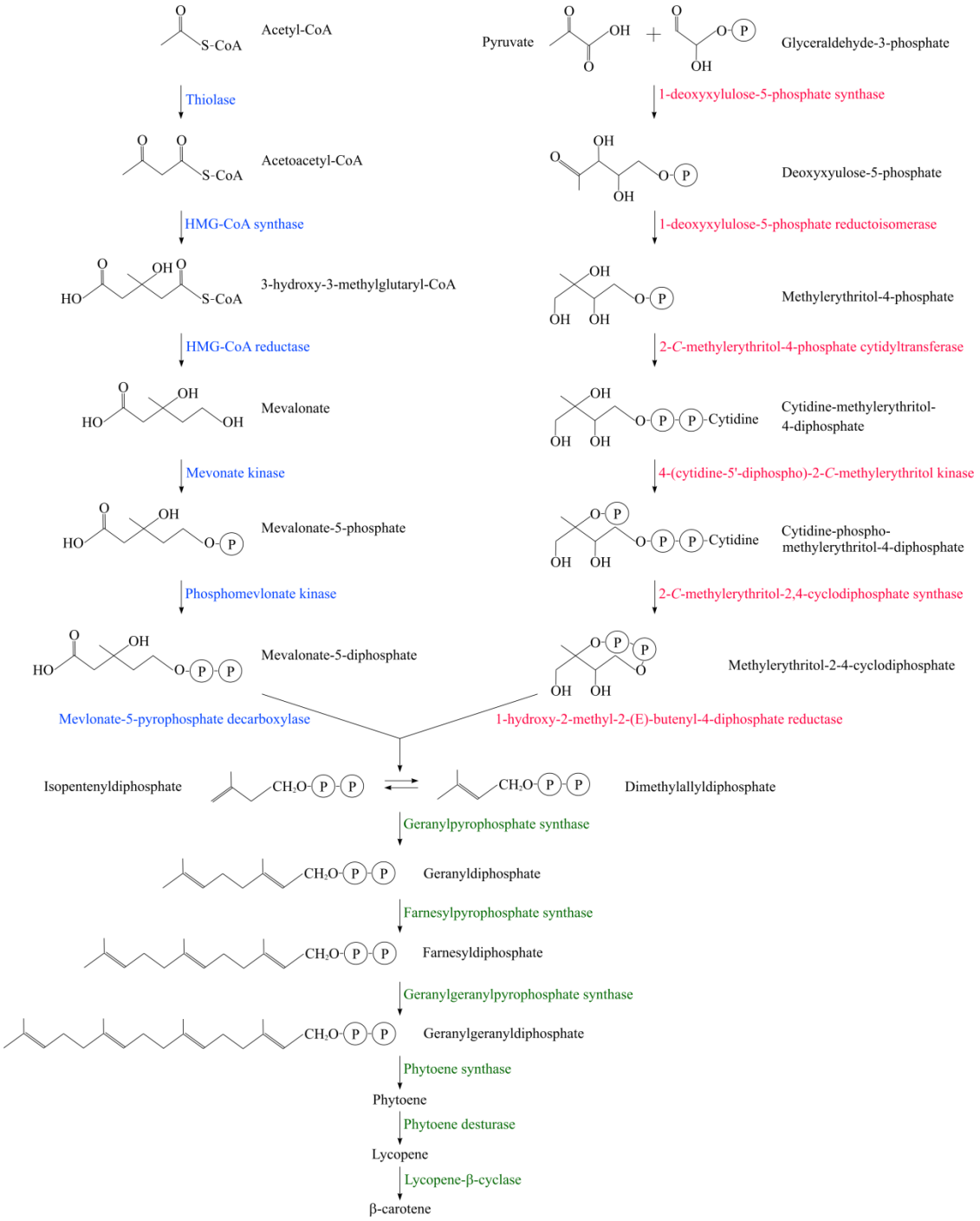

Figure 1.31 The isoprenoid pathway

Two routes for isopentyl diphosphate (IPP) occur. The mevalonate pathway (A) found in mammals and yeasts and the deoxy-xylulose-5-phosphate pathway (B) found in bacteria, algae and higher plants. The intermediate geranylgeranyl-diphosphate (GG-PP) is a substrate for the final stages of chlorophyll biosynthesis, when GG-PP is esterifyed to chlorophyllide and reduced to phytol to form the hydrophobic tail of chlorophyll. B-carotene also acts as a photo-pigment (Kajiwara et al., 1997; Lange et al., 2000).

2. Materials and methods

2.1 Standard buffers, reagents and media

All buffers and media were prepared using deionised Milli-Q water (Millipore). Growth media and solutions used for DNA work were sterilised by autoclaving at 15 *psi* and 121°C for a minimum of 20 minutes, or by filtration through 0.2 µm filters. Heat sensitive reagents, such as antibiotics and vitamins were added to medium cooled to below 50°C.

A list of all growth media used can be viewed in Table 2.1.

2.2 *E. coli* strains and plasmids

All *E. coli* stains were grown up in LB medium in the presence of any relevant antibiotics: kanamycin 50 µg ml^{-1}, neomycin 50 µg ml^{-1}, streptomycin 50 µg ml^{-1}, ampicillin 50 µg ml^{-1}, chloramphenicol 34 µg ml^{-1}. *E. coli* strains were stored at -80°C in 20 % (v/v) glycerol.

Plasmids used are listed in Table 2.2 and *E. coli* strains are listed in Table 2.3.

2.2.1 Chemically competent *E. coli* cells

JM109 chemically competent cells were obtained from Promega and stored as 25 µl aliquots at -80°C. All other chemically competent cells were made as follows.

Cell lines to be made competent were removed from storage at -80°C and streaked onto LB-agar plates containing the appropriate antibiotics and incubated at 37°C overnight. A single colony was then used to inoculate 5 ml of LB media, which was grown shaking overnight at 37°C and 250 rpm. 1 ml of the overnight culture was used to inoculate 50 ml of LB media in a 250 ml shake flask, which was incubated at 37°C and 250 rpm until an optical density at 600nm (OD_{600}) of 0.4 – 0.5 was reached. The cells were then pelleted in an ice-cold Falcon tube in a pre-chilled centrifuge at 4000 *xg* for 10 minutes. Following removal of the supernatant, the pellet was washed with 25 ml ice cold 0.1 M $MgCl_2$ solution and the cells re-pelleted by centrifuging at 4000 *xg* for 10 minutes. The supernatant was then decanted and the pellet washed with 25 ml ice cold 0.1 M $CaCl_2$ solution, before being re-pelleted and re-suspended in 1 ml ice-cold 0.1M $CaCl_2$, 20% glycerol solution. 50 µl of competent cell-suspension was then aliquotted into pre-chilled Eppendorfs, flash frozen in liquid nitrogen and stored at -80°C until required.

2.2.2 Transformation of chemically competent *E. coli* cells

An aliquot of competent *E. coli* cells was thawed on ice. Strains DH5α and JM109 were used for cloning transformations, whilst BL21 (DE3) pLysS, ROSETTA (DE3) PlysS and ROSETTA GAMI (DE3) PlysS were used for protein over-expression. ~100 ng of DNA was added to thawed competent cells. Cells were incubated on ice for 15 minutes, followed by heat shock at 42°C for 45 seconds. Cells were incubated on ice for a further 2 minutes, followed by re-suspension in 450 µl SOC medium heated to 37°C. The suspension was incubated at 37°C, 250 rpm for 1 hour, followed by micro-centrifugation at 3000 rpm for 2 minutes. The pellet formed was re-suspended in a smaller volume of LB (~100 µl) and was plated out on an agar plate containing the relevant antibiotic. Agar plates were incubated overnight at 37°C.

2.3 Synechocystis strains

All strains of *Synechocystis* sp. described in this thesis were derived from a glucose tolerant strain of *Synechocystis* sp. PCC6803 obtained from Professor Wim Vermass (Arizona State University) (Williams, 1988), and are detailed in Table 2.4. Strains were routinely grown in BG11 media, supplemented with glucose and TES KOH pH 8.2, grown under medium (40 µmol photons m^{-2} s^{-1}) or low light conditions (4 µmol photons m^{-2} s^{-1}) at 150 rpm and 30°C.

When antibiotic selection was required, the following concentrations were used: 30 µgml^{-1} kanamycin/ neomycin, 5 µg ml^{-1} zeocin, 15 µg ml^{-1} erythromycin and 34 µg ml^{-1} chloramphenicol.

Stocks of *Synechocystis* strains were stored at -80°C in 3:1 BG11 medium: 40 % DMSO (v/v).

2.3.1 Transformation of *Synechocystis* sp PCC6803

A shake flask containing 100 ml of Bg11 culture was inoculated with a scraping of *Synechocystis* cells from a Bg11 agar plate and incubated with gentle shaking until an OD_{750} = 0.6 – 0.7 was reached. 1 ml aliquots of cell culture were then transferred to a sterile eppendorf and the cells pelleted by centrifugation at 6,000 rpm for 5 minutes. Following removal of the supernatant, the cells were gently re-suspended in 100 µl of fresh Bg11 medium. 10 – 50 ng of plasmid DNA or linear DNA was then added to the cell suspension, and gently mixed by inversion. The cells were then incubated under medium light conditions for 30 minutes at 30°C. Following transformation, the cells were transferred to a Bg11 agar plate and the suspension was allowed to dry, before incubation at medium light intensity, overnight at 30°C.

Initial selection was performed by transferring the dried cell suspension to a Bg11 agar plate containing a low concentration of the appropriate antibiotic: 10 µg ml^{-1} kanamycin/ neomycin, 2 µg ml^{-1} zeocin, 5 µg ml^{-1} erythromycin or 10 µg ml^{-1} chloramphenicol. These agar plates were incubated under medium or low light conditions (depending upon the severity of the mutation introduced) until colonies appeared after approximately 8 – 12 days. Colonies were then transferred to new Bg11 agar plates, containing double the concentration of the original antibiotic and incubated as before. Like many cyanobacteria, *Synechocystis* contains many copies of its genomic DNA (the latest estimate being 60 copies per cell); therefore a period of segregation on exponentially increasing concentrations of antibiotic is required, to ensure the foreign DNA is introduced into every copy of the genome, to prevent desegregation upon removal of antibiotic selection. The transfer of colonies to double the concentration of antibiotic was continued until a fully segregated mutant was obtained, as confirmed by colony PCR.

2.3.2 Construction of *Synechocystis* knock-out mutants

Replacement of a *Synechocystis* gene with an antibiotic resistance cassette was achieved using a modification of the mega primer method of Key and Madison (1997). Sequences of ~500 bp up- and downstream of the gene of interest were amplified to generate primary megaprimers. The upstream reverse and downstream forward primers contained overhang sequences able to amplify a resistance cassette from a commercially available plasmid. The upstream megaprimer along with a reverse primer and the downstream megaprimer with a forward primer were used to amplify large over-lapping portions of the resistance cassette. These secondary megaprimers were then used for overlap extension PCR to generate a large DNA fragment able to integrate into the *Synechocystis* genome by homologous recombination after transformation.

2.4 Nucleic acid manipulation

2.4.1 Preparation of plasmid DNA

Mini-preps were performed using a QIAGEN Plasmid mini-prep kit following the manufacturer's instructions. Midi-preps were performed using a QIAGEN Plasmid midi-prep kit following the manufacturer's instructions.

2.4.2 Polymerase chain reaction (PCR)

Synechocystis colony PCR reactions were performed using 2x ACCUZYME mix (Bioline). A small scraping of *Synechocystis* cells was re-suspended in 10 µl QH_2O and 1 µl of this

suspension was added to 5 µl 2x ACCUZYME mix, 25 ng F primer, 25 ng R primer, 0.2 µl DMSO and made up to a final volume of 10 µl with QH_2O. PCR reactions were then carried out using an initial denaturing step of 4 minutes at 95°C, followed by 30 reaction cycles of (1 min at 96°C, 30 sec at 58°C and 2 min kb^{-1} at 72°C) and a final extension for 10 minutes at 72°C in a PHC-3 Thermal Cycler (Techne).

PCR reactions to obtain linear DNA for cloning were performed using 2x ACCUZYME mix (Bioline). 100 ng of genomic or plasmid DNA was mixed with 25 µl 2x ACCUZYME mix, 125 ng of F primer, 125 ng of R primer and 1 µl of DMSO. All reactions were topped up to a final volume of 50 µl using QH_2O. PCR reactions were carried out as above, for 25 reaction cycles.

Site-directed mutagenesis of DNA plasmids was performed using the QuickChange (Stratagene) method, according to the manufacturer's instructions.

Following amplification PCR reactions were either resolved by gel electrophoresis, or cleaned up using the PCR purification kit (QIAgen) according to the manufacturer's instructions.

A list of primers used in PCR reactions can be found in Table 2.5.

2.4.3 Restriction enzyme digests

All restriction enzymes were purchased from Promega or New England Biolabs. Restriction digests were carried out in a final volume of 20 µl containing 1 µl of each restriction enzyme, 2 µl 10x enzyme buffer, 1 µl BSA (1 mg ml^{-1}), 5 µl mini-prep DNA and made up to 20 µl with QH_2O. These were then incubated at 37°C for 1 hour.

2.4.4 Agarose gel electrophoresis of DNA

Analytical gels: restriction digests and PCR reactions were analysed by electrophoresis on a 1 % (w/v) agarose gel containing 0.5 µg ml^{-1} ethidium bromide, with Tris Acetate EDTA (TAE) running buffer (0.04M Tris acetate, 0.001M EDTA, pH 7.2). Samples were mixed with 6x loading buffer (Invitrogen) and loaded onto the gel (maximum volume of 30 µl was used). 5 µl of DNA marker (BIOLINE Hyperladder I) was run in tandem to estimate the size of DNA fragments. Gels were run at 80mV for 1 hour. DNA was visualised under 254 nm ultraviolet (UV) light.

2.4.5 Recovery of DNA from agarose gels

This was carried out using the QIAQuick (QIAGEN) gel extraction kit, following the manufacturer's instructions. Where the DNA being eluted was for cloning the elution volume used was 30 µl to produce a higher DNA concentration.

For small DNA fragments (< 300 bp) gel fragments were frozen in liquid nitrogen, then micro-centrifuged at 13,000 rpm for 5 minutes. Liquid separated from the gel was transferred to a fresh Eppendorf. DNA was precipitated by adding 1/10th volume of 3 M sodium acetate, pH 8.2 and 2 volumes of absolute ethanol and incubating at -20°C for two hours. DNA was pelleted by micro-centrifugation at 13,000 rpm for 5 minutes. DNA pellet was washed with 200 µl of 70% ethanol (v/v) and re-pelleted by micro-centrifugation at 13,000 rpm for 5 minutes. All of the ethanol solution was pipette off and the DNA pellet dried. DNA was re-suspended in 20 µl milli-Q water.

2.4.6 Ligation of DNA into vectors

Ligations were carried out in a final volume of 10 µl, comprising 1 µl of 10x ligation buffer, 100-200ng cut vector, 1 µl T4 DNA ligase (New England Biolabs) and varying ratios of purified cut DNA/vector (e.g., 3:1, 1:1). These were incubated at room temperature for 15 to 30 minutes. Ligations were then transformed into competent cells.

2.4.7 DNA sequencing

A 15 µl sample of mini-prepped plasmid DNA was mixed with 15 µl of QH_2O and sent to GATC-Biotech for sequencing. Returned results were analysed by CodonCode Aligner and CLUSTALW2.

2.4.8 Reverse transcription PCR

Total RNA was extracted from 25 ml of pelleted *Synechocystis* cells grown to an OD_{750} = 0.6 - 0.7. The cell pellet was re-suspended in 1 ml fresh BG11 media and mixed with an equal volume of RNA protect reagent (QIAgen). Following incubation and centrifugation, according to the manufactures instructions, the pellets were stored at -80°C until the RNA extraction protocol was performed using the RNeasy kit (QIAgen) according to the manufacturer's instructions. Any DNA bound to the spin columns was removed with RNAse-free DNAase (QIAgen). At the end of the protocol, the RNA was eluted from the spin columns using 50 µl of RNAse free water. The concentration of RNA was then determined by UV-VIS spectroscopy from the absorbance at A260 (for a sample of pure RNA in QH_2O, an absorbance of 1 unit at 260 nm corresponds to 40 µg ml^{-1} of RNA) and individual RT-PCR reactions were set up using RNA starting concentrations of 50 ng, 100 ng, 200 ng and 400 ng. RT-PCR reactions were performed using SuperScript® II Reverse Transcriptase (Invitrogen) in 20 µl reactions containing 50 – 400 ng RNA, 250 ng random primers (Invitrogen), 10 mM each of ATP, GTP, CTP and TTP, 4 µl 5x First-strand buffer, 2 µl 0.1 M DTT and 1 µl SuperScript II reverse

transcriptase. First stand cDNA synthesis was performed according to the manufacturer's instructions.

Following reverse transcription, 1 µl of cDNA was used as a template in PCR reactions with gene specific primers, using MyTaq DNA polymerase. Each PCR reaction volume was 20 µl and contained: 125 ng F primer, 125 ng R primer, 4 µl 4x reaction buffer, 1 µl cDNA template, 12.9 µl QH_2O and 0.1 µl MyTaq DNA polymerase. PCR reactions were performed with an initial 1 minute at 95°C, followed by 30 reaction cycles of, 15 seconds at 95°C, 15 seconds at 58°C and 15 seconds kb^{-1} at 72°C. Each PCR reaction was visualised by DNA gel electrophoresis.

2.5 Analysis of protein

2.5.1 Polyacrylamide gel electrophoresis (PAGE)

2.5.1.1 SDS-PAGE

To separate protein samples SDS-Polyacrylamide Gel Electrophoresis was used. 12 % Bis-Tris pre-packed gels were obtained from Invitrogen and run in 1x MES running buffer (Invitrogen, NP0002). 75 µl of sample were mixed with 25 µl 4x SDS gel loading buffer, heated to 100°C for 10 minutes and micro-centrifuged at 4000 *xg* for 5 minutes, 10 - 40 µl of sample was loaded onto the gel. Molecular weight markers were obtained from BioRad**.** The gels were run at 185 mV for 1 hour and the protein bands visualised by staining with Coomassie Magic stain or Silver stain.

2.5.1.2 Clear Native (CN)-PAGE

To separate protein complexes, clear native PAGE was used. 4 – 16 % non-denaturing gels were obtained from Invitrogen. Gels were run using the cathode buffer 0.02 % β-DDM, 0.05 % DOC, 50 mM Tricine and 7.5 mM imidazole pH 7.0 and the anode buffer 25 mM imidazole pH 7.0 at 4°C and a constant current of 10 mA overnight.

2.5.2 Western blot analysis of protein

Western blots were performed using the XCell II blot module (Invitrogen). The nitrocellulose membrane was washed with distilled water and allowed to equilibrate in Western Transfer Buffer (20 mM Tris, 150 mM glycine, 0.016 % SDS (V/V), 20 % methanol (V/V), pH 8.3) for 10 minutes with two sheets of Whatmann 3 MM filter paper and the blot pads. The blot module was prepared according to the manufacturer's instructions; two blot pads were placed at the bottom of the blot module followed by one sheet of filter paper, the SDS-PAGE gel, the nitrocellulose membrane, the second sheet of filter paper and four more blot pads. The blot

module was placed inside the power pack and the transfer run at 4°C for 60 minutes at 350 mA. The transfer membrane was removed from the cassette and placed in blocking buffer (5 % non-fat milk protein, 20 mM Tris-HCl, 0.05 % Tween20 pH 7.2). For more information see (Invitrogen, 2009)

2.5.3 Immunodetection

After Western blotting the PVDF membrane was placed in blocking buffer and incubated, rocking, at room temperature for 60 minutes or at 4°C overnight. The membrane was washed twice for 5 minutes in Western wash buffer (0.1 % Tween20, 20 mM Tris-HCl, pH 7.2). The membrane was then incubated with primary antibody, at the correct dilution, for 120 minutes in blocking buffer, at room temperature, whilst shaking, or overnight at 4°C. This was followed by two washes as before. The membrane was then incubated with secondary antibody at the required dilution in Western wash buffer, at room temperature, shaking, for 60 minutes. Finally the membrane was washed twice more in Western wash buffer. Immunodetection was performed using the ECL Detection System (Amersham), according to the manufacturer's instructions. The image was viewed using the GENEGNOME system.

2.5.4 Determination of protein concentration

The protein concentration of pure recombinant proteins was determined by absorbance at 280 nm. For protein concentrations in mg ml^{-1}, the following equation was applied to the absorbance readings:

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

where n_{Trp}, n_{Tyr} and n_{Cys} are the numbers of tryptophan, tyrosine and cysteine residues respectively, and Mr is the predicted molecular weight of the protein (Gill and Von Hippel, 1989).

Protein concentrations of mixed samples were determined spectroscopically using the equation described by Kalb and Bernlohr (1977), using the equation:

$$Protein\ concentration\ \mu g\ ml^{-1} = (183 \cdot A_{230}) - (75.8 \cdot A_{260})$$

2.5.5 Analysis of proteins by mass spectrometry: whole solution digest

2.5.5.1 In solution tryptic digest

25 µl of FLAG-pulldown eluate was added to 3.5 µl of each of 1 M triethylammonium bicarbonate (TEAB) pH 8.5 (Sigma), 10 % SDS (v/v) and 50 mM DTT and the sample was sonicated for 60 seconds, followed by incubation at 60°C for 30 minutes. After the sample was cooled to room temperature, 3.5 µl 100 mM iodoacetamide was added and the sample was incubated in the dark for 30 minutes. 311 µl of 50 mM TEAB was then added to the sample, followed by 4 µl of trypsin solution (Invitrogen) (1 g L^{-1} trypsin in 1 mM HCl). The mixture was then incubated at 37°C overnight, followed by drying down in a vacuum centrifuge overnight.

2.5.5.2 Peptide clean-up by SCX spin columns

SCX spin columns were prepared by adding 40 µl Poros20 SP slurry in 50 % methanol to the top of a barrier tip spin column and packed by centrifuging at 1000 rpm for 3 minutes. The column was then washed with two applications of 100 µl of cleaning buffer (25 % acetonitrile (v/v), 1 % 1 M potassium phosphate pH 3.0 (v/v) and 50 % 2 M potassium chloride (v/v)), followed by centrifugation at 1000 rpm for 3 minutes. The column was then equilibrated with two applications of 100 µl of loading buffer (25 % acetonitrile (v/v) and 1 % 1 M potassium phosphate pH 3.0 (v/v)), followed by centrifugation at 1000 rpm for 3 minutes.

The tryptic peptides prepared in section 2.5.5.1 were re-dissolved in 200 µl loading buffer, by sonicating for 5 minutes. 7 µl of 0.5 M phosphoric acid was then added to the sample (sample pH should be between pH 2.5 – 3.0) and the sample was centrifuged at 13,000 rpm for 10 minutes. The tryptic peptides were then applied to the equilibrated spin column in 2x 100 µl aliquots, each followed by centrifugation at 1000 rpm for 3 minutes.

The SCX column was then washed with 3x 100 µl applications of loading buffer and the tryptic peptides were eluted into a clean low-bind Eppendorf tube with 80 µl elution buffer (25 % acetonitrile (v/v), 1 % 1 M potassium phosphate pH 3.0 (v/v) and 25 % 2 M potassium chloride (v/v)). The eluate containing the tryptic peptides was then dried down in the vacuum centrifuge overnight.

2.5.5.3 Peptide clean-up by C_{18} spin columns

C_{18} SpinTip columns (Proteabio) were used to desalt the SCX eluted tryptic peptides prepared in section 2.5.5.2. The dried peptides were prepared by re-suspending them in 100 µl reconstitute and rinse solution and sonicating for 5 minutes. This was followed by centrifugation at 13,000 rpm for 10 minutes. The sample was then applied to the equilibrated C_{18} SpinTip columns according to the manufacturer's instructions. Following C_{18}

chromatography the eluate containing the tryptic peptides was dried down overnight in low-bind Eppendorfs in a vacuum centrifuge.

2.5.5.4 NanoLC-MS/MS analysis

Following C_{18} desalting (section 2.5.5.3), the dried tryptic peptides were re-dissolved in 7 µL 0.1% (v/v) TFA and 3% (v/v) acetonitrile. 5 µL was analysed by LC-MS/MS using an Ultimate 3000 RSLCnano liquid chromatography system (Dionex) with 5 mm x 300 µm trapping and 75 µm x 15 cm analytical PepMap C_{18} reverse-phase columns. Tryptic peptide elution was by a 60-min linear gradient from 94% solvent A (0.1% (v/v) formic acid) to 40% solvent B (0.1% (v/v), 80% (v/v) acetonitrile) at a flow rate of 300 nl/min. Mass spectra were acquired online using a Maxis UHR-TOF instrument (Bruker Daltonics, Bremen, Germany) operating in in-line format with automated dependent MS/MS scans. Peak lists for database searching, in the form of Mascot Generic Files (MGFs) were created from the datafiles using a processing script supplied by Bruker. The MGFs were submitted for database searching using Mascot Daemon v. 2.2.0 running with Mascot Server v. 2.2.01 (Matrix Science, London, UK) against the *Synechocystis* complete proteome database (http://www.uniprot.org/uniprot/?query=organism%3a1111708+keyword%3a1185&format=fasta). Search parameters were set to allow for one missed cleavage with S-carbamidomethyl-cysteine as a fixed modification and oxidised methionine as a variable modification. Tolerances were ±0.5 Da for both peptide (MS) and product (MS/MS) ions.

2.5.6 Analysis of proteins by mass spectrometry: in-gel digestion

2.5.6.1 In-gel tryptic digest

Protein sample of interest was separated by SDS-PAGE (see section 2.5.1). Following Coomassie staining, the gel was soaked in keratin away solution (0.1 % acetic acid, 10 % isopropanol) for 30 minutes, to remove any contaminates from the surface of the gel. Gel bands of interest were then excised and the volume of each band estimated in µl. The bands were then diced into 1 mm x 1 mm pieces and placed in pre-cleaned Eppendorf tubes. The gel pieces were then incubated with 100 µl 50 mM ammonium bicarbonate for 5 minutes at room temperature. 100 µl acetonitrile was added and the gel pieces incubated for a further 10 minutes at room temperature. The solvent was then discarded and the ammonium bicarbonate/ acetonitrile step repeated 2-3 more times until the gel pieces became colourless.

50 µl 10 mM DTT was then added to the gel pieces and the sample was incubated at 56°C for 30 minutes. Following reduction, the DTT was discarded and the gel pieces were incubated with 50 µl acetonitrile for 10 minutes at room temperature. The solvent was then discarded

and the gel pieces were incubated with 50 µl 55 mM iodoacetamide in 50 mM ammonium bicarbonate in the dark at room temperature for 20 minutes. Following alkylation, the iodoacetamide reagent was discarded and the sample incubated with 100 µl 50 mM ammonium bicarbonate for 10 minutes. 100 µl acetonitrile was then added and the sample was incubated for a further 10 minutes at room temperature. The incubation step with ammonium bicarbonate/ acetonitrile was then repeated and the gel pieces dried down in a vacuum centrifuge.

The dried gel pieces were then rehydrated with 20 µl digestion buffer (12.5 ng μl^{-1} Trypsin (Invitrogen), 5 mM calcium chloride and 50 mM ammonium bicarbonate) and incubated on ice for 20 minutes. At this point more digestion buffer was added if required and the sample was incubated for a further 45 minutes on ice. Following incubation, excess digestion buffer was removed, and a sufficient volume of 50 mM ammonium bicarbonate with 5 mM calcium chloride was added to cover the gel pieces. The sample was then digested overnight at 37°C.

After digestion, the buffer surrounding the gel pieces was transferred to a low-bind Eppendorf tube on ice. The gel pieces were then mixed with 15 µl 25 mM ammonium bicarbonate and incubated at room temperature for 10 minutes. The buffer surrounding the gel pieces was transferred to the low-bind Eppendorf and 2x the estimated gel volume of acetonitrile was added to the gel pieces and incubated at 37°C for 15 minutes. This buffer was also removed from the gel pieces and transferred to the low-bind Eppendorf. 50 µl of 5 % formic acid was then added to the gel pieces and the sample was incubated at 37°C for 15 minutes. The buffer surrounding the gel pieces was then transferred to the low-bind Eppendorf. The pooled extracts were then dried in the vacuum centrifuge overnight and stored at -20°C.

2.5.6.2 NanoLC-MS/MS analysis

The dried tryptic peptides prepared in section 2.5.6.1 were re-dissolved in 7 µl 0.1 % TFA and 3 % acetonitrile (v/v). 5 µl of sample was then loaded onto the nanoLC-MS/MS using an Ultimate 3000 RSLCnano liquid chromatography system (Dionex, Camberley, UK) with 5 mm x 300 µm trapping and 75 µm x 15 cm analytical PepMap C18 reverse-phase columns. Tryptic peptide elution was by a 40-min linear gradient from 94% solvent A (0.1% (v/v) formic acid) to 40% solvent B (0.1% (v/v), 80% (v/v) acetonitrile) at a flow rate of 300 nL/min. Mass spectra were acquired online using a Maxis UHR-TOF instrument (Bruker Daltonics, Bremen, Germany) operating in line format with automated dependent MS/MS scans. Peak lists for database searching, in the form of Mascot Generic Files (MGFs) were created from the datafiles using a processing script supplied by Bruker. The MGFs were submitted for

database searching using Mascot Daemon v. 2.2.0 running with Mascot Server v. 2.2.01 (Matrix Science, London, UK) against the Synechocystis complete proteome database (http://www.uniprot.org/uniprot/?query=organism%3a1111708+keyword%3a1185&format=fasta). Search parameters were set to allow for one missed cleavage with S-carbamidomethyl-cysteine as a fixed modification and oxidised methionine as a variable modification. Tolerances were ±0.5 Da for both peptide (MS) and product (MS/MS) ions.

2.5.7 Analysis of protein complexes by sucrose density gradient ultra centrifugation

Cultures of *Synechocystis* in the logarithmic growth phase were harvested by centrifugation at 15,000 × *g* for 10 min. The pellet was re-suspended in three volumes of thylakoid buffer (50 mM HEPES/NaOH, pH 7.0, 5 mM $MgCl_2$, 25 mM $CaCl_2$, 10% glycerol (v/v)) and the cells were disrupted in a bead beater using 0.1 mm glass beads. Glass beads and unbroken cells were removed by centrifugation for 5 min at 8,000 × *g* and the supernatant containing the thylakoid membranes was ultra-centrifuged for 60 min at 150,000 × *g*. Thylakoid membranes were re-suspended in thylakoid buffer to a concentration of 10 mg ml^{-1} of protein.

The thylakoid membranes were solubilised with 1 % β-DDM in the dark using a homogeniser for 5 minutes. Non-solubilised material was removed by centrifugation at 30,000 × *g* for 30 min. The supernatant was loaded onto a 12 ml continuous sucrose density gradient (10–30% (w/v) sucrose) in thylakoid buffer without glycerol but containing 0.04 % (w/v) β-DDM. Ultracentrifugation was performed in a SW28 rotor for 16 h at 150,000 × *g* and 4°C.

2.6 Protein expression and purification

2.6.1 Protein expression in *E. coli*

The plasmid containing the gene for over-expression of the desired protein was transformed into the desired *E. coli* strain (see Table 2.3 for a list of *E. coli* strains). A single colony was used to inoculate 20 ml of LB containing the appropriate antibiotics and grown at 37°C overnight. This starter culture was used to inoculate 500 ml LB + antibiotic and incubated at the desired temperature until the desired OD_{600} for induction was achieved (see Table 2.6 for a list of over-expression conditions for each recombinant protein). Cells were then induced with isopropyl-beta-D-thiogalactopyranoside to a final concentration of 0.04 mM. The cells

were either induced for two hours or overnight, shaking at 250 rpm, incubated at desired induction temperature. Cells were harvested by centrifugation at 4,200 *xg* for 20 minutes, supernatant was discarded and the pellet stored at -20°C.

2.6.2 Lysis of *E. coli* cells using sonication

Pelleted cells are re-suspended in the appropriate re-suspension buffer (see Table 2.6 for details) containing EDTA-free protease inhibitor, DNAase and lysozyme. The lysis mixture was incubated on rocker at room temperature for 20 minutes and cells were lysed by sonication (30 seconds on, 30 seconds off) for 5 – 6 minutes. Cell lysate was then clarified by centrifugation at 18,500 rpm for 30 minutes.

2.6.3 Purification of His_6-tagged proteins

2.6.3.1 Preparation of the nickel affinity column

Fast flow chelating Sepharose (GE Healthcare) was packed into an Econo-Pac® gravity flow column (BioRad) to the desired bed height with 2 column volumes (CV) of QH_2O. The column was then equilibrated with 2 CV of 100 mM $NiSO_4$ and washed with 3 CV of binding buffer (50 mM Tris pH7.4, 0.5 M NaCl, 10 mM Imidazole), ready for binding.

2.6.3.2 Nickel-affinity chromatography

Cell lysate (pre-filtered through a 0.45 µm filter) was loaded onto the column, and the flowthrough re-applied once. The column was then washed with His-binding buffer until no protein could be detected in the eluate (BioRad reagent), followed by His-Wash I buffer (50 mM Tris pH 7.4, 0.5 M NaCl, 50 mM Imidazole), also until no protein could be detected in the eluate. The column was then washed with 1 CV of His-Wash II buffer (50 mM Tris pH 7.4, 0.5 M NaCl, 100 mM Imidazole) and the protein eluted with His-elution buffer (50 mM Tris pH 7.4, 0.5 M NaCl, 400 mM Imidazole), with the elute collected in 2ml fractions.

Each fraction was analysed for the presence of the desired protein by SDS-PAGE. Fractions containing the protein of interest were pooled and either frozen at -80°C or further processed.

2.6.3.3 Removal of the His_6 tag

Recombinant His-tagged protein purified as in section 2.6.3.2, was buffer exchanged into PBS using a PD-10 desalting column (GE Healthcare). 80 U of thrombin (GE Healthcare) was then added to $\leq$ 30 mg recombinant protein and allowed to digest overnight, on an end-to-end rocker at room temperature. The recombinant protein, with the His_6-tag removed by digestion, was then passed back through a nickel affinity column (equilibrated as in section

2.6.3.1), which was washed with 1 CV of His-binding buffer. The recombinant protein, minus the His_6 tag was collected in the flowthrough.

2.6.4 Purification of soluble GST-tagged proteins

2.6.4.1 Preparation of the GST affinity column

Glutathione Sepharose 4 fast flow resin (GE Healthcare) is packed and equilibrated in a Poly-Prep® gravity flow column (BioRad) to the desired bed height (normally 1 ml) with 10 CV of PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3).

2.6.4.2 Preparation of the benzamidine affinity column

Benzamidine Sepharose 4 fast flow resin (Amersham Biosciences) is packed and equilibrated in a Poly-Prep® gravity flow column (BioRad) to the desired bed height (normally 1 ml) with 10 CV of QH_2O. The column was then equilibrated with 10 CV of thrombin binding buffer (20 mM sodium phosphate buffer, 0.15 M NaCl, pH 7.4).

2.6.4.4 GST-affinity chromatography

Cell lysate (pre-filtered through a 0.45 µm filter) was loaded onto the column prepared as detailed in section 2.6.4.1, and the flowthrough re-applied once. The column was then washed with PBS until no protein could be detected in the eluate (BioRad reagent). The resin was allowed to run until almost dry and the column stopped. The resin was then re-suspended in 1 ml of PBS, transferred to a 7 ml bijou containing 80 U Thrombin (GE Healthcare) and allowed to digest overnight, on an end-to-end rocker at room temperature. The recombinant protein, minus the GST affinity tag was then washed off the GST resin with PBS and the flow through collected.

2.6.4.5 Benzamidine-affinity chromatography

The flowthrough from the GST-affinity column (from section 2.6.4.4) was loaded to a Benzamidine column, equilibrated as detailed in section 2.6.4.2. The column was then washed with thrombin binding buffer until no protein was detected in the flowthrough by BioRad reagent. Purified recombinant protein was then eluted with up to 10 CV of thrombin elution buffer (20 mM sodium phosphate buffer, 1.0 M NaCl, 0.04 % β-DDM, pH 7.4) and the elute collected in 1 ml fractions. Fractions containing the desired recombinant protein, as determined by SDS-PAGE analysis, were pooled, concentrated and stored at -80°C in buffer containing 10 % glycerol (v/v).

2.6.5 Purification of insoluble GST-tagged proteins

Insoluble proteins were purified as detailed in section 2.6.4 with the addition of 0.04 % β-DDM (v/v) in all buffers used.

Following clarification of the cell lysate after sonication (section 2.6.2), the pellet obtained is retained for solubilisation. This was re-suspended in a minimal volume of PBS (normally 5 ml PBS for a pellet generated from 1 L original cell culture) and transferred to 25 ml glass beaker containing a magnetic stirrer bar. 20 % β-DDM was then added very slowly, whilst sample is continually mixed, to a final concentration of 2 % (v/v). Solubilisation was allowed to proceed for 1 hour, before the solubilised pellet was clarified by centrifugation at 30,000 × *g* for 20 minutes to remove any remaining unsolubilised material. The resulting supernatant was diluted with PBS to reduce the β-DDM concentration from 2 % to 0.4 % and loaded onto the equilibrated GST column (as detailed in section 2.6.4.4).

2.6.6 Protein expression in *Synechocystis*

Genes encoding proteins to be expressed in *Synechocystis* were cloned into the pFLAG vector between regions of DNA homologous to those found up- and downstream of one of the *Synechocystis* genes encoding the photosystem II D1 protein (*slr1311*). These vectors were transformed into *Synechocystis* as detailed in section 2.3.1. The resulting transformants expressing FLAG-tagged proteins were grown photoautotrophically in 1 L, 8 L or 20 L cultures of BG11 media at 30°C. Protein synthesis was initiated by growth under high light conditions (100 µM photons m^{-2} s^{-1}). Cultures were grown to an OD_{750} = 0.8 – 1.0 before being harvested by centrifugation at 4,000 rpm for 1 hour at 4°C. Cell pellets were stored at -80°C until required.

2.6.7 Breakage of *Synechocystis* cells by bead beating

Cell pellets containing over-expressed FLAG-tagged proteins (produced as detailed in section 2.6.6) were washed and re-suspended in a minimal volume of FLAG-buffer (25 mM sodium phosphate buffer, 10 mM $MgCl_2$, 10 % glycerol (w/v), 0.05 M NaCl, pH 7.4) containing EDTA free protease inhibitor (BioCompare), lysozyme and DNase. This cell suspension was then transferred to 6 ml bijous half filled with 0.1 mm glass beads, to create a 1:1 ratio of cell suspension to glass beads.

From this point on, all steps were performed in the dark, or under dim green light.

Cells were broken by a series of eight bead-beating cycles, consisting of 55 seconds of breakage, followed by 5 minutes cooling on ice. Following breakage the cell lysate was removed from the beads by washing with 2 volumes of FLAG-buffer. The cell lysate was then centrifuged at 3,000 rpm for 15 minutes to remove any unbroken cells.

2.6.8 Separation of thylakoid membrane and soluble cell fractions

The clarified cell lysate obtained in section 2.6.7 was separated by centrifugation at 64,000 *xg* for 45 minutes. The soluble fraction (blue) was then decanted from the insoluble fraction.

The thylakoid membrane pellet was then transferred to a new container using a small soft paint brush and re-suspended in a minimal volume of FLAG-buffer. These were then solubilised with a β-DDM at a final concentration of 1.5 % (v/v), as described for insoluble proteins in section 2.6.5. The resulting solubilised supernatant was either loaded directly onto a FLAG-affinity column or frozen in liquid N_2 stored at -80°C until required.

2.6.9 Purification of FLAG-tagged proteins

Note all solubilised material was purified using buffers containing 0.04 % β-DDM.

2.6.9.1 Preparation of FLAG-affinity resin

ANTI-FLAG-M2-Agarose from Mouse (Sigma) was flow packed in a Poly-Prep® gravity flow column (BioRad) to the desired bed volume (usually 100 – 300 µl) with 1 ml QH_2O. The resin was then equilibrated with 5 ml FLAG buffer.

2.6.9.2 FLAG-affinity purification

Solubilised *Synechocystis* membrane and soluble fractions were loaded onto separate FLAG columns and the flow-through re-applied once. The column was then washed with 10 x 1 ml applications of FLAG-buffer.

Following washing, the anti-FLAG resin was re-suspended in 300 µl wash buffer, containing 150 µg ml^{-1} FLAG peptide (sequence: DTKDDDDKDTKDDDDKDTKDDDDK) (Invitrogen), transferred to an Eppendorf and incubated on a rocker for 1 hour at 4°C. The elute was then separated from the FLAG-resin by passage through a Costar Spin-X column containing a 0.22 µM cellulose acetate membrane, frozen in liquid nitrogen and stored at -80°C.

2.7 Pigment extraction

2.7.1 Extraction of chlorophyll precursors

Standardised samples of *Synechocystis* (OD_{750}) were pelleted by centrifugation at 7,000 *xg* for 5 minutes. Each pellet was washed with water and re-pelleted in a 2 ml screw-cap micro-centrifuge tube at 7,000 *xg* for 5 minutes. The supernatant was discarded and a volume of methanol + 0.2 % ammonium (100 – 500 µl) was added to the cell pellet. This was mixed in the bead beater for 15 seconds and incubated on ice for 15 minutes. The cell debris was pelleted by centrifugation at 15,000 *xg* for 5 minutes and the supernatant transferred to a

new micro-centrifuge tube. The extraction was repeated on the pellet, with half the original volume of methanol + 0.2 % 35 % ammonium (Fisher) (v/v). The two supernatants were pooled.

Hydrophobic components (chlorophylls and carotenoids) were removed by two extractions with hexane. A half volume of hexane was added to the methanol extract, mixed by vortexing for 10 seconds and incubated on ice for 5 minutes to allow for phase separation. The upper hexane phase was discarded.

Precursor extracts were either analysed immediately, or dried down in a vacuum centrifuge overnight and stored at -20°C until required. All steps were performed in the dark.

2.7.2 Determination of chlorophyll concentration

Standardised samples of *Synechocystis* (OD_{750}) were pelleted by centrifugation and the pigments extracted by methanol + 0.2 % ammonium (v/v) as described in section 2.7.1. Samples were not extracted with hexane, which removes chlorophyll. The concentration of chlorophyll was then determined using the method of Porra et al., (1989), represented in Equation 2.1, below.

Equation 2.1 Calculation of chlorophyll *a* concentration

$$[\mathrm{Chl}a]\ m\mathrm{g\ L^{-1}} = (16.29 \cdot \mathrm{A}_{665}) - (8.54 \cdot \mathrm{A}_{652})$$

2.7.3 Large scale purification of the mystery chlorophyll pigment

It is important to note the mystery pigment sticks to plastic, therefore all steps must be carried out using glassware (including using glass pipettes) to prevent a dramatic loss of yield.

1L cultures of *Synechocystis* strain *Δslr1780* and R82A were grown in BG11 media in glass flasks to an OD_{750} = 1.5, which allowed for a significant amount of pigment to accumulate. Pigments accumulating in the supernatant were then extracted twice by phase partitioning by adding a ¼ volume of diethyl ether. The diethyl ether extractions were then pooled and dried down under vacuum in a rotary evaporator. The dried pigment was re-suspended in a minimal volume of methanol using a glass pipette and transferred to a glass HPLC sample vial. The vial was then placed in a black Eppendorf and pigment was dried down in a vacuum centrifuge overnight at 30°C.

The mystery pigment was then further purified using a preparative HPLC column (UniverSil C_{18}) and the chlorophyll precursors program (described in Section 2.8 and Table 2.7), with the HPLC modified with the addition of a 0.5 ml sample injection loop. The HPLC program was set

to collect fractions every 15 seconds to ensure the mystery pigment could be collected in highly pure fractions. The fractions containing the mystery pigment were pooled and diluted with an equal volume of QH_2O. The pigments were then loaded onto a C18 SEP-PAK column (Sigma), washed with 1 ml 100 % MeOH and equilibrated with 1 ml 50:50 MeOH and QH_2O. Following loading, any ammonium acetate remaining from the HPLC buffer was removed by washing with 2 ml QH_2O. The concentrated pigment was then eluted in 100 % MeOH, transferred to a glass HPLC sample vial and dried down in a vacuum centrifuge overnight at 30°C. The dried pigment was stored at -20°C until required.

2.8 High performance liquid chromatography

Analytical HPLC was performed on an Agilent 1200 high-performance liquid chromatograph. All buffers were filtered through a 0.22 µm membrane and degassed before use.

Details of the HPLC programs and the molecules separated on each can be found in Table 2.7.

Porphyrin elution was detected using a multichannel diode array detector (DAD) (Agilent) set to record absorbance's as follows: 400 nm (protoporphyrin), 416 nm (Mg-protoporphyrin IX & MgPME), 433 nm (Mystery pigments), 440 nm (Pchlide, Chlide & Chlorophyll), 632 nm (Pchlide) and 665 nm (Chlide). Porphyrin elution was also monitored with a fluorescence detector (Agilent) using an excitation wavelength of 440 nm and monitoring emission at 632 nm.

2.9 Spectroscopy

2.9.1 Room temperature absorbance spectroscopy

Room temperature absorbance spectra were recorded on a Cary 50 UV-VIS spectrophotometer. Whole cell spectra were taken between 350 nm and 750 nm and base-line correction was performed. Samples were diluted as appropriate so that the spectrophotometer readings were within the linear range.

2.9.2 77 K low temperature fluorescence spectroscopy

All whole cell emission spectra were recorded in solutions containing 80 % glycerol (v/v). UV-VS fluorescence spectroscopy was performed in a SPEX Fluorolog spectrofluorometer (SPEX Industries Inc.) with a xenon light source. Samples were re-suspended to an OD_{750} = 0.1 and cooled to 77 K in an OpstistatDN nitrogen bath cryostat (Oxford Instruments, Oxford, UK). The temperature of the sample was monitored with a thermocouple sensor (Comark,

Stevenage, UK). Emission spectra were recorded from cells excited at 435 nm and 580 nm with 5 nm slit widths, scanning between 450 – 900 nm and 600 – 900 nm, respectively, with an integration time of 1 second. For each experiment, a three scan average was recorded.

2.9 Protein crystallography and data processing

For all crystallisation trials protein samples were concentrated to 10 mg ml^{-1} in 100 mM $NaHCO_3$, 50 mM NaCl, pH 8.0. Crystals were obtained by the vapour diffusion method and all crystallisation trays were incubated at 17°C. Initial screens were performed using a Hydra II robot, which set down sitting drops that contained 200 nl protein sample and 200 nl crystallisation solution. These trials were performed using the broad QIAgen NeXtal screening suites PACT, JCSG, Classics, Ammonium Sulphate, PEG and MPD.

Crystals selected for Diamond were looped in a solution of the mother liquor that contained 50 % ethylene glycol and transported in a cryostat. In all cases, test images were taken at angles of 0° and 90° and the diffraction patterns were analysed by the Mosflm auto indexing routine. Based on this information, a data collection strategy was implemented; two data sets, each of 900 diffraction images taken with a rotation of 0.2 degrees were collected for each crystal. Following data collection, MOSFLM was used to index the data, providing an estimate of the mosaicity, Laue group and unit cell; the data collected was processed at Diamond using Xia2 (Winter, 2009).

Data was processed using the CCP4i suite of packages (1994). Initial indexing, cell refinement and integration were carried out using iMOSFLM. MATTHEWS was used for solvent content calculation (Matthews, 1968).These data were scaled and merged using SCALA (Evans, 2006), before molecular replacement was carried out using PHASER and ARP/wARP Classic. REFMAC5 (Murshudov et al., 1997) was used for refinement; model building and structure validation was carried out in COOT (Chen et al., 2010).

2.10 Nuclear magnetic resonance spectroscopy on chlorophyll precursor pigments

Pigment NMR experiments were conducted using a 600 MHz Brucker spectrometer. Experiments were done at room temperature, using a cryoprobe to improve sensitivity. Pigments were re-suspended in deuterated MeOH to a final volume of 0.5 ml and transferred to an opaque sample vial. All data was processed in Felix 2007 using in-house macros, which can be found at /1d10/home1/dave/felix970/macs. All experiments were processed with a

90° sine bell window function over the first 1024 points and a 90° sine bell window function over all the nitrogen points.

Table 2.1 Growth media

Growth Media	Reagents for 1 L media
LB/ LB agar and auto-induction LB	Ready mixed LB media from FORMEDIUM, prepared following the manufacturer's instructions. Auto-induction LB grows IPTG-inducible expression strains without induction to a high cell density, and then induces the production of the target protein automatically. Inclusion of a small amount of glucose and lactose in the media regulates expression from IPTG inducible promoters.
Super LB	40 % yeast extract, 20 % tryptone and 5 % sodium chloride
SOC	0.5 % (w/v) bacto yeast extract, 2 % (w/v) bacto tryptone, 10 mM NaCl, 25 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM Glucose, adjusted to pH7.0
BG11	**BG-11 stock solutions:** Trace Minerals: 2.86 g boric acid, 1.81 g manganese chloride, 0.22 g zinc sulphate, 0.39 g sodium molybdate, 0.079 g copper sulphate, 0.049 g cobaltous nitrate. 100X BG-11: 149.6g sodium nitrate, 7.49 g magnesium sulphate, 3.6 g calcium chloride, 0.6 g citric acid, 0.56 ml of 0.5M stock pH 8.0 EDTA (disodium salt), 100 ml trace minerals. 1000X Iron stock: 6 g ferric ammonium citrate 1000X Phosphate stock: 30.5 g dipotassium hydrophosphate 1000X Carbonate stock: 20 g sodium carbonate 1M Glucose stock: 180 g glucose 1M TES stock: 229.2 g TES pH 8.2 (KOH) **BG11 liquid media** 10 ml 100X BG-11, 1 ml 100x Iron, 1 ml 1000x Phosphate, 1 ml 1000x Carbonate, 10 mM TES, 5 mM glucose. **BG11 agar** 10 ml 100X BG-11, 1 ml 100x Iron, 1 ml 1000x Phosphate, 1 ml 1000x Carbonate, 10 mM TES, 5 mM glucose, 15 g Bactoagar.

Table 2.2 Plasmids constructed

Plasmid	Resistance	Manufacturer/ Source	Constructs
pET-14b	Amp^R	Invitrogen	pET14b-POR pET14b-Ycf54 pET14b-A9G pET14b-D39A pET14b-F40A pET14b-R82A pET14b-CPO III
pGEX-4T1	Amp^R	GE-Healthcare	pGEX-Sll1214 pGEX-Sll1874 pGEX-ChlP pGEX-DVR pGEX-Pitt
pFLAG (N-terminal FLAG)	Amp^R Kan^R	Dr Paul Davison, Dr Dan Canniffe.	pFLAG-Pitt pFLAG-POR pFLAG-Ycf54 pFLAG-Sll1214 pFLAG-Sll1874 pFLAG-ChlP pFLAG-ChlM pFLAG-DVR pFLAG-Ycf54-A9G pFLAG- Ycf54-D39A pFLAG- Ycf54-F40A pFLAG- Ycf54-R82A pFLAG-slr1702 pFLAG-slr0483
pFLAG (C-terminal FLAG)	Amp^R Kan^R	Dr Roman Sobotka.	pFLAG-POR pFLAG-Ycf54 pFLAG-Sll1214 pFLAG-Sll1874 pFLAG-ChlP pFLAG-ChlM pFLAG-DVR
pET3a	Amp^R	Invitrogen	pET3a-Ycf54 point mutations (A9G, F13A, E22A, E26A, D39A, F40A, R82A)
pET32b	Amp^R	Invitrogen	pET32b-Sll1214 pET32b-Sll1874

Table 2.3 *E. coli* strains

Strain	Properties	Source
DH5α	*fhu*A2 *lac*(del)U169 *phoA glnV*44 Φ80' *lacZ*(del)M15 *gyrA*96 *recA1 relA1 endA1 thi-1 hsdR17*	Novagen
BL32 (DE3) pLysS	F–, *ompT, hsdSB* (rB^-, mB^-), *dcm, gal*, λ(DE3), p*LysS*, Cam^R	Novagen
ROSETTA (DE3) pLysS	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm*(DE3) pRARE (Cam^R)	Merck Millipore
ROSETTAGAMI (DE3) pLysS	Δ(*ara-leu*)7697 Δ*lacX74* Δ*phoA PvuII phoR araD139 ahpC galE galK rpsL* (DE3) F'[lac^+ $lacI^q$ pro]*gor522*:Tn10 *trxB* pLysSRARE2 (Cam^R, Str^R, Tet^R)	Merck Millipore
JM109	*endA1, recA1, gyrA96, thi, hsdR17* (rk–, mk+), *relA1, supE44,* Δ(*lac-proAB*), [F', *traD36, proAB, laqIqZ*ΔM15]	Promega

Table 2.4 *Synechocystis* strains

Strain	Properties	Source
Glucose tolerant *Synechocystis*	Generated from wild type *Synechocystis* sp PCC6803	Prof. Wim Vermaas, Arizona State University (Williams, 1988)
N-terminal FLAG tagged constructs: ChlM Sll1214 Sll1874 Ycf54 POR DVR ChlP Slr0483 Pitt Ycf54-A9G Ycf54-D39A Ycf54-F40A Ycf54-R82A.	Mutant of wild type, in which the *slr1311* gene is replaced with DNA encoding an N-terminally 3x FLAG-tagged gene and a Nm^R cassette	This thesis
C-terminal FLAG tagged constructs: ChlM Sll1214 Sll1874 Ycf54 POR DVR ChlP	Mutant of wild type, in which the *slr1311* gene is replaced with DNA encoding an C-terminally 3xFLAG-tagged gene and a Nm^R cassette	This thesis
Δslr1780	Mutant of wild type, in which the *slr1780* gene is replaced with a zeocin resistance cassette.	This thesis
Δsll1751	Mutant of wild type, in which the *sll1751* gene is replaced with a zeocin resistance cassette.	This thesis
Δslr1702	Mutant of wild type, in which the *slr1702* gene is replaced with a zeocin resistance cassette.	This thesis
Δslr1135	Mutant of wild type, in which the *sll1135* gene is replaced with a zeocin resistance cassette.	This thesis
Δslr1906	Mutant of wild type, in which the *slr1906* gene is replaced with a zeocin resistance cassette.	This thesis
Ycf54 point mutants: Ycf54-A9G Ycf54-F13A Ycf54-E22A Ycf54-E26A Ycf54-D39A Ycf54-F40A Ycf54-R82A	Mutant of wild type, in which the *slr1780* gene is replaced with *slr1780* gene containing a single point mutation and a Cm^R cassette	This thesis

Table 2.5 PCR primers

Primer	Sequence
C-terminal FLAG-DVR R:	GCGCTAGCTTGCTGGGGAAGTTTATACTGCTCAAC
QC remove NheI DVR F:	CCAGCCGAAGTGCTCGCCGCCAAGGTTAAT
QC remove NheI DVR R:	ATTAACCTTGGCGGCGAGCACTTCGGCTGG
C-terminal FLAG-ChlP R:	ATTGCTAGCAGGGGCTAAAGCGTTACCCCG
C-terminal FLAG-ChlM R:	ATTGCTAGCAGAGCGCACCGCCTCTAAAATACG
C-terminal FLAG-Sll1874 R:	ATTGCTAGCACACACCATCCCCCGACGGGC
C-terminal FLAG-ycf54 R:	TTAGCTAGCATCCAGGGATGCAAGGGGG
C-terminal FLAG-1214 R:	TTAGCTAGCGCGCACAGCTCCAGCCAAC
Sll1214 (NcoI) F:	GCCCATGGTTAATACCCTCG
Sll1874 (NcoI) F:	GCCCATGGTATCCACT
Sll1874 (XhoI) R:	GCCTCGAGTTAACACACCATCCC
COP III F (NdeI):	ATACATATGACCGTCTCTCCC
FLAG-slr0483 NotI F:	ATAAGAATGCGGCCGCAAACATTCAATCC
FLAG-slr0483 BglII R:	GCGAGATCTCTAACCGCCAAAAATTTGC
Sll1135 KO Rz1:	acattaattgcgttgcgctcactgcCCAACACGGTGACCTCC
Sll1135 KO Fz2:	caacttaatcgccttgcagcacatGGCTTCCCACTTGCATC
Slr0483 KO Rz1:	acattaattgcgttgcgctcactgcCCATCTTCAACCAGTCC
Slr0483 KO Fz2:	caacttaatcgccttgcagcacatGGCATGGGCTACACC
Slr1702 KO Rz1:	acattaattgcgttgcgctcactgcGGATACGACGGTAGCC
Slr1702 KO Fz2:	caacttaatcgccttgcagcacatCGGAATTGCCTGAAAAGCC
Sll1214 (XhoI) R:	GCCTCGAGTTAGCGCACAGCTCC
COPIII R (XhoI):	CGCTCGAGCTAACTATTAACCC
POR HindIII QC F:	ATACGGAGCTAAGGCCTTAATTGAC
POR HindIII QC R:	GTCAATTAAAGCCTTAGCTCCGTAT
Sll1751 KO F:	GGTTCAGTGGAAGTGGG
Sll1751 KO R:	GCAAATTTATGCCGCC
Sll1751 KO 1zF:	CCACAAGTGTGACCTTTGTTTCC
Sll1751 KO 2zR:	acattaattgcgttgcgctcactgcGGAATAAATTCTTGATTAAGC
Sll1751 KO 3zF:	caacttaatcgccttgcagcacatGATCTAATCTTAGAAAGTTTAGG
Sll1751 KO 4z:	GCAGTAAACTAAGCAGTTTAACCATGC

Primer	Sequence
FtsZ F:	CCTGCAATGACGCTCAATAATG
FtsZ R:	CTAACGGCGGGGAAAACG
Ala9Gly F:	CCTATTATTATGCTTTGGGAAGTCAAAAATTTCTCTTGG
Ala9Gly R:	CCAAGAGAAATTTTTGACTTCCCAAAGCATAATAATAGG
Phe13Ala F:	GGCAAGTCAAAAAGCTCTCTTGGAAGAGG
Phe13Ala R:	CCTCTTCCAAGAGAGCTTTTTGACTTGCC
Glu22Ala F:	GGAACCCTTTGAAGCGGTGCTGAAAGAACG
Glu22Ala R:	CGTTCTTTCAGCACCGCTTCAAAGGGTTCC
Glu26Ala F:	GGTGCTGAAAGCACGGCGACGGG
Glu26Ala R:	CCCGTCGCCGTGCTTTCAGCACC
Asp39Ala F:	ATAAGGAAATTGCCTTTTGGCAGGTG
Asp39Ala R:	CACCTGCCAAAAGGCAATTTCCTTAT
Phe40Ala F:	GGAAATTGACGCTTGGCAGGTG
Phe40Ala R:	CACCTGCCAAGCGTCAATTTCC
Arg82Ala F:	GGGTAAAACTCGCGTTGGAATATGTC
Arg82Ala R:	GACATATTCCAACGCGAGTTTTACCC
Ycf54 Upstream F1:	GGGAGATCTGGGGCTAAGGATCAACGCTGG
Ycf54 Upstream R1:	GGGTCTAGATTACTAATCCAGGGATGCAAGGGGG
ChlmR F2:	GGGTCTAGACATCATCGTTGATCGGCACGTAAGAGG
ChlmR R2:	CCCCATATGACTAGTTACGCCCCGCCCTGCCACTC
Ycf54 Downstream F3:	GGC ACTAGTGTTCAATAACCATGGCGATCG
Ycf54 Downstream R3:	GGGGGATCCGGGGATTTAGAAGTGATTCTC
QC XbaI Pitt F Primer:	GGGGGCACACCTCTGGAAAGTCAGGCGTTGG
QC XbaI Pitt R Primer:	CCAACGCCTGACTTTCCAGAGGTGTGCCCCC
Slr1780 KO 1zF:	GGAATTGGGGTCCCTGGCGATCG
Slr1780 KO 2zR:	acattaattgcgttgcgctcactgcCTTCAAAGGGTTCCTCTTCC
Slr1780 KO 3zF:	caacttaatcgccttgcagcacatGCTCCCTCCGATGCCATTCC
Slr1780 KO 4zR:	CCAGCACCTATGGTACCTTCAGC
Sll1906 KO 1zF:	CGGATCAAACTGGCGATCGCC
Sll1906 KO 2zR:	acattaattgcgttgcgctcactgcCCCAGACGGAACATAGTCACC
Sll1906 KO 3zF:	caacttaatcgccttgcagcacatCCTGCTCAACAAAGTCAATGTGC

Primer	Sequence
Sll1906 KO 4zR:	GCATGATGAAATCCCTGGCTCGGG
YCF54-GEX F2:	GCCGGAATTCATGGAAAGTTGGGCATTGAC
L&L pET-Ycf54 F:	GCCGGCGGCCGCGAAAGTTGGGCATTGAC
L&L pET-Ycf54 R:	GCCGCCATGGATGGAAAGTTGGGCATTGAC
pGEX-YCF54 F:	GCCGGAATTCATGGCTACCTATTATTATGC
pGEX-YCF54 R:	CCGGCGGCCGCCTAATCCAGGGATGCAAGGG
pFLAG-YCF54 F:	GCCGGCGGCCGCTACCTATTATTATGC
pFLAG-YCF54 R:	CCGAGATCTCTAATCCAGGGATGCAAGGG

Table 2.6 Plasmids constructed

Plasmid	Cell line	Over-expression conditions	Re-suspension buffer
pET14b-POR	ROSETTA	Grown at 37°C in TY media to an OD_{600} = 1.0 – 1.2 Induced with IPTG for 3 – 4 hours at 20°C	50 mM Tris, 0.5 M NaCl, 10 mM Imidazole, pH 7.4
pET14b-Ycf54 pET14b- Ycf54-A9G pET14b- Ycf54-R82A	BL21 (DE3) pLysS	Grown at 28°C in Super LB overnight with no induction.	50 mM Tris, 0.5 M NaCl, 10 mM Imidazole, pH 7.4
pET14b-D39A pET14b-F40A	BL21 (DE3) pLysS	Grown at 37°C in LB to an OD_{600} = 0.6 – 0.8 Induced with IPTG overnight at 20°C	50 mM Tris, 0.5 M NaCl, 10 mM Imidazole, pH 7.4
pET14b-CPO III	BL21 (DE3) pLysS	Grown at 37°C to an OD_{600} = 0.6 – 0.8 Induced with IPTG overnight at 20°C	50 mM Tris, 0.5 M NaCl, 10 mM Imidazole, pH 7.4
pET32b-Sll1214 pET32b-Sll1874	ROSETTA GAMI (DE3) pLysS	Grown at 37°C to an OD_{600} = 1.0 – 1.2 Induced with IPTG for 3 – 4 hours at 20°C	50 mM Tris, 0.5 M NaCl, 10 mM Imidazole, pH 7.4
pGEX-Sll1214 pGEX-Sll1874 pGEX-ChlP pGEX-Ycf54	BL21 (DE3) pLysS	Grown at 25°C in auto-induction media overnight.	PBS
pGEX-Pitt	ROSETTA	Grown at 37°C in LB to an OD_{600} = 0.6 – 0.8 Induced with IPTG for 3 – 4 hours at 20°C	PBS

Table 2.7 HPLC programs

Molecules separated	Column used	Buffers used	HPLC program
Chlorophyll precursors	150 mm x 4.6 mm C18 5 µm NovaPak (Waters), flowrate = 0.9 ml min^{-1}. 250 mm C18 x 4.6 mm 5 µm (Phenominex), flowrate = 1.0 ml min^{-1}. 150 mm x 10 mm C18 5 µm Universil (Fortis), flowrate = 3.5 ml min^{-1}.	Buffer A: 350 mM ammonium acetate pH 6.9 Buffer B: 100 % methanol	0-35 minutes, linear gradient from 65 % to 75 % buffer B. 35-45 minutes linear gradient from 75 % - 100 % buffer B. 45-50 minutes 100 % buffer B. 50-60 minutes 65 % buffer B.

3. Investigating Protein-Protein Interactions in the Chlorophyll Biosynthesis Pathway

3.1 Introduction

Higher plants and cyanobacteria contain at least three classes of tetrapyrrole molecules: chlorophylls, phycobilins and haems. All of these pigments strongly absorb light and upon illumination the free pigments may act as photosensitisers, which can lead to pigment bleaching and cell death (Rebeiz et al., 2014). To minimise oxidising damage inside the cell, these pigments are normally found bound to proteins, where their light energy is either dissipated by use in cellular process, such as light harvesting and charge separation, or safely dissipated by non-photochemical quenching mechanisms (Muller et al., 2001). However, the biosynthetic precursors of these pigments are thought to be freely available in the cell, where they have a greater potency for photo-oxidative damage (Matringe et al., 1989c; Matringe et al., 1989d; Mock et al., 1998). Therefore it is not unreasonable to hypothesise that photosynthetic organisms may have developed substrate channelling to enable the efficient passage of photosynthetic intermediates though the chlorophyll, haem or phycobilins biosynthesis pathways to their final destination as mature pigments. An additional argument for substrate channelling is that the chemical composition of the chlorophyll precursors means these molecules are innately hydrophobic and in danger of forming enzyme-inaccessible aggregates if released directly into aqueous solution. Thus, if the chlorophyll biosynthesis enzymes formed a large membrane-associated super-complex, hydrophobic substrates could be channelled from protein to protein until the mature chlorophylls are inserted into the protein complexes where they are required; this would prevent the aggregation of the precursors, reduce photo-oxidative damage and facilitate the efficiency of the chlorophyll biosynthesis pathway.

There are several existing examples of super-complexes that form in order to aid substrate channelling and facilitate pigment biosynthesis. For example, carotenoid biosynthesis is known to occur in a multi-enzyme complex containing the phytoene synthase, phytoene desaturase, lycopene-β-cyclase and β-carotene hydroxylase enzymes (Cunningham and Gantt, 1998; Monaco et al., 2001; Lopez et al., 2008). Like chlorophylls, carotenoids are hydrophobic molecules and substrate channelling between enzyme active sites is thought to aid the efficiency of their synthesis. In the initial stages of tetrapyrrole biosynthesis, two of the three enzymes responsible for ALA formation, GluTR and GSAT, have been found to form a functional complex that allows the channelling of GSA, a highly reactive intermediate,

between the two enzymes (Nogaj and Beale, 2005b). Further down the tetrapyrrole biosynthesis pathway, at the branch point of chlorophyll and haem biosynthesis, a complex is known to form between PPO and FeCH; here the channelling of protoporphyrin IX between the two enzymes is thought to protect the cell by minimising the production of reactive-oxygen species (Masoumi et al., 2008). Another interaction has been identified at the beginning of the committed chlorophyll biosynthesis pathway between the ChlH subunit of MgCH and ChlM. Here, the ChlH interaction with ChlM was found to stimulate methyltransferase activity (Shepherd et al., 2005b).

In *Arabidopsis*, a membrane associated complex composed of FLU and the chlorophyll biosynthesis enzymes CHL27 (cyclase), POR and ChlP has been reported (Kauss et al., 2012a). The FLU protein, which does not have a homologous counterpart in *Synechocystis*, is a membrane associated regulatory protein that has a C-terminal tetratricopeptide repeat (TPR) domain. FLU is thought to act in a metabolic feedback loop that mediates GluTR activity according to the flux of the chlorophyll branch of the tetrapyrrole pathway (Meskauskiene et al., 2001a; Meskauskiene and Apel, 2002; Kauss et al., 2012b). Interestingly, in *Synechocystis*, POR has been identified as forming a complex with Pitt (*slr1644*), an integral membrane protein, which, like FLU, contains a TPR domain (Sato et al., 2007; Schottkowski et al., 2009). Pitt is thought to be localised in a specialised fraction of *Synechocystis* membranes, known as the PDM or PratA-defined membrane, where early steps in the biogenesis of PSII are thought to occur (Rengstl et al., 2011). It would be interesting to see if Pitt also co-localises with ChlP, the cyclase or any other components of the chlorophyll biosynthesis apparatus. Although FLU and Pitt share no homology, parallels can be drawn between the two; they both associate with the photosynthetic membranes and contain a C-terminal TPR domain.

Evidence from the literature suggests it would not be unreasonable to hypothesise that the enzymes of the chlorophyll biosynthesis interact with each other to form a large membrane associated multi-enzyme complex. Therefore, the aim of this chapter is to investigate whether the chlorophyll biosynthesis enzymes interact with each other and/or enzymes higher up the tetrapyrrole biosynthesis pathway. This project has focused on the use of *in vivo* and *in vitro* pulldown assays to ascertain if the enzymes of the chlorophyll biosynthesis pathway interact. Many of the genes involved in the committed chlorophyll biosynthesis pathway in *Synechocystis* have been FLAG-tagged *in vivo* and used as bait in pulldown experiments; this *in vivo* pulldown approach was used to determine if the enzymes in the chlorophyll biosynthesis pathway form a multi-enzyme complex.

During the course of these experiments, an unknown hypothetical protein, Slr0483, was identified as a potential candidate with a role in chlorophyll biosynthesis. mRNA transcripts of a *Chlamydomonas reinhardtii* homologue of *slr0483* were found to be significantly up-regulated in a *Chlamydomonas reinhardtii* GUN4 mutant (Formighieri et al., 2012), and Slr0483 was additionally found to be pulled down by FLAG-tagged FeCH, ChlM, Sll1214 and ChlP (data reported in this chapter). This chapter reports the initial steps taken to identify what role Slr0483 may play in photosynthesis.

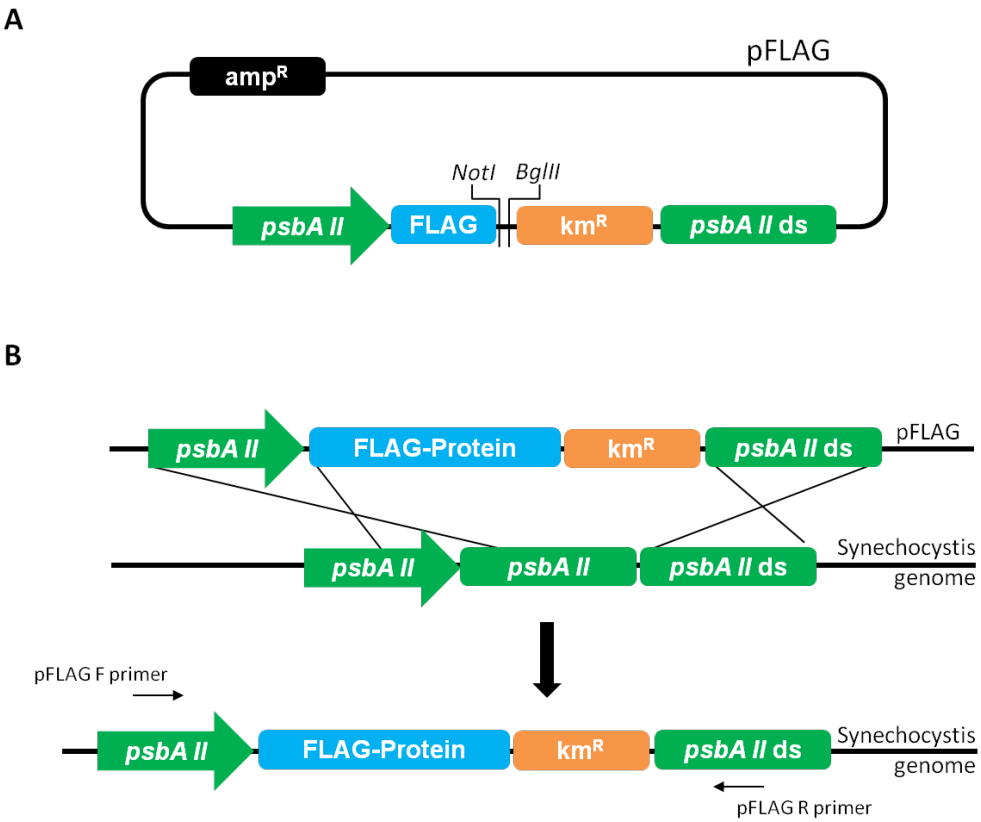

Figure 3.1 Method of FLAG-tagging proteins in *Synechocystis*

pFLAG vector contains two regions of homology (~500bp) to the *psbAII* promoter region (green arrow) and the psbAII II gene (green rectangle), situated between these is a 3xFLAG tag, a kanamycin resistance cassette and a NotI/BglII cloning site for the insertion of a gene of choice (A). Upon transformation into Synechocystis the two regions of homology align with their compatriot regions in the Synechocystis genome. A homologous recombination event occurs, inserting the FLAG-tagged protein of interest into the Synechocystis genome (B). pFLAG F primer and pFLAG R primer, represent the PCR primers used to amplify the *psbAII* region to confirm full segregation of the FLAG-construct.

3.2 Results

3.2.1 Construction of FLAG-tagged chlorophyll biosynthesis enzymes in *Synechocystis*

A pFLAG vector was constructed by Dr Daniel Canniffe, which enables the integration of a FLAG-tagged recombinant gene at the locus of the light activated promoter, *psbAII*, in *Synechocystis* (**Figure 3.1**). All N-terminal FLAG-tagged constructs were made by amplifying the recombinant gene of interest, ChlM, Sll1214, Sll1874, POR, ChlP or Pitt, with primers that introduced a *5'* NotI site and a *3'* BglII site. These gene products were then ligated downstream of a 3xFLAG coding region; clones were confirmed by restriction digests and sequencing analysis. N-terminal FLAG constructs ChlM, Sll1214, Sll1874, POR and ChlP were constructed and fully segregated *Synechocystis* strains were obtained by Dr Daniel Canniffe.

A C-terminal pFLAG vector, identical to the N-terminal FLAG vector, was kindly gifted by Dr Roman Sobotka (University of South Bohemia); this vector enables the C-terminal FLAG tagging of the recombinant protein of interest. All C-terminal FLAG-tagged constructs were made by amplifying the recombinant gene of interest, ChlM, Ycf54, Sll1214, Sll1874, DVR, ChlP, with primers that introduced a *5'* NdeI site and a *3'* NheI site. These gene products were ligated upstream of a 3xFLAG coding region; clones were confirmed by restriction digests and sequence analysis.

Synechocystis contains approximately 60 genome copies per cell; therefore to create a strain where a FLAG-tagged gene of interest is unable to revert to wild type, a fully segregated mutant must be obtained. This is achieved by increasing the antibiotic selection pressure, until all copies of the genome contain a copy of the *in vivo* FLAG-tagged protein. Initial transformants were selected on BG11 agar plates supplemented with a low concentration (5 mg ml^{-1}) of kanamycin, incubated under normal light conditions (30 µmol photons m^{-2} s^{-1}) at 30°C. Single colonies were transferred to fresh plates containing double the concentration of antibiotic, this transfer of colonies to plates supplemented with double the previous concentration of antibiotic was continued until a fully segregated strain was isolated. To confirm segregation, colony PCR was performed using primers pFLAG F and pFLAG R, designed to amplify the *psbAII* locus (**Figure 3.2**).

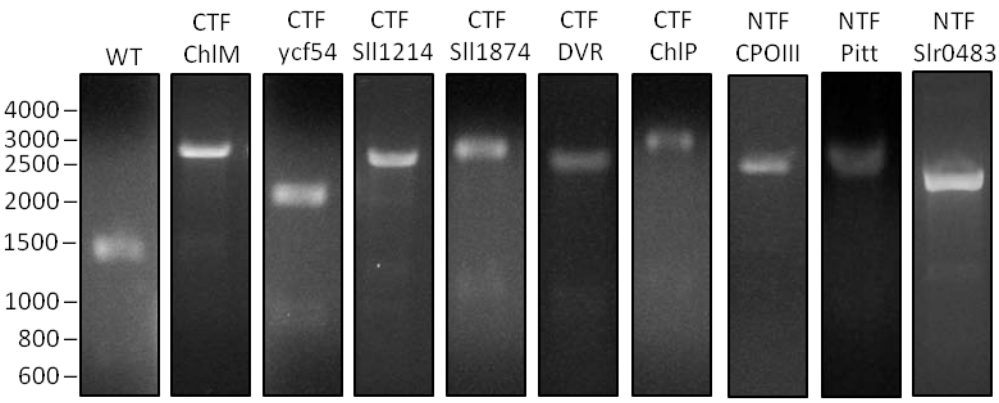

Figure 3.2 Colony PCR analyses to confirm complete segregation of N- and C-terminal FLAG tagged proteins in *Synechocystis*

PCR analysis to show complete segregation of C-terminal FLAG-tagged (CTF) proteins ChlM, Ycf54, Sll1214, Sll1874, DVR and ChlP and N-terminal FLAG-tagged (NTF) proteins CPOIII, Pitt and Slr0483. Primers were used to amplify the psbAII locus to confirm the complete segregation of the FLAG-tagged strains. Full segregation is shown by the absence of PCR product in the WT strain and the presence of a new, higher molecular weight PCR product in the mutant strain. PCR reactions were resolved on a 1 % agarose gel, visualised with ethidium bromide. WT, 1490 bp; CTF-ChlM, 2750 bp; CTF-Ycf54, 2360 bp; CTF-Sll1214, 3111 bp; CTF-Sll1874 3111 bp, CTF-DVR 3120 bp; CTF-ChlP 3140 bp; NTF-CPOIII, 3110 bp, NTF-Pitt, 2872 bp ; NTF-Slr0483 .

3.2.2 Production of antibodies to detect *Synechocystis* POR, ChlP, Pitt and Ycf54

1 mg ml^{-1} of purified protein is required for the production of a rabbit primary antibody (BioServe, Sheffield, UK). This protein was produced by N-terminally tagging ChlP, Pitt, POR and Ycf54 with GST, which was achieved by cloning the gene of interest into the pGEX-4T-1 vector. All four genes were obtained by amplifying the gene of interest (ChlP, Pitt, POR or Ycf54) from *Synechocystis* genomic DNA, with primers that introduced a *5'* EcoRI site and a *3'* XhoI site. Following digestion with the appropriate restriction enzymes, the genes were ligated into pGEX-4T-1 and the resulting constructs confirmed by restriction digests and sequence analysis (data not shown).

To determine the optimal expression conditions for the GST-tagged constructs, expression trials were performed in two different *E. coli* cell lines, BL21 (DE3) pLysS or the ROSETTA (DE3) pLysS, at two different induction temperatures (20°C and 25°C). SDS-PAGE analysis showed GST-Pitt and GST-POR were more highly expressed in the ROSETTA (DE3) pLysS cell line at 20°C and GST-ChlP and GST-Ycf54 expressed the best in the BL21 (DE3) pLysS cell line at 20°C and 25°C respectively (data not shown).

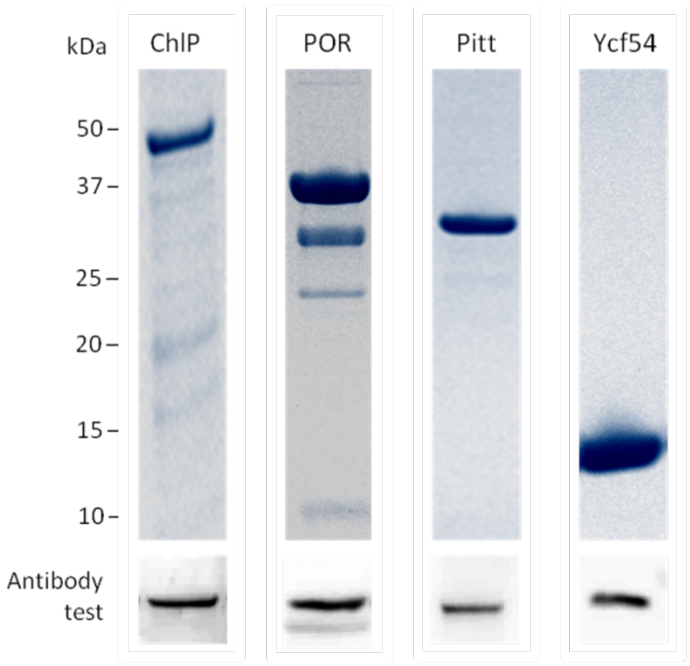

Figure 3.3 Purification of ChlP, POR, Pitt and Ycf54 for primary antibody production and analysis of primary antibody

SDS-PAGE analysis of the purified recombinant proteins used for production of rabbit primary antibodies and Western blot analysis of the primary antibody produced, probed against wild type *Synechocystis* whole cell lysate. ChlP (*sll1091*) is 42 kDa, Pitt (*slr1644*) is 32 kDa, POR (*slr0506*) is 37 kDa and Ycf54 (*slr1780*) is 12 kDa.

Purification of GST-tagged proteins requires purification over two columns. The GST-tagged protein is first separated from the cell lysate by a GST-affinity column (thrombin cleavage separates the protein of interest from the GST-tag), and the thrombin is removed by passage over a benzamidine ion-exchange column (purifies the thrombin used for cleavage away from the protein of interest). Details of the methods used can be found in Section 2.6.4, with the exception of Pitt. It was found the Pitt protein did not bind to the benzamidine under the buffer conditions recommended by the manufacturer, therefore modified buffer conditions, with reduced NaCl concentrations, were used. Each round of purification was performed using cell pellets harvested from 6 litres of over-expressed cell culture, with GST-POR and GST-Ycf54 purified from the soluble fraction and GST-ChlP and GST-Pitt purified from the β-DDM solubilised insoluble fraction (see section 2.6.4 for details).

Several contaminants remained after benzamidine chromatography; therefore all proteins were further polished by purification over an S200 size exclusion column. The size exclusion elution fractions containing the proteins of interest were pooled and concentrated to 1mg ml^{-1}, ready for antibody production. An SDS-PAGE gel of the purified proteins is shown in **Figure 3.3**. Rabbit primary antibodies were produced by BioServe (Sheffield, UK).

All four antibodies were tested for their specificity by probing against a Western blot of wild type *Synechocystis* whole cell lysate. As shown in **Figure 3.3**, all four antibodies are reactive to the denatured *Synechocystis* protein. These antibodies are now able to be used in immuno-precipitation experiments with *Synechocystis* cell lysates, to investigate interactions in the chlorophyll biosynthesis pathway (not included in this thesis). Additionally, the antibodies will be used to investigate the phenotypes of various *Synechocystis* mutants.

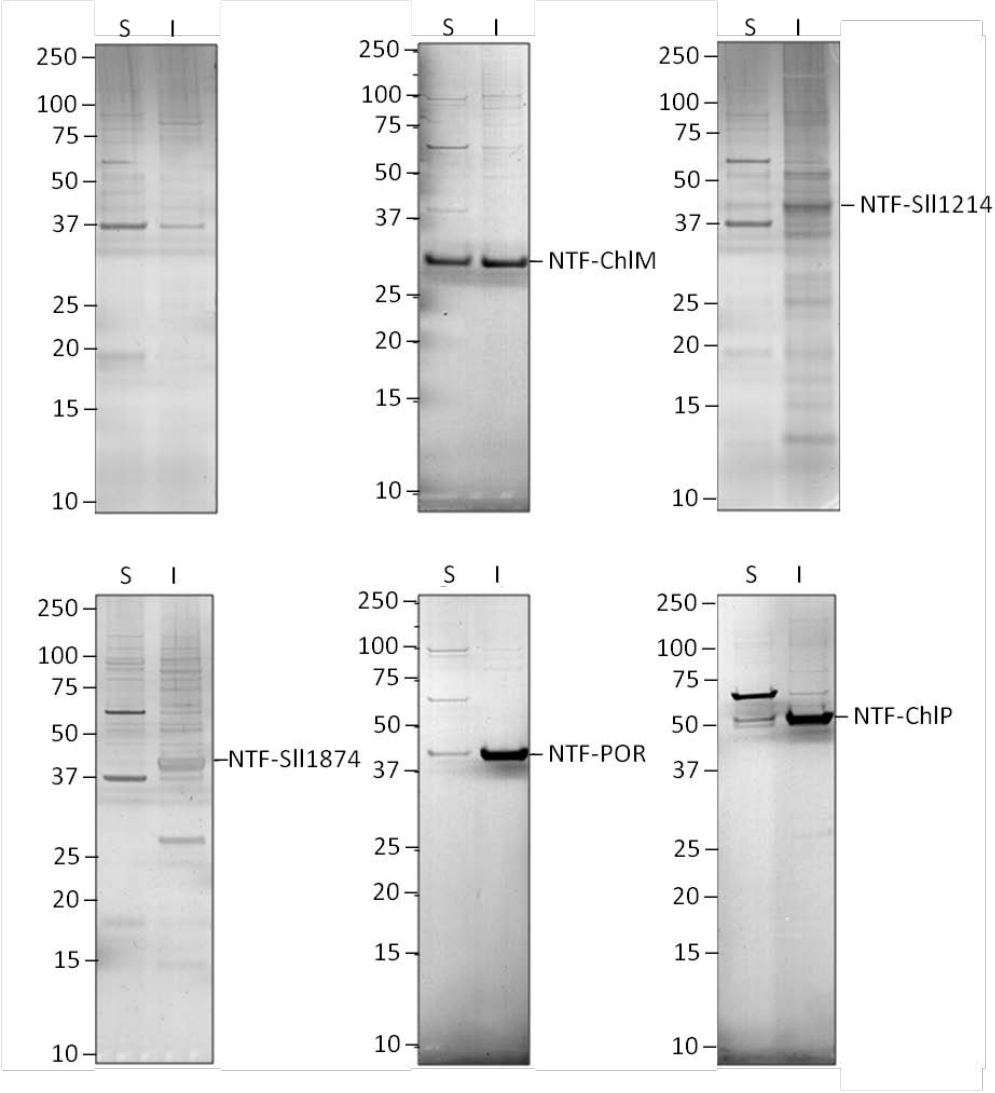

Figure 3.4 Silver stained SDS-PAGE analyses of elution fractions from N-terminal FLAG-affinity pulldowns

FLAG-affinity purification of *Synechocystis* wild type control and N-terminal FLAG-tagged ChlM (28 kDa), Sll1214 (45 kDa), Sll1874 (45 kDa), POR (39 kDa) and ChlP (48 kDa). FLAG-tagged proteins were purified from *Synechocystis* soluble cytoplasmic (S) and β-DDM solubilised thylakoid (I) fractions using anti-FLAG M2 affinity gel. Captured proteins were eluted using anti-FLAG peptide, separated by SDS-PAGE and silver stained. The molecular marker is shown in kDa.

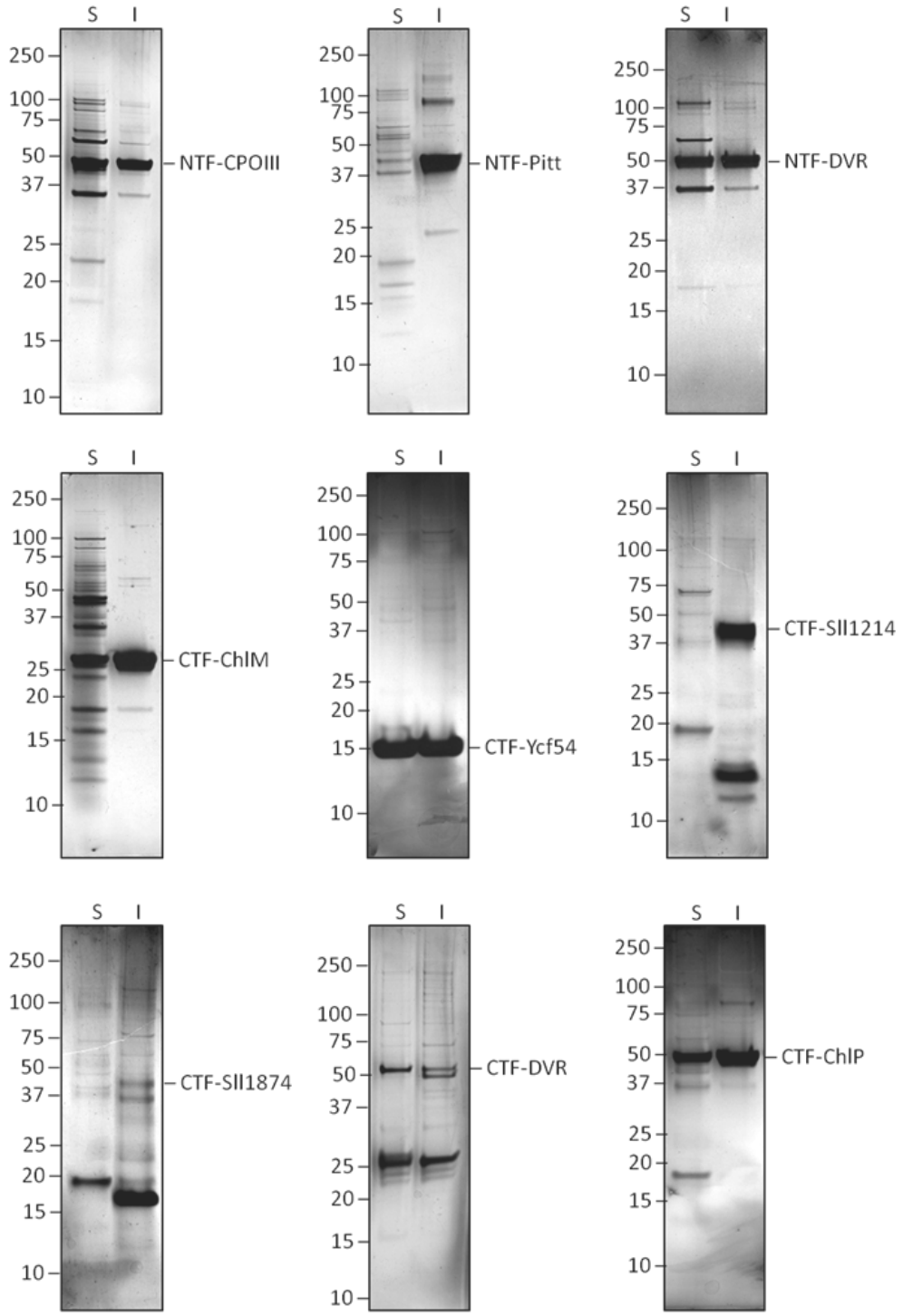

Figure 3.5 Silver stained SDS-PAGE analyses of elution fractions from N-terminal and C-terminal FLAG-affinity pulldown experiments

FLAG-affinity purification of *Synechocystis* N-terminal-FLAG-tagged CPOIII (42 kDa), Pitt (35 kDa) and DVR (48 kDa) and C-terminal-FLAG-tagged ChlM (28 kDa), Ycf54 (15 kDa), Sll1214 (45 kDa), Sll1874 (45 kDa), DVR (48 kDa) and ChlP (48 kDa). FLAG-tagged proteins were purified from *Synechocystis* soluble cytoplasmic (S) and β-DDM solubilised thylakoid (I) fractions using anti-FLAG M2 affinity gel. Captured proteins were eluted using anti-FLAG peptide, separated by SDS-PAGE and silver stained. The molecular marker is shown in kDa.

3.2.3 FLAG-pulldown experiments

FLAG-pulldown experiments were performed in triplicate with fully segregated N-terminally FLAG-tagged chlorophyll biosynthesis enzymes: CPOIII, ChlM, Sll1214, Sll1874, DVR, POR, Pitt, ChlP and Slr0483, and C-terminally FLAG-tagged chlorophyll biosynthesis enzymes: ChlM, Sll1214, Sll1874, DVR, Ycf54 and ChlP, as described in sections 2.6.8 and 2.6.9. 25 µl of the elution fractions from a WT control and the FLAG-tagged proteins were separated by SDS-PAGE and silver stained; images of these can be viewed in **Figure 3.4** and **Figure 3.5**.

3.2.4 Identification of FLAG-eluted proteins by mass spectrometry

To confirm the presence of the FLAG-tagged bait proteins and identify their interaction partners, the FLAG-eluted proteins were analysed by nano-flow liquid chromatography coupled to mass spectrometry (nanoLC-MS), as detailed in section 2.5.5. NanoLC-MS was used in preference to Western blotting to eliminate the requirement for antibodies to all the potential interaction partners. In all cases, proteins were derivatised by S-carbamidomethylation and digested by trypsin in the presence of SDS. The tryptic peptide fragments were prepared for nanoLC-MS by solid phase extraction using both cation exchange and C18 reverse-phase media. MS data acquisition was by a combination of survey scans to detect tryptic peptide ions and automated product ion (MS/MS) scans of the most intense peptide ions to obtain sequence-diagnostic spectra.

The survey scans and associated MS/MS spectra were converted to Mascot Generic Files to enable submission for searching against the complete *Synechocystis* proteome database. The *Synechocystis* proteins identified in each pulldown experiment were then entered into the Prohits database. The control experiments, using WT cell lysates, generated a list of 16 proteins that bound non-specifically to the anti-FLAG affinity resin. These include tryptophanyl-tRNA synthetase, elongation factor Tu and several phycobilisome subunits, these were then subtracted from the lists of potential interaction partners generated for each FLAG-tagged protein. Tables were then generated of the proteins present in each pulldown experiment and their associated Mascot score, which can be viewed in the appendix in Section 3.4. Proteins that appeared in $\geq$ 2 pulldown experiments with a Mascot score > 30, which did not appear in any of the control pulldown experiments, are listed as possible interaction partners.

The Mascot score represents a probability scoring based algorithm, whereby the actual mass of the MS/MS fragments detected are compared to an *in silo* database of MS/MS fragments generated for every tryptic fragment in an organisms proteome; in this case, the

Synechocystis PCC60803 proteome. From this, the detected ions score for an MS/MS match that is based on the probability of the observed match between the experimental and *in silico* data being a random event (Perkins et al., 1999). Generally, a Mascot score above 67 is accepted as the tryptic fragment being present. However, in this work, a Mascot score of >30 was used as the cut-off. This lower score has been used in consideration of the nature of the pulldown samples, which contain lowly abundant proteins and membrane proteins, which are difficult to ionise and whose hydrophobic tryptic fragments are lost through the sample preparation process.

Table 3.1 Summary of the potential interaction partners found using FLAG-pulldown experiments

Bait	Hits
Wild type control	Phycobiliprotein subunits: Sll1577 (CpcB), Sll1578 (CpcA), Slr2067 (ApcA) Aminoacyl tRNA synthetases and tRNA modification: Slr1884 (TprS) Ribosomal proteins: RplL (Sll1746) RNA synthesis, modification, and DNA transcription: Pnp (Sll1043), RpoA (Sll1818) DNA replication, restriction, modification, recombination, and repair: Slr0925 (Ssb), Sll1712 (DBH) Transport and binding proteins: Sll1204 Hypothetical proteins: Sll0839, Slr1102, Sll0033 (CrtH), Slr0455, Sll1687 (Hik17), Sll1665
ChlM	Tetrapyrrole biosynthesis: Sll1184 (Haem oxygenase), Slr0506 (POR), Slr1780 (Ycf54), Sll1091 (ChlP) Carotenoid biosynthesis: Slr1254 (Phytoene desaturase) Ribosomal proteins: Sll1260 (RpsB) Purine ribonucleotide biosynthesis: Slr1226 (PurC) Adaptations and atypical conditions: Sll0947 (IrtA) Photosystem/ light harvesting subunits: Sll1867 (PsaA), Slr1835 (PsaB), Ssl0563 (PsaC), Slr0737 (PsaD), Ssr2831 (PsaE), Sll0819 (PsaF), Slr1655 (PsaL), Slr1181 (PsbA), Ssl2598 (PsbH), Slr2067 (ApcA), Slr1963 (OCP) ATP synthase: Sll1326 (AtpA), Slr1329 (AtpB) NADH dehydrogenase: Sll0520 (NdHI) Transport and binding proteins: Slr0447 (UrtA), Sll1450 (NrtA) Proteases: Sll0020 (ClpC) Glutamate family / Nitrogen assimilation: Ssl0707 (GlnB) Chemotaxis: Sll1695 (HofG), Slr0161 (PilT) Biosynthesis: Sll1945 (1-deoxyxylulose-5-phosphate synthase) Circadian clock: Slr0757 (KaiB) Unknown: Sll0837 (member of TPR superfamily), Sll1769, Ssl0294, Sll0982, Slr1220,

	(PRC superfamily), Sll1218 (Ycf39), Slr0483, Slr1417 (Ycf57), Sll1528, Slr1128
Sll1214	Tetrapyrrole biosynthesis: Sll1185 (CPOIII), Sll1874, Slr1780 (Ycf54), Slr0505 (POR), Slr1923 (DVR), Sll1091 (ChlP) Elongation factors: Sll1099 (Tuf), Slr1463 (Fus) Adaptations and atypical conditions: Sll0947 (IrtA) Chaperones: Slr2076 (GroEL), Sll0416 (GroEL-2), Sll0058 (DnaK), Slr2075 (Cpn10) Ribosomal proteins: Sll1744 (RplA), Sll1808 (RplE), Sll1810 (RplF), Sll1274 (RplI), Sll1801 (RplW), Slr1356 (Rps1), Sll1097 (Rps7), Sll1812 (RpsE), Sll1821 (RpsM), Ssl3432 (RpsS), Ssr2799 (RpmA), Ssl3436 (RpmC) RNA synthesis, modification, and DNA transcription: Sll1789 (RpoC2) Photosystem/ light harvesting subunits: Sll1867 (PsaA), Slr1835 (PsaB), Ssl0563 (PsaC), Slr0737 (PsaD), Ssr2831 (PsaE), Sll0819 (PsaF), Slr1655 (PsaL), Sll0629 (PsaK2), Slr1181 (PsbA), Slr0906 (PsbB), Sll0851 (PsbC), Sll0849 (PsbD), Smr0006 (PsbF), Ssl2598 (PsbH), Sll1194 (PsbU), Sll0226 (Ycf4), Slr0947 (Ycf27) B6F complex: Sll1317 (PetA), Sll1316 (PetC) Phycobiliprotein subunits: Slr1986 (ApcB), Ssr3383 (ApcC) Sll0928 (ApcD), ApcE (Slr0335), Sll1580 (CpcC), Ssl3093 (CpcD), Sll1471 (CpcG), Slr1878 (CpcE) ATP synthase: Sll1326 (AtpA), Slr1329 (AtpB), Sll1327 (AtpC), Slr1329 (AtpD), Slr1330 (AtpE), Sll1324 (AtpF), Pyruvate dehydrogenase: Slr1934 Glycolysis: Slr0394 (PGK) CO_2 fixation: Sll1342, Slr0009 (RbcL) Serine family / Sulphur assimilation: Sll1908 (SerA) Transport and binding proteins: Slr0447 (AmiC), Sll1450 (NrtA) Cell division: Slr1390 (FtsH) Glutamate family / Nitrogen assimilation: Ssl0707 (GlnB) Antioxidant: Slr1198 (Rehydrin), Sll1621 Chemotaxis: Sll1695 (HofG) Unknown: Sll1106 (DUF1269), Sll0296 (DUF0820), Slr0552, Sll1530 (methyltransferase), Slr0909 (methyltransferase), Slr0483, Slr1506 (hydrolase), Slr1188 (Im30), Sll1130, Sll1528
Sll1874	Tetrapyrrole biosynthesis: Sll1214, Slr1780 (Ycf54) Chaperones: Slr2076 (GroEL), Sll0416 (GroEL-2), Slr2075 (Cpn10) Ribosomal proteins: Sll1808 (RplE) Photosystem/ light harvesting subunits: Slr1963 (OCP) Phycobiliprotein subunits: Slr1986 (ApcB) Chemotaxis: Sll1695 (HofG)
Ycf54	Tetrapyrrole biosynthesis: Sll1214, Sll1874 Chaperones: Slr2076 (GroEL)

	Elongation factors: Sll1099 (Tuf) Ribosomal proteins: Sll1821 (RplM) CO_2 fixation: Slr0009 (RbcL) Regulatory: Slr1666 (DegT), Slr1140 Photosystem/ light harvesting subunits: Ssr2831 (PsaE), Slr1655 (PsaL) ATP synthase: Slr1329 (AtpB) Antioxidant: Slr1198 (Rehydrin) Proteases: Slr0619 Unknown: Sll1106 (DUF1269), Slr5058 (AAA+ domain), Sll1130
POR	Tetrapyrrole biosynthesis: Slr1923 (DVR) Elongation factors: Sll1099 (Tuf) Chaperones: Slr2076 (GroEL), Sll0416 (GroEL-2) Ribosomal proteins: Sll1744 (RplA), Sll1802 (RplB), Sll1800 (RplD), Sll1808 (RplE), Sll1810 (RplF), Sll1274 (RplI), Sll1821 (RplM), Sll1805 (RplP), Sll1740 (RplS), Slr1678 (RplU), Ssl3436 (RpmC), Sll1097 (Rps7), Sll1817 (Rps11), Sll1804 (RpsC), Sll1812 (RpsE), Sll1101 (RpsJ), Sll1821 (RpsM), Ssr1399 (RpsR) Photosystem/ light harvesting subunits: Slr1835 (PsaB) Unknown: Sll1665, Sll1106 (DUF1269)
DVR	Unknown: Slr2121
ChlP	Tetrapyrrole biosynthesis: Slr0506 (POR), Slr1923 (DVR) Chaperones: Slr2076 (GroEL), Sll0416 (GroEL-2), Slr2075 (Cpn10) Photosystem/ light harvesting subunits: Sll1867 (PsaA), Slr1835 (PsaB) Phycobiliprotein subunits: Slr1986 (ApcB) ATP synthase: Slr1329 (AtpB) Proteases: Sll0020 (ClpC) Chemotaxis: Sll1695 (HofG) Biosynthesis: Sll1536 (MoeB) Unknown: Sll1751, Slr0483, Slr0244 (universal stress protein family), Slr1220, Slr6021
Pitt	Transcription/ Translation machinery: RpmC, RpsL, RplI, RpoC2 Chaperones: Slr2076 (GroEL), Sll0416 (GroEL-2), Slr2075 (Cpn10) Photosystem/ light harvesting subunits: Slr1986 (ApcB), Ssl3093 (CpcD), Slr2051 (CpcG) Unknown: Sll1217 (DBH)
Key for highlighted unknown proteins: Conserved in green lineage Conserved in non-green lineage	

Conserved in cyanobacteria
No homology

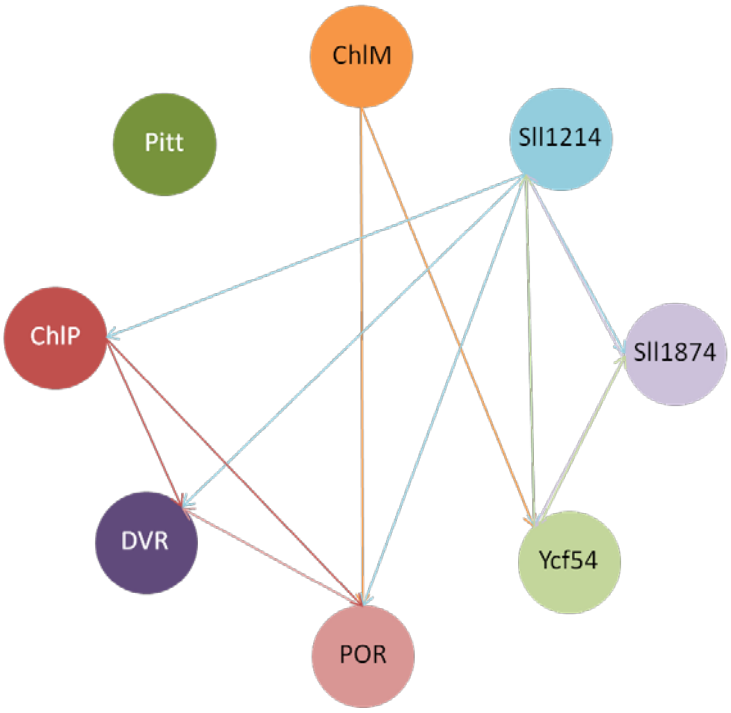

Figure 3.6 Interactions between chlorophyll biosynthesis enzymes identified from *in vivo* FLAG-tag pulldown experiments

The interactions found between the chlorophyll biosynthesis enzymes are shown for *in vivo* pulldown experiments with N-terminal and/or C-terminal FLAG-tagged ChlM, Sll1214, Sll1874, Ycf54, POR, DVR, ChlP and Pitt. Interactions are shown for hits that had a Mascot score over 30 and appeared in two or more pulldown experiments.

Figure 3.6 displays a schematic for the interactions that potentially occur between the chlorophyll biosynthesis enzymes and their putative interaction partners and **Table 3.1** summarises the hits found with each FLAG-tagged bait protein. These data showed that the FLAG-tagged chlorophyll biosynthesis enzymes do indeed interact with each other. For example, the two cyclase proteins Sll1214 and Sll1874 are both found to interact with each other and with the unknown hypothetical protein Slr1780 (Ycf54) (found to be required for cyclase activity, which is demonstrated in Chapters 4, 5 and 6); Sll1214 was additionally found to interact with ChlP, POR and Slr0483.

Several of the pulldown experiments contain several translation associated proteins, such as ribosome subunits, chaperones and elongation factors. This would be expected for proteins tagged at the N-terminus, as a partially translated protein would still be attached to the translation machinery and be capable of binding to the FLAG-affinity column.

The interaction partners that fall into the unknown category have all been subjected to BLAST searches, to determine if these are conserved in the green lineages (plants, algae and cyanobacteria, highlighted in green), in non-green lineages (highlighted in grey), or conserved in cyanobacteria (highlighted in blue); those with no homology to proteins in other species are not highlighted.

A *Synechocystis* knockout mutant of Sll1106, a potential interaction partner of POR and Sll1214, was found to exhibit wild-type growth and no defects in the chlorophyll biosynthesis pathway (Hollingshead S, unpublished data, 2012). Similarly, knockout mutants of Sll1751 and Sll1091, potential interaction partners of ChlP and ChlM respectively, were also found to have no defects in the chlorophyll biosynthesis pathway (Hollingshead S., unpublished data)

3.3 Discussion

3.3.1 Interactions are found to occur between the chlorophyll biosynthesis enzymes

Substantial conclusions can be drawn regarding the interactions observed between the FLAG-tagged chlorophyll enzymes and their protein partners. It is also clear from the FLAG-pulldown experiments and subsequent analysis of knockout mutants (reported in Chapter 4 and Chapter 7) that the *in vivo* FLAG-pulldown approach can be used to successfully identify proteins required for chlorophyll biosynthesis (Ycf54), and for the accumulation of chlorophyll-photosynthetic complexes (Slr0483).

The approach of *in vivo* FLAG pulldowns was used to investigate if the chlorophyll biosynthesis enzymes formed a large membrane-associated super-complex to enable the channelling of photo-active and hydrophobic chlorophyll intermediates. This would enable the efficient synthesis of chlorophyll, facilitating its arrival at the membrane, where it is inserted into the photosynthetic chlorophyll binding proteins (Sobotka, 2014). The pulldown data have enabled the identification of several interactions between the chlorophyll biosynthesis enzymes (see **Figure 3.7** and **Table 3.1** for a complete list). A schematic showing how these proteins may hypothetically form a complex associated with the membrane can be viewed in **Figure 3.7**. This shows that Slr0483 could be a transmembrane protein scaffold upon which its membrane-associated interaction partners (ChlP, ChlM and Sll1214) congregate; it should be noted that this is a tentative suggestion, drawn from the pulldown data gathered in this chapter.

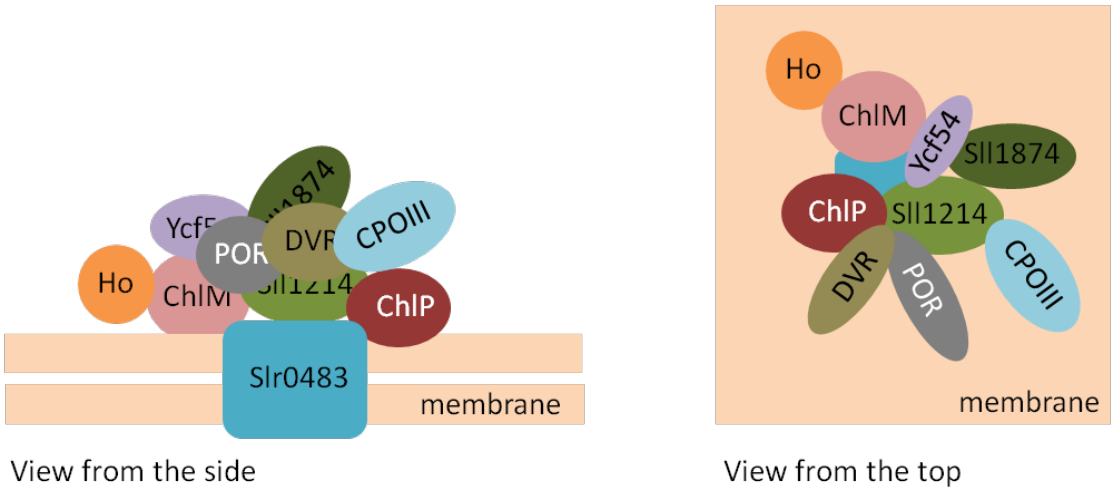

Figure 3.7 Schematic showing the possible interactions between the chlorophyll biosynthesis enzymes

Hypothetical diagram as to how the chlorophyll biosynthesis may arrange themselves on the membrane surface. Interactions are shown from the side and from the top, looking down onto the membrane. Ho (Haem oxygenase), Ycf54 (Slr1780), CPOIII (oxygen-dependent coproporphyrinogen III oxidase), ChlP (geranylgeranyl reductase), DVR (di-vinyl reductase), POR (protochlorophyllide oxio-reductase), ChlM (protoporphyrin IX methyl transferase).

In *Arabidopsis*, the tetrapyrrole pathway is known to be controlled by a negative feedback loop, where GluTR activity is regulated independently by intracellular haem (Pontoppidan and Kannangara, 1994; Vothknecht et al., 1998) and the FLU protein (Meskauskiene et al., 2001b; Meskauskiene and Apel, 2002). The FLU protein is tightly associated with the chloroplast membranes and has been shown to interact with the coiled-coil domain located at the end of the C-terminal of GluTR (seen in **Figure 1.12**) (Goslings et al., 2004) via a TPR domain located at the C-terminus of FLU (Meskauskiene and Apel, 2002). Analysis of the migration of the FLU protein on blue native gels showed that this protein forms higher complexes with molecular weights of ~200 kDa and ~700 kDa (Kauss et al., 2012a). The authors subsequently analysed these complexes by mass spectrometry and found that the FLU protein co-migrates with 117 proteins, including the chlorophyll biosynthesis enzymes ChlM, CHL27 (Sll1214 homologue), PORB, PORC, DVR, ChlP and ChlG. They then performed immuno-precipitation experiments using an *in vivo* expressed HA-tagged FLU protein, and found that FLU forms a stable complex with CHL27, PORB, PORC and ChlP. From these experiments, Kauss *et al* (2012) deduced that the FLU protein formed a chloroplast membrane complex, linking the final stages of chlorophyll biosynthesis with the initial stages of tetrapyrrole biosynthesis via an interaction with GluTR. Although the interactions found through the pulldown experiments described in this chapter need further validation, it would appear the chlorophyll biosynthesis enzymes in *Synechocystis* also form a higher order complex, which is associated with the membrane. A

BLAST search of the *Synechocystis* genome for FLU homologues found no significant matches, so the arrangement of the chlorophyll biosynthesis complex is likely to vary between the two organisms. Nevertheless, the data presented by Kauss *et al* (2012) and the data presented in this chapter do complement each other, and taken together they show it is very likely the chlorophyll biosynthesis enzymes form a super-complex, enabling the channelling of photoactive chlorophyll intermediates.

A few anomalies with previously reported data were also observed. In a previous study, ChlH was reported to stimulate methyltransferase activity (Shepherd et al., 2005b), however no ChlH was identified in the ChlM pulldown assay. There are several reasons why this may be, for example the FLAG-tag on ChlM may interfere with the ChlM-ChlH interaction or the interaction may be exceedingly transient and unable to survive the 100 column-volume wash step. Unfortunately, it has not been possible to obtain a FLAG-tagged ChlH inserted at the *psbAII* locus, (the ChlH gene is ~4,000 bp in length, which may exceed the limits of the recombination machinery that inserts the DNA fragments into the genome), in order to perform the reciprocal pulldown experiment. To address this problem, a construct has been designed to N and C-terminally FLAG-tag ChlH at the native locus, which requires the insertion of a short length of DNA. In addition to the ChlM interaction, ChlH has been implicated in several cellular events. For example, in *Arabidopsis*, ChlH was found to bind abscisic acid and act as a possible ABA receptor (Shen et al., 2006b; Wu et al., 2009b), and be cyclically expressed in response to the circadian rhythm (Harmer et al., 2000); in *Synechocystis* it apparently regulates sugar catabolism via the anti-sigma factor SigE (Osanai *et al*., 2005; Osanai *et al*., 2009),. Therefore, a FLAG-tagged ChlH protein could be essential for exploring the possible interaction partners ChlH has in *Synechocystis*, and ascertaining the presumed global role of this protein.

The second observed anomaly was the lack of POR in the FLAG-Pitt pulldown experiment, and the lack of Pitt in the POR pulldown experiments. An interaction between POR and Pitt was initially reported in a global yeast two hybrid study, which aimed to characterise the interactome of *Synechocystis* (Sato et al., 2007), which was then validated in more detail (Schottkowski et al., 2009). Schottkowski and co-authors observed that the levels of POR were reduced in the Pitt knockout mutant and that the two enzymes co-migrated on BN-PAGE and sucrose density gradients. However, they also reported that the complex is likely to be transient, as they could not pull down POR in anti-Pitt immuno-precipitation experiments and only a fraction of Pitt is found in higher-molecular weight complexes, where it presumably forms a complex with POR. Presumably, the transient nature of this interaction is why Pitt and POR are not detectable in the FLAG-POR and FLAG-Pitt pulldown experiments.

An additional advantage of these FLAG-tagged pulldown experiments is the insights into unknown proteins and/ or co-factors that the tagged target protein(s) interact with. For example, only one subunit of the cyclase, the enzyme which catalyses the formation of the chlorophyll fifth isocyclic ring, is known, when biochemical evidence suggests this enzyme consists of two or more subunits (Walker et al., 1988; Walker et al., 1989; Walker et al., 1991; Bollivar and Beale, 1996; Rzeznicka et al., 2005). Therefore, it could be possible to identify additional components of the cyclase complex using the FLAG-pulldown approach (see Chapter 4).

3.3.4 Further work

3.3.4.1 Further work for mapping the protein-protein interactions in the chlorophyll biosynthesis pathway

With regard to the FLAG-pulldown experiments, a number of approaches could be applied to further verify and understand the interactions observed. A combination of gel filtration and native-PAGE could be used to ascertain the molecular weights of the complexes formed and fractions and/or gel slices could be analysed by immuno-blotting and mass spectrometry to determine the protein content of these complexes.

Many of the chlorophyll biosynthesis enzymes that interact can be produced recombinantley by over-expression in *E. coli*. These could be incubated together in various combinations and resolved by gel filtration or native-PAGE and the complexes analysed by mass spectrometry, to determine whether these interactions can be recreated *in vitro*. Alternatively, the *Synechocystis* supernatant could be resolved using native-PAGE and the bands containing the high molecular weight complexes of interest could be excised and analysed by mass spectrometry to determine their composition, as described in Kauss *et al* (2012).

Ongoing work involves the use of immuno-precipitation experiments with antibodies raised against the chlorophyll biosynthesis enzymes, ChlH, ChlI, ChlD, GUN4, ChlM, Ycf54, POR, ChlP and DVR. The aim of this approach is to find out if the interactions observed with the FLAG-pulldown experiments are also observed in immuno-precipitation experiments, which would further confirm the existence of a chlorophyll biosynthesis enzyme super-complex.

Many of the cellular interactions may be too transient to survive the mechanical disruption and solubilisation process the FLAG-tagged protein containing strains are subjected to. Therefore, it is likely the proteins identified as potential interaction partners in the FLAG-pulldown eluates are only a fraction of those that occur within the cell. To identify these interactions a very low concentration of a chemical cross-linker could be used on whole cells *in vivo* and the proteins interacting with the FLAG-tagged protein of interest would be

isolated via the FLAG-pulldown protocol. Recently, this *in vivo* cross-linking approach was used to identify a phycobilisome-PSI-PSII mega-complex in *Synechocystis* (Liu et al., 2013); this interaction could not previously be consistently isolated owing to its weak and easily disrupted interactions.

4. Identification of Ycf54 (Slr1780) as an interaction partner of the Mg protoporphyrin IX monomethylester oxidative cyclase in *Synechocystis* PCC 6803

4.1 Introduction

All chlorophylls are modified tetrapyrrole molecules that possess a unique isocyclic ring, known as the "fifth" or "E" ring. This isocyclic ring arises from the cyclisation of the methyl-propionate side-chain at C-13 to the C-15 bridge carbon between rings C and D, in a reaction catalysed by the Mg-cyclase. **Figure 4.1**, depicts a schematic by which the conversion of MgPME to Pchlide is thought to proceed (Wong et al., 1985; Bollivar and Beale, 1996a). The two proposed intermediates, Mg-protoporphyrin IX 6-methyl-β-hydroxypropionate and Mg-protoporphyrin IX 6-methyl-β-ketopropionate, have both been shown to be catalytically active substrates of the cyclase, as demonstrated using whole cell assays on cucumber cotyledons (Wong et al., 1985).

The cyclisation of the C-13 side-chain is thought to be preceded by the addition of a keto-group at C-13^1 (Mg-protoporphyrin IX 6-methyl-β-hydroxypropionate), possibly arising from the addition of a hydroxyl group at C-13^1, which is then further oxidised to a keto group (Mg-protoporphyrin IX 6-methyl-β-ketopropionate) (Figure 4.1). This oxygen is known to originate from two sources, oxygen or water, depending on whether the host organism possesses an oxidative or anaerobic cyclase. Aerobic photosynthetic organisms, including higher plants and cyanobacteria, have a mono-oxygenase Mg-cyclase, whereby molecular oxygen is the source of the C-13^1 keto group. In the unrelated hydratase-type Mg-cyclase, found in anaerobic photosynthetic organisms such as *Rhodobacter* sphaeroides, oxygen from a molecule of water is incorporated (Chereskin et al., 1982; Porra et al., 1996). Some organisms contain both types of Mg-cyclase, of which two examples are *Rubrivivax gelatinosus* and *Rhodovulum sulfidophilum* (Porra et al., 1998; Pinta et al., 2002; Ouchane et al., 2004). *Synechocystis sp*. is thought to contain only the aerobic Mg-cyclase, despite possessing three homologues (Sll1242, Slr0906, Slr0309) to the anaerobic cyclase (Minamizaki et al., 2008a).

Perhaps the most striking feature of the Mg-cyclase is that it is solely responsible for producing the characteristic green colour of chlorophyll. During the course of the Mg-cyclase reaction, the colour of the tetrapyrrole molecule changes from the reddish pink of MgPME to the bright green of Pchlide. This colour change arises from a shift in the Soret peak from 416

nm (Mg-PME) to 440 nm (Pchlide), brought about by an extension of the conjugating double bonds (**Figure 1.26 A**).

Huge advances have been made in characterising chlorophyll biosynthesis and most enzymes in the pathway have now been characterised in detail; however, the Mg-cyclase remains the notable exception and several components of this enzyme remain to be discovered. So far enzymology analyses have been confined to assays using extracts from cucumber chloroplasts (Fuesler et al., 1984; Wong and Castelfranco, 1984b), wheat etioplasts (Nasrulhaq-Boyce et al., 1987b) and cell extracts from *Chlamydomonas reinhardtii*, *Synechocystis sp PCC6803* (Bollivar and Beale, 1996b) and *Arabidopsis thaliana* (Tottey et al., 2003b). The assays in cucumber and *Synechocystis* showed the cyclisation reaction was catalysed by a multi-subunit complex that consisted of at least one membrane bound and one soluble component (Bollivar and Beale, 1996a). Partial biochemical purification of cucumber extracts suggested the soluble component had a molecular mass of over 30 kDa (Walker et al., 1991). The only known candidate for a catalytic cyclase subunit was identified by mutational analysis of *Rubrivivax gelatinosus*, where inactivation of the *AcsF* (aerobic cyclisation system Fe-containing protein) gene resulted in accumulation of MgPME under aerobic conditions (Pinta et al., 2002). Subsequently AcsF homologues were identified across all oxygenic photosynthetic organisms, examples of which are *crd1* and *crd2* in *Chlamydomonas reinhardtii* (Moseley et al., 2002), *chl27* in *Arabidopsis* (Tottey et al., 2003a) and *sll1214* and *sll1874* in *Synechocystis* (Minamizaki et al., 2008a). Of the two AcsF genes in *Synechocystis* Sll1214 is absolutely required for cyclase activity under aerobic conditions and Sll1874 for activity under micro-oxic conditions (Minamizaki et al., 2008a; Peter et al., 2009).

Despite these advances in solving the unknown components of the Mg-cyclase complex, at least one soluble component that is absolutely essential for activity remains unidentified, as well as possibly another membrane bound subunit, as analysis of two barley mutants *viridis-k-23* and *xantha-l-35* showed that two different membrane components were required for reconstitution of cyclase activity (Rzeznicka et al., 2005).

In Chapter 3 FLAG pulldown experiments using *Synechocystis* strains transformed with FLAG-Sll1214 and FLAG-Sll1874 identified the Ycf54-like protein (slr1780) as an interaction partner in each case. A BLAST search revealed Ycf54 to be highly conserved among oxygenic photosynthetic organisms, including all plants and cyanobacteria. As the Ycf54-like protein was designated as a hypothetical protein with unknown function, it was decided to investigate the effects of interrupting the *slr1780* gene in *Synechocystis*. This chapter reports the construction of a *slr1780* knockout mutant and investigation of the resulting phenotype.

4.2 Results

4.2.1 Identification of Ycf54 forming a complex with Sll1214 and Sll1874

In Chapter 3, *in vivo* pulldown experiments using N-terminal FLAG-tagged Sll1214 and Sll1874 identified Ycf54 (slr1780) as an interaction partner of both proteins. To further strengthen this hypothesis, it was decided to perform the reciprocal pulldown experiment, with N-terminal FLAG-tagged Ycf54 as bait, to determine if Sll1214 and Sll1874 could be captured by Ycf54.

A N-terminal FLAG-tagged Ycf54 construct was created by amplifying the *slr1780* gene from *Synechocystis* genomic DNA and ligating it into the vector pFLAG (described in Section 2.4); successful construction of pFLAG-Ycf54 was confirmed by restriction digests and sequencing. The resulting pFLAG-Ycf54 vector was transformed into wild type *Synechocystis* cells and a fully segregated mutant was obtained by sequential selection on increasing concentrations of kanamycin (described in Section 2.3.1).

Three sets of FLAG-affinity pulldown experiments were then performed with the FLAG-Ycf54 mutant (detailed in Section 2.6.9). The elution fractions from each of the pulldown experiments were analysed by mass spectroscopy (**Table 4.1**), western blotting (**Figure 4.2 C**) and silver stained SDS-PAGE (**Figure 4.2 B**). Unlike Sll1214 and Sll1874, FLAG-Ycf54 was found to be highly over-expressed and present at significant levels in the pulldown elution fraction, indicating FLAG-Ycf54 is able to accumulate at significantly greater levels than FLAG-Sll1214 or FLAG-Sll1874 proteins (**Figure 3.5**).

The western blot of the FLAG-Ycf54 pulldown elution fractions probed with anti-Chl27 (**Figure 4.2 C**) demonstrates that the cyclase subunits Sll1214/Sll1874 interact with FLAG-Ycf54, whereas the WT control fraction does not. This is further strengthened by the mass spectrometry data (**Table 4.1**), which identifies FLAG-Ycf54 as interacting with Sll1214 in two separate pulldown experiments, lending further support to the proposal that Ycf54 is required for the cyclase step in chlorophyll biosynthesis. Sll1874 was not identified by mass spectroscopy as interacting with Ycf54, but this is not surprising as Sll1874 is expressed under micro-oxic conditions (Minamizaki et al., 2008a; Peter et al., 2009) and the FLAG-Ycf54 cultures used in the pull-downs were grown aerobically.

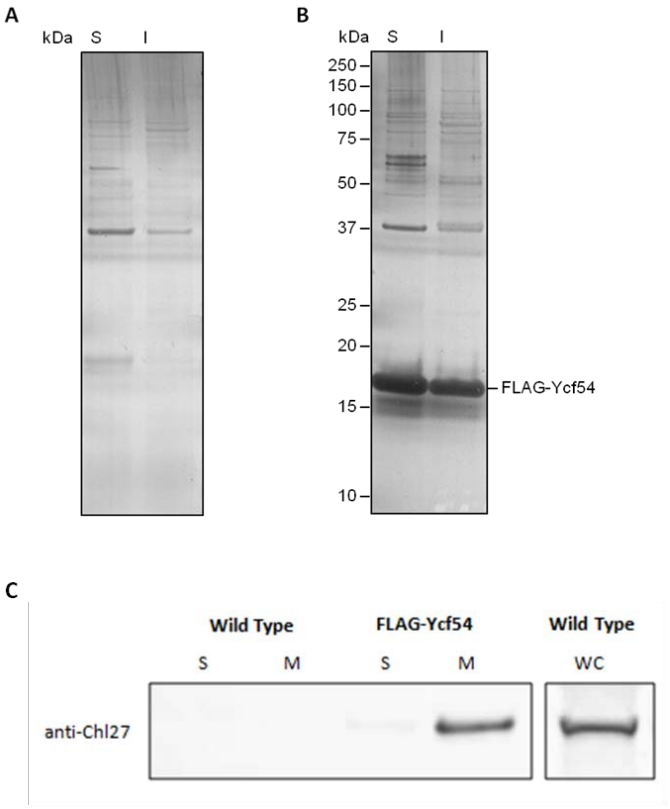

Figure 4.2 Silver stained SDS-PAGE and western blot analysis of a FLAG-Ycf54 pulldown

Elution fractions from the FLAG-affinity purification of *Synechocystis* wild type control (**A**) and FLAG-tagged Ycf54 (**B**) were separated by SDS-PAGE and silver stained. FLAG-tagged proteins were purified from *Synechocystis* soluble cytoplasmic (S) and β-DDM solubilised thylakoid (I) fractions using anti-FLAG M2 affinity gel and eluted using the anti-FLAG peptide. (**C**) Western blot of Soluble (S) and membrane (M) elution fractions from FLAG-Ycf54 and wild type control FLAG-affinity pulldown experiments probed with anti-CHl27 to show Ycf54 interacts with Sll1214.

4.2.2 Construction of a *Synechocystis Δslr1780* mutant

To ascertain whether Ycf54 has a role in cyclase activity, it was decided to construct a *Synechocystis* strain in which the *slr1780* (Ycf54) gene is interrupted with a zeocin resistance cassette. Using the mega primer method adapted from Ke and Madison (1997) a length of DNA consisting of a zeocin resistance cassette flanked by ~500 bp of the upstream and downstream regions of the *slr1780* gene locus was generated though sequential PCR reactions. A fully segregated Δ*slr1780 Synechocystis* strain was created by transforming the *slr1780* resistance cassette into wild type *Synechocystis* and sequentially selecting on increasing concentrations of zeocin until full segregation was achieved (Section 2.3.1). Confirmation of full segregation by amplification of the *slr1780* region is shown (**Figure 4.3 C**). During segregation it was noted this mutant had a long doubling time and would only grow at low light intensities under mixotrophic conditions.

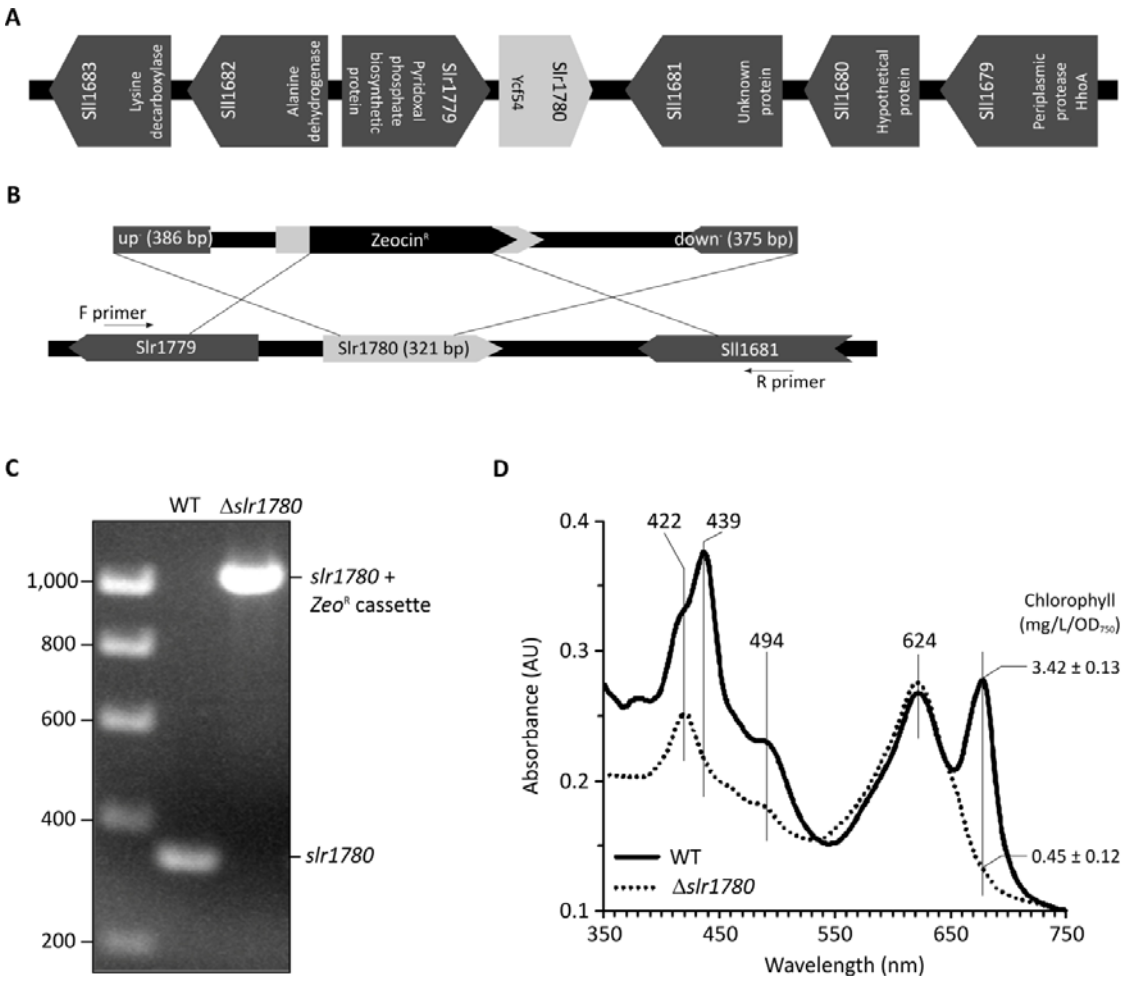

Figure 4.3 Location of Slr1780 in the *Synechocystis* genome and production of a Δ*slr1780* mutant

(**A**) Location of the Ycf54 coding gene slr1780 in the *Synechocystis* genome. (**B**) Knock out cassette, generated by fusion PCR, used to create Δ*slr1780* and the region of the *Synechocystis* genome into which it integrates. (**C**) PCR amplification of the slr1780 region to confirm full segregation of the zeocin resistance cassette, using the primers indicated in 1B. (**D**) Whole cell absorbance spectra of wild type and Δ*slr1780* mutant *Synechocystis* cells.

4.2.3 Investigating the role of Ycf54 in *Synechocystis*

Initial observations of Δ*slr1780* suggested the Ycf54 protein had an important function in the cell, as Δ*slr1780* cells exhibited significantly retarded growth and had a "blue" hue to their pigmentation when compared to wild type cells. To further investigate the function of Ycf54 several methods were employed to analyse how the cell physiology of *Synechocystis* was affected in Δ*slr1780*.

4.2.3.1 Whole cell absorbance spectrum of *Δslr1780*

Examination of the absorption spectra from cells normalised for OD_{750}, shows the chlorophyll absorbance maxima at 439 and 679 nm and the carotenoid absorbance maxima at 494 nm are severely depleted in Δ*slr1780*, whilst the absorbance maxima of the phycobiliproteins at

624 nm remains unchanged when compared to the wild type (**Figure 4.3 D**). Upon quantification of cellular chlorophyll (**Figure 4.3 D**), which was related to the same biomass as assessed by light scattering at 750 nm, it was found Δ*slr1780* accumulated 87 % less chlorophyll than wild type cells grown under identical conditions. The absence of chlorophyll in Δ*slr1780* is clearly seen in photographs of cell lysate fractions (**Figure 4.5 A**); in wild type the whole cell lysate and membrane fraction are green, and in Δ*slr1780* whole cell lysate and membrane fraction are blue and orange respectively.

4.2.3.2 HPLC analysis of *Δslr1780* methanol extracted pigments

To investigate whether a deficiency in the chlorophyll biosynthesis pathway is the origin of the significant reduction in chlorophyll in *Δslr1780*, HPLC was used to analyse the chlorophyll precursor pigments present in *Δslr1780* and wild type cells. Cellular pigments from equal quantities of cells were extracted with methanol (described in Section 2.7.1) and separated on a Phenomenex C18 column according to the method described in Table 2.7. **Figure 4.4 A** shows DAD absorbance spectra recorded at 432 nm. As observed with the whole cell absorbance spectrum, the HPLC of Δ*slr1780* is noticeably different to that of wild type cells. Two "new" peaks are observed at 16.7 and 29.9 minutes and the peak at 26.4 minutes in the wild type spectrum is absent in Δ*slr1780*. The elution time and absorbance spectrum of the peak at 29.9 minutes correspond to that of Mg-PME, the substrate of the cyclase and the absent peak at 26.4 minutes corresponds to Pchlide, the product of the cyclase (for standards see **Figure 5.4**). The accumulation of Mg-PME and absence of Pchlide suggest that the chlorophyll biosynthesis pathway is blocked at the cyclase step, as the cells appear to be incapable of converting Mg-PME to Pchlide.

The HPLC spectra also show a small peak with an elution time of 16.7 minutes (**Figure 4.4 A**) that has an absorbance maximum of 432 nm (**Figure 4.4 B**) and does not relate to any known photosynthetic precursor pigment; from here on this pigment will be referred to as pigment A432. Of interest is that the Soret band of pigment A432 falls between the Soret bands of Mg-PME (416 nm) and Pchlide (440 nm), as demonstrated in **Figure 4.4 B**. Similar behaviour is observed with the red-end peak, which is located at 616 nm in pigment A432, at 589 nm in Mg-PME and 631 nm in Pchlide. These spectra suggest pigment A432 may be an intermediate in the cyclase reaction, or a breakdown product of MgPME, which is able to accumulate in the absence of a working cyclase complex. To further investigate the identity of pigment A432, a large scale purification protocol was devised to purify the pigment to near homogeneity to allow for structural analyses using a combination of NMR and mass spectrometry (**Section 4.3.4**).

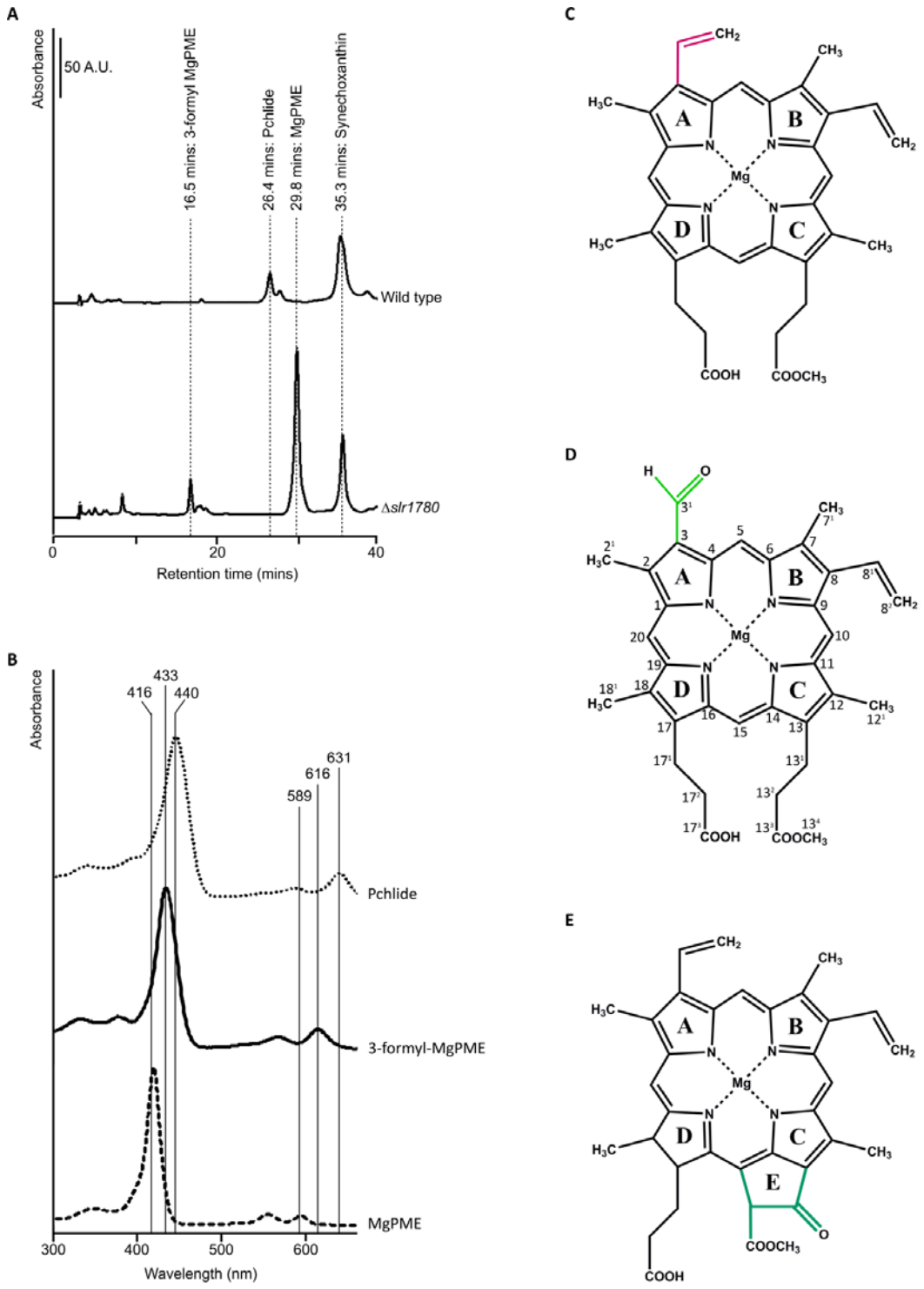

Figure 4.4 HPLC analyses of methanol extracted pigments from wild type and Δslr1780 cells

(A) Chlorophyll precursors were extracted with methanol containing 0.2% (v/v) ammonium from an equal volume of cells at an OD_{750} = ~0.7 and analysed on a Phenomenex C18 column. Separation of precursors was detected by a diode array detector set to 432 nm, the Soret peak of 3-formyl MgPME, which is observed in Δ*slr1780* cells. The elution times of the pigments of interest, which differ between wild type and Δ*slr1780* cells, are indicated. (B) Absorbance spectra of MgPME, Pchlide and 3-formyl-MgPME (C) Mg-protoporphyrin IX monomethylester (D) Mg-3-formyl-protoporphyrin IX monomethylester (E) Mg-protochlorophyllide.

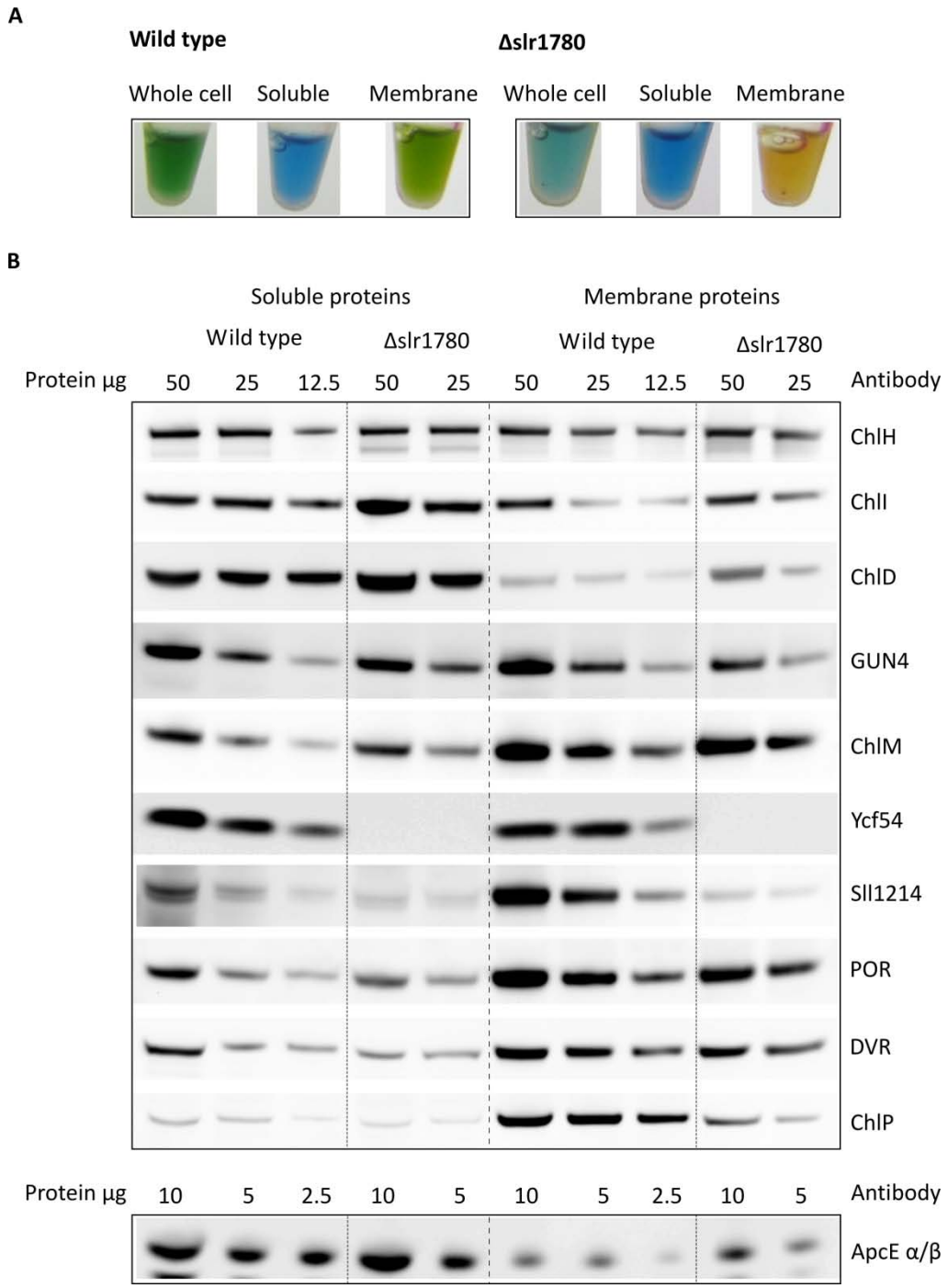

Figure 4.5 Accumulation of chlorophyll biosynthesis enzymes in wild type and Δ*slr1780* cells

Photographs of the whole cell, soluble and solubilised membrane lysate fractions, illustrating the changes in pigmentation between wild type and Δ*slr1780* (**A**). Western blot analysis of the soluble and membrane fractions from wild type and Δ*slr1780 Synechocystis* strains (**B**). *Synechocystis* cells were disrupted by bead beating and whole cell samples were taken; soluble and insoluble samples were generated by ultracentrifugation. Samples of known protein concentration (calculated using the method of Kalb and Bernlohr (1977)) were separated by SDS-PAGE and transferred to a nitrocellulose membrane, which was probed with antibodies to the magnesium chelatase subunits (ChlH, ChlI, ChlD and GUN4), Sll1214/Sll1874 (CHL27), ChlM, Ycf54, POR, DVR and ChlP.

4.2.3.3 Accumulation of the chlorophyll biosynthesis enzymes in *Δslr1780*

HPLC analyses showed that chlorophyll biosynthesis in a mutant lacking Ycf54 was blocked at the Mg-cyclase step. To investigate whether the absence of Ycf54 affects the accumulation of the chlorophyll biosynthesis enzymes, specifically the AcsF catalytic subunit of the Mg-cyclase Sll1214/Sll1874, western blot analyses were performed on equal quantities of cell lysates from wild type and Δ*slr1780* cells.

Cell lysates were obtained from cultures grown photomixotrophically under low light conditions to an OD_{750} of 0.6-0.8 (the method for preparation of cell lysates is described in Section 2.6.7). These were then separated into the soluble and insoluble fractions by high speed ultra-centrifugation and protein content was calculated using the method of Kalb and Bernlohr (1977). Samples containing known quantities of protein were separated by SDS-PAGE, then transferred to a nitrocellulose membrane and probed with antibodies to the chlorophyll biosynthesis enzymes ChlH, ChlI, ChlD, GUN4, ChlM, Ycf54, Sll1214/Sll1874 (CHL27), POR, DVR and ChlP (**Figure 4.5 B**).

Figure 4.5 B shows there to be differences in the accumulation of several chlorophyll biosynthesis enzymes between wild type and Δ*slr1780* cells. Importantly, no Ycf54 is observed in any of the Δ*slr1780* fractions. Greater levels of the MgCH subunits ChlI and ChlD are observed in the soluble and membrane fractions of Δ*slr1780* in comparison to wild type, whilst the levels of ChlH, another sununit of MgCH, remain unchanged. Additionally, the accumulation of GUN4, a stimulator of MgCH activity, is reduced in the membrane fraction of Δ*slr1780* in comparison to wild type (**Figure 4.5 B**). Δslr1780 also has significantly reduced accumulation of Sll1214/Sll1874 and ChlP, and, to a lesser extent, reduced accumulation of POR, in comparison to wild type. The observed decrease in Sll1214/Sll1874 in *Δslr1780* compared to wild type is not surprising given the two proteins are known to associate with each other. More surprising is the effect Δ*slr1780* has on the accumulation of ChlP, an enzyme that does not directly precede or carry out the cyclase reaction.

No significant differences in the levels of ChlH, ChlM and DVR were found when probing these blots. The amounts of these proteins appear to be similar across the three sets of samples in all four strains.

4.2.4 Investigating the global cellular effects caused by the depletion of Ycf54

It has been widely reported that a reduction in cellular chlorophyll accumulation has wider implications for the cell, affecting the synthesis and stability of chlorophyll binding proteins

like those found in the photosystem complexes PSI and PSII. In turn, a reduction in photosystem complex assembly is known to have knock-on effects on thylakoid formation and cell size (Shimada et al., 2008; Fuhrmann et al., 2009). To investigate the effects caused by depleting YCF54, likely to have wide-ranging consequences over the photosynthetic machinery, the whole cell architecture was imaged using electron microscopy and photosystem accumulation was assessed using a combination of CN-gels, sucrose density gradients and low temperature fluorescence microscopy.

4.2.4.1 Comparison of *Δslr1780* and wild type cellular architecture using electron microscopy

Given the dramatic reduction of chlorophyll observed in *Δslr1780*, it was decided to utilise electron micrographs of thin layer sections to see if the composition of the thylakoid membranes is affected. Electron micrographs of thin sections of both wild type and *Δslr1780* cells are shown in **Figure 4.6**. These micrographs were kindly prepared by Lenka Bucinska using the method described in Kopecna et al. (2012).

Thin sections of wild type and mutant cells show significant changes in the ultrastructure of the mutant (**Figure 4.6**). In the wild type the thylakoids are observed as parallel stacks of 2 – 5 membranes that closely follow the contour of the cell membrane (**Figure 4.5**). In Δ*slr1780* no such organised thylakoid membranes are visible. Instead membrane-like structures are dispersed throughout the cytoplasm of the cell, whose curvature loosely follows that of the plasma membrane. A similar phenotype is observed for the *Synechocystis vipp1* (vesicle-inducing protein in plastids 1) mutant. In *Arabidopsis* disruption of the *VIPP1* gene severely affects the ability of the plant to form structured thylakoids (Kroll et al., 2001) and in *Synechocystis* partial deletion of the *vipp1* gene leads to almost complete loss of thylakoid formation (Westphal et al., 2001).

In Δ*slr1780* the abnormal biogenesis of the thylakoid membranes is likely to be a cumulative effect stemming from the disruption of chlorophyll biosynthesis, leading to the reduced assembly of photosystems, which form the majority of the protein component of the thylakoids. To investigate this theory further, the accumulation of the photosystem complexes photosystem I (PSI) and photosystem II (PSII), as well as their light harvesting antennae, the phycobilisomes, was investigated.

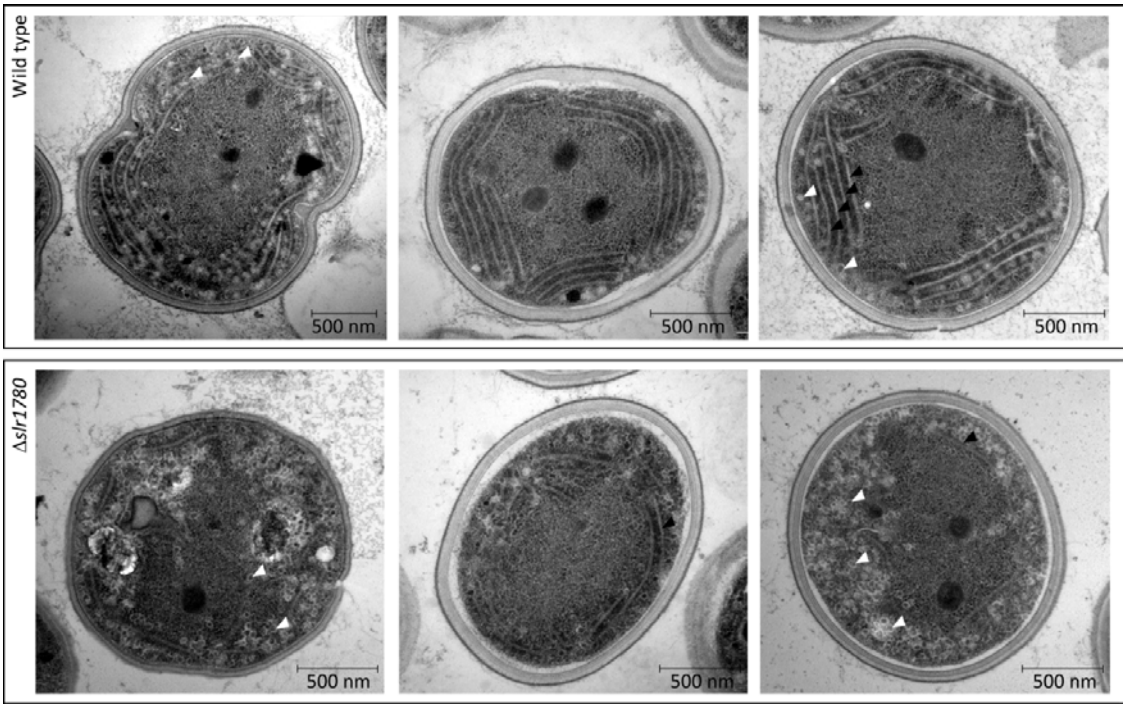

Figure 4.7 Thin layer electron micrographs of wild type and Δ*slr1780* Synechocystis cells

Transmission electron microscopy of strained ultrathin sections from *Synechocystis* wild type and Δ*slr1780* cells, grown photomixotrophically under low light conditions. Black arrows indicate thylakoid membranes and white arrows indicate glycogen granules.

4.2.4.2 Analysis of photosystem and phycobilisome accumulation in *Δslr1780*

In wild type *Synechocystis*, the phycobilisome light harvesting antennae connect with the photosystems and transfer light energy to the chlorophylls in PSI, which in turn transfer their energy to a special pair of chlorophyll molecules located within the photosystem reaction centres, where charge separation takes place (Renger and Holzwart, 2005). However, this process is likely to be interrupted in Δ*slr1780* owing to the 93 % reduction in chlorophyll accumulation, as the presence of chlorophyll is necessary for the assembly of functional photosystem complexes (Mullet et al., 1990; Eichacker et al., 1992). A previous study found a *Synechocystis* mutant lacking the light-independent protochlorophyllide reductase gene *chlL* contained negligible levels of chlorophyll when grown in the dark, but still retained wild type levels of the phycobilisome light harvesting antennae (Yu et al., 1999). From this previous study it could be interpreted that the biosynthesis of chlorophyll containing complexes, such as the photosystems, and their associated light harvesting antennae is "uncoupled", with synthesis of each occurring irrespective of their photosynthetic energy transfer partner. To ascertain whether this is the case for other chlorophyll biosynthesis mutants, and not just a phenomenon confined to dark grown chlorophyll-less mutants, the accumulation of the photosystems and the phycobilisome light harvesting antennae in Δ*slr1780* was determined.

Initial analysis of photosystem accumulation was performed using two dimention CN-PAGE/SDS-PAGE analysis. Purified membranes were prepared from Δ*slr1780* and wild type cells harvested at mid-log phase as described in Section 2.6.8. Isolated membranes, normalised to the concentration of chlorophyll per OD_{750} unit, were solubilised with β-DDM at a final concentration of 1 % (v/v) (described in Section 2.6.5).

The 2D PAGE analysis (**Figure 4.7 A**) of the wild type sample clearly resolves the PSI trimers (PSI [3]), PSII dimers (PSII [2]), PSI monomers (PSI [1]) and PSII monomers (PSII [1]). In the *slr1780* mutant, the bands representing PSI [3] and PSI [1] are present, but are far less abundant than the wild type; the PSII [2] and PSII [1] complexes were undetectable on the silverstained 2nd dimention SDS-PAGE. A Western blot of the second dimention was probed with antibodies to ChlP, Sll1214, POR and Ycf54. The ChlP and Sll1214 blots show these two proteins to be significantly reduced in Δ*slr1780*, which is concurrent with the Western blot shown in **Figure 4.6 B**. A higher-order complex between POR and Sll1214, marked by an asterisk, is present in the wild type, but absent in Δ*slr1780*. This complex is supported by the observation that Sll1214 interacts with POR, as discovered in Chapter 3.

To gain an understanding of the levels of PSII and phycobilisome light harvesting antennae in Δ*slr1780*, low temperature fluorescence emission spectra were recorded after excitation of the chlorophylls at 435 nm and the phycobilisomes at 580 nm (described in Section 2.9.2). For ease of comparison the 435 nm spectra were normalised for PSII fluorescence at 682 nm and the 580 nm spectra were normalised to phycocyanin fluorescence at 646 nm. In wild type cells after excitation at 435 nm, four clearly defined maxima are observed at 646 nm, 664 nm, 682 nm and 722 nm; an additional peak at 693 nm in the 682 nm shoulder is also visible. The maxima at 682 nm, 693 nm and 722 nm arise from the chlorophyll molecules in PSII (CP43), PSII (CP47) and PSI, respectively, and the peaks at 646 nm and 664 nm represent fluorescence from phycocyanin and allophycocyanin, respectively (Andrizhiyevskaya et al., 2005).

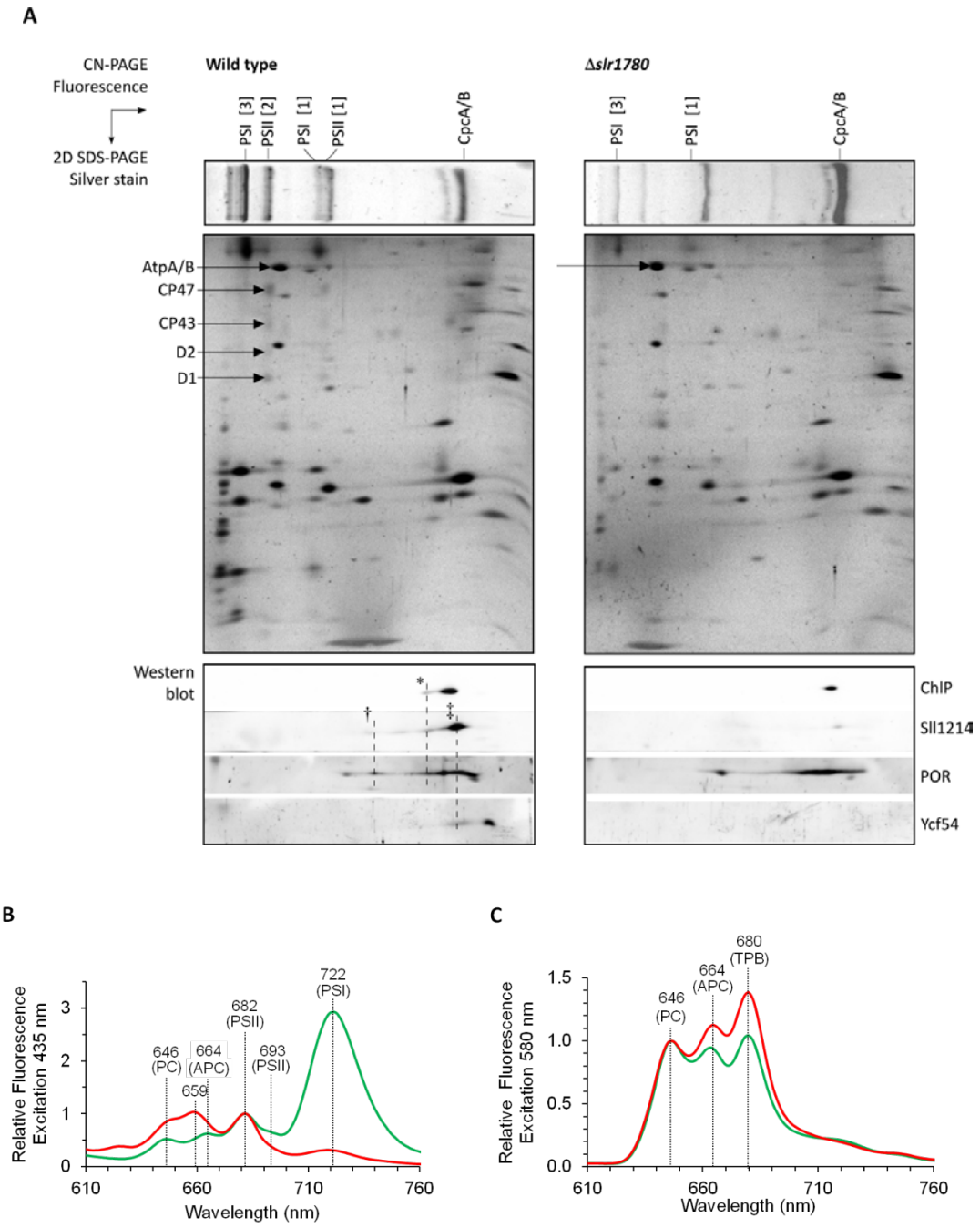

Figure 4.7 Analysis of photosystem accumulation in wild type and *Δslr1780 Synechocystis* cells

(A) For CN-gel electrophoresis isolated membranes were solubilised with 1 % β-DDM (v/v) and samples were normalised for chlorophyll concentration/ OD_{750}. The fluorescence from the first dimension native gel was photographed and after SDS-PAGE in the second dimension the gel was silver stained. Alternatively, the gel was blotted and probed with antibodies specific to ChlP, Sll1214, POR and Ycf54. (B) For the 77 K emission spectra, all samples were re-suspended in 80 % glycerol to an OD_{750} = 0.1. Fluorescence emission after excition at 435 nm (B) and 580 nm (C) was used to moitor PSI and the phycobilisomes. For comparability the 435 nm emission spectra were normalised to the PSII emission peak at 682 nm and the 580 nm emission spectra were normalised to the peak at 646 nm. PC = phycocyanin, APC = allophycocyanin and TPB = terminal phycobiliprotein.

The differences after excitation at 435 nm between wild type and Δ*slr1780* cells are considerable (**Figure 4.7 B**) and include increased phycocyanin fluorescence at 646 nm, a blue shift in the allophycocyanin peak from 664 nm to 659 nm, the loss of the PSII shoulder at 693 nm and a dramatic reduction in the PSI fluorescence maximum at 722 nm. The reduced intensities of the PSI and PSII maxima derive from the reduced chlorophyll levels in Δ*slr1780* and the blue-shifted allophycocyanin peak can be explained by a disruption in energy transfer between the terminal phycobiliprotein pigments and the photosystems, possibly suggesting the phycobiliprotein light harvesting antennae are no longer attached to photosystem complexes (Williams et al., 1980; Shimada et al., 2008).

After excitation at 580 nm (**Figure 4.7 C**) three highly defined maxima at 646 nm, 664 nm and 681 nm are observed. These represent phycocyanin, allophycocyanin and PSII respectively. Although these peaks are all present in the wild type and Δ*slr1780* emission spectra, their intensities differ. The emission maxima from allophycocyanin and PSII are greater in Δ*slr1780* cells than wild type cells. As mentioned earlier, the populations of the major chlorophyll containing complexes, including PSII are significantly decreased in Δ*slr1780*, thus, the increased intensity in the PSII fluorescence maximum is unlikely to represent an increased number of PSII complexes. More plausible is that the increased fluorescence intensity is generated from terminal phycobiliprotein pigments not associated with photosystem reaction centres that lose their excess energy as fluorescence or increased vibration (Shimada et al., 2008; Fuhrmann et al., 2009).

For further characterisation of the relative levels of phycobilisomes and photosystem complexes, the photosynthetic complexes associated with the membrane were separated by sucrose density gradient ultracentrifugation. Cultures of wild type and Δ*slr1780* cells were harvested at mid-log phase (OD_{750} = 0.5 – 0.8), the cells broken and the membrane fractions isolated by ultra-centrifugation. Following normalisation of the membrane fractions to a protein concentration of 20 mg ml^{-1}, the membranes were solubilised with β-DDM at a final concentration of 1 % (v/v) using a homogeniser. The solubilised material was then diluted to a final protein concentration of 1 mg ml^{-1} and 1 mg of protein was separated by ultracentrifugation on a continuous 10 – 30 % sucrose gradient (experiments were adapted from Dühring et al (2006) and performed as described in Section 2.5.7). Images of the thylakoid membranes fractionated by sucrose density gradient ultracentrifugation for wild type and Δ*slr1780* membrane complexes are presented in **Figure 4.8**. Both gradients have three clearly defined coloured bands, which, moving from the slow to fast sedimenting complexes, contain the free pigments, the PSI/PSII monomers and the PSI trimers respectively.

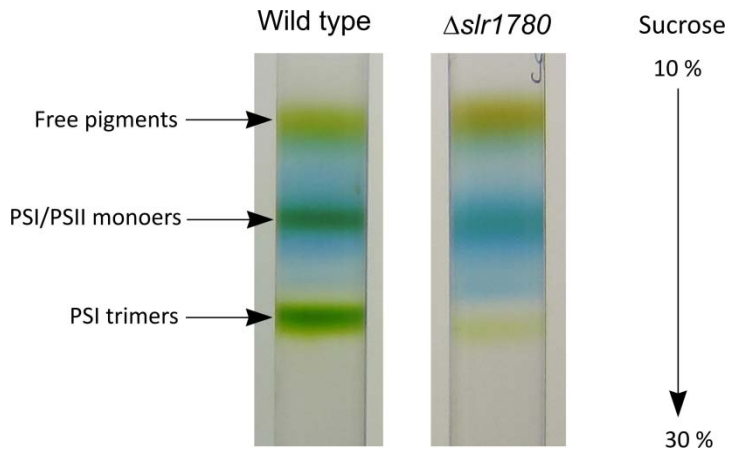

Figure 4.8 Separation of membrane protein complexes in wild type and *Δslr1780* Synechocystis cells via sucrose density gradients

10-30 % continuous sucrose gradients were used to separate membrane complexes, at a protein concentration of 1 mg ml^{-1}, from β-DDM solubilised membranes isolated from wild type and *Δslr1780 Synechocystis* cells. Three clearly coloured bands were resolved, which represent free pigments, PSI/PSII monomers and PSI trimers.

4.2.4 Purification and identification of pigment A432 from Δ*slr1780*

HPLC analysis of methanol extracted pigments from Δ*slr1780* reveals an unknown pigment (pigment A432) that could be a putative intermediate of the cyclase reaction. The pigment was found to elute at 16.7 minutes (**Figure 4.4 A**) and has a Soret peak at 433 nm (**Figure 4.4 B**). A large scale purification procedure was designed to obtain ~1 -2 mg of the pigment A432 in order to solve its identity through a combination of MS and NMR. However, this purification was somewhat hindered by the low abundance of pigment A432. Therefore, an alternative mutant, which accumulates the pigment A432 at a three-fold higher concentration than Δ*slr1780* was used. This mutant (constructed and described in Chapter 5), has a point mutation in the *slr1780* gene and exhibits a similar phenotype to Δ*slr1780*.

Several methods to extract pigment A432 from the culture supernatant or pelleted cells were tested, which ranged from using a solvent-based extraction system or a DEAE cellulose ion-exchange column. However, the best method for extracting pigment A432 was to use phase partitioning with diethyl ether to extract the pigment directly from a cell suspension of R82A, grown to an OD_{750} ~1.5 (described in Section 2.7.3), avoiding all contact with plastic (the pigment is highly sticky and the yield was dramatically reduced if the sample came into contact with plastic). It was found the cells required growth to a relatively high OD in order to accumulate significant amounts of pigment for extraction.

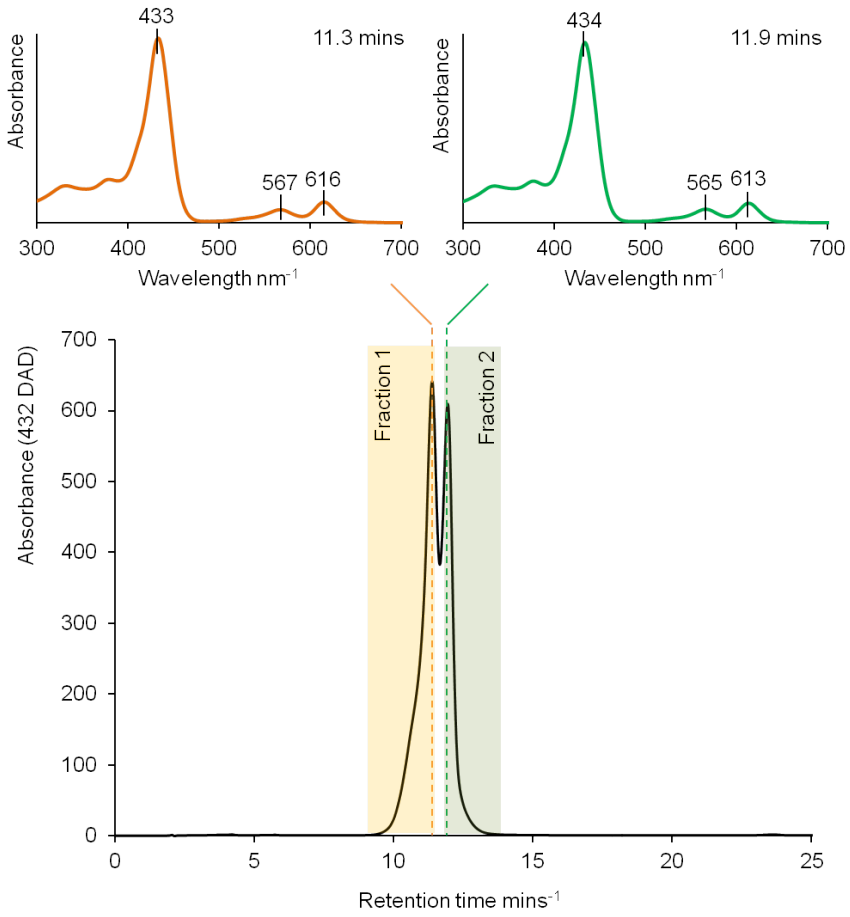

Figure 4.9 HPLC analysis of large scale purification of pigment A432 from *Synechocystis* strain R82A

R82A cells were grown photomixotrophically under low light conditions to an OD_{750} = ~ 1.5. Pigment A432 was extracted from whole cell suspensions with diethyl ether via phase partitioning. The pigments were further purified on a preparative HPLC column using a UniverSil C18 preparative column. The HPLC chromatogram of the purification, recorded at 432 nm, is shown above. The two forms of pigment A432 and their associated absorbance spectra are highlighted. Two samples were collected for NMR and MS analysis, fraction 1 (highlighted in gold) and fraction 2 (highlighted in green).

Analysis of the pigment purified using the large scale protocol revealed the accumulation of two peaks, which were only partially resolved by HPLC, each with the characteristic absorbance spectra of pigment A432 (**Figure 4.9**). Given the similarity of the absorbance spectra and the difficulty of separating these peaks with the use of different HPLC protocols, it is likely these could be monovinyl and divinyl versions of the same pigment, as the DVR enzyme is known to act upon multiple substrates in the chlorophyll biosynthesis pathway (Tripathy and Rebeiz, 1986; Tripathy and Rebeiz, 1988b; Whyte and Griffiths, 1993; Rebeiz et al., 1994). Owing to the close elution times of the two peaks, only the outer regions, highlighted in yellow and green, were collected for further experimental work, to help

minimise cross-contamination of the two samples. From here on in fraction 1 will be referred to as peak 1 and fraction 2 as peak 2.

For MS analysis, ~350 µg of dried pigment A432 was dissolved in 50 µl of LCMS grade methanol and infused, via a syringe pump into a MaXis UHR-TOF instrument (Bruker Daltronics), at a flow rate of 3 µl min^{-1}. For NMR analysis ~0.75 – 1 mg of dried pigment was re-suspended in 500 µl of deuterated methanol, placed in a sample vessel and the 1D spectra acquired at 298 K in a 600 MHz Bruker spectrometer fitted with a cryoprobe.

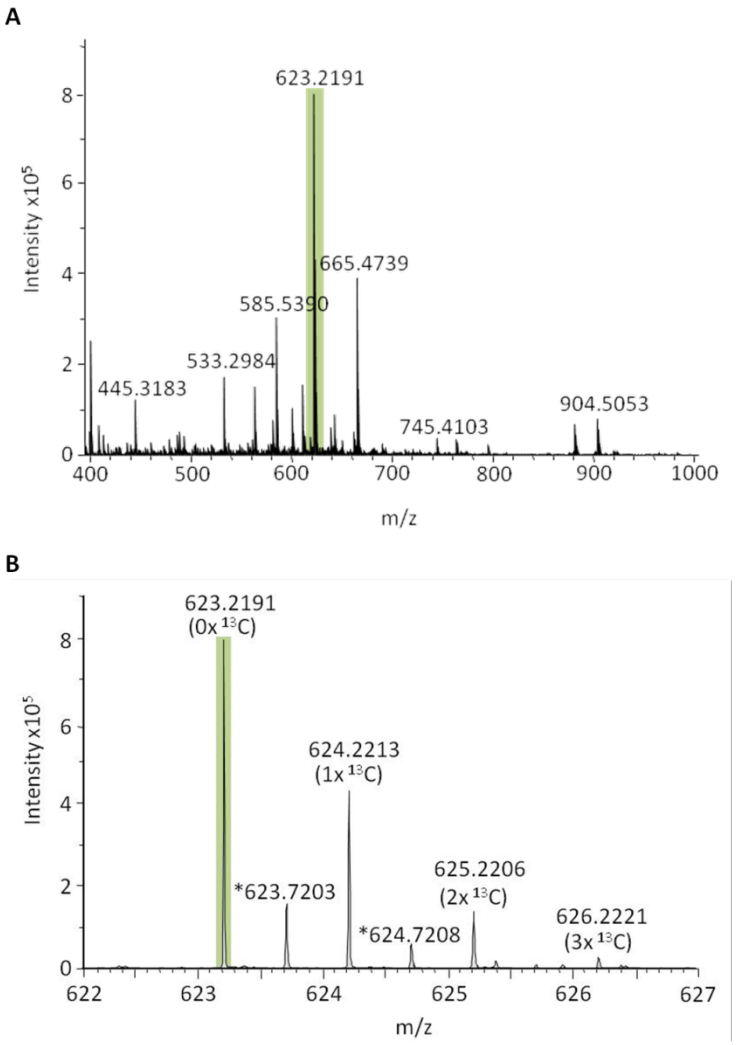

Figure 4.10 MS analysis of pigment A432

Mass spectrum showing the radical cations generated from HPLC purified fraction 1, containing pigment A432 (peak 1) (**A**). Close up of the isotopic cluster peaks of the major product ion at 623.2191 m/z (**B**). MS analysis was performed by Dr Phil Jackson.

Mass analysis using electrospray ionisation (ESI) MS showed that the molecular ion peak of pigment A432 has an m/z of 623.2191 (**Figure 4.10 A**), a mass value that is compatible with the expectations of an m/z around 600 with a fractional mass of 0.20 - 0.24 Da. Upon

magnification of the spectra (**Figure 4.10 B**), the isotopic cluster peaks correspond to a molecule with up to three ^{13}C atoms. To elucidate the chemical composition of pigment A432, the closest matching empirical formula was estimated for a radical cation with a molecular mass of 623.2191 Da, using the MgPME empirical formula as a backbone. The closest match to this mass corresponds to an empirical formula of $C_{36}H_{31}N_4O_5Mg$, which has a molecular mass of 623.2139 Da, with an error of 8 ppm. This error is reasonable given the MS instrument used. The mass difference between MgPME ($C_{35}H_{34}N_4O_4Mg$ = 598.9753 Da) and pigment A432 is 24.9709 Da, which translates to the addition of one carbon and one oxygen, as well as the loss of three hydrogen atoms from MgPME to form to pigment A432. Additionally, the estimated m/z of pigment A432 does not correspond to the mass of either of the proposed cyclase intermediates, β-hydroxypropyl monomethylester or β-ketopropyl monomethylester, whose molecular weights are 614.9748 Da and 612.9589 Da, respectively. This finding suggests pigment A432 may not be a direct intermediate of the cyclase reaction, but rather the product of further, as yet unknown chemical transformation, likely to be driven by the high intracellular concentration of MgPME.

To further probe the structural identity of pigment A432 MS-MS analysis was performed. The target ion at 623.2191 Da (**Figure 4.13 A**) was selected for MS-MS, which is represented by the precursor ion peak at 623.2295 Da in **Figure 4.11 B** (the mass difference between the two ions is due to reduced mass sensitivity during MS-MS analysis). The MS-MS did not yield many product ions, a commonly reported phenomenon when analysing the MS-MS of tetrapyrrole molecules (Alvarez et al., 2012), arising from the innate stability of the tetrapyrrole ring system. Nevertheless, some interesting ions, representing neutral losses from the macrocycle could be detected and are highlighted in **Figure 4.11 B**. Possible fragment ions, representing the loss of side chains from the main macrocycle are shown in **Figure 4.11 A**, along with the structures and molecular weights of the C-13 side chains from the two predicted cyclase reaction intermediates, β-hydroxypropyl monomethylester and β-ketopropyl monomethylester. The neutral loss of 72 Da (highlighted in blue) corresponds to the radical cation generated from the loss of the functional group (CH_2CH_2COOH) attached at C17, indicating this side chain is unaltered in the conversion of MgPME into pigment A432. A neutral loss of 36 Da is also observed, which corresponds to the removal of two H_2O molecules, one of which is likely to be lost from the carboxylic acid side chain attached at C17; the other may arise from the as yet unidentified C13 functional group. Two smaller fragments with masses at 343.2319 Da and 371.2474 Da, with a mass difference of 28 Da are also observed in the MS-MS. It is not clear where these peaks originate from, but they possibly represent fragments of the broken macrocycle, the neutral loss of 28 Da is likely to

represent the loss of a CO moiety. No losses were observed that correspond to the C13 functional groups of the proposed intermediates of the cyclase reaction. This suggests that whilst parts of the macrocycle remain unchanged in comparison to MgPME the functional group attached at C13 has been modified in a manner that is not consistent with the proposed intermediates of the cyclase reaction.

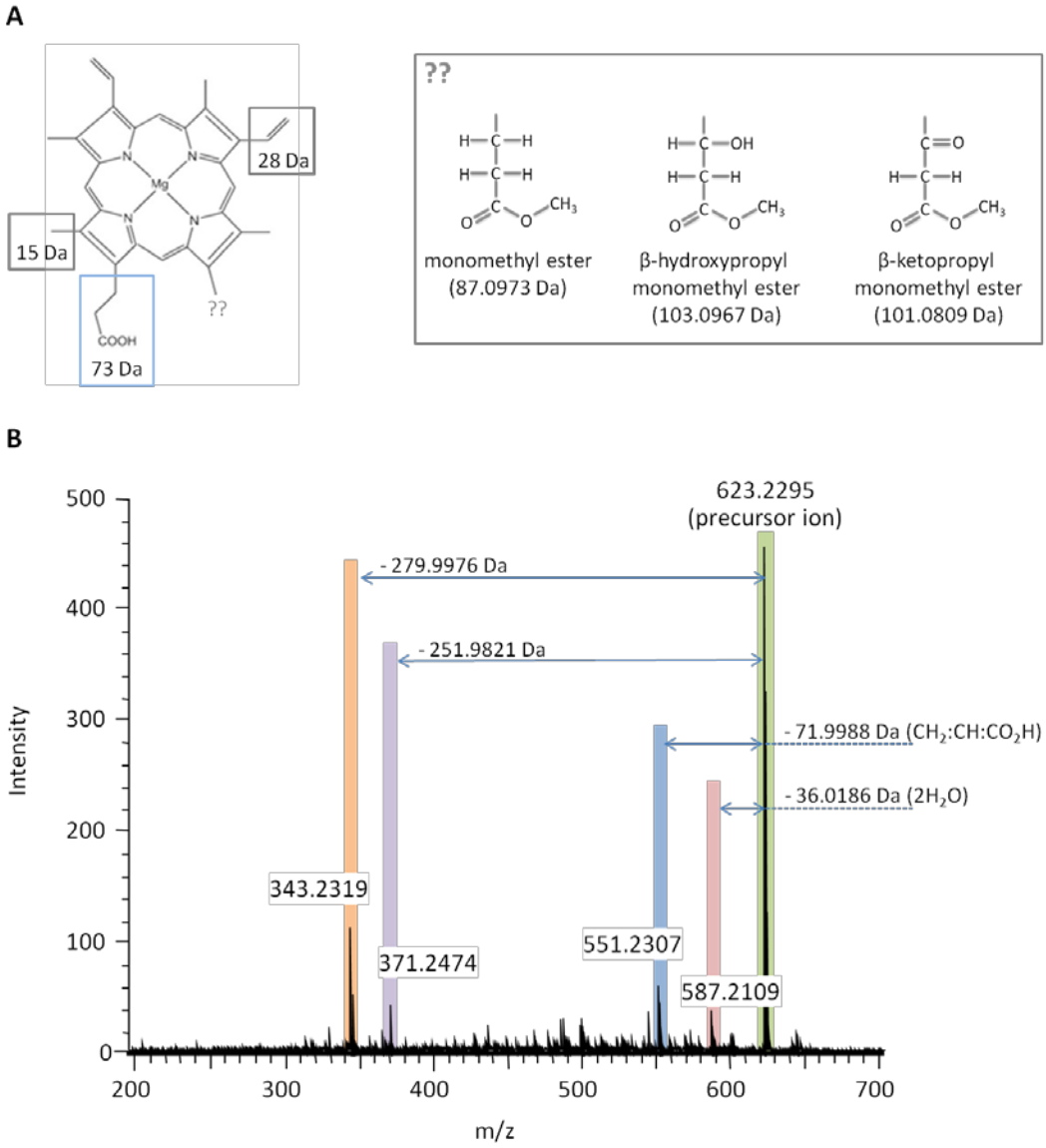

Figure 4.11 MS-MS analysis of 623.2191 Da target ion

Mass spectrum showing the radical cations generated from MS-MS analysis of the 623.2191 product ion. (**A**) Structure of the tetrapyrrole backbone common to both MgPME and Pchlide. The three different side chains likely to be lost during MS-MS are highlighted, and their corresponding molecular weights are given. ?? represents the three potential side chains likely to be attached at position 13C according to existing literature. (**B**) Spectra of product ions generated after MS-MS of the 623.2191 product ion highlighted in **Figure 4.13 A**. Highlighted are the major product ions and their mass difference from the precursor ion (623.2295 Da).

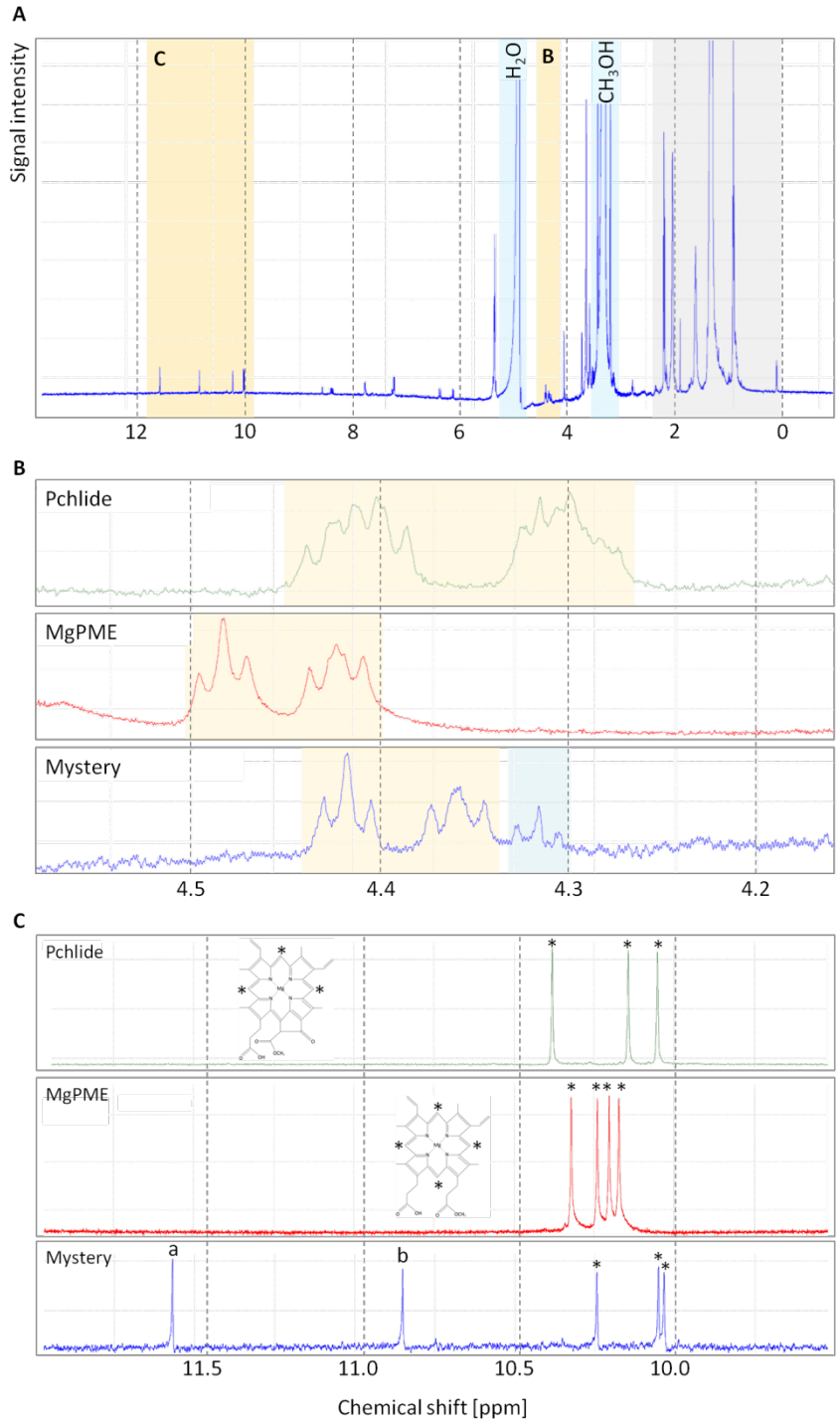

Figure 4.12 1D NMR analysis of pigment A432

1D NMR spectra showing the protons in pigment A432 (**A**), highlighted in blue are the signals generated from methanol and water, highlighted in grey are the signals from contaminants within the buffer and highlighted in orange are the regions of spectra shown in **B** and **C**. (**B**) proton signals from MgPME, Pchlide and pigment A432 with a chemical shift between 4.2 – 4.6 ppm, highlighted in blue is the additional triplet signal from the mystery pigment. (**C**) mesoproton signals from MgPME, Pchlide and pigment A432 with a chemical shift between 9.5 – 12.0 ppm, asterisked on MgPME and Pchlide are the mesoprotons to which the signals correspond; peaks a and b in the pigment A432 spectra represent two additional proton peaks. NMR analysis was performed by Dave Armstrong.

In order to gather a clearer picture as to the structure of pigment A432, 1D NMR was performed and the spectra of pigment A432 was compared to those generated by MgPME and Pchlide. Also present in the 1D NMR spectra of pigment A432 (**Figure 4.12 A**) are several signals from other sample components, such as the intense shifts from methanol and water, at 3.5 ppm and 4.7 ppm, respectively, and the shifts in the low field region, between 0 and 2.5 ppm, which are most likely from remaining column material or remaining impurities from *Synechocystis*. Two regions of interest, in which pigment A432 spectra differs from the 1D spectra of MgPME and Pchlide, are highlighted in yellow (B and C) and magnified for further clarification in **Figure 4.12 B** & **C**.

Figure 4.12 B displays a magnification of the region between 4.2 and 4.6 ppm, which shows a set of three triplets in the spectra of pigment A432, a set of two triplets in the MgPME spectra and a doublet of triplets in the Pchlide spectra. The differences observed between the MgPME and Pchlide spectra are due to the proximity of the Pchlide C17 side chain protons to the cyclised C13 side chain, which in MgPME is not cyclised, hence the signals from the MgPME-C17 side chain protons are not split, as the protons in the C15 and C17 side chains are not close enough in space to exert an effect on each other. These spectra suggest the C13 side chain is not cyclised in pigment A432, as the triplets observed have not been split into a doublet of triplets. However, the presence of a third triplet in this region, highlighted in blue, may indicate an additional CH or CH_2 group on the C13 side chain.

The singlet signals generated by the meso-CH's in MgPME, Pchlide and pigment A432 are magnified in **Figure 4.12 C**; the signals from these protons are shifted into the region between 10 -10.5 ppm due to their proximity to the delocalised tetrapyrrole ring. Typically these appear in the order of C5, C10, C15 and C20 moving up-field (Fookes and Jeffrey, 1989; Alvarez et al., 2012). The four singlet signals from MgPME are from the protons located on C5, C10, C15 and C20, whilst the four singlet signals from Pchlide are from the protons located on C5, C10 and C20. The loss of a signal at C15 in Pchlide is due to the loss of the proton at C15 upon formation of the isocyclic ring. The pigment A432 spectra clearly show three meso-CH singlets between 10 – 10.5 ppm, most probably representing the protons located at C5, C10 and C20. The spectra also shows two additional singlets, signals a and b, located further downfield at 11.6 ppm and 10.9 ppm, respectively. This downfield shift could be caused by two chemical environments, either the presence of an additional ring structure, which magnifies the effect of delocalisation on the proton shift, or the presence of an electron withdrawing substituent, such as the addition of an oxygen. This suggests that the pigment A432 C13 side chain could consist of an isolated proton group, situated next to an additional ring structure or oxygen functional group, which would account for one of the

singlet peaks. The second singlet peak is likely to be from the mesoproton at C15, as the signals shown in **Figure 4.12 B** show the isocyclic ring is unformed, with the further shift arising from the proximity of this proton to the delocalised ring structure or oxygen containing functional group situated on the C13 side chain.

Although the structure of pigment A432 has not yet been fully elucidated, several advances have been made in determining its identity. ESI-MS has enabled the pigment's molecular weight to be determined as 623.2 Da and both ESI-MS and NMR have suggested the presence of an additional oxygen atom, most likely located on the C13 side chain. It is also known that the C17 side chain remains unchanged from MgPME and that the C13 side chain has not cyclised to form the isocyclic group. However, to allow for further elucidation of the structure of pigment A432, additional NMR studies are required, incorporating the use of further structural diagnosis techniques. Similar studies on the structures of tetrapyrrole molecules have shown ROESY (Rotating-frame Overhauser Effect SpectroscopY) analyses to be a powerful tool in determining the distances between protons (Xiliang et al., 2005; Alvarez et al., 2012) and this technique will therefore be used in future studies to obtain the structural identity of pigment A432.

4.3 Discussion

The aim of the work presented in this chapter was to determine if the protein Ycf54 had a functional role in the Mg-cyclase. Based on the available biochemical data, the Mg-cyclase is expected to be active as a multi-subunit complex. This evidence comes from reconstitution of cell extracts from cucumber, barley, *Chlamydomonas reinhardtii* and *Synechocystis*, where both soluble and insoluble fractions of the cell lysate were required for conversion of MgPME to Pchlide to be observed (Walker et al., 1991; Bollivar and Beale, 1996a; Rzeznicka et al., 2005). In addition, genetic analyses show the membrane component is likely to be composed of two proteins; one is a homologue of *AcsF* (Pinta et al., 2002), a known gene, and the other is an as yet unknown protein so far only identified in the barley mutant *viridis-k* (Rzeznicka et al., 2005). Current knowledge of the individual components of the Mg-cyclase is limited to homologues of the *Rubrivivax gelatinosus* AcsF protein. Previous work identified two genes in *Synechocystis*, *sll1214* and *sll1874*, as *acsF* homologues, which encode a membrane associated component of the Mg-cyclase (Minamizaki et al., 2008a; Peter et al., 2009), which contains a putative di-iron site and, thus, is viewed as the true catalytic subunit of the Mg-cyclase (Tottey et al., 2003a).

4.3.1 Ycf54 interacts with the Mg-cyclase catalytic components Sll1214 and Sll1874

To identify other Mg-cyclase components, the N-termini of Sll1214 and Sll1874 was tagged *in vivo* with a 3xFLAG peptide. Pulldown experiments identified the unknown hypothetical protein Ycf54 as an interaction partner in each case (Chapter 3). A reciprocal pulldown assay using N-terminal FLAG-tagged Ycf54 as bait, trapped Sll1214 (**Figure 4.2**). Although Sll1874 was absent from the FLAG-Ycf54 elute, this was not surprising as Sll1874 is reported to only accumulate under micro-oxic conditions (Minamizaki et al., 2008a) and the FLAG-Ycf54 strain was cultured under aerobic conditions.

4.3.2 *Synechocystis* cells deficient in Ycf54 are blocked at the Mg-cyclase step

Analysis of the cellular localisation of Ycf54 (**Figure 4.5**) shows the protein to be highly hydrophilic, with a small proportion associating with the membranes, where it presumably forms an interaction with Sll1214/Sll1874 (**Figure 4.2 C**). Thus, Ycf54 is presumably a candidate protein for the "soluble" component of the Mg-cyclase. To test this hypothesis and to further understand the function of Ycf54, a strain of *Synechocystis* was generated, where the Ycf54 encoding gene *slr1780* was replaced with a zeocin resistance cassette. The resulting strain *Δslr1780* produced a strong phenotype, which greatly impeded cell growth and reduced the level of chlorophyll accumulation to 13 % of wild type (**Figure 4.4**). Most importantly, *Δslr1780* accumulated very high levels of MgPME, the substrate of the Mg-cyclase, and no Pchlide, the product of the Mg-cyclase, could be detected (**Figure 4.4 A**); this is the first evidence for Ycf54 being a newly identified component of the Mg-cyclase.

The absence of Ycf54 has consequences for the accumulation of other chlorophyll biosynthesis enzymes, most notably the AcsF homologue Sll1214 and the geranylgeranyl reductase, ChlP (**Figure 4.5**). The decrease in Sll1214 could itself (to an extent) explain the observed 93 % reduction in chlorophyll content, as a *sll1214*$^{-}$ *Synechocystis* mutant possessing about 50 % of wild type Sll1214 displays a similar decrease in chlorophyll content (Minamizaki et al., 2008a). One possible conclusion to be drawn is that Ycf54 may not be a true catalytic component of the Mg-cyclase, but rather required for the stability and accumulation of the catalytic Sll1214 protein. Alternatively, Ycf54 could facilitate the formation of a catalytic complex between the Mg-cyclase and the chlorophyll biosynthesis enzymes that precede and carry out the Mg-cyclase reaction. Evidence for the formation of a large multi-enzyme chlorophyll biosynthesis complex was found in Chapter 3 of this thesis.

Both AcsF homologues, Sll1214 and Sll1874, were found to interact with CPOIII, POR and ChlP, chlorophyll biosynthesis enzymes found up- and downstream of the cyclase reaction.

4.3.3 Reduction in chlorophyll accumulation in Δ*slr1780* has global effects on chlorophyll containing complexes and the thylakoid membranes

The knock-on effects of reduced chlorophyll accumulation are dramatic and have severe consequences for the accumulation of "green" proteins and the internal cellular architecture of *Synechocystis*. Chlorophyll is known to be essential for the assembly and innate stability of the PSI and PSII complexes (Kim et al., 1994; Kopecna et al., 2013) and 80 % of cellular chlorophyll is associated with PSI when *Synechocystis* is grown under low light conditions (Mimuro and Fujita, 1977; Rogner et al., 1990). In the case of *Δslr1780*, cellular chlorophyll is reduced by 87 % (**Figure 4.3 D**) and both low temperature fluorescence spectroscopy (**Figure 4.7 B** & **C**) and sucrose density gradient analyses (**Figure 4.8**) show PSI is proportionally reduced. Defects in thylakoid membrane formation are known to arise from different limitations, including depletion of photosystem components and assembly factors (Voelker and Barkan, 1995; Meurer et al., 1998; Fuhrmann et al., 2009), galactolipids (Kobayashi et al., 2007), carotenoids (Mohamed et al., 2005) and chlorophylls (Eichacker et al., 1992; Falbel and Staehelin, 1994). *Δslr1780* has reduced accumulation of chlorophyll and PSI, two factors which have been shown to result in undeveloped or a reduced number of thylakoid membranes (seen in **Figure 4.6**). The reduced number of thylakoid membranes could also be a factor in the decreased levels of chlorophyll biosynthesis enzymes ChlP and Sll1214. Both are non-transmembrane proteins that localise with the thylakoid membranes (Srivastava et al., 2005) so it is plausible these proteins could be degraded soon after translation if they cannot properly localise with the thylakoid membrane.

4.3.4 Pigment A432, which accumulates in Δ*slr1780* cells is not one of the proposed cyclase intermediates

In addition to accumulating MgPME, *Δslr1780* was also found to accumulate pigment A432, an unidentified pigment (**Figure 4.4 A**), whose absorption spectra (**Figure 4.4 B**) revealed some intriguing spectroscopic properties. At 432 nm, the Soret peak of pigment A432 falls directly between MgPME (416 nm) and Pchlide (440 nm), suggesting this pigment is an intermediate between the substrate and product of the Mg-cyclase, thus indicating we may have identified an intermediate in the cyclase reaction. However, the absorption characteristics of this pigment do not correspond to the spectroscopic properties of the proposed intermediates β-hydroxypropyl- and β-ketopropyl- monomethylester; the former has absorbance characteristics very similar to those of MgPME (Walker et al., 1988) and the

latter has a similar absorbance spectrum that is red-shifted by a few nanometers (Smith and Goff, 1986). Therefore it was decided to structurally identify the pigment using a combination of mass spectrometry and NMR.

The molecular mass of the pigment had an m/z of 623.21 Da (**Figure 4.8**), which is slightly larger than the molecular masses of MgPME and Pchlide, whose molecular weights are 598.2430 Da and 610.94 Da respectively. This correlates to the addition of one each of carbon and oxygen and the loss of three hydrogen's from MgPME, assuming that pigment A432 is a product originating from MgPME, which is likely given pigment A432 still retains the propanoic side chain situated at C17. NMR analyses further revealed the pigment to retain a non-cyclised side chain at C13, which, according to the downfield shift of the C15 mesoproton (**Figure 4.10 C**) and predicted empirical formula, is likely to contain an oxygen atom and a possible delocalised ring system. Whilst the exact structure of pigment A432 has yet to be fully elucidated, two important conclusions can be drawn. Firstly, pigment A432 appears to retain much of the MgPME backbone, but with a modification at the C13 side chain and secondly, pigment A432 is not either of the proposed intermediates, β-hydroxypropyl- or β-ketopropyl- monomethylester.

4.3.5 Conclusions and future work

Overall the data gathered within this chapter lead to the conclusion that Ycf54 is required for cyclase activity *in vivo* and for the accumulation of chlorophyll and chlorophyll requiring complexes. However, whether Ycf54 has a role in catalytic Mg-cyclase activity, or if the protein plays a critical role in the synthesis/maturation of Sll1214/Sll1874 or in the assembly of a cyclase complex is unclear and is further investigated through a series of *in vivo* point mutations in Chapter 5.

The work performed in this chapter complements the work published by Hollingshead S, et al. (2012), which was performed using an unstable partially segregated *slr1780*$^{--}$ mutant obtained from Professor T Ogawa. Analysis of the two *slr1780* mutant strains showed that both strains synthesised very low levels of chlorophyll (*Δslr1780* has a 93 % reduction and *slr1780*$^{--}$ has a 70 % reduction compared with wild type cells) and accumulated significant levels of MgPME, indicating a "blockage" at the cyclase step. However, differences are observed in the accumulation of the chlorophyll biosynthesis enzymes. Whilst both strains exhibited a reduction in the AcsF subunits Sll1214/Sll1874, the *slr1780*$^{-}$strain was found to have reduced levels of ChlM and POR, whilst *Δslr1780* retained wild type levels of these proteins. Additionally, *Δslr1780* was found to have reduced levels of ChlP, whilst *slr1780*$^{-}$did not. It is likely these differences result from the *slr1780*$^{-}$strain was constructed and grown;

Slr1780^{-}was constructed by the insertion of an erythromycin resistance cassette between 175 bp and 176 bp of the *slr1780* gene (data obtained by sequencing the *slr1780* region), which means half of the 321 bp gene will be expressed, and could possibly interfere with cellular mechanisms and *slr1780*$^{-}$was grown in medium containing 10 µg ml^{-1} erythromycin, an antibiotic that works by binding to the 50s ribosome subunit, preventing the transfer of tRNA from the A site of the ribosome to the P site and protein synthesis (Weisblum, 1995).

These data suggest the biosynthesis of the chlorophyll-containing complexes PSI and PSII occur independently of the biosynthesis of the phycobilisomes light harvesting antennae with which they are associated. This is an interesting phenomenon as it is not known why a cell would synthesise complex protein complexes that are surplus to requirements. Additionally, the phycobiliproteins not associated with photosystem reaction centres will absorb actinic light whose energy cannot be dissipated through charge separation, but rather will be lost as heat or down other pathways likely to be harmful to the cell. Therefore, future work is likely to focus on understanding the relationship between phycobilisome and PS biosynthesis and how these pathways are regulated with respect to each other.

Analyses of the chlorophyll biosynthesis enzymes and light harvesting complexes showed a huge difference in their accumulation between wild type and *Δslr1780* cells. However, these experiments were limited to probing for known enzymes with a limited number of available antibodies. In order to assess the global effects caused by the Ycf54 deficient phenotype, quantitative MS will be performed comparing the proteome of *Δslr1780* and N_{15} labelled wild type cells. Such a study may result in the identification of previously undiscovered proteins involved in the regulation of chlorophyll biosynthesis, photosystem assembly and possibly the unidentified subunit(s) of the cyclase.

Although several steps have been taken to solve the identity of pigment A432, further experiments are required to obtain full structural identity. As discussed earlier, future work to fully solve the structure of pigment A432 is likely to involved analysing a larger quantity of pigment, cleaned up to contain fewer impurities in the up-field region, by ROSEY NMR.

5. Introduction of point mutations into the native *slr1780* (Ycf54) gene to determine the role of Ycf54 in the Mg-protoporphyrin IX monomethylester oxidative cyclase complex

5.1 Introduction

The previous chapter described how Ycf54, a protein required for cyclase activity, was identified through pulldown experiments with the known oxidative cyclase components Sll1214 and Sll1874. A *Synechocystis* Ycf54 knockout mutant (*Δslr1780*) was created and the effects of this mutation on the cell were characterised. These analyses showed that the *Δslr1780* strain to be severely impaired at the step of the Mg-protoporphyrin IX monomethylester oxidative cyclase, which in turn led to the *Δslr1780* cells being almost devoid of chlorophyll. Additionally, it was found *Δslr1780* had reduced accumulation of the chlorophyll biosynthesis enzymes Sll1214 and ChlP and that this mutant contained virtually no thylakoid membranes, when thin sections were imaged using electron microscopy. Taken together these results lead to the conclusion that Ycf54 is required for cyclase activity *in vivo* and for the accumulation of chlorophyll and chlorophyll-requiring complexes. However, the exact function of Ycf54 has not been established. For example, it is not known if Ycf54 is required for the assembly and stability of an active cyclase complex, or if the protein is essential for catalytic activity of the cyclase. If Ycf54 is not an essential component of the cyclase, it may be required to promote stability of the cyclase or a higher chlorophyll biosynthesis complex. This hypothesis stems from the significant reduction in Sll1214 and ChlP observed in the *Δslr1780* mutant. Furthermore, Sll1214 is a critical component of the cyclase and a reduction in this protein alone would explain the blockage at the cyclase step.

To determine whether Ycf54 has a catalytic or complex stabilisation/formation role it was decided to identify highly conserved residues within the Ycf54 gene and systematically mutate these to alanine *in vivo*. In the absence of a working cyclase complex, *in vitro* enzyme assays cannot be performed, therefore the observation of point mutations on the *Synechocystis* phenotype is a valid approach for determining the role of Ycf54 in cyclase function. Additionally, many groups have successfully used *in vivo* point mutations to study the function of proteins that cannot easily be performed *in vitro* (Chamovitz et al., 1993; Shen et al., 1993; Vermaas, 1993; Giese and Vierling, 2002).

In this chapter the role of seven absolutely conserved residues on Ycf54 function is investigated through a variety of techniques including, absorbance and fluorescence spectroscopy, HPLC analyses, FLAG-pulldown affinity experiments and western blots.

5.2 Results

5.3.1 Identification of conserved residues in the Ycf54 gene

A BLASTp search of the *Synechocystis* Ycf54 gene was performed to identify homologues in other photosynthetic organisms. Ycf54 genes from fourteen different organisms, representing a diverse spectrum of oxygenic photosynthesisers, were aligned and conserved residues identified using ClustalW2. The alignments reveal a conserved core domain of ninety residues, seven of which are absolutely conserved (**Figure 5.1**). The N-terminal domain extension (highlighted in green) found in the plant and algal proteins is a plastid localisation sequence required to target the nuclear encoded Ycf54 to the site of photosynthesis inside the chloroplasts.

All of the conserved seven residues are found within discrete groups of high similarity, suggesting these regions may be important for Ycf54 function. With reference to the *Synechocystis* Ycf54 primary sequence the seven conserved residues are: Ala9 (A9), Phe13 (F13), Glu22 (E22), Glu26 (E26), Asp39 (D39), Phe40 (F40) and Arg82 (R82). To determine if these residues are required for Ycf54 function *in vivo* a construct was made to individually mutate six of these residues to alanine (F13, E22, E26, D39, F40 and R82A) and A9 to a glycine. Currently, there is no reproducible assay to monitor cyclase activity *in vitro*, therefore one of the only methods available for probing the function of these residues is by *in vivo* point mutations. By placing each of these point mutations in the native gene under the control of the native *slr1780* promoter, rather than the light driven *psbAII* promoter used for expression of FLAG-tagged proteins, the true effect of these point mutations on the whole cell physiology will be observed.

5.2.2 Construction of a vector to introduce point mutations into *slr1780*

Homologous recombination in *Synechocystis* requires the alignment of ~300 bp of homologous DNA. To allow for a recombination event at the *slr1780* locus, a plasmid was constructed containing the *slr1780* gene and the 300bp directly upstream of the start codon, followed by a chloramphenicol resistance cassette and the 500 bp directly downstream of *slr1780* (**Figure 5.2**). The plasmid was constructed by amplifying the two regions of *Synechocystis* DNA from *Synechocystis* genomic DNA (using primers Slr1780 Upstream F +

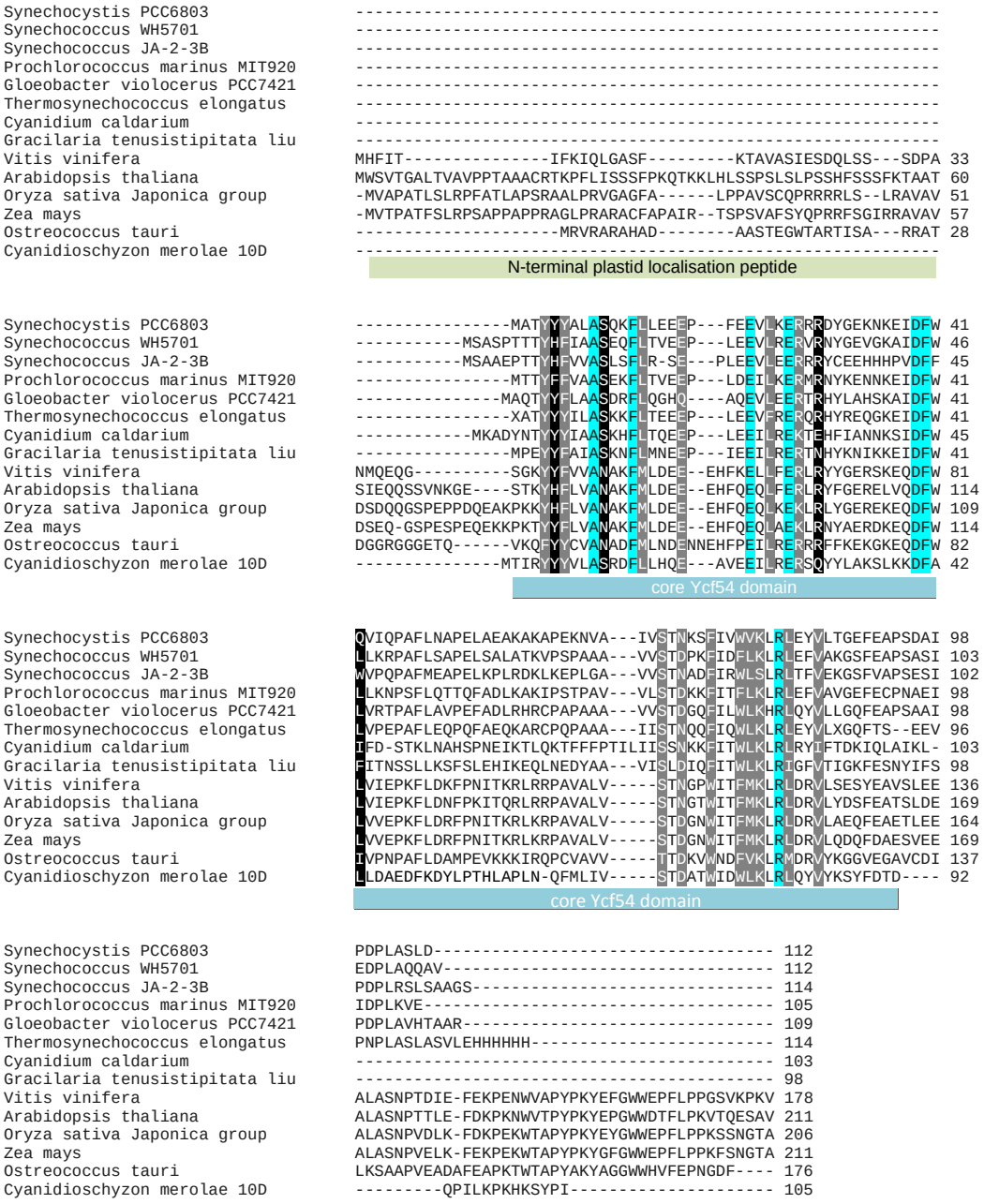

Figure 5.1 Sequence alignment showing the evolutionary conserved residues of the Ycf54 domain

Amino acid sequence alignments of proteins predicted to contain the Ycf54 domain, generated using ClustalW2. Sequences were obtained from: primordial cyanobacterium, *Gloeobacter violoceus*; cyanobacteria, *Synechocystis* sp PCC6803, *Synechococcus* sp WH5701 and JA-2-3B, *Prochlorococcus marinus* MIT920 and *Thermosynechococcus elongatus*; green plants, *Vitis vinifera*, *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays*; red algae *Cyanidium caldarium*, *Cyanidioschyzon merolae* and *Gracilaria tenusistipitata* and the green alga *Ostreococcus tauri*. Conserved, highly similar and similar residues are highlighted in blue, dark grey and black respectively.

Slr1780 Upstream R and Slr1780 Downstream F + Slr1780 Downstream R) digesting them with appropriate restriction enzymes and ligating them into pET3a, so as to flank a chloramphenicol resistance cassette (PCR amplified using primers Cm^R F + Cm^R R) inserted at the multiple cloning site. The resulting construct was confirmed by restriction digests and sequencing and named pPM-*slr1780*.

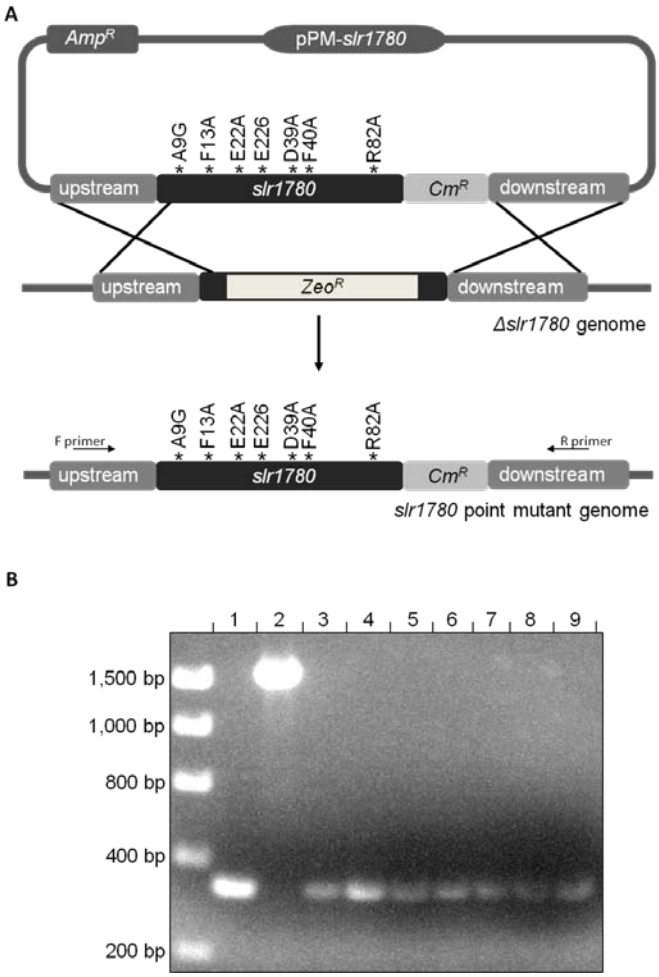

Figure 5.2 Construction of a *slr1780 in vivo* point mutation cassette

(**A**) Diagram of the cassette constructed to insert individual point mutations into the *Δslr1780* genome, via homologous recombination of the up- and down-stream regions. (**B**) Isolation of fully segregated *Δslr1780* knockout mutant and *slr1780* point mutation mutants, confirmed by colony PCR by amplifying the *slr1780* locus (shown by F primer and R primer). Lane 1, WT; lane 2, *Δslr1780*; lane 3, A9G; lane 4, F13A; lane 5, E22A; lane 6, E26A; lane 7, D39A; lane 8, F40A; lane 9, R82A.

The Stratagene QuickChange site directed mutagenesis kit was used to introduce the seven point mutations of interest into the *slr1780* gene located on pPM-*slr1780*. The primers used to introduce these point mutations are listed in Table 2.5. Individual point mutations were confirmed by sequencing of pPM-*slr1780* before their transformation into the *Synechocystis* genome.

To ensure the point mutations are fully recombined into the *Synechocystis* genome the constructs were transformed into the *Synechocystis Δslr1780* mutant, which lacks the native *slr1780* gene (transformations and selection on chloramphenicol were performed as described in Section 2.3.1). The transformants were serially plated onto BG-11 agar plates containing 20 μg ml^{-1}, 40 μg ml^{-1}, 80 μg ml^{-1} and 160 μg ml^{-1} chloramphenicol. Complete segregation was confirmed by PCR analysis (**Figure 5.2**) using the primers Ycf54 F and Ycf54 R.

To confirm the point mutations were in all copies of the genome and in the correct location, the *slr1780* region was amplified with primers *Δslr1780* KO F1 and *Δslr1780* KO R2 and sequenced, with the sequencing analysis confirming all the point mutation transformations as successful.

5.2.4 Analysis of the slr1780 point mutants

Several methods were employed to determine if any of the seven point mutations had an aberrant effect on cyclase activity or other components of the photosynthetic apparatus. Analyses of the whole cell spectra and photosynthetic precursor pigments revealed disruptions in chlorophyll accumulation, whilst native PAGE and western blot analyses revealed differences in the assembly of the photosynthetic apparatus and expression of the chlorophyll biosynthesis enzymes.

5.2.4.1 Spectroscopic analysis of *slr1780* point mutant strains

In order to determine any differences in the physiology of the *Synechocystis slr1780* point mutants, whole cell spectra were recorded and compared to wild type and *Δslr1780* strains. Each of the strains was grown photomixotrophically under low light conditions in BG-11 medium supplemented with 5 mM glucose to an OD_{750} of 0.5-0.6. The cells were normalised for OD_{750} = 0.6 and whole cell spectra for each sample was measured by monitoring absorbance between 350 nm and 800 nm on a UV-VIS spectrophotometer (**Figure 5.3 A**).

Differences in the whole cell pigments were observed by comparing the spectra of methanol extracted pigments. Methanol extracts were obtained from cell pellets of *slr1780* mutants normalised for OD_{750} = 1.0. These were washed in distilled water, before the pigments were extracted as described in Section 2.7.1. The methanol extracted pigments were analysed by UV-VIS spectroscopy by monitoring absorbance between 350 nm and 750 nm (**Figure 5.3 B**).

As a control, the pPM-*slr1780* plasmid containing the wild type *slr1780* gene and no point mutations was transformed into the *Δslr1780* mutant, enabling any effects caused by the insertion of the chloramphenicol resistance cassette to be determined. This strain is referred to as the Cm^R control. Comparison of whole cell and methanol extracted spectra from the

Cm^R control and wild type strains show the introduction of a chloramphenicol resistance cassette has no adverse effect on either of the absorbance spectra. Therefore, any variances between the spectra of the point mutants and wild type *Synechocystis* will be a result of introducing a specific point mutation and not an effect of the chloramphenicol resistance cassette (**Figure 5.3**).

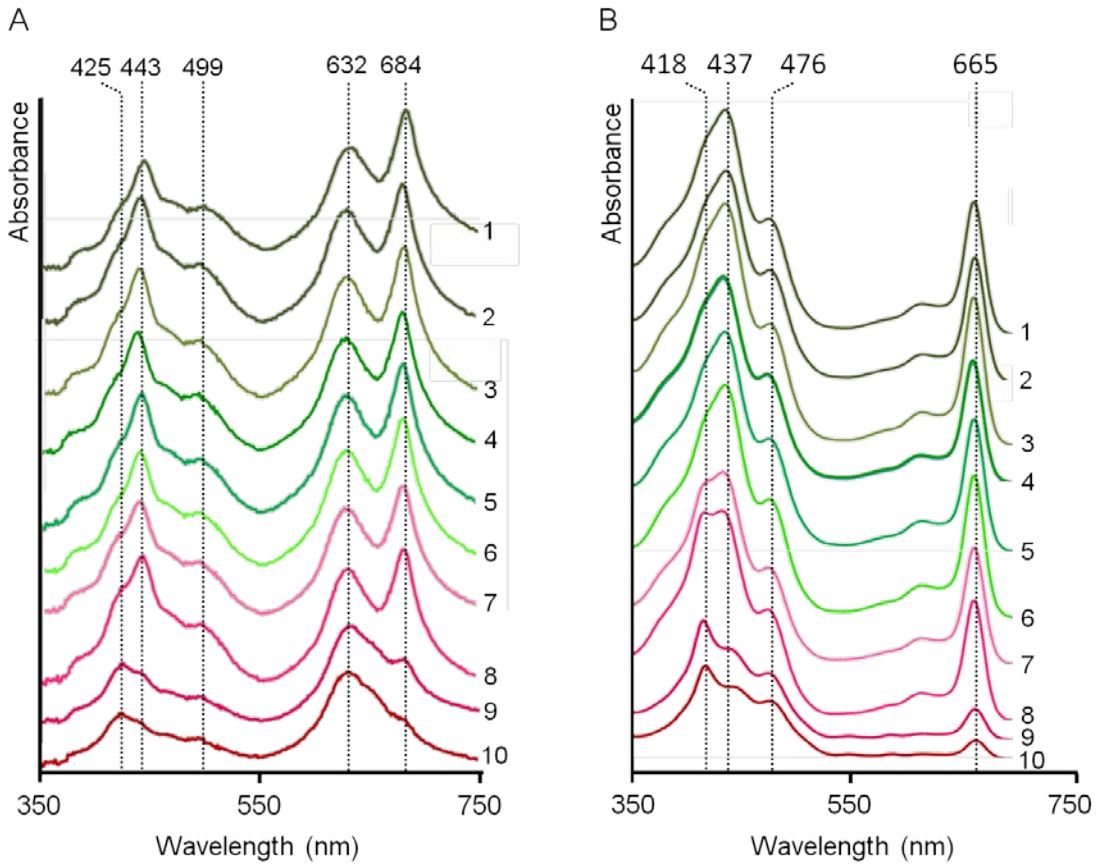

Figure 5.3 Whole cell and methanol extracted pigment spectra of *slr1780* mutant strains

Synechocystis strains were grown photomixotrophically under normal light conditions and standardised for OD_{750} = 0.6. (**A**) Whole cell absorbance spectra recorded for samples normalised to light scattering at 750 nm. (**B**) Absorbance spectra of pigments extracted in 90 % methanol from cell pellets. 1, WT; 2, *Cm*^{R} control; 3, A9G; 4, F13A; 5, E22A; 6, E26A; 7, D39A; 8, F40A; 9, R82A; 10, *Δslr1780*.

The whole cell spectra of the majority of *slr1780* point mutants (A9G, F13A, E22A, E26A, D39A and F40A) are virtually identical to wild type indicating these point mutations have little or no discernible effect on whole cell physiology. Also, the methanol extracted pigment spectra of mutants A9G, F13A, E22A and E26A are identical to wild type, indicating pigment biosynthesis in these strains is unaffected by any of these point mutations. Interestingly, the methanol extracted pigment spectra of D39A and F40A both have an additional peak at 420 nm, which is not found in wild type and not detectable in the whole cell spectra. This

suggests that D39A and F40A have an aberration in the biosynthetic pathway of a visible pigment, which most likely translates into an aberration in the chlorophyll biosynthesis pathway.

Of all the point mutants, R82A exhibits the most interesting phenotype; both the spectra taken of R82A are very similar to the spectra of *Δslr1780*. The whole cell spectrum shows that whilst the absorbance peak of the phycobilisomes at 632 nm is barely affected, there is a significant reduction in the absorbance peaks of the photosystems at 443 nm and 684 nm. The methanol extract spectrum also shows peaks at 416 nm and 442 nm, which differ from the blue-end peaks found in the wild type spectra at 437 nm and 476 nm. Additionally, the chlorophyll *a* absorption peak at 665 nm is significantly reduced in R82A to the level observed in *Δslr1780*, illustrating chlorophyll is unable to accumulate in this mutant, an early sign of an aberration in chlorophyll biosynthesis.

5.2.4.2 Growth rates and chlorophyll *a* content of *Synechocystis slr1780* point mutant strains

To determine the effects of each point mutation on photoautotrophic growth and chlorophyll *a* biosynthesis, the doubling time and chlorophyll *a* content of cells in mid-log phase were ascertained.

Photoautotrophic growths were performed in triplicate under normal light conditions in BG-11 medium supplemented with 10 mM TES pH 8.2. Growths were set up from starter cultures of all strains grown under low light conditions in BG-11 medium supplemented with 10 mM TES and 5 mM glucose. The starter cultures were standardised by OD_{750}, washed in BG-11 medium to remove the glucose, and used to inoculate the photoautotrophic growth cultures to a starting OD_{750} = 0.1. The starter cultures used were grown in medium supplemented with glucose, as *Δslr1780* and R82A were unable to achieve significant growth without the aid of an external carbon source. The doubling time of each mutant was monitored by taking cell density readings at OD_{750} every 12 hours for a total of 108 hours. The doubling times (**Table 5.1**) were calculated (**Equation 5.1**) for each strain by taking the exponent (u) from the equation for the trend-line plotted of each growth curve.

Equation 5.1 Calculation of doubling times (hours)

$$y = n \cdot e^{ux}$$

$$\text{Doubling time (hours)} = \frac{u}{\ln 2}$$

The chlorophyll *a* content was calculated in triplicate from methanol extracts of cultures grown photoautotrophically, under normal light conditions (40 µmol photons m^{-1} s^{-1}),

normalised for cell density at OD_{750} = 0.6, using **Equation 2.1** (Porra et al., 1989), as described in Section 2.7.2.

The doubling times and chlorophyll *a* content calculated for the wild type and Cm^R control strains are almost identical, demonstrating that insertion of a chloramphenicol resistance cassette directly downstream of the *slr1780* gene has no effect on cell growth and accumulation of chlorophyll *a*.

Table 5.1 Photosynthetic growth rates and chlorophyll *a* content of *slr1780* mutants

Strain	Doubling time		Chlorophyll *a*	
	Hours	fold increase compared to WT	$mg \cdot L^{-1} \cdot OD_{750}^{-1}$	% reduction compared to WT
WT	32 (± 1)	_	3.4 (± 0.1)	_
Δslr1780	120 (± 12)	4	1.0 (± 0.4)	72
control	32 (± 1)	1	3.2 (± 0.4)	6
A9G	33 (± 3)	1	3.0 (± 0.1)	15
F13A	33 (± 3)	1	3.1 (± 0.3)	10
E22A	33 (± 2)	1	3.3 (± 0.2)	3
E26A	33 (± 1)	1	3.0 (± 0.1)	12
D39A	33 (± 1)	1	2.0 (± 0.1)	40
F40A	34 (± 1)	1	2.0 (± 0.1)	40
R82A	79 (± 4)	3	0.7 (± 0.1)	81

A9G, F13A, E22A, E26A, D39A and F40A, have doubling times ranging between 32 and 34 hours, which are similar to the doubling time of 31 hours calculated for the wild type. With a doubling time of 79 hours, R82A was the only mutant to have a significantly retarded growth rate in comparison to wild type, although this is still significantly less than the doubling time of 120 hours calculated for *Δslr1780* (**Table 5.1**).

Comparison of the chlorophyll *a* levels reveals that chlorophyll *a* accumulation is reduced by 40 % in both D39A and F40A, and by 81 % in R82A (**Table 5.1**). The other point mutants exhibit less significant reductions in chlorophyll *a*, which range between 3 % for E26A and 15 % for A9G. Taken together these data imply that although D39A and F40A are unable to accumulate chlorophyll *a* at the levels of wild type, the growth of these strains under photoautotrophic conditions is not significantly affected, suggesting these mutants are capable of synthesising enough chlorophyll to meet the demands of the photosynthetic apparatus under the growth conditions studied.

These results also indicate that the critical chlorophyll level per cell, i.e. the amount of chlorophyll *a* required for normal growth, falls between 2.04 $mg/L/OD_{750}$ and 0.97

mg/L/OD_{750}, the amount of chlorophyll *a* found in D39A/F40A and R82A/ *Δslr1780*. Interestingly, the amount of chlorophyll *a* was calculated to be lower in R82A (0.65 mg/L/OD_{750}) than in *Δslr1780* (0.97 mg/L/OD_{750}), which is unexpected given *Δslr1780* has a longer doubling time than R82A, suggesting it is less capable of growing under photoautotrophic conditions.

5.2.4.3 HPLC analysis of chlorophyll precursor pigments in *slr1780* point mutants

The methanol extracted pigments were separated by HPLC to determine if chlorophyll photosynthetic precursor pigments were responsible for the unusual absorbance spectra observed for mutants D39A, F40A and R82A (**Figure 5.3 B**). HPLC analyses were performed as described (Section 2.8 and Table 2.7) on a 250 mm x 4.16 mm C_{18} Phenomenex reverse phase column. The retention times and absorbance spectra (obtained from the online DAD) of the two pigments Pchlide (27.7 mins) and Mg-protoporphyrin IX (17.6 mins) are shown for reference (**Figure 5.4**).

Figure 5.5 displays the retention times and absorbance spectra of each of the photosynthetic precursor pigments found in the slr1780 point mutant strains. Generally, each mutant falls into one of two classes, wild type-like spectra or *Δslr1780*-like spectra. The point mutants that had methanol extracted pigment spectra similar to wild type (A9G, F13A, E22A and E26A) have HPLC absorbance spectra virtually identical to that of wild type, where the major pigments observed are monovinyl (26.4 mins) and divinyl (27.7 mins) Pchlide and synechoxanthin. The point mutants that fall into the *Δslr1780*-like spectra are D39A, F40A and R82A. Each of these mutants exhibits two additional peaks, located at 16.9 mins and 29.8 mins, which are not observed in wild type. The 16.9 minute pigment corresponds to pigment A432 whose Soret peak is located between that of MgPME and Pchlide (see Chapter 4 for details) and the 29.8 minute peak is MgPME, the substrate of the cyclase. The R82A mutant was found to accumulate greater amounts of pigment A432 than the *Δslr1780* mutant, therefore this mutant was used as the source of pigment A432 for structural studies using NMR and mass spectrometry

.

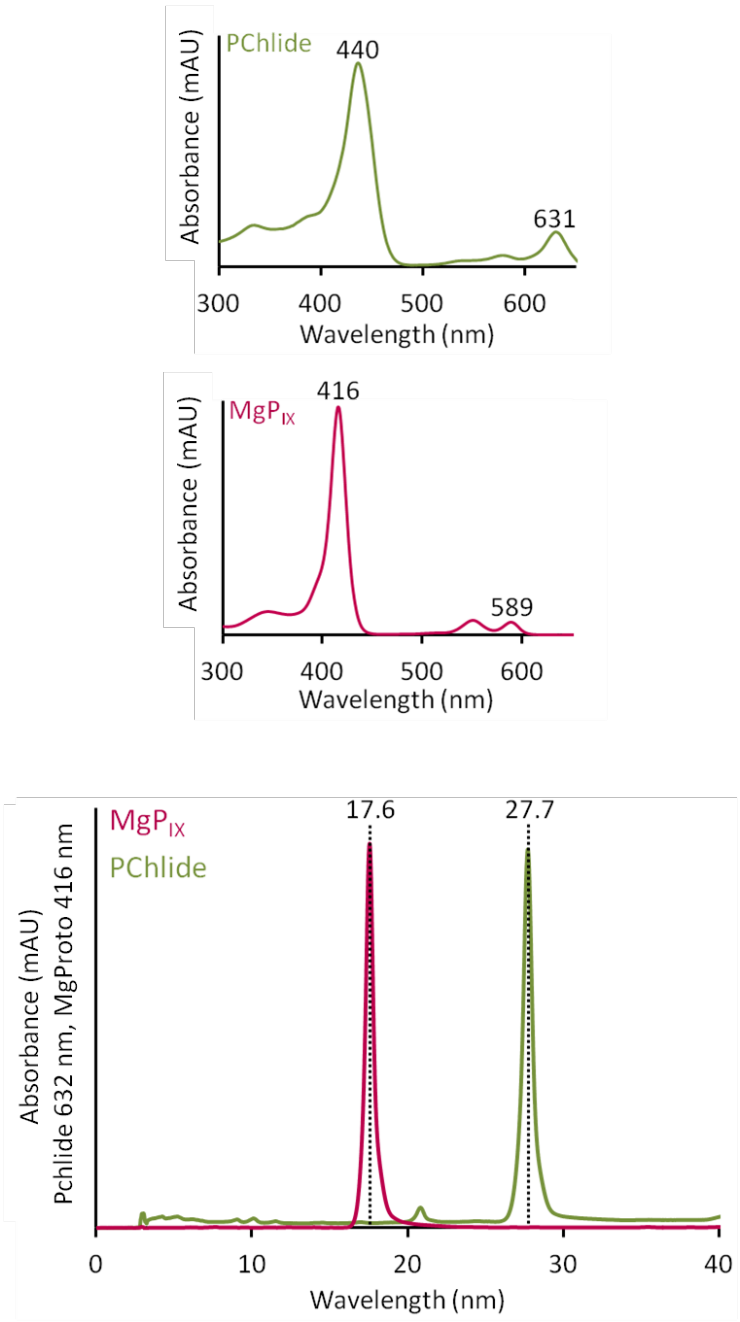

Figure 5.4 Retention times of MgProto and Pchlide standards

Retention times and absorbance spectra of pigment standards; MgProto, retention time of 17.6 mins, recorded at DAD 416; Pchlide, retention time = 27.7 mins, recorded at DAD 632. Pigments were separated on a Phenomenex 250 mm C_{18} column, using the precursors method detailed in Table 2.7.

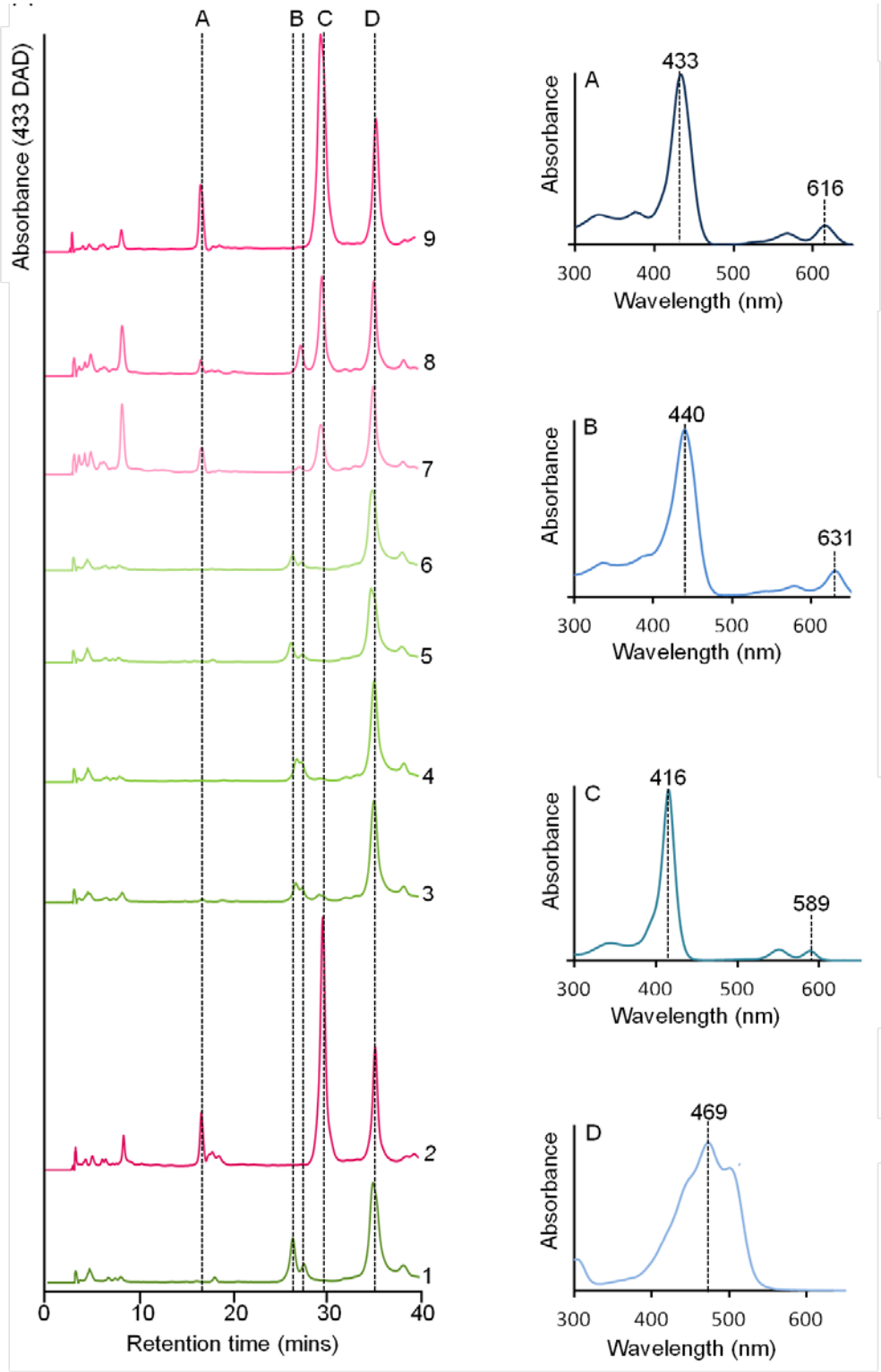

Figure 5.5 HPLC analyses of methanol extracted pigments from wild type and slr1780 mutant strains

HPLC analyses of methanol extracted pigments from wild type and slr1780 mutants grown under photomixotrophic conditions and normalised for absorbance at OD_{750}. 1, WT; 2, *Δslr1780*; 3, A9G; 4, F13A; 5, E22A; 6, E26A; 7, D39A; 8, F40A; 9, R82A. The retention times of peaks of interest and their respective absorbance spectra are; A, pigment A432, 16.5 mins; B, MV and DV Pchlide, 26.4 and 27.7 minutes respectively; C, MgPME, 29.8 mins; D, synechoxanthin, 35.3 mins. Pigments were separated on a Phenomenex 250 mm C_{18} column, using the precursors method detailed in Table 2.7.

5.2.4.4 Analysis of the levels of chlorophyll biosynthesis enzymes and photosystem I accumulation in the *slr1780* point mutants

In order to elucidate the hypothesis that Ycf54 is a factor required for the accumulation of chlorophyll biosynthesis enzymes or the formation of a higher chlorophyll biosynthesis complex, the accumulation of chlorophyll biosynthesis enzymes in the slr1780 mutants was determined by western blot and the accumulation of photosystem I (PSI) trimers was determined by clear native-PAGE (CN-PAGE).

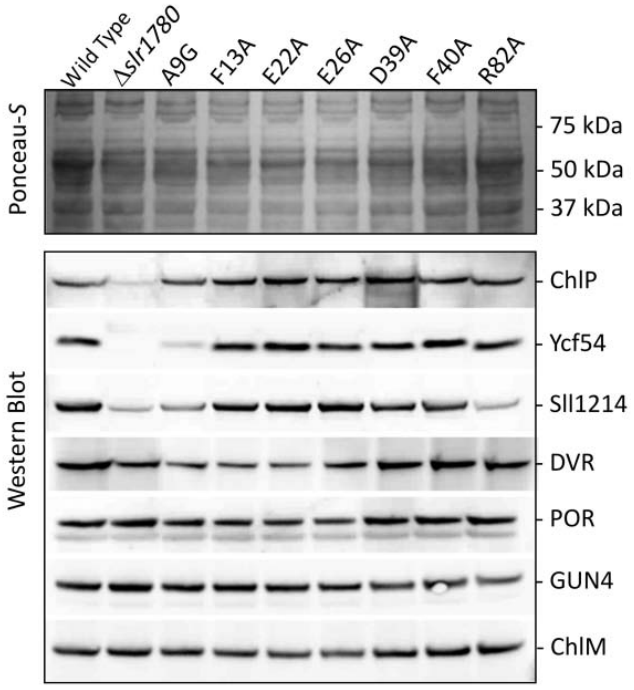

Figure 5.6 Analysis of levels of chlorophyll biosynthesis accumulation in slr1780 mutant strains

Western blot analysis of the levels of chlorophyll biosynthesis enzymes in *slr1780* mutants. Whole cell lysates of each of the Δ*slr1780* mutant strains, grown under photomixotrophic conditions, were normalised for protein concentration and the proteins separated by SDS-PAGE, before transfer to a nitrocellulose membrane. The blots were probed with antibodies to the chlorophyll biosynthesis enzymes, ChlP, Ycf54, Chl27 (Arabidopsis homologue to Sll1214), DVR, POR, GUN4 and ChlM. Lane 1, WT; lane 2, *Δslr1780*; lane 3, A9G; lane 4, F13A; lane 5, E22A; lane 6, E26A; lane 7, D39A; lane 8, F40A; lane 9, R82A.

Cultures of each of the *slr1780* mutants were grown under photomixotrophic conditions (due to the inability of R82A and the *Δslr1780* mutant to grow under photoautotrophic conditions) and whole cell lysates, normalised for protein concentration, were prepared for western blotting as described in Section 2.6.7. For each mutant 40 µg protein was separated by SDS-PAGE and transferred to a nitrocellulose membrane (Section 2.5.2), which was sequentially probed with antibodies to the chlorophyll biosynthesis enzymes: ChlP, Chl27 (Sll1214), Ycf54,

DVR, POR, GUN4 and ChlM. The membrane was also stained with Ponceau-S to confirm equal protein loadings (**Figure 5.6**).

The levels of the chlorophyll biosynthesis enzymes, POR, GUN4, ChlP and ChlM, remain constant in all of the mutant strains when compared to wild type *Synechocystis*. However, the accumulation of Sll1214/Sll1874 is significantly reduced in strains R82A and A9G (**Figure 5.6**, lanes 3 and 9). This reduction in Sll1214/Sll1874 in R82A (lane 9) and *Δslr1780* (lane 2) could be responsible for the near abolition of cyclase activity observed, as indicated by levels of the MgPME substrate (**Figure 5.5**). However, this hypothesis is thrown into doubt when taking into account the impediment of cyclase activity in the F40A and D39A strains, which accumulate Sll1214 at levels comparable to wild type, and the wild type phenotype exhibited by the A9G strain, which also has reduced levels of Sll1214. This suggests Ycf54 may be required for catalytic activity of the cyclase, rather than as an assembly, regulatory or stability subunit of Sll1214.

Interestingly the levels of ChlP are not affected in any of the point mutants, which is surprising given the significant reduction of ChlP observed in *Δslr1780*. One reason for this could be that the stability of ChlP is dependent upon the presence of thylakoid membranes (ChlP is a membrane localised protein) and the lack of thylakoid membranes in *Δslr1780* leads to the degradation of ChlP, hence its absence in *Δslr1780*. A similar theory could also be put forward for the disappearance of Sll1214/Sll1874, which is also a membrane localised protein. However, this would not explain the loss of Sll1214/Sll1874 in A9G and R82A, both of which contain levels of ChlP akin to wild type.

CN-PAGE was used to determine the effect of the point mutations on PSI assembly. Accumulation of PSI trimer complexes was investigated by separating solubilised thylakoid membranes by CN-PAGE (Figure 5.7 A), as described in Section 2.5.7. Cultures of each of the mutant strains were grown photomixotrophically under low light conditions to an OD_{750} = 0.7 - 0.8. Membranes were then isolated from the cells and solubilised as described in Section 2.6.7; 500 µg of protein was then separated on a 4-16 % Bis-Tris native gel (Invitrogen). The high molecular weight PSI was separated from the rest of the cellular complexes by electrophoresis performed overnight at 4°C.

The majority of the mutants, with the exceptions of R82A and *Δslr1780*, were found to accumulate PSI trimers at a level comparable to wild type (Figure 5.7 A). Although R82A only accumulated very low levels of PSI trimers, this is a slight improvement on *Δslr1780*, which appears to accumulate no PSI trimers.

Low temperature fluorescence microscopy was used to further probe the accumulation of PSI and the phycobiliprotein light harvesting antennae in wild type, A9G, D39A, F40A and R82A *Synechocystis* strains. The fluorescence emission spectra of cells re-suspended to an OD_{750} = 0.1 in 80 % glycerol were recorded for excitation at 435 nm (analysis of chlorophyll fluorescence) and 580 nm (analysis of phycobiliprotein fluorescence). These spectra were then normalised to the PSII maxima at 680 nm (**Figure 5.7 B**) or the phycocyanin maximum at 464 nm (**Figure 5.7 C**); of these spectra, only A9G was found to exhibit PSI chlorophyll fluorescence akin to wild type cells. The three mutant strains with deficiencies in chlorophyll biosynthesis all exhibit varying degrees of reduction in PSI fluorescence. D39A and F40A both have a small reduction in PSI fluorescence, which is to be expected in cells whose chlorophyll accumulation is reduced by a third, but was not reflected in the CN-PAGE analysis. R82A has a severe decrease in PSI fluorescence, which is almost, but not quite, as severe a reduction as observed in *Δslr1780*; this corresponds to the visualisation of a small amount of PSI trimer in R82A, but not *Δslr1780* on the CN-PAGE (**Figure 5.7 A**). Thus, when chlorophyll is depleted, a knock-on effect is observed in the PSI/PSII ratio, with all the mutants containing a lower ratio of PSI to PSII. Given that 80 % of cellular chlorophyll is associated with PSI, it is to be expected that a depletion in chlorophyll will have a directly proportional effect on the amount of chlorophyll available to be sequestered into PSI complexes and therefore on the number of PSI complexes that can be assembled. As discussed in Chapter 4, the PSII fluorescence at 693 nm is seen only as a shoulder of the PSII maximum at 682 nm. These two peaks are normally of equal height in cells grown mixotrophically under constant light conditions.

The fluorescence emission spectra from the phycobiliproteins, obtained by exciting at 580 nm, show several changes in the mutant cells in comparison to wild type, with the exception of A9G, which has similar phycobilisome fluorescence to wild type. In all spectra, there are fluorescence maxima at 646 nm, 664 nm, 680 nm and 722 nm, which correspond to the phycobiliproteins phycocyanin, allophycocyanin, PSII and PSI, respectively. The phycobiliprotein vibrational bands can also been seen as a small maximum at 745 nm (Szalontai et al., 1985; Szalontai et al., 1994; Yokono et al., 2008).

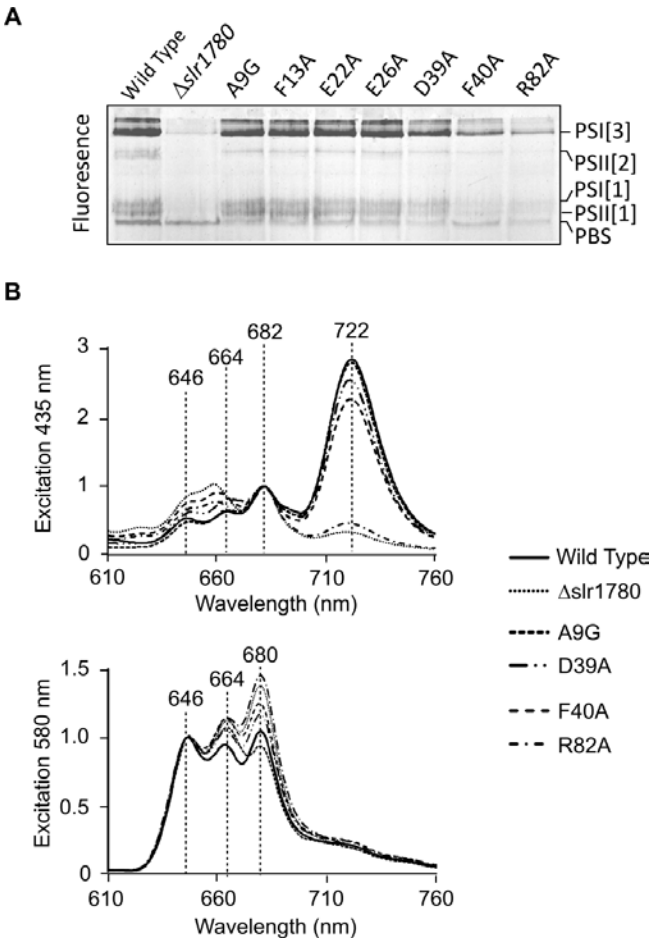

Figure 5.7 Analysis of photosystem accumulation in wild type and *slr1780* mutant *Synechocystis* cells

(**A**) Membrane fractions isolated from each of the *slr1780* mutant strains were separated by CN-PAGE and visualised by recording fluorescence emission at 680 nm and 549 nm. Indicated are the PSI and PSII complexes. (**B**) 77K whole cell emission spectra from *Synechocystis* strains wild type, Δ*slr1780*, A9G, D39A, F40A, R82A. Spectra were recorded for emission at 435 nm and 580 nm. For comparability the 435 nm spectra were normalised to PSII emission at 682 nm and the 580 nm spectra were normalised to phycocyanin emission at 646 nm.

A general trend can be observed with the chlorophyll deficient mutants with the fluorescence emission peak from PSII at 680 nm gradually increasing in proportion to the decreasing levels of chlorophyll in the mutant, with the order of fluorescence being A9G < WT < D39A < F40A < *Δslr1780* < R82A. Due to the limiting levels of chlorophyll in these cells, the increase in PSII fluorescence in comparison to PC fluorescence is unlikely to represent an increase in PSII molecules. In a previous study on *Synechocystis* cells depleted in the CP43 protein Shimanda and co-authors (2008) observed a large emission maximum at 680 nm, proposed to arise from the lowered ratio of PSII to phycobilisomes, resulting in uncoupled phycobiliprotein complexes that lose their energy via fluorescence from the terminal pigment.

5.2.5 Is the expression of Ycf54 altered in mutant A9G?

In **Figure 5.6** it was noted Ycf54 accumulation in the point mutant strain A9G was reduced to less than a quarter of the level observed in wild type, which is of particular interest considering the cell spectra, cellular pigments, growth rate and chlorophyll *a* content of A9G were not significantly different to wild type. To rule out this observation as an artefact of the point mutation affecting the Ycf54 antibody epitope, recombinant A9G (produced as described in Chapter 6) was tested for antigenicity to the Ycf54 antibody (**Figure 5.8 A**). The western blot, which compares the antigenicity of wild type and A9G recombinant Ycf54 proteins, shows the Ycf54 antibody recognises A9G and wild type Ycf54 proteins equally well, proving the reduction of Ycf54 observed in A9G is a true physiological result.

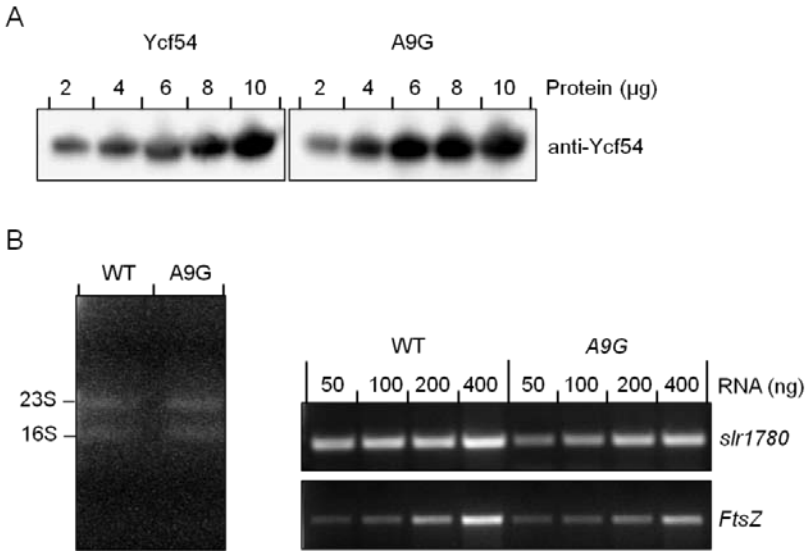

Figure 5.8 RT-PCR analysis of *slr1780* expression in wild type and A9G *Synechocystis* strains

(**A**) Comparison of the Ycf54 antibody reactivity to Ycf54 and A9G. Five concentrations (2 µg, 4 µg, 6 µg, 8 µg and 10 µg) of purified recombinant Ycf54 protein were probed with the Ycf54 antibody. (**C**) 100 ng of total RNA extracted from wild type and A9G cells grown to D_{750} ~ 0.6 under photomixotrophic conditions. (**D**) RT-PCR analysis of *slr1780* and *FtsZ* expression levels in WT and A9G strains.

Next, the mRNA expression level of A9G was ascertained using RT-PCR and compared to that of wild type. The total RNA was extracted from each strain (rRNA can be seen in **Figure 5.8 B**) and the mRNA converted to cDNA, using random primers and reverse transcriptase (as described in Section 2.4.8). As a control the expression levels of the housekeeping gene *ftsZ* were also determined. **Figure 5.8 C** displays the mRNA expression levels of both *slr1780* and *ftsZ*. Both genes were found to be expressed at slightly lower levels in A9G than wild type. As the slightly reduced mRNA levels are comparable and consistent between *ftsZ* and *slr1780*, this is probably not enough of a reduction to explain the dramatic decrease in Ycf54

expression. This leads to the conclusion that the transient levels of *slr1780* are not significantly affected in A9G, and is therefore not likely to be the root cause of the significantly reduced accumulation of Ycf54.

5.2.5 Can the Ycf54 mutants with disrupted chlorophyll biosynthesis interact with Sll1214/Sll1874?

The data gathered so far indicates the conserved residues D39, F40 and R82 are crucial for the accumulation of chlorophyll at wild type levels, whilst the conserved A9 is essential for accumulation of Ycf54. If either D39, F40 or R82 are replaced with an alanine, aberrant chlorophyll biosynthesis is observed via a blockage at the cyclase step, whereas if A9 is replaced with a glycine, a significant reduction in Ycf54 is observed, but cyclase activity appears to remain unaffected. To determine if cyclase activity is dependent upon Ycf54 interacting with Sll1214/Sll1874, each of these mutants was N-terminally FLAG-tagged and pulldown experiments were performed.

Synechocystis strains containing the FLAG-tagged Ycf54 protein with one of the following point mutations, A9G, D39A, F40A or R82A were generated as described in Section 2.3.1. These strains were constructed using QuickChange mutagenesis to insert one of the four point mutations into the pFLAG-Ycf54 vector (constructed in Chapter 4, Section 4.3.1). These were then transformed into *Δslr1780*, so as to remove the native Ycf54 protein, which could potentially out-compete the FLAG-tagged Ycf54 mutant for the binding site on Sll1214/Sll1874. Full segregation was achieved by sequential selection on increasing concentrations of kanamycin and confirmed by PCR. These strains were named FLAG-A9G/ *Δslr1780*, FLAG-D39A/ *Δslr1780*, FLAG-F40A/ *Δslr1780* and FLAG-R82A/ *Δslr1780*.

FLAG-pulldown experiments were performed using cultures grown under photomixotrophic conditions as described in Section 2.6.9. The resulting elution fractions were analysed for the presence of Sll1214/Sll1874 and FLAG-tagged Ycf54 proteins by western blot (**Figure 5.9**). Equal volumes of the elution fraction from the FLAG-ycf54 mutant pulldowns were separated by SDS-PAGE and the proteins were transferred to a nitrocellulose membrane, which was probed with anti-Chl27 (Agrisera) and anti-FLAG (AbCam).

Figure 5.9 shows that all the FLAG-Ycf54 mutant proteins, with the exception of FLAG-A9G, are expressed at levels comparable to wild type FLAG-tagged Ycf54, whereas FLAG-A9G, like its untagged counterpart, accumulates at a reduced level. According to this blot the root cause of the disruption in cyclase activity is the inability of FLAG-D39A, FLAG-F40A and FLAG-R82A to interact with the Sll1214/Sll1874 subunits. Although FLAG-A9G is present in reduced

quantities, the protein is still able to interact with Sll1214/ Sll1874 at levels comparable to wild type Ycf54, suggesting there is enough Ycf54 available for the cyclase to function. Thus, it would appear the interaction of Ycf54 with Sll1214/Sll1874 is absolutely essential for normal function of the cyclase. Whether Ycf54 has a role in catalysis of the cyclase reaction, or is required to activate the cyclase complex remains unclear and requires further elucidation.

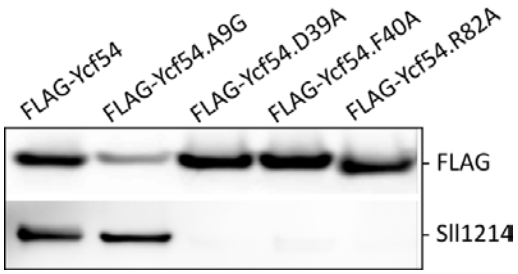

Figure 5.9 Western blot analyses of FLAG-tagged mutant pulldown experiments

FLAG-Ycf54 was purified from the dodecyl-β-maltoside solubilised membrane fractions from *Synechocystis* strains, FLAG-Ycf54, FLAG-Ycf54.A9G, FLAG-Ycf54.D39A, FLAG-Ycf54.F40A and FLAG-Ycf54.R82A. Eluted proteins from the FLAG-pulldown assays were resolved by SDS-PAGE and transferred by Western blot to a nitrocellulose membrane. The membrane was probed with anti-Chl27 (Agrisera), which detects Sll1214 and anti-FLAG (Sigma), which detects FLAG-tagged Ycf54.

5.3 Discussion

In the previous chapter, which was dedicated to investigating the effects of a *Synechocystis* mutant deficient in Ycf54, the Ycf54 protein was found to be essential for Mg-cyclase function *in vivo*, although its precise role remained unclear; two hypotheses, as to this protein's function, were proposed. The first is that Ycf54 is a catalytic component of the Mg-cyclase and thus, could be the soluble component of the enzyme, which was first identified in assays using whole cell lysates, but whose identity is as yet unknown (Walker et al., 1991; Bollivar and Beale, 1996a). The second postulates that Ycf54 plays a critical role in the synthesis/maturation of Sll1214/Sll1874 or in the assembly of the Mg-cyclase complex or a higher order chlorophyll biosynthesis complex, as the accumulation of Sll1214 and chlorophyll biosynthesis enzyme ChlP were dramatically reduced in *Δslr1780*. To further investigate the role of Ycf54, a series of *Synechocystis* mutants, with targeted point mutations within the *slr1780* gene, were generated. These point mutations are all at the location of absolutely conserved residues, identified by aligning Ycf54 sequences from fourteen diverse oxygenic photosynthetic organisms (**Figure 5.1**). Of the seven mutant strains made, four had a detectable phenotype; these were mutants A9G, D39A, F40A and R82A.

5.3.1 Ycf54 point mutants are "blocked" at the Mg-cyclase

The three mutant strains D39A, F40A and R82A, all accumulated MgPME (**Figure 5.5**), the substrate of the cyclase, and all had reduced chlorophyll accumulation (**Table 5.1**), evidence that all three of these residues are required for normal function of the Ycf54 protein. In contrast the mutant strain A9G accumulated ~ 75 % less Ycf54 protein than wild type cells (**Figure 5.6**), but showed no other detectable phenotype (**Figure 5.7** and **5.9**), evidence which suggests Ycf54 is present in excess, as only a small proportion of the synthesised protein appears to be required by the Mg-cyclase. Without a working *in vitro* assay for the Mg-cyclase, probing the exact effects of these point mutations of Mg-cyclase activity is difficult and analysis has therefore focused on examining the physiological effects these point mutations have on the whole cell.

Of the three mutant strains found to have reduced chlorophyll accumulation, R82A exhibited the most severe phenotype. Under photoautotrophic conditions, the doubling time of R82A was 2.5 fold greater than wild type and chlorophyll was reduced by 81 %, whilst the doubling time of *Δslr1780* was 3.8 fold longer than wild type, but chlorophyll accumulation was reduced to a lesser extent by 72 %. Why chlorophyll accumulation is reduced by a further 10 % in R82A than *Δslr1780*, when the doubling time of *Δslr1780* is significantly longer than that of R82A is not known and requires further investigation. Interestingly, the chlorophyll content of *Δslr1780* grown under mixotrophic conditions (measured in Chapter 4) is reduced by 93 %, suggesting the mode of growth can impact on chlorophyll accumulation. In mutants D39A and F40A, chlorophyll accumulation was reduced by 40 %, suggesting that although chlorophyll biosynthesis is impaired in these two strains, the impediment is not as severe as in R82A. Indeed, the growth rates of these mutants under photoautotrophic conditions and normal light intensity were not significantly altered in comparison to wild type (**Table 5.1**). Furthermore, the HPLC traces (**Figure 5.5**) show that while both D39A and F40A accumulate significant amounts of MgPME, both strains are capable of synthesising detectable amounts of Pchlide, whereas R82A and *Δslr1780* are not. Thus, D39A and F40A could be considered to exhibit "leaky" cyclase function, as the enzyme is capable of operating, but at reduced levels. Further elucidation of the effect these mutations have on Mg-cyclase activity could be resolved by detailed kinetic studies; however, this is not yet possible, as not all of the components for a working Mg-cyclase complex are known.

Perhaps the most interesting point mutation is A9G as, although this strain was found to accumulate wild type levels of chlorophyll (**Table 5.1**) and photosynthetic precursor pigments (**Figure 5.5**) and exhibit wild type growth rates (**Table 5.1**), western blot analysis showed

Ycf54 accumulation to be reduced by ~75 % (**Figure 5.6**) and Sll1214/Sll1874 to be reduced to the levels found in R82A and *Δslr1780*. This result is significant, as it disproves an earlier hypothesis that the reduction in Sll1214/Sll1874 in *Δslr1780* is the sole agent causing the severe impediment at the Mg-cyclase step. Additionally, these data show Sll1214/Sll1874 accumulation is dependent on the level of Ycf54, further supporting the proposition that Ycf54 may play a critical role in the assembly/stability of the Mg-cyclase complex and its constituents. It may be the case that Ycf54 plays a similar role to that of the GUN4 protein and the MgCH complex, where GUN4 not only acts to lower the Mg^{2+} threshold required for MgCH activity (Larkin et al., 2003a; Reid and Hunter, 2004; Davison et al., 2005b), thus stimulating MgCH activity under physiologically relevant conditions, but also has a wider role in regulating the chlorophyll biosynthesis enzymes and flux down the tetrapyrrole biosynthesis pathway (Sobotka et al., 2008b; Peter and Grimm, 2009; Formighieri et al., 2012).

5.3.2 Ycf54 interaction with Sll1214/Sll1874 is essential for normal Mg-cyclase function

One of the most significant results deduced from the study of *slr1780* point mutations, is that the interaction of Ycf54 with Sll1214/Sll1874 is essential for normal chlorophyll biosynthesis. This observation was deduced from experiments with FLAG-tagged Ycf54 constructs of each of the point mutations A9G, D39A, F40A and R82A. Analyses of the pulldown elution fractions from each of these strains showed the three mutants with deficiencies in chlorophyll biosynthesis, D39A, F40A and R82A, were unable to interact with the catalytic cyclase component Sll1214/Sll1874, whereas the FLAG-A9G construct, which like the A9G point mutation was present at lower levels, pulled down Sll1214/Sll1874 in similar quantities to those observed in the FLAG-Ycf54 control pulldown (**Figure 5.9**). These results explain to some extent the observable lack of a phenotype in A9G, as they suggest the level of Sll1214/Sll1874 capable of interacting with Ycf54 is consistent between FLAG-Ycf54 and FLAG-A9G. Indeed, pulldown experiments (**Figure 4.2**), showed only the FLAG-Ycf54 located within the insoluble fraction was capable of interacting with Sll1214/Sll1874 and analyses of the sub-cellular localisation of Ycf54 showed only a small minority of this protein was located in the insoluble membrane fraction (**Figure 4.8**). Therefore it is reasonable to conclude that although A9G is only capable of accumulating a small amount of Ycf54, there is more than enough protein available to interact with Sll1214/Sll1874, thus, Mg-cyclase activity is allowed to proceed unhindered. It is not unusual for an enzyme within the chlorophyll biosynthesis pathway to be present in excess, as *Synechocystis* mutants with 5 – 10 % of wild type levels of FeCH are not deficient in haem and have no other discernible phenotype (Sobotka et al.,

2005; Sobotka et al., 2011); similarly *Synechocystis* mutants with reduced levels of ChlG have no growth or pigmentation defects (Sobotka 2012, unpublished data). The nature of the interaction of Ycf54 with Sll1214/Sll1874 is unknown and can only be truly elucidated with structural studies of these proteins. It may be the case that residues D39A, F40A and R82A do not form the specific interactions with Sll1214/Sll1874 that normally enable the proteins to form a complex, but rather that these mutations cause a global structural alteration in Ycf54 that prevents these proteins from interacting. Without structures of the native Ycf54 and the associated point mutations, it cannot be hypothesised as to the kind of effect these mutations have on the intrinsic make-up of Ycf54. Therefore, one of the aims of this project was to solve the structure of Ycf54 as well as the structures of the point mutants that have a detectable phenotype, hence, the crystal structures of several Ycf54 proteins have been solved and are reported in Chapter 6.

5.3.3 Ycf54 mutants affect accumulation of chlorophyll biosynthesis proteins and photosystems

In Chapter 4, it was shown the accumulation of ChlP was differentially regulated by varying expression and transcription of *slr1780* (**Figure 4.6 & 4.8**); however, A9G and R82A were found to accumulate ChlP at levels akin to wild type (**Figure 5.6**). The cause for the differences in ChlP accumulation is as yet unresolved and could be influenced by several different factors, such as the number of thylakoid membranes or the quantity of Ycf54/Sll1214/Sll1874 proteins. ChlP is known to be a non-transmembrane protein, associated with the thylakoid membranes (Shpilyov et al., 2005), therefore it could be hypothesised that association of ChlP with the thylakoid membranes is a requirement for this protein's stability. In *Δslr1780* the thylakoid membranes were found to be reduced (**Figure 4.10**), which was reflected in the dramatic reduction in PSI fluorescence (**Figure 5.7**). It is entirely possible the reduction of thylakoid membranes in *Δslr1780* adversely affects the stability of ChlP. However, the point mutant R82A also exhibits a similarly dramatic reduction in PSI fluorescence and accumulates noticeably fewer chlorophyll complexes than wild type (**Figure 5.7**), suggesting the number of thylakoid membranes in this mutant are similarly decreased, but western blot analyses (**Figure 5.6**) show ChlP accumulation to be akin to wild type. Therefore, it would appear the state of the thylakoid membranes may not be an influencing factor on ChlP accumulation. To further investigate this observation, electron microscopy analysis of R82A thin layer sections would be required to ascertain the number and condition of the thylakoid membranes in this mutant. Additionally, the accumulation of ChlP was similarly unaffected by the reduction of Sll1214/Sll1874 in R82A and A9G, and Ycf54 in A9G. In Chapter 3, ChlP was identified as an interaction partner of Sll1214 and thus, it was

suggested in Chapter 4 that the reduction of Sll1214/Sll1874 or Ycf54 in *Δslr1780* may also influence the accumulation of ChlP, as the stability of these proteins may be dependent upon them forming a "higher" multi-enzyme complex. However, as shown by the western blot in **Figure 5.6**, this appears not to be the case, as the point mutants with decreased Sll1214/Sll1874/Ycf54 all accumulate ChlP to the levels observed in wild type; therefore the exact cellular process, by which the levels of ChlP are altered by differential expression of Ycf54 (**Figure 4.6**) remains to be elucidated.

5.3.4 Ycf54 may have a wider role in regulating components of the photosynthetic apparatus

Generally it is known synchronisation exists between the biosynthesis of chlorophyll and the expression of chlorophyll binding proteins, which allows for chlorophyll to be titrated into chlorophyll-binding complexes upon synthesis, thus avoiding the harmful accumulation of chlorophylls and chlorophyll precursors (Thomas, 1997; Yaronskaya et al., 2003). Additionally, the pathway is further complicated by the need to direct intermediates down the alternative haem pathway, a step known to be regulated by interplay between the Mg-CH and Fe-CH enzymes, which sit at the branch point and whose expression/activities are controlled by light (Hihara et al., 2001), the redox state of the photosynthetic electron transport machinery (Hihara et al., 2003) and diurnal/ circadian rhythms (Papenbrock et al., 2000; Matsumoto et al., 2004). Whether the variable disruption at the Mg-cyclase step has further reaching consequences with regard to porphyrin flux down the chlorophyll and haem pathways and on the expression/activities of enzymes located upstream of the branch-point on this pathway is not yet known. Tight post-transcriptional regulation of the Mg-cyclase catalytic proteins Sll1214 and Sll1874 is known to exist, presumably representing a control point in the chlorophyll pathway, although a depletion of their interaction partner, Ycf54, is known not to affect the expression of the genes encoding these two proteins (**Figure 4.6**), but is known to decrease expression of *chlP* (**Figure 4.6**) and affect the accumulation of Sll1214/Sll1874 and ChlP (**Figure 5.6**). However, the R82A point mutation in the Ycf54 protein, which results in a matching phenotype to *Δslr1780,* does not result in a decrease in ChlP accumulation (**Figure 5.6**), suggesting the role Ycf54 plays in Mg-cyclase activity and its effects on ChlP expression/accumulation are unrelated. Thus, it is possible Ycf54 may have a role in the expression or stability of other enzymes in the chlorophyll biosynthesis pathway, including the unknown entity of the Mg-cyclase complex; whether this is the case could be further elucidated by comparing the relative transcriptome or proteome of *Δslr1780* verses wild type cells. A study of the transcriptome isolated from a *Chlamydomonas reinhardtii* mutant deficient in *GUN4* revealed significant alterations in expression of protein encoding

genes for light-harvesting complexes, tetrapyrrole biosynthesis, signal transduction, transcription and chromatin remodelling (Formighieri et al., 2012). One of the differentially expressed genes from this study, which corresponds to the unknown-hypothetical gene *slr0483* in *Synechocystis*, was found to interact with protoporphyrinogen oxidase, FeCH and Sll1214 using pulldown experiments (Sobotka 2013, personal communication) and the subsequent *Synechocystis* knock-out mutant had reduced PSI, PSII and significantly retarded growth, indicating an important, but as yet unidentified, role in the photosynthetic machinery (see Chapter 7 for more details). There is plausible evidence that such genome/proteome wide approaches can be used to unearth unknown, yet important, regulatory or enzymatic factors required for normal photosynthetic function, such as the missing Mg-cyclase subunit.

6. Crystallisation and Structure Determination of Ycf54

6.1 Introduction

In Chapter 4 it was discovered that the Ycf54 protein is required for conversion of MgPME to Pchlide, a reaction catalysed by the cyclase enzyme complex. Large amounts of MgPME were found to accumulate in *Synechocystis* cells devoid of Ycf54 and no Pchlide could be detected (**Figure 4.6**). As would be expected, for a mutant severely hampered at one of the steps in chlorophyll biosynthesis, the amount of chlorophyll *a* synthesised by Δ*Ycf54* was dramatically reduced; upon quantification Δ*Ycf54* was found to contain 92 % less chlorophyll than an equivalent amount of wild type cells. In Chapter 5, the role of Ycf54 was further investigated by analysing the importance of seven absolutely conserved residues in the Ycf54 primary sequence. Seven *Synechocystis* strains were generated, each with an individual point mutation in the Ycf54 locus, in which one of the seven conserved residues was substituted with an alanine or a glycine. These analyses revealed that three residues, D39, F40 and R82, are required for Pchlide formation (**Figure 5.5**). The *Synechocystis* strains in which each of these residues were individually converted to an alanine all accumulated considerable levels of MgPME and synthesised less Pchlide than the wild type strain. Additionally, when FLAG-tagged and used in *in vivo* pulldown assays, none of these Ycf54 mutants was able to interact with Sll1214/Sll1874, the catalytic component(s) of the cyclase (**Figure 5.10**). In addition to the three mutants deficient in Pchlide formation, another mutant, A9G, was found to exhibit a wild type phenotype, but accumulate ~75 % less Ycf54 and ~50 % less Sll1214/Sll1874 than wild type cells (**Figure 5.6**). Subsequent analysis of the *Ycf54* mRNA transcripts showed similar transcription levels of *Ycf54* mRNA in wild type and A9G *Synechocystis* strains (**Figure 5.9**), indicating that altered expression of Ycf54 is not the root cause of this phenotype.

To further investigate the *Synechocystis* Ycf54 protein and its potential role in the cyclase reaction, the crystal structure of the *Synechocystis* wild type protein was solved to a resolution of 1.3 Å. This crystallography was greatly facilitated by the availability of two other Ycf54 structures in the PBD, one from *Thermosynechococcus elongatus* (PDB 3ZHE), which was solved to a resolution of 2 Å and the other from *Nostoc* sp PCC7120 (PDB 3JSR), which was solved to a resolution of 1.8 Å. In order to understand the changes wrought on the Ycf54 protein structure by the point mutations A9G and R82A, the crystal structures of these two proteins were also solved. The structures showed that whilst the interaction between Ycf54 and Sll1214 may be interfered with via a change in the electrostatic charge surrounding residue 82, no structural change was visible to account for the reduced accumulation of A9G.

6.2 Results

6.2.1 Production of Recombinant Ycf54

An over-expression vector for the production of recombinant N-terminally His-tagged Ycf54 was constructed by cloning the *slr1780* (*Ycf54*) gene, amplified from *Synechocystis* genomic DNA, into pET14b (Novagen). This was achieved by designing primers which introduced an N-terminal NdeI site and a C-terminal BamHI site at the 5' and 3' ends of the *Ycf54* gene, respectively. This insert was ligated into pET14b, digested with identical restriction enzymes and the final construct confirmed by further restriction digests and sequencing (GATC-Biotech). This vector was named pET14b-Ycf54.

The pET14b-Ycf54 construct was transformed into BL21(P*lys*S) competent cells and optimum expression of recombinant His-Ycf54 was determined to be in Super Luria Broth with cultures incubated for 24 hours at 25°C, shaking at 200 rpm with no IPTG induction. Following over-expression the cells were pelleted, re-suspended in a minimal volume of HIS-binding buffer (50 mM Tris-HCl, 500 mM NaCl, 10 mM Imidazole, pH 7.4) and stored at -20°C until required. Recombinant His-Ycf54 was then purified by nickel affinity chromatography as described in Section 2.6.3, buffer exchanged into PBS and the His-tag removed by thrombin cleavage. Following removal of the His-tag and passage of the mixture back through an equilibrated nickel affinity column, the recombinant Ycf54 was found to be pure enough to be used directly in crystal trials. The highly purified Ycf54, shown in **Figure 6.1**, was buffer exchanged into 100 mM $NaHCO_3$, 50 mM NaCl pH 8.2 buffer and concentrated to 10 mg ml^{-1} for use in crystal trials. Generally 5-10 ml of purified Ycf54 at a concentration of 5-10 mg ml^{-1} was purified from 500 ml of cell culture.

6.2.2 Crystallisation of wild type Ycf54 and data collection

Initial crystal trials were performed using purified recombinant Ycf54 protein at a concentration of 10 mg ml^{-1} in 100 mM $NaHCO_3$ buffer at pH 8.2. The preliminary crystallisation conditions were screened for using the sitting-drop vapour diffusion technique with the NeXtal® commercial crystallisation screens (QIAgen®) and a Hydra II Plus I robot. As it is not possible to predict the conditions under which a protein will crystallise, four separate screens, each consisting of 96 different conditions, were trialled: $AmSO_4$ suite, classics suite, JCSG suite and the PACT suite. Each sitting drop contained 100 nl of protein and a 100 nl of crystallisation buffer. After dispensing the buffer and protein, the trays were incubated at 17°C and checked for crystal growth after 7-10 days.

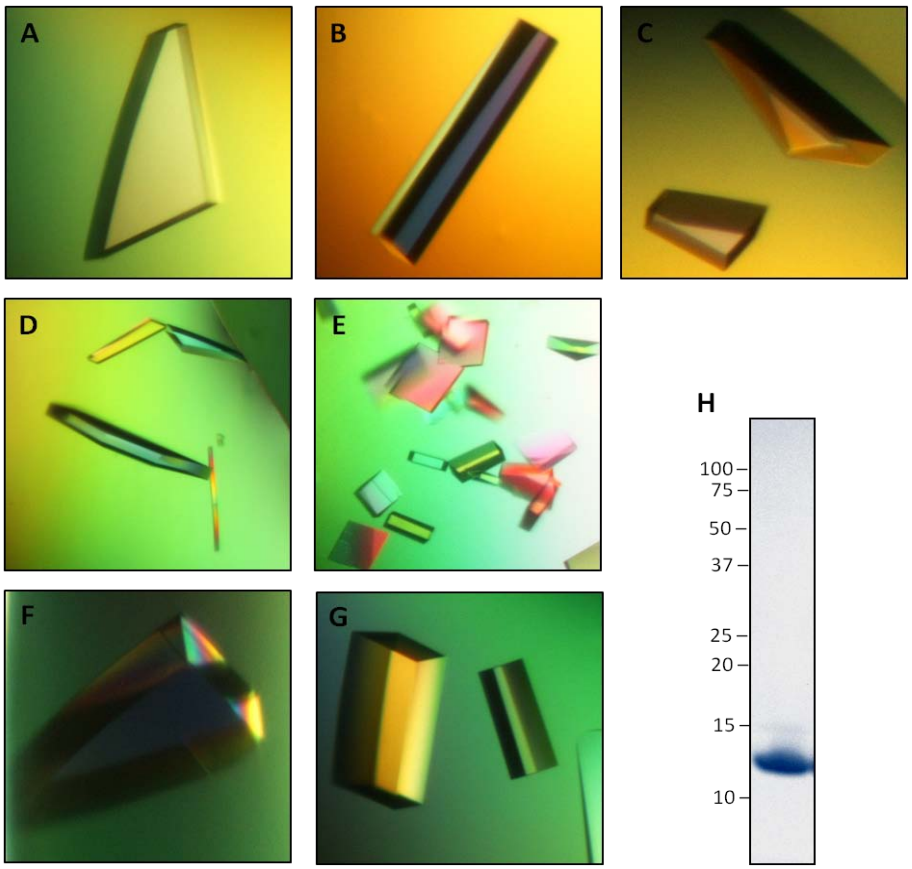

Figure 6.1 Crystals of *Synechocystis* wild type Ycf54 and purified Ycf54 protein used in crystal trials
Photographs of the wild type Ycf54 crystals formed during the initial screening trials and the crystallisation conditions in which they grew. These trials were performed using the sitting drop vapour diffusion technique in 96 sitting-drop well plates. (**A**) 0.2 M ammonium fluoride, 2.2 M ammonium sulphate, (**B**) 0.2 M ammonium iodide, 2.2 M ammonium sulphate, (**C**) 0.2 M caesium sulphate, 2.2 M ammonium sulphate, (**D**) 0.2 M potassium sodium tartrate, 2.2 M ammonium sulphate, (**E**) 0.2 M potassium sulphate, 2.2 M ammonium sulphate, (**F**) 2.2 M ammonium sulphate, (**G**) 0.1 M tri-sodium citrate, 2.4 M ammonium sulphate, (**H**) SDS-PAGE gel of purified Ycf54, molecular weight marker is in kDa.

Wild type Ycf54 was found to easily form crystals under many different ammonium sulphate based conditions, photographs of a selection of these can be viewed in **Figure 6.1**. These crystals grew to a sufficient size in order to be looped, thus several crystals were looped and cryogenically stored in a solution of the mother liquor containing ethylene glycol at a final concentration of 30 % (v/v), ready for transportation to the Diamond Light Source (Oxford). Once there, diffraction data were collected from the I04 beam line. Each crystal was screened by taking three diffraction images at 0.2 degrees oscillation, with a 45 degree separation between each image. An exposure time of 0.15 seconds was sufficient to give diffraction patterns that could be processed for data quality and resolution. From this, it was found the crystal that grew in 0.2 M ammonium fluoride and 2.2 M ammonium sulphate provided the

best diffraction. Thus, two data sets, each of 900 diffraction images taken with a rotation of 0.2 degrees, were collected from this crystal. 900 images with a 0.2 degree oscillation will give 180 degrees coverage of the crystal; this strategy was chosen because all unique reflections will be assessed over this range regardless of crystal symmetry.

Following data collection, the diffractions were processed by Diamond into MTZ files, encoding the information about electron density. Initial analysis of the diffraction data indicated wild type Ycf54 had unit cell dimensions of: a = 42.73 Å, b = 45.93 Å, c = 119.29 Å, α = β = γ = 120°, the space group C 2 2 2_1 and a resolution of 1.3 Å. The data collection statistics for wild type Ycf54 crystals are listed in **Table 6.1**.

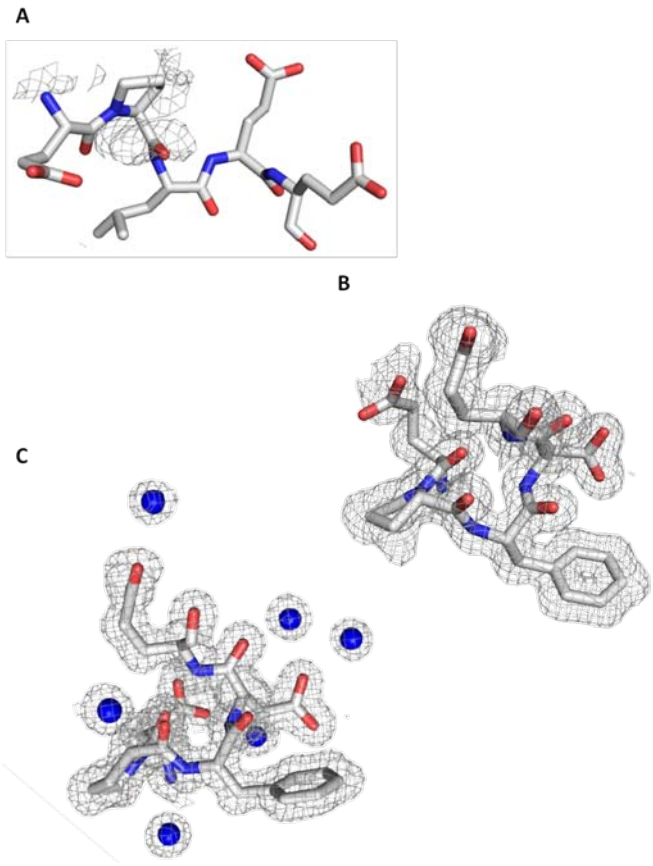

Figure 6.2 Sample region from the electron density map taken from the beginning, middle and end of the refinement process

Electron density maps of wild type Ycf54, showing the region encompassing residues E18 P19 F20 E21 E22, taken from the initial model after molecular replacement using the *Thermosynechococcus elongatus* Ycf54 (**A**), the refinement model following mutagenesis to *Synechocystis* Ycf54 primary sequence (**B**) and the final refinement model, which includes the water molecules (**C**). In white are carbon and hydrogen atoms, oxygen atoms are in red, nitrogen atoms are in blue and waters are blue spheres.

6.2.3 Data processing, model building and structure validation for *Synechocystis* wild type Ycf54

The MTZ file of the combined datasets was fed into the Cell Content Analysis module in the CCP4 Molecular replacement processing program (1994) to estimate the solvent content and number of molecules within the unit cell. This revealed the unit cell to be 47.8 % solvent, with a Matthews's coefficient of 2.35 $Å^3$/Da, containing a single protein molecule. Following calculation of the number of molecules in the unit cell, the MolRep Molecular Replacement Program was used to solve the *Synechocystis* Ycf54 structure. The *Synechocystis* Ycf54 MTZ data were input, along with the *Thermosynechococcus elongatus* Ycf54 .pdb as a template (PDB 3HZE) and the search parameter of one monomer per asymmetric unit. This generated the .pdb file of the initial model, which underwent several rounds of rebuilding and refinement using the Refmac5 module in the CCP4 processing suite (1994) and COOT (Emsley et al., 2010). **Figure 6.2** shows a representative region of the model and electron density map before refinement, in the middle of refinement and the final model.

After rounds of rebuilding and refining all of the Ycf54 protein chain could be built into the density; no residues were truncated to Cα or un-modelled. Refinement statistics from Refmac5 and the Validate Model module in the CCP4 processing suite are presented in **Table 6.2**. The R-factor and R_{free} are 0.189 and 0.22 respectively, and the Ramachandran plots displayed in **Figure 6.3** show all residues are within the favoured range for psi (ψ) and phi (ϕ) angles. B-factors were also calculated for each residue within the Ycf54 model, the majority of which have a B-factor less than 30 (**Figure 6.5**). A few of the lysine and glutamic acid side chains have a B-factor greater than 60, most likely resulting from these flexible residues crystallising in multiple conformations. Indeed four residues, E56, K73, R82 and E84, were found to have two distinct conformations well defined in the electron density (**Figure 6.4)**.

The electron density map also contained density that did not correspond to the main chain or waters. Subsequent analysis of these suggested they belonged to a phosphate group and an ethylene glycol molecule, which would have been introduced by the purification conditions and cryoprotectant used (**Figure 6.6**).

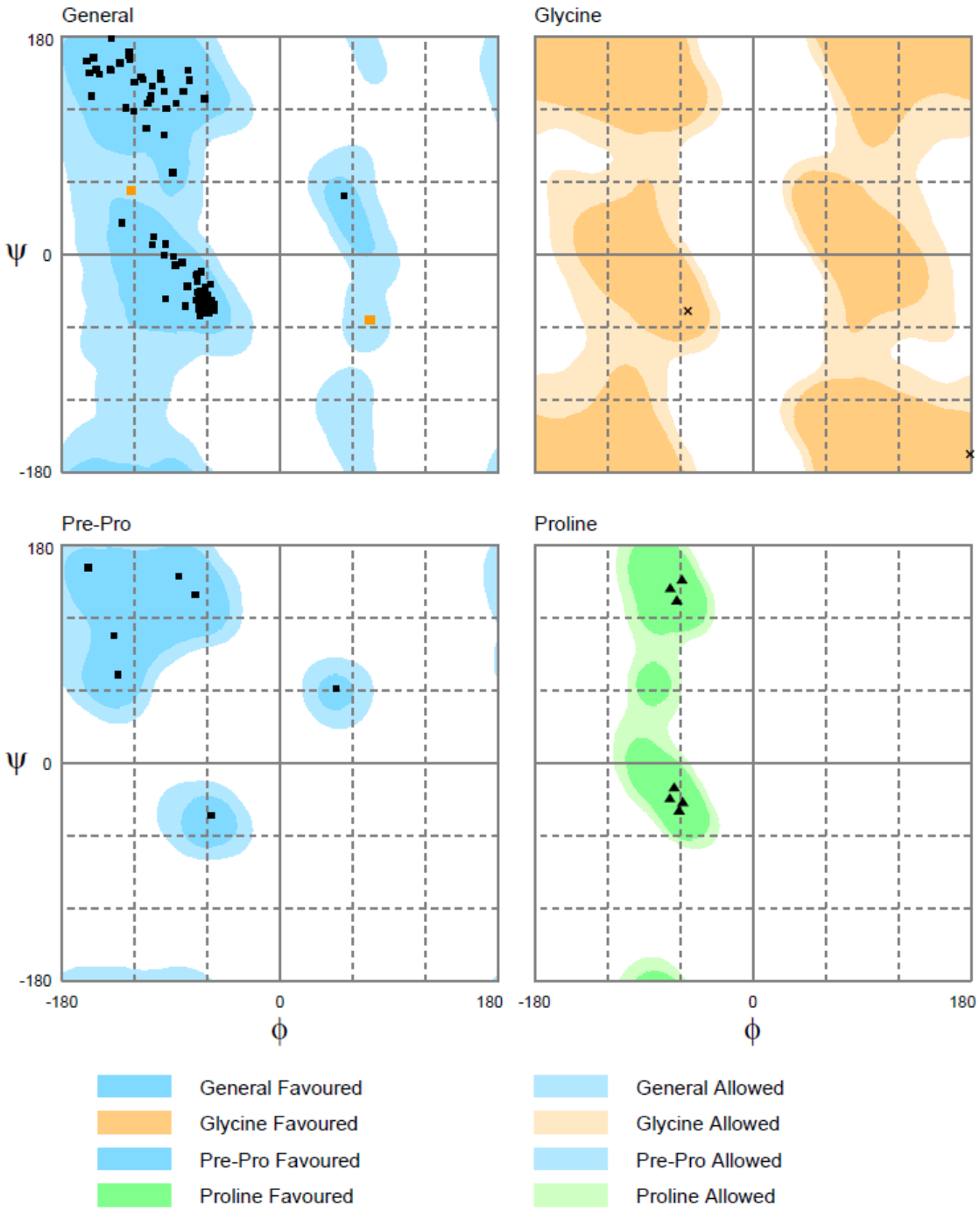

Figure 6.3 Ramachandran plots of the wild type Ycf54 structure

Ramachandran plots generated from the Richardson's data using RAMPAGE in the CCP4 suite. All residues are within the favoured regions, with 98.1 % of residues located within the favoured region and 1.9 % of residues located within the allowed region.

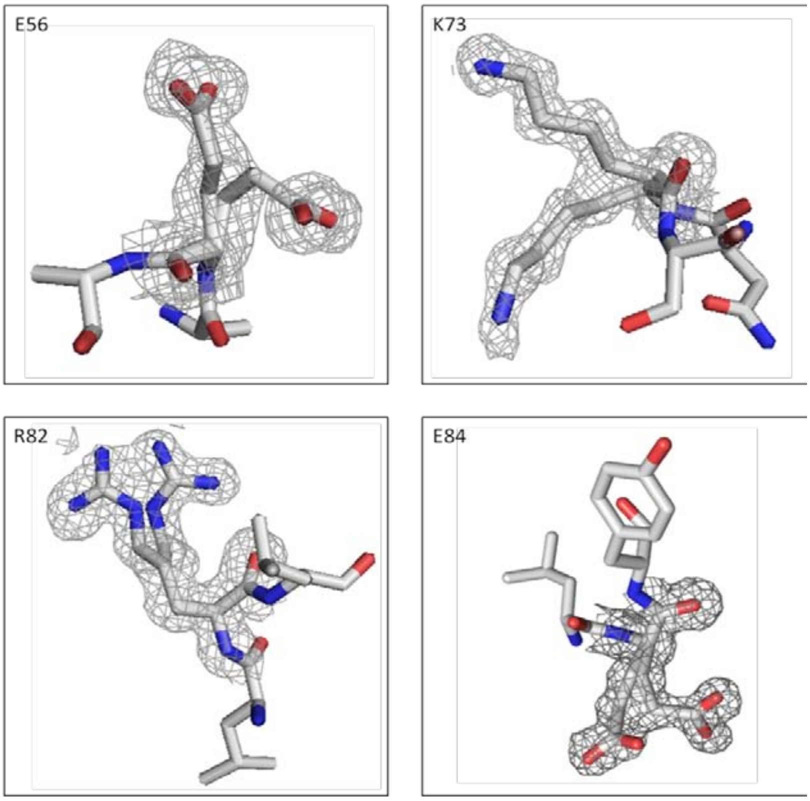

Figure 6.4 Residues in YCF54 found to adopt dual conformations

Electron density maps of wild type Ycf54, showing the four residues that were found to adopt dual conformations in the crystal structure.

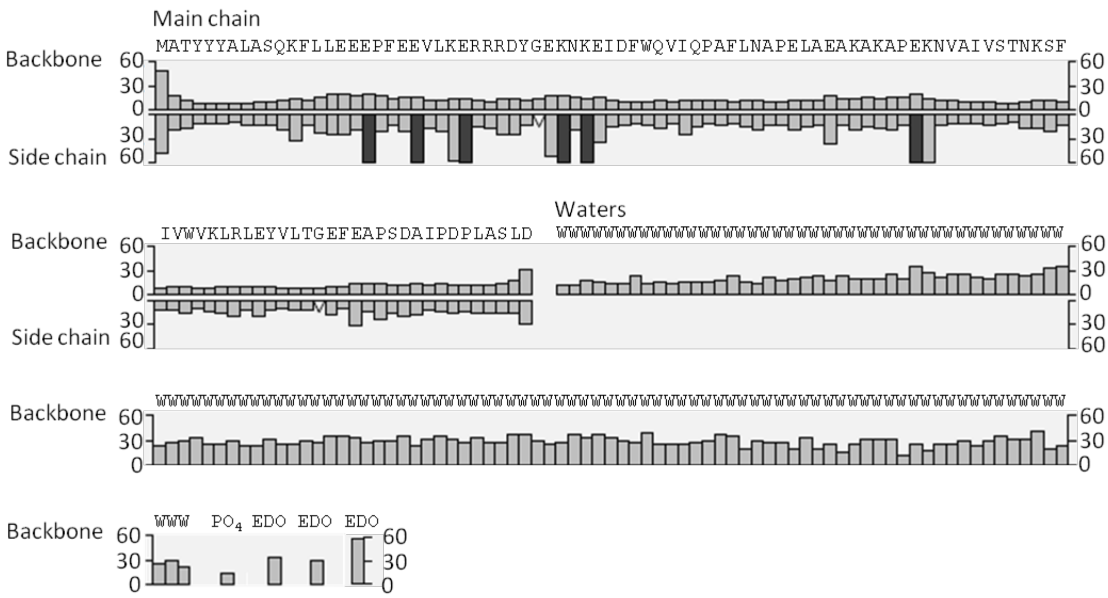

Figure 6.5 B-factors calculated for each residue in the Ycf54 model

Graphical representation of the B-factors calculated for individual residues in the Ycf54 model, including waters. B-factors were calculated using the Validate Model module in Refmac5. Light grey, B-factors < 60, dark grey, B-factors > 60.

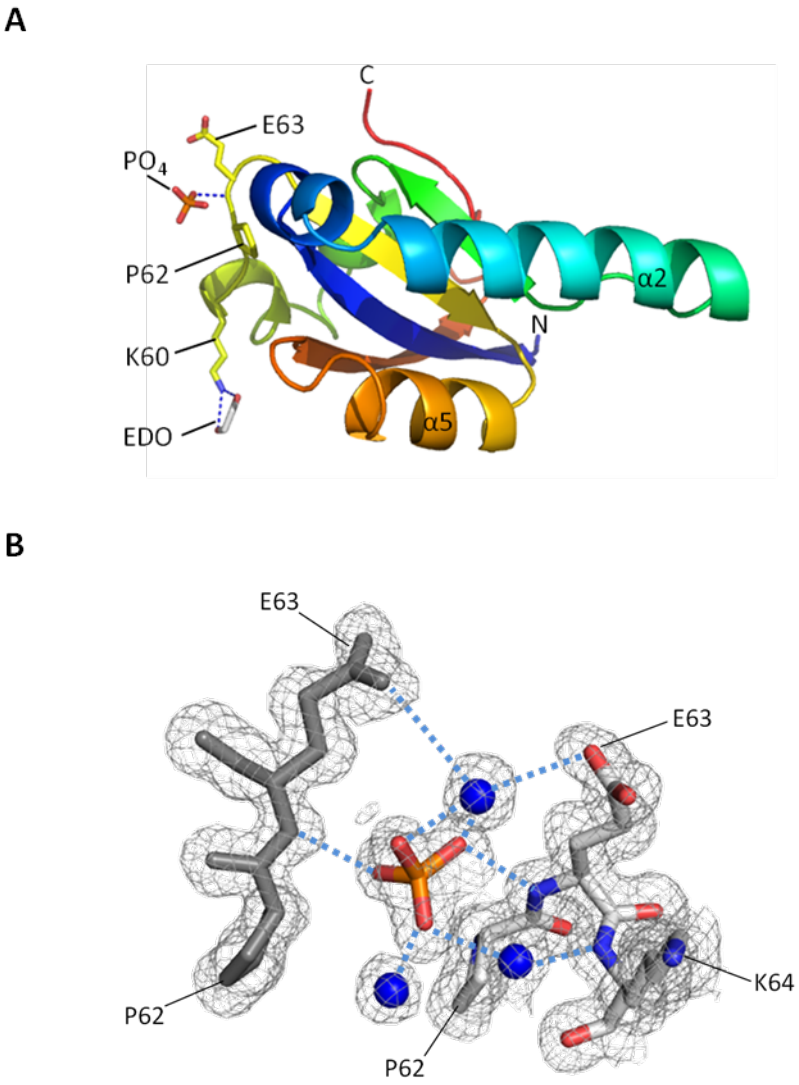

Figure 6.6 Location of the phosphate and ethylene molecules and map of the electron densities surrounding the phosphate group in the Ycf54 crystal structure

Location of the ethylene glycol (EDO) and phosphate (PO_4) molecules with reference to the wild type Ycf54 molecule (**A**). Electron density maps of the phosphate group found to crystallise with wild type Ycf54, showing the neighbouring residues (**B**). The phosphate group is co-ordinated by hydrogen bonding with waters and the main chain along the symmetry axis. In white are carbon and hydrogen atoms, oxygen atoms are in red and nitrogen atoms are in blue, the waters are shown as blue spheres and hydrogen bonds as blue dashed lines. The symmetry related molecule in the neighbouring unit cell is shown in grey.

6.2.4 Structure description

As anticipated from the structures of *Thermosynechococcus elongatus* (PDB 3HZE) and *Nostoc* sp PCC 7120 (PDB 3JSR) Ycf54 proteins, the *Synechocystis* YCF54 structure is composed of one domain (annotated as Ycf54 domain in PFAM) in which a central four-stranded anti-parallel β-sheet is flanked on both sides by α-helices. These α-helices were named α1 – α5, and the β-sheets were designated β1 – β4; α-helices 1, 2 and 5 are partitioned from α-helices 3 and 4 by the central β-sheet. These are all connected by eight short loop regions, designated L1 – L8, which range in length from 1 to 5 residues. Interestingly, the sequences of L5 and L6 are highly similar, comprising of APEL and APEK respectively. There is also a long C-terminal extension, L9, which is likely to be flexible that encompasses residues P94 – D106. An annotated diagram of YCF54 is presented in **Figure 6.7** and shows the "front" and "back" views of the protein.

Despite its small size (12.1 kDa), YCF54 displays several interesting surface features, which are emphasised upon mapping the surface electrostatics (**Figure 6.8**). The protein forms a disc-like conformation, with a concave and a convex surface, which is covered with a number of charged nodes that point out from the main body (annotated in **Figure 6.9**). Although these nodes look interesting, upon examination of the conserved residues (**Figure 6.11** and **Figure 5.1**), none of these was found to be conserved among the YCF54 homologues examined.

Undoubtedly the most notable surface feature is the negatively charged ridge created by a procession of evenly spaced glutamate residues arranged along the outer surface of α2 (**Figure 6.10**). Unlike the charged nodes mentioned earlier, several of these residues (E18, E22 and E26) are conserved among Ycf54 homologues. However, when individual *in vivo* point mutations of E22 and E26 were created in *Synechocystis* (Chapter 5), there was no detectable phenotype. Nevertheless, after viewing the locations of these residues, it is probable the single mutations constructed did not reduce the charge of this region significantly enough to affect any functional role α2 may have.

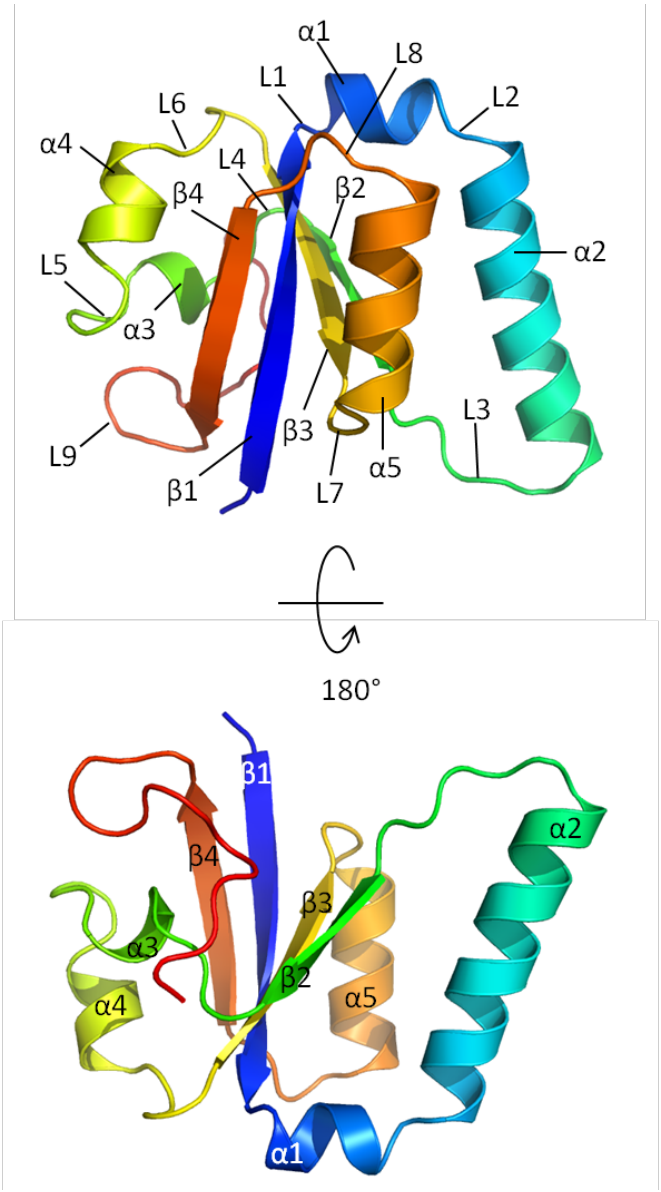

Figure 6.7 Ribbon diagrams of *Synechocystis* Ycf54

Ribbons diagrams show the "front" and "back" view of Ycf54 rotated 180° around a horizontal axis. The rainbow colours track the progression from the N-terminus (blue) to the C-terminus (red). The α-helices (α1 – α5), β-sheets (β1 – β4) and loops (L1 – L9) are labelled.

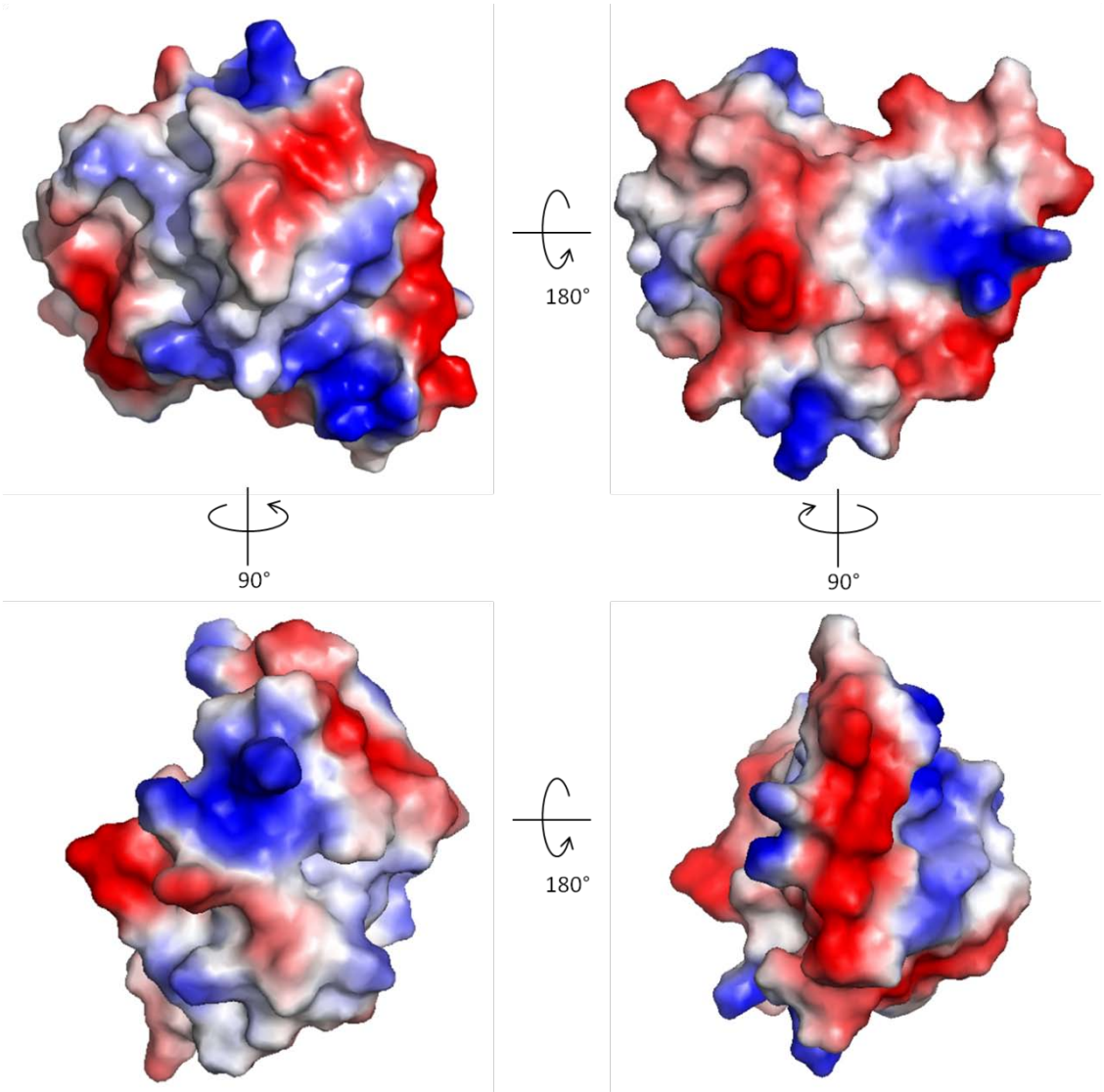

Figure 6.8 Surface electrostatics of Ycf54

Surface electrostatics of Ycf54 from all angles, calculated using the Adaptive Poisson-Boltzman Solver (APBS) in PyMOL. Blue: positive charge, Red: negative charge, White: neutral.

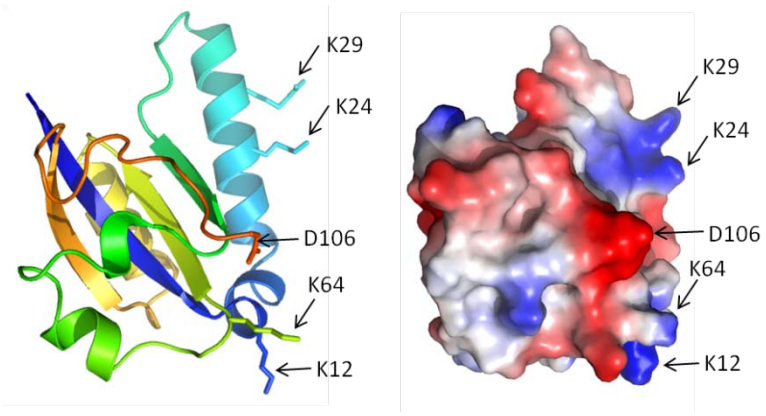

Figure 6.9 Residues that form charged nodes that project from the surface

The YCF54 surface features several charged node that project from the surface like prongs. Several of these are annotated on the ribbon structure and on the surface electrostatic diagram.

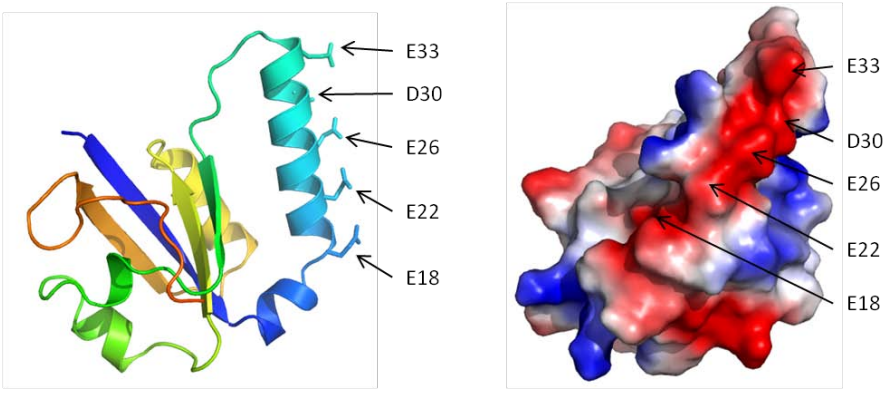

Figure 6.10 Glutamic acid residues on α-helix 2 that form a negative ridge

Highlighted are the five glutamic acid residues located on α-helix 2 that form the negatively charged ridge.

6.2.5 Phylogeny of YCF54 and analysis of conserved residues

YCF54 is a highly conserved domain found in all oxygenic photosynthetic organisms examined. In the PFAM database 65 cyanobacterial and 36 eukaryote species are reported to contain YCF54 (viewed on 01 October 2013). Of the 65 YCF54 sequences located in cyanobacteria, the majority of sequences (40) are from the Chrococcales and a small number of sequences are from the Oscillatoriales (9), the Prochlorales (7), the Nostocacese (7) and the Stigonematales (1). In the eukaryotes the YCF54 sequences are located in green plants (18), green algae (8), Steamenophiles or brown algae (5), red alga (5) and one amoeboid (1). The phylogenetic relationship between all of these YCF54 sequences was calculated using the maximum likelihood method (**Figure 6.12**). Interestingly, the mosses *Physcomitrella patens* and *Selaginella moelienforffii* contain isoforms of distantly related YCF54 genes, and the

cyanobacterium *Synechococcus* genus, contains many distantly related YCF54 genes, perhaps representing independent evolution of these bacteria over a wide geographical range. The amoeboid *Paulinnella chromatophora*, which recently enslaved a *Prochlorococcus*-like cyanobacterium, has (as would be expected) an YCF54 gene closely related to those found within the *Prochlorococcles*. Also of interest is the location of *Synechocystis* YCF54, whose sequence appears to be more closely related to those of plants than the majority of cyanobacteria, and as such is located on a node with *Arabidopsis* and Poplar (**Figure 6.12**).

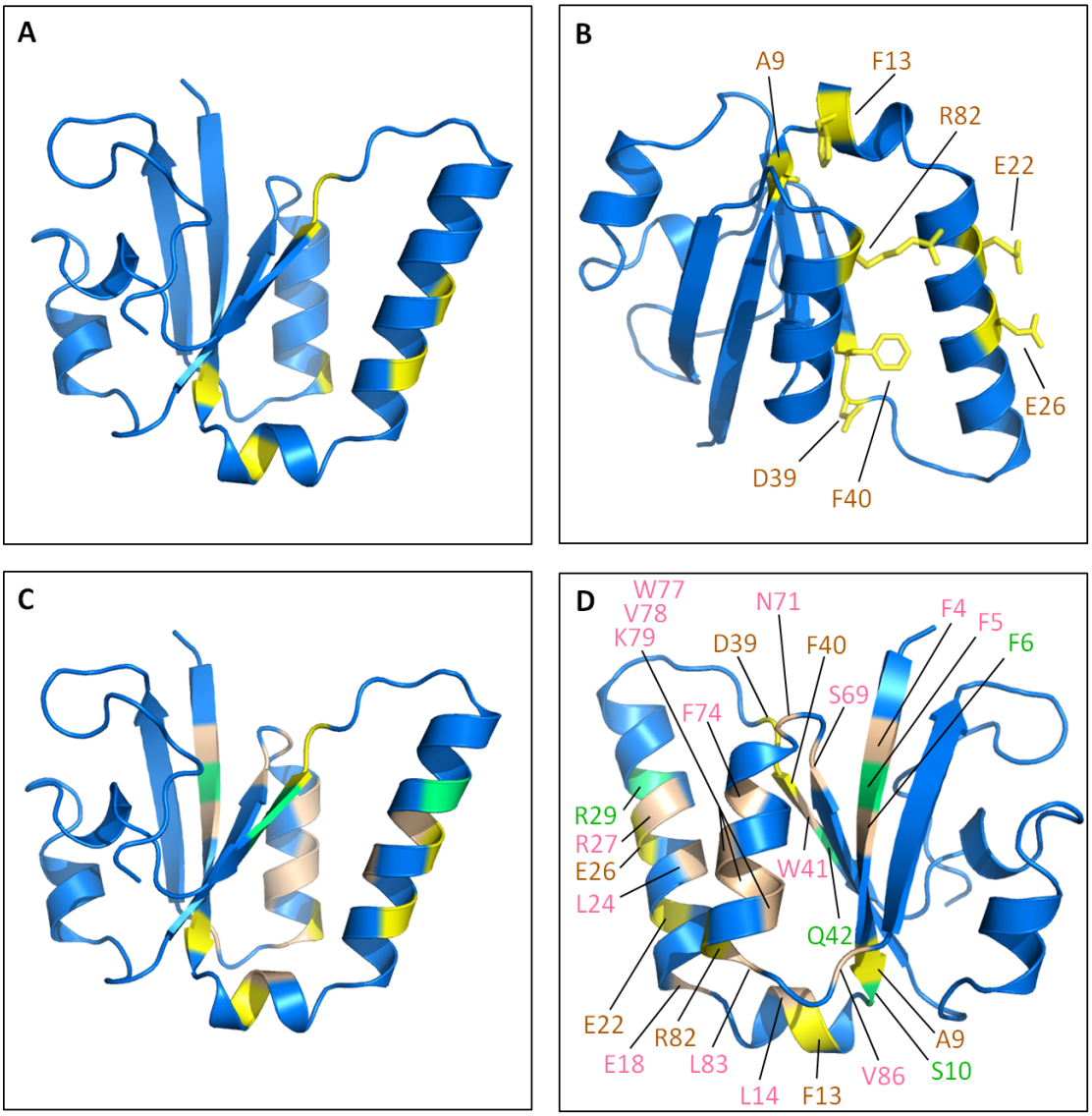

Figure 6.11 Highly conserved residues mapped onto *Synechocystis* Ycf54

The seven absolutely conserved residues are highlighted in yellow (**A**) and labelled in (**B**). The highly similar conserved residues (according to the sequence alignment conducted in ClustalW2, shown in Figure 5.1) are highlighted in green and the similar conserved residues are highlighted in tan, shown in (**C**) and labelled in (**D**). Images (**A**) and (**C**) are taken from the same view point and (**B**) and (**D**) are rotated 180°C to show the back of the protein.

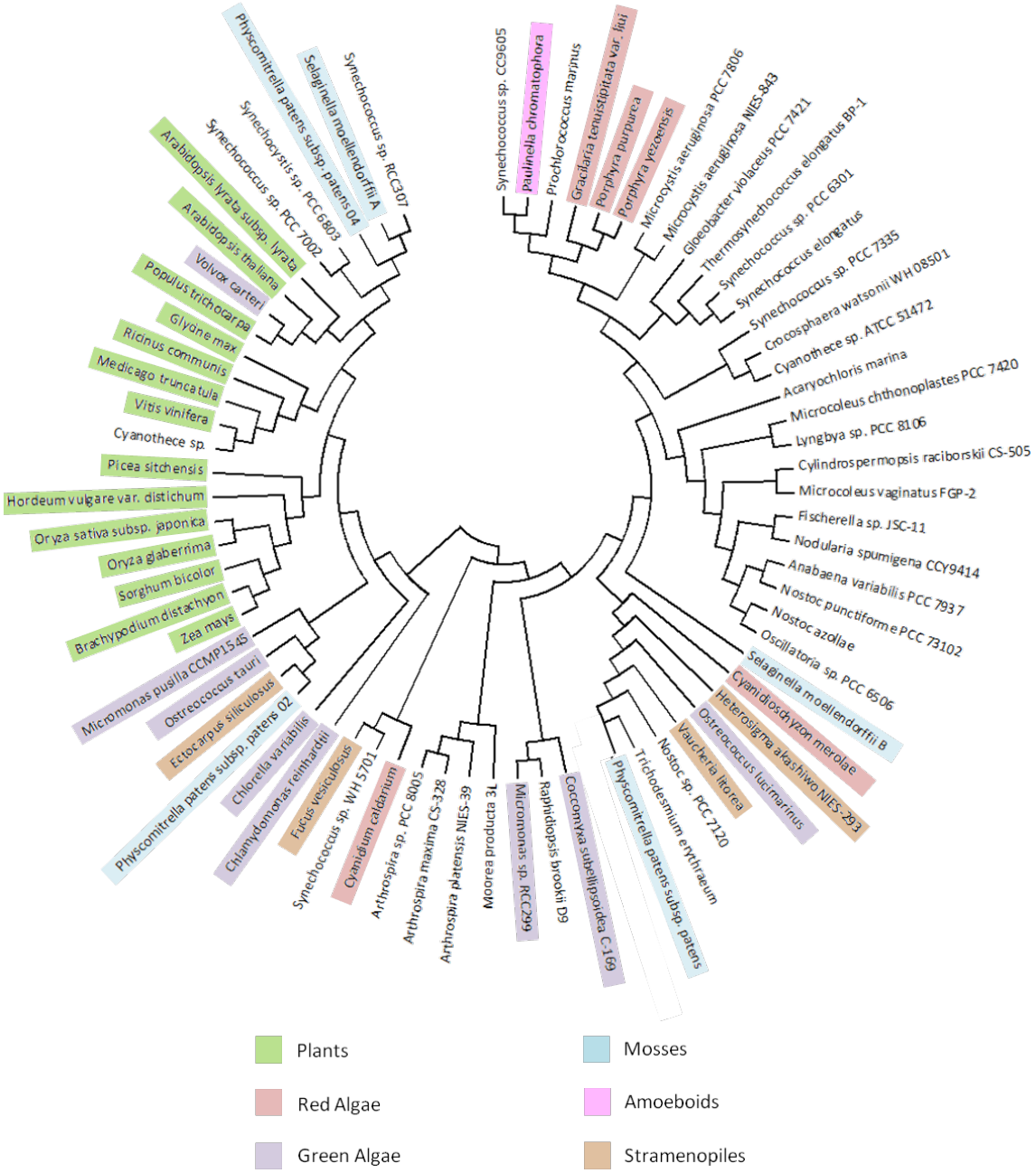

Figure 6.12 Molecular phylogenetic analysis of YCF54 distribution among oxygenic photosynthetic organisms by the maximum likelihood method

The evolutionary history of Ycf54 was inferred by using the maximum likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-12712.1676) is shown. An initial tree for the heuristic search was obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model. The tree is drawn to scale, with the branch lengths measured in the number of substitutions per site. The analysis involved 72 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 91 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).

There are many residues that are conserved across the different YCF54 homologues; all of these are highlighted in **Figure 6.11 D**. No conserved residues were found in α3, α4, β3 or β4. In Chapter 5, it was noted seven of these residues were absolutely conserved. When the locations of these residues are highlighted (**Figure 4.9 B**), they are all visibly concentrated on one half of the structure, which may represent a highly conserved structural motif. Three of these seven residues, D39, F40 and R82, were found to be important for YCF54 to interact with its binding partner, the catalytic component of the cyclase, Sll1214,whilst another mutation, A9G, was found to reduce the accumulation of Ycf54 by ~75 % in comparison to wild type. To investigate whether any of these residues result in structural changes in Ycf54, recombinant proteins of each Ycf54 mutant were produced and crystal trials were performed (Section 6.3.8).

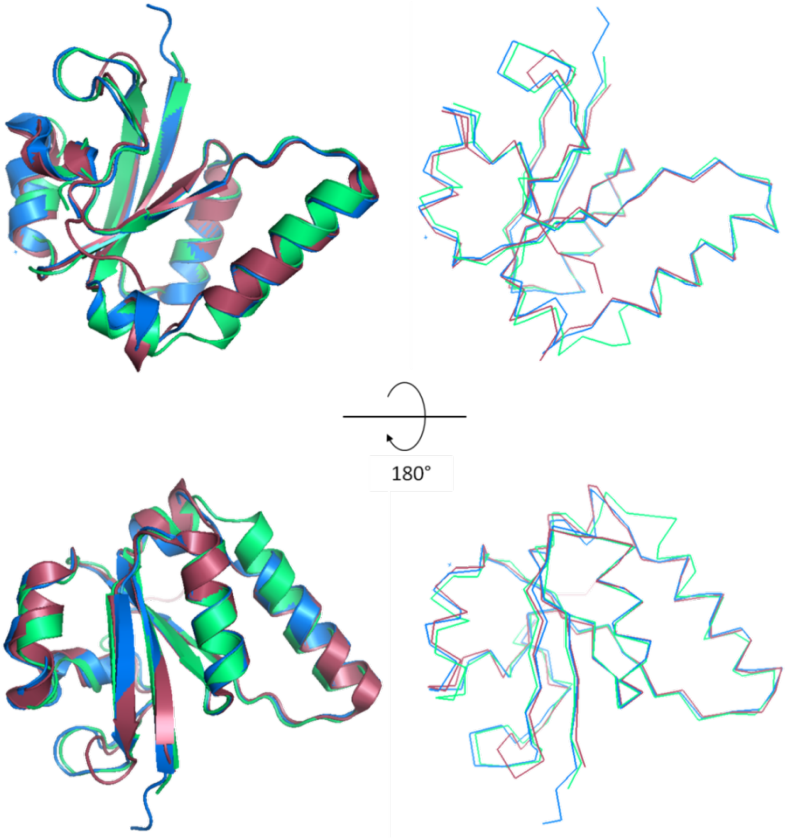

Figure 6.13 Superimposition of *Synechocystis*, *Thermosynechococcus elongatus* and *Nostoc* sp. Ycf54 proteins

Secondary structure alignment and superimposition of the Ycf54 proteins from *Synechocystis* (green), *Thermosynechococcus elongatus* (pink) and *Nostoc* sp (blue) generated using Pymol. The PDB references for the *Thermosynechococcus elongatus* and *Nostoc* sp structures are 3HZE and 3JSR respectively. Diagrams to the left show how the secondary structures align and the diagrams to the right show the alignment of the Cα backbone. RMSD = 0.537 Å and is inclusive of all atoms.

6.2.6 Structural homologues

A structural alignment of the *Thermosynechococcus elongatus* (shares 63 % sequence identity), *Nostoc* sp PCC 7120 (shares 70 % sequence identity) and *Synechocystis* wild type Ycf54 proteins can be viewed in **Figure 6.13**. The *Themosynechococcus elongatus* structure was used over the *Nostoc* sp. structure for molecular replacement, as the *Nostoc* sp. PDB entry is annotated as hypothetical protein all0216, and was only identified in later Dali searches as an Ycf54 homologue. These alignments show the Ycf54 proteins to all retain highly similar secondary structures with a RMSD for all α carbons of 0.537 Å. The major significant differences occur at the loop regions, which are likely to be flexible and form different conformations in the crystal.

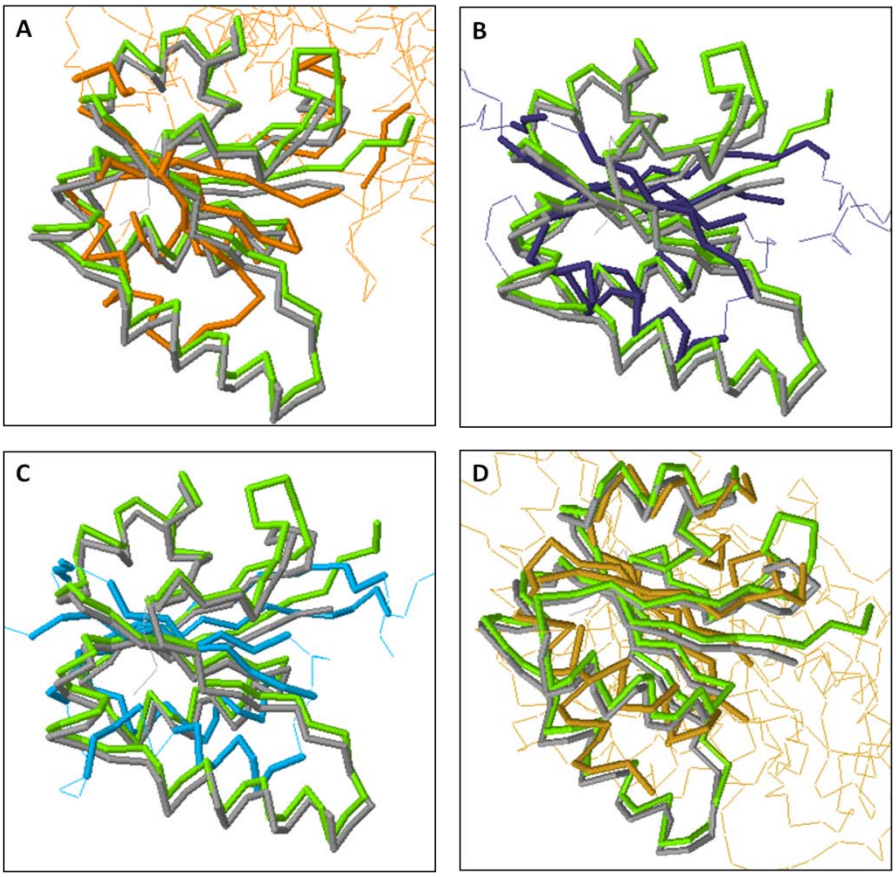

Figure 6.14 Superimposition of *Thermosynechococcus elongatus* and *Nostoc* sp PCC 7120 Cα backbones with structural homologues

Superimposition of structural homologues identified in a Dali database search with the Cα backbones of *Thermosynechococcus elongatus* (grey) and *Nostoc* sp PCC 7120 (green). (**A**) *E. coli* DNA polymerase II, residues E39 to M115 (**B**) *Azospirillum brasilense* PII like protein, residues M1 to E98 (**C**) *Pyrococcus horikoshii* transcriptional regulation factor HP1062, residues H6 to F52 and D69 to V106 (**D**) *E. coli* D-lactate dehydrogenase, residues L377 to H433.

The Dali server (Dietmann et al., 2001) was used to search the Protein Data Bank for proteins with structural homology to *Thermosynechococcus elongatus* Ycf54 (PDB 3HZE, Chain A) and *Nostoc* sp. PCC 7120 Ycf54 (PDB 3JSR). The top results are displayed in Table 6.3 and 6.4 respectively. In general, all the hits found Ycf54 to have structural similarity to sub-sections of larger proteins including DNA polymerase II and the nitrogen metabolism regulatory protein PII. **Figure 6.14** displays superimpositions of the *Thermosynechococcus elongatus* and *Nostoc* sp. Ycf54 proteins with the regions of structural similarity identified in DNA polymerase II, PII, transcription regulation factor HP1062 and D-lactate dehydrogenase.

Analysis of the *E. coli* DNA polymerase II sequence, the top non-Ycf54 hit for both 3HZE and 3JSR, revealed that the region of structural similarity (residues 39 –115) is situated upstream of the exonuclease domain (residues 130 – 370), with no particular associated function (Pham et al., 2001). Of more interest are the P_{II} protein and D-lactate dehydrogenase hits.

P_{II} is a highly conserved signal transduction protein that co-ordinates assimilation of nitrogen and carbon metabolism in bacteria, archaea and plants. Regulation of nitrogen assimilation and photosynthesis is of critical importance in non-heterocyst forming cyanobacteria, as nitrogenase, the enzyme responsible for reducing atmospheric N_2 to NH_3, is highly sensitive to O_2 damage (Masephol and Forchhammer, 2007). The region to which Ycf54 displays homology (residues 1-98) encompasses the conserved residue S49, which is modified by phosphorylation in response to intracellular 2-oxoglutarate levels and the conserved residues K58 and Q39, which are involved in 2-oxoglutarate binding (Fokina et al., 2010).

FAD-binding D-lactate dehydrogenase is a 65 kDa peripheral membrane protein involved in electron transfer that catalyses the oxidation of D-lactate to pyruvate. The protein has three domains, the cap domain, the membrane binding domain and the FAD-binding domain. Ycf54 displays structural homology to the C-terminal half of the cap domain (residues 377-433), which is composed of residues 269-310 and 388-425 (Dym et al., 2000). As its name suggests, the cap domain forms a lid over the FAD co-factor binding site; although this domain is not directly involved in flavin binding, the FAD molecule does lie at the junction between all three domains (Dym et al., 2000).

6.2.7 Production of recombinant Ycf54 point mutants

To ascertain whether any of the Ycf54 point mutants with detectable phenotypes (described in Chapter 5) have any structural alterations, recombinant versions of these proteins were produced with the aim of producing material for crystal trials.

The point mutations A9G, D39A, F40A and R82A were introduced into pET14b-Ycf54 (constructed in Section 6.3.1) using the QuickChange kit (Stratagene) and the primers listed in Table 5.1. Insertion of the correct point mutation was confirmed by sequencing analysis (GATC-Biotech). These plasmids were named pET14b-A9G, pET14b-D39A, pET14b-F40A and pET14b-R82A.

Each of the Ycf54 point mutation plasmids was transformed into BL21 (DE3) pLysS *E. coli* cells, ready for recombinant protein over-expression. Strains pET14b-A9G and pET14b-R82A were over-expressed as described for pET14b-Ycf54 in Section 6.3.1. The over-expression of strains pET14b-D39A and pET14b-F40A was modified, as these proteins were liable to precipitate at the high concentrations resulting from this method of over-expression. Therefore an alternative overexpression condition, producing a lower level of recombinant protein, was determined (growth in LB media at 37°C to an OD_{600} 0.7 – 0.8, followed by induction with IPTG and overnight overexpression at 25°C).

All four Ycf54 point mutants were purified as described for wild type Ycf54 in section 6.3.1 and SDS-PAGE analysis of the purified Ycf54 mutant proteins can be viewed in **Figure 6.15**.

6.2.8 Crystallisation of Ycf54 mutants and data collection

Initial crystal trials were performed using purified recombinant Ycf54 mutant proteins at a concentration of 10 mg ml^{-1} in 100 mM $NaHCO_3$ buffer at pH 8.2, as described for the wild type Ycf54 protein in section 6.3.2. The preliminary crystallisation conditions were screened for using the sitting-drop vapour diffusion technique with the NeXtal® commercial crystallisation screens (QIAgen®) and a Hydra II Plus I robot. Six separate screens, each consisting of 96 different conditions, were trialled: $AmSO_4$ suite, classics suite, JCSG suite, PACT suite, pH Clear suite and the MPD suite. Each sitting drop contained 100 nl of protein and a 100 nl of crystallisation buffer. After dispensing the buffer and protein, the trays were incubated at 17°C and checked for crystal growth after 7-10 days.

A9G, F40A and R82A were found to form crystals in a small subset of the conditions trialled, pictures of the A9G and R82A crystals are displayed in **Figure 6.16**. There was no evidence of crystallisation from D39A. A9G and F40A were found to form large cubic crystals in ammonium sulphate based conditions and R82A was found to form needle-like clusters under a variety of different conditions. Some of the crystallisation conditions were identical between the mutant and wild type Ycf54s, with A9G and wild type both forming crystals in 0.1 M tri-sodium citrate and 2.4 M ammonium sulphate and R82A and wild type both forming crystals in 2.2 M ammonium sulphate.

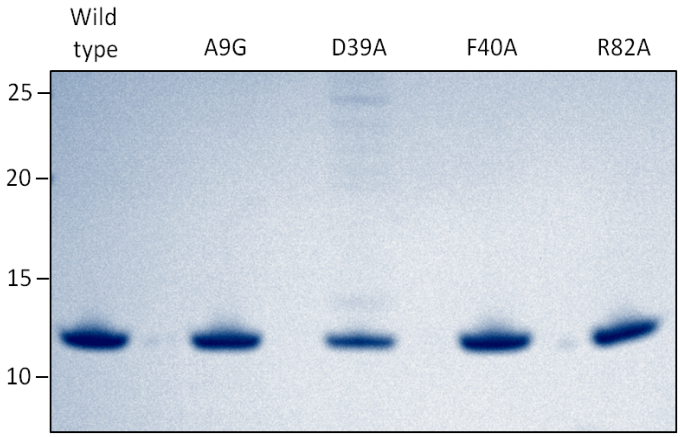

Figure 6.15 Purified recombinant *Synechocystis* Ycf54 proteins

20 µg of purified wild type Ycf54 protein and Ycf54 mutant proteins A9G, D39A, F40A and R82A were analysed by SDS-PAGE and strained with Coomassie blue to confirm the final purity of the recombinant protein before crystal trials. Molecular weight marker is in kDa.

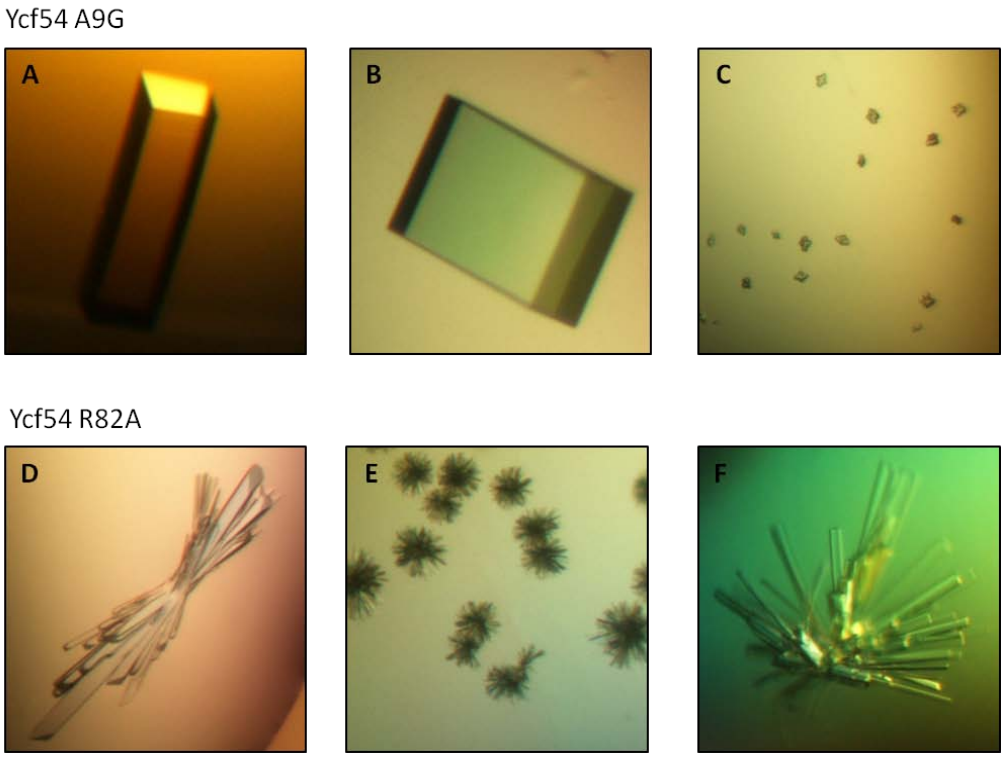

Figure 6.16 Crystals of Synechocystis mutant Ycf54 proteins A9G and R82A

Photographs of the crystals formed by Ycf54 mutants A9G and R82A during the initial screening trials and the crystallisation conditions in which they grew. These trials were performed using the sitting drop vapour diffusion technique in 96 sitting-drop well plates. (A) 0.2 M di-ammonium tartrate and 2.2 M ammonium sulphate, (B) 0.1 M tri-sodium citrate and 2.4 M ammonium sulphate, (C) 0.2 M lithium sulphate, 0.1 M Tris pH 8.5 and 40 % (v/v) PEG 400, (D) 2.2 M ammonium sulphate, (E) 0.1 M tri-sodium acetate pH 4.5, 0.1 M Bis-Tris pH 5.5 and 25 % (w/v) PEG 3350, (F) 0.1 M PCB buffer pH 4 and 25 % (w/v) PEG 1500.

The mutant Ycf54 crystals were looped and cryogenically stored in the mother liquor supplemented with ethylene-glycol to a final concentration of 30 % (w/v) and cooled in liquid nitrogen, ready for transportation to the Diamond light source (Oxford). Diffraction data were collected from the I02 beam-line, where three diffraction images at a rotation of 0.2 omega degrees were collected for each crystal. It was found the A9G crystal that grew in 0.2 M di-ammonium tartrate and 2.2 M ammonium sulphate and the R82A crystal that grew in 2.2 M ammonium sulphate provided the best diffraction, whereas both the F40A crystals tested only diffracted to a resolution of 3.8 Å. Two data sets, each of 900 diffraction images taken with a rotation of 0.2 omega degrees, were collected from the best diffracting A9G and R82A crystals. Following data collection, the diffractions were processed by Diamond into MTZ files, encoding the information about electron density. Initial analysis of the diffraction data indicated A9G had unit cell dimensions of: a = 42.20 Å, b = 46.13 Å, c = 119.66 Å, α = β = γ = 90°, the space group C 2 2 2_1 and a resolution of 1.5 Å, whilst R82A had unit cell dimensions of: a = 55.52 Å, b = 91.15 Å, c = 120.30 Å, α = β = γ = 90°, the space group P 2_1 2_1 2_1 and a resolution of 2.2 Å. The data collection statistics for all the Ycf54 crystals are listed in **Table 6.1**.

Following the results from the initial data collection, trials were set up to optimise the crystallisation of F40A and R82A with the aim of obtaining better diffracting crystals, however no crystals were found to grow in any of the conditions tested and no further optimisation was performed.

6.2.9 Data processing, model building and structure validation for *Synechocystis* Ycf54 mutants A9G and R82A

The MTZ file of the combined datasets was fed into the Cell Content Analysis module in the CCP4 Molecular replacement processing program (1994) to estimate the solvent content and number of molecules within the unit cell. This revealed the A9G unit cell to be 42.84 % solvent, with a Matthews's coefficient of 2.17 $Å^3$/Da, containing a single protein molecule and the R82A unit cell to be 58.53 % solvent, with a Matthews coefficient of 2.99 $Å^3$/Da, containing four protein molecules per unit cell. After calculating the number of molecules in the unit cell, the MolRep Molecular Replacement Program was used to solve the structures of A9G and R82A. The MTZ data were input, along with the wild type *Synechocystis* Ycf54.pdb as a template, and the search parameter of one monomer (A9G) or four monomers (R82A) per asymmetric unit. This generated the .pdb file of the initial models, which underwent several rounds of rebuilding and refinement using the Refmac5 module in the CCP4 processing suite and COOT. **Figure 6.17** and **Figure 6.19** show a representative region of the model and

electron density map before refinement, after processing through the ArpWarp module and the final refinement model of A9G and R82A respectively.

After rounds of rebuilding and refining all of the A9G protein chain could be built into the electron density; no residues were truncated to Cα or un-modelled. Refinement statistics from Refmac5 and the Validate Model module in the CCP4 processing suite are presented in **Table 6.2**. The R-factor and R_{free} are 0.21 and 0.26 respectively, and the Ramachandran plots displayed in **Figure 6.18** show all residues are within the favoured range for psi (ψ) and phi (ϕ) angles. The B-factors calculated for each residue within the A9G model shows all of the waters and backbone residues to have a B-factor less than 30. Some of the lysine, argenine and glutamate side chains have a B-factor greater than 60, most likely resulting from these flexible residues having multiple conformations.

Following rounds of rebuilding and refining the majority of the R82A protein chains could be built into the density. Of the four monomer chains in each unit cell, chains A and C were found to have the better electron density, enabling modelling of the protein chains into the density, whilst the density in chains B and D was markedly poorer. This disparity in density is represented by the higher percentage of residue side chain B-factors over 60 in chains B and D, with the most problematic region to refine situated between residues 10 - 24 and residues 47 - 65. Additionally, the Ramachandran plots displayed in **Figure 6.20** show the greatest number of outlier residues belong to chain D. Nevertheless an overlay of all four chains (**Figure 6.21**) shows the majority of the Cα backbone to conform to the same secondary structure. The refinement statistics for R82A are presented in **Table 6.2**, and include an R-factor of 0.19 and an R_{free} of 0.26.

No significant electron density was observed in the A9G or R82A density maps that did not correspond to the main chain and its associated water molecules, suggesting the phosphate molecule observed in the wild type structure (**Figure 6.6**) is an artefact of the purification or crystallisation conditions used. Additionally, no ethylene glycol molecules were observed in the density maps of either mutant Ycf54 protein, which is unsurprising as the presence of these molecules in the wild type electron density maps is a result of the cryoprotectant in which the wild type Ycf54 crystals were stored.

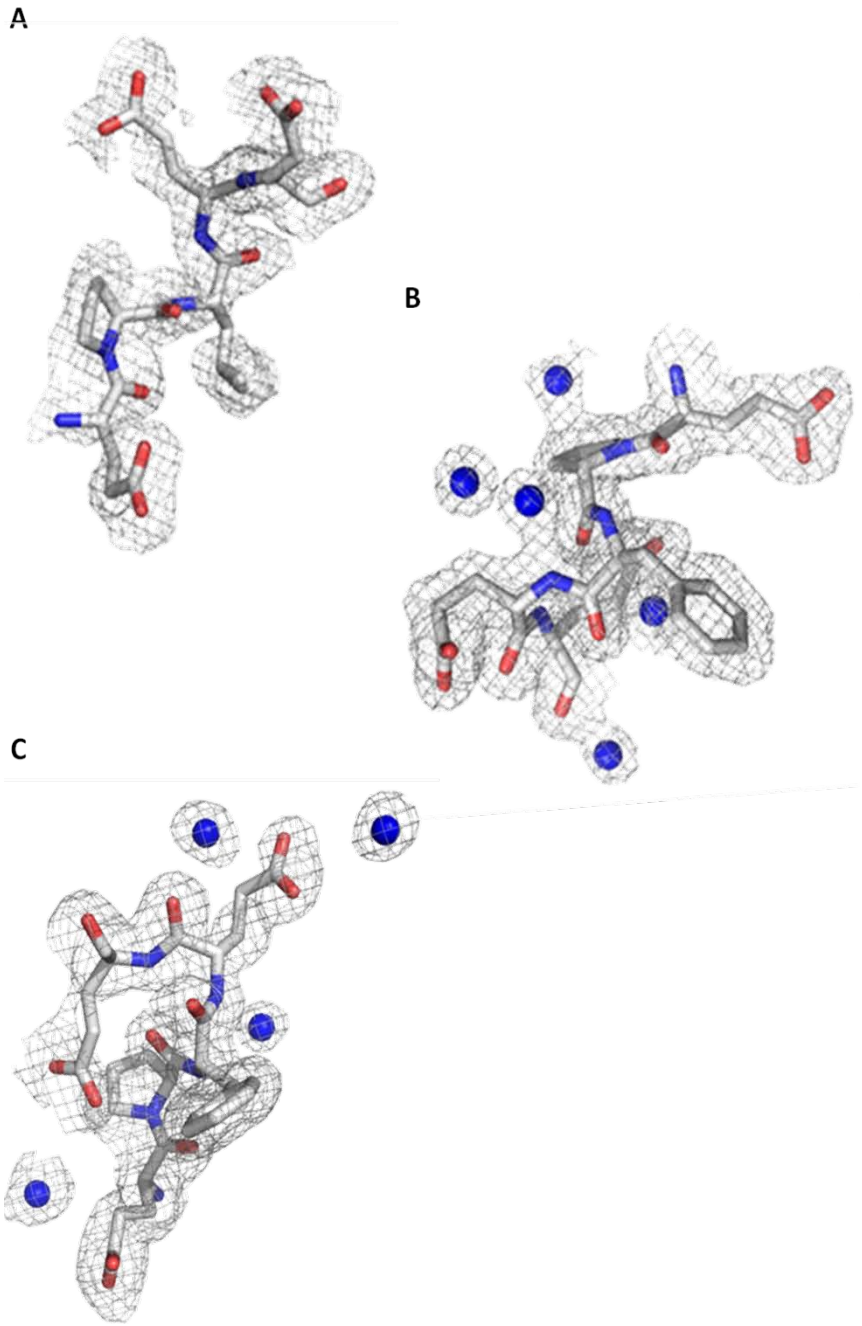

Figure 6.17 Sample regions from the A9G electron density map taken from the beginning, middle and end of the refinement process

Electron density maps of the Ycf54 mutant A9G, showing the region encompassing residues E18 P19 F20 E21 E22, taken from the initial model after molecular replacement using the *Thermosynechococcus elongatus* YCF54 (**A**), the refinement model following processing using the Refmac5 module Arp Warp against the wild type Ycf54 structure (**B**), and the final refinement model (**C**). In white are carbon atoms, in red are the oxygen atoms, in blue are the nitrogen atoms and the waters are shown as blue spheres.

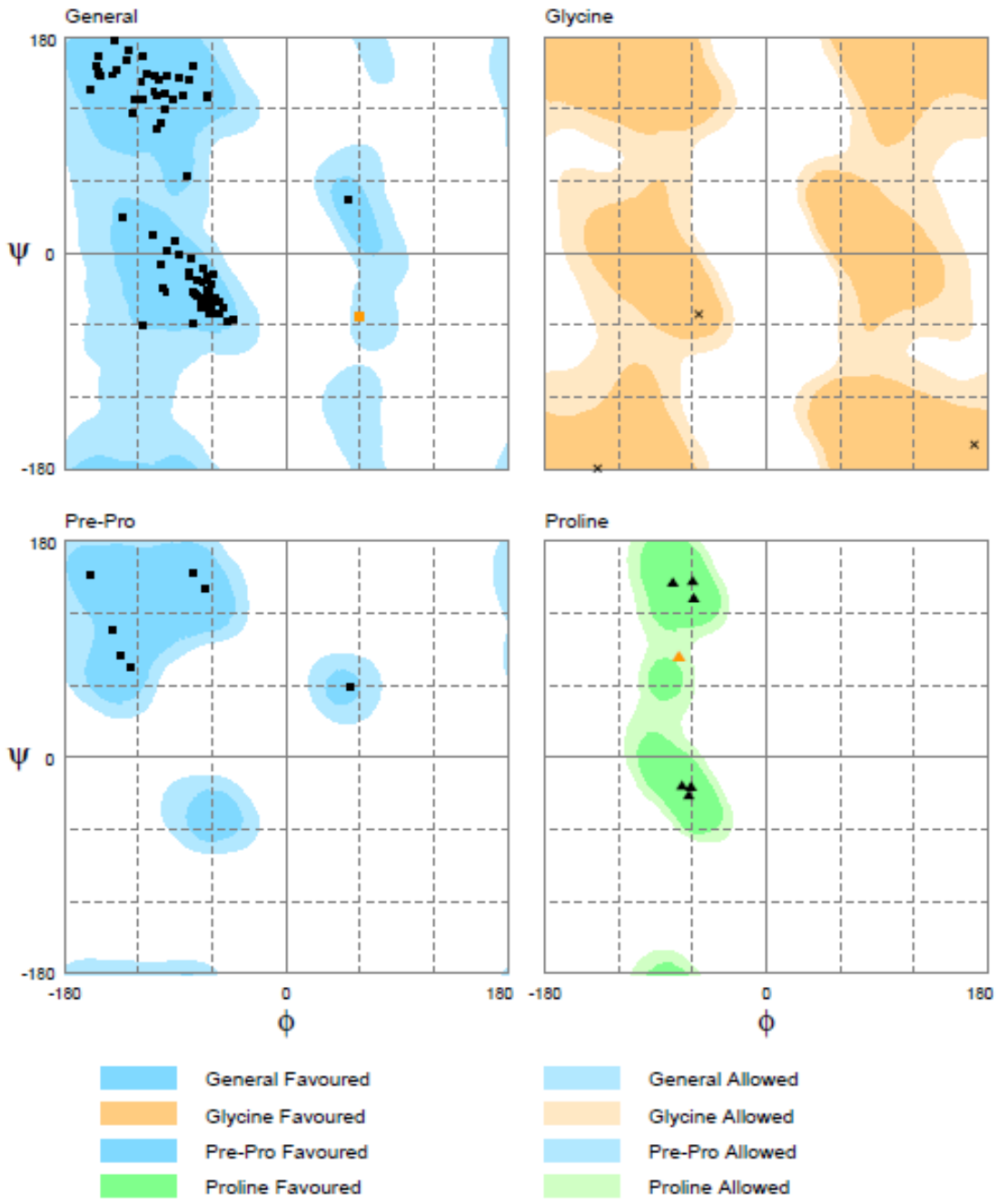

Figure 6.18 Ramachandran plots of the A9G Ycf54 mutant structure

Ramachandran plots generated from the Richardson's' data using RAMPAGE in the CCP4 suite. All residues are within the favoured regions, with 98.1 % of residues located within the favoured region and 1.9 % of residues located within the allowed region.

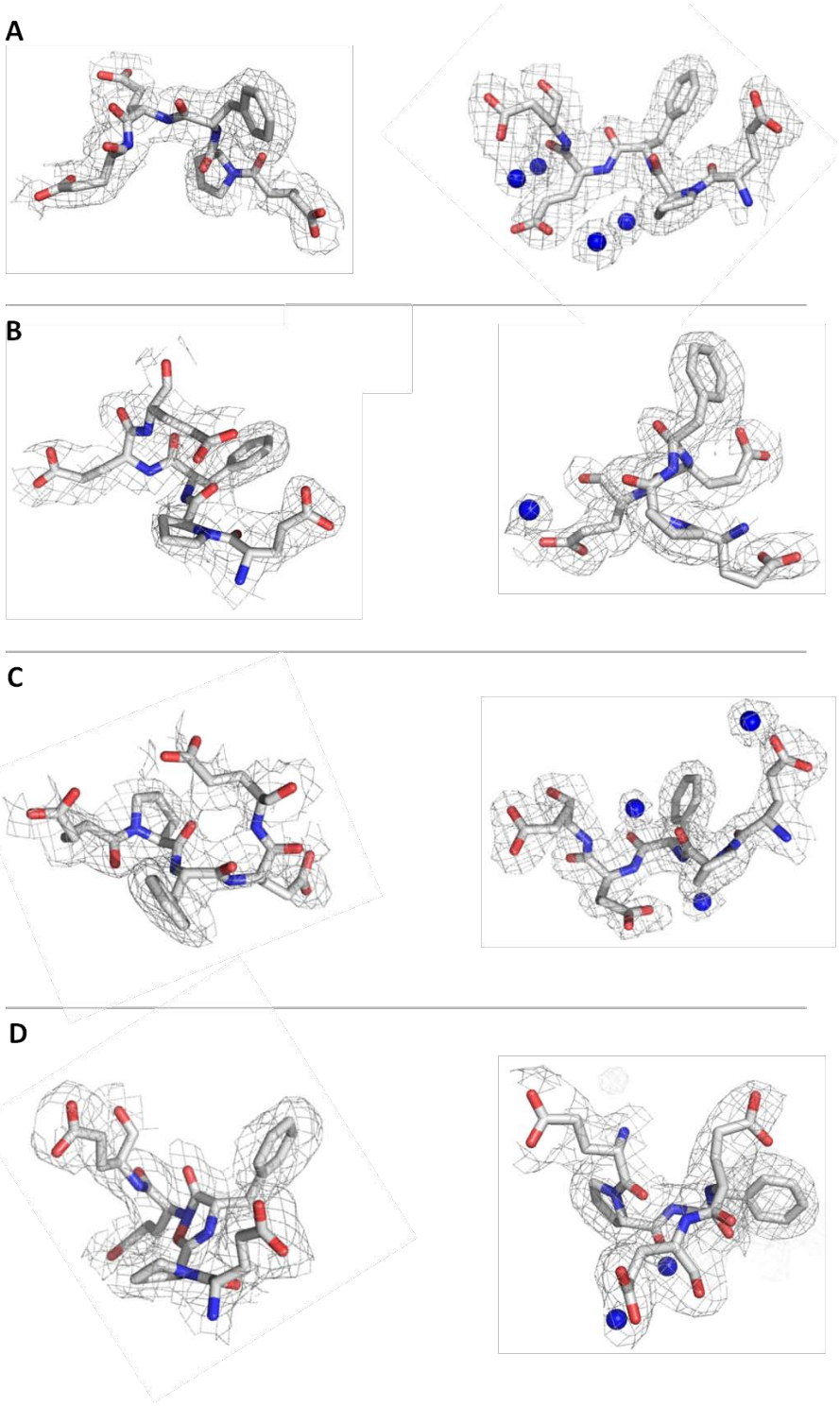

Figure 6.19 Sample regions from the R82A electron density map from the beginning and end of the refinement process

Electron density maps of the Ycf54 mutant R82A, showing the region encompassing residues E18 P19 F20 E21 E22, for each of the four chains in the unit cell; Chain A (**A**), Chain B (**B**), Chain C (**C**) and Chain D (**D**). The images to the left represent the electron density from the initial model generated from molecular replacement using the wild type *Synechocystis* YCF54 and the images to the right represent the final refinement model. In white are carbon atoms, in red are the oxygen atoms, in blue are the nitrogen atoms and the waters are shown as blue spheres.

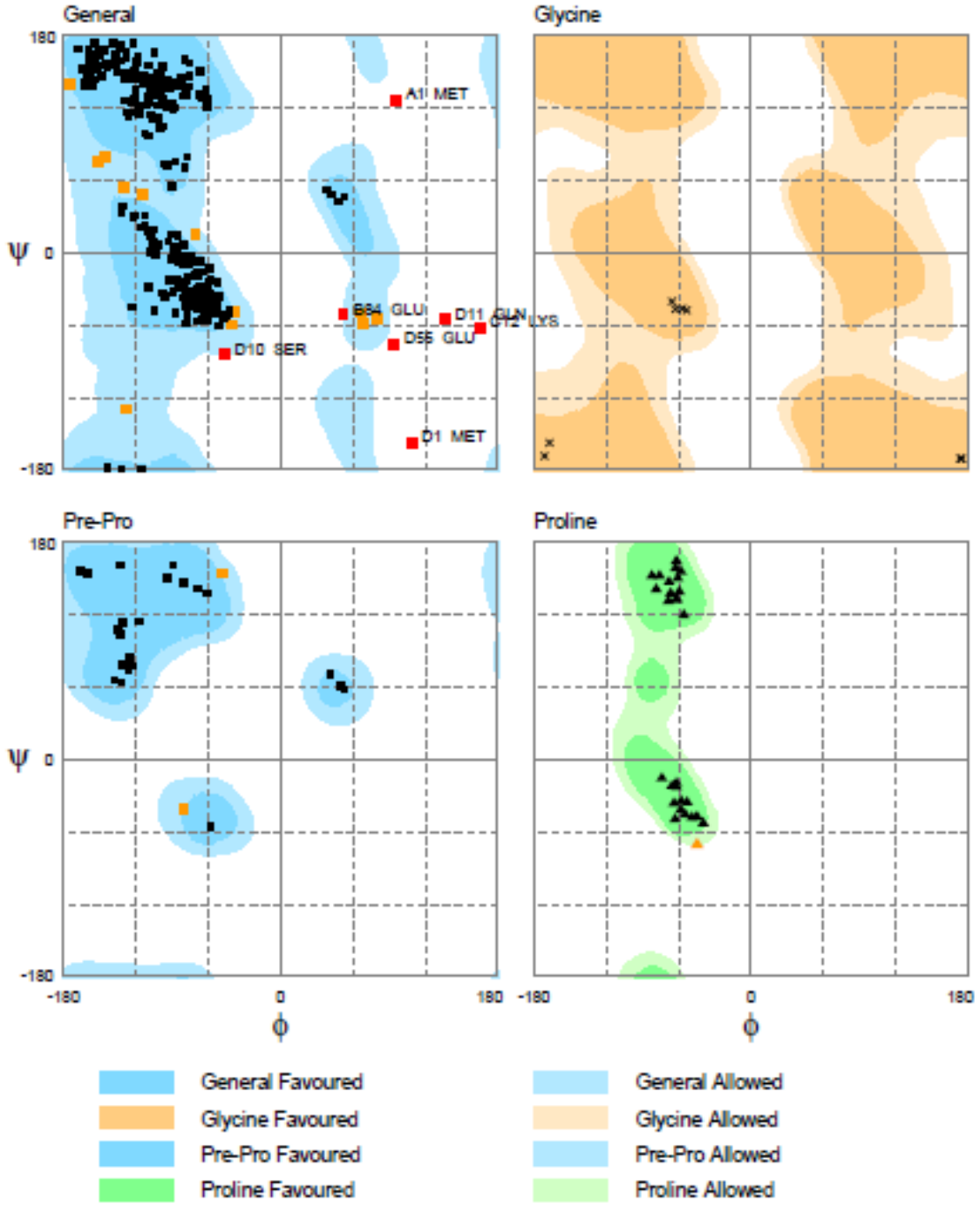

Figure 6.20 Ramachandran plots of the R82A Ycf54 mutant structure

Ramachandran plots generated from the Richardson's data using RAMPAGE in the CCP4 suite for all four chains of R82A present within the unit cell. The majority of residues are within the favoured regions, with 94.8 % of residues located within the favoured region, 3.5 % of residues located within the allowed region and 1.7 % of residues located within the outlier region.

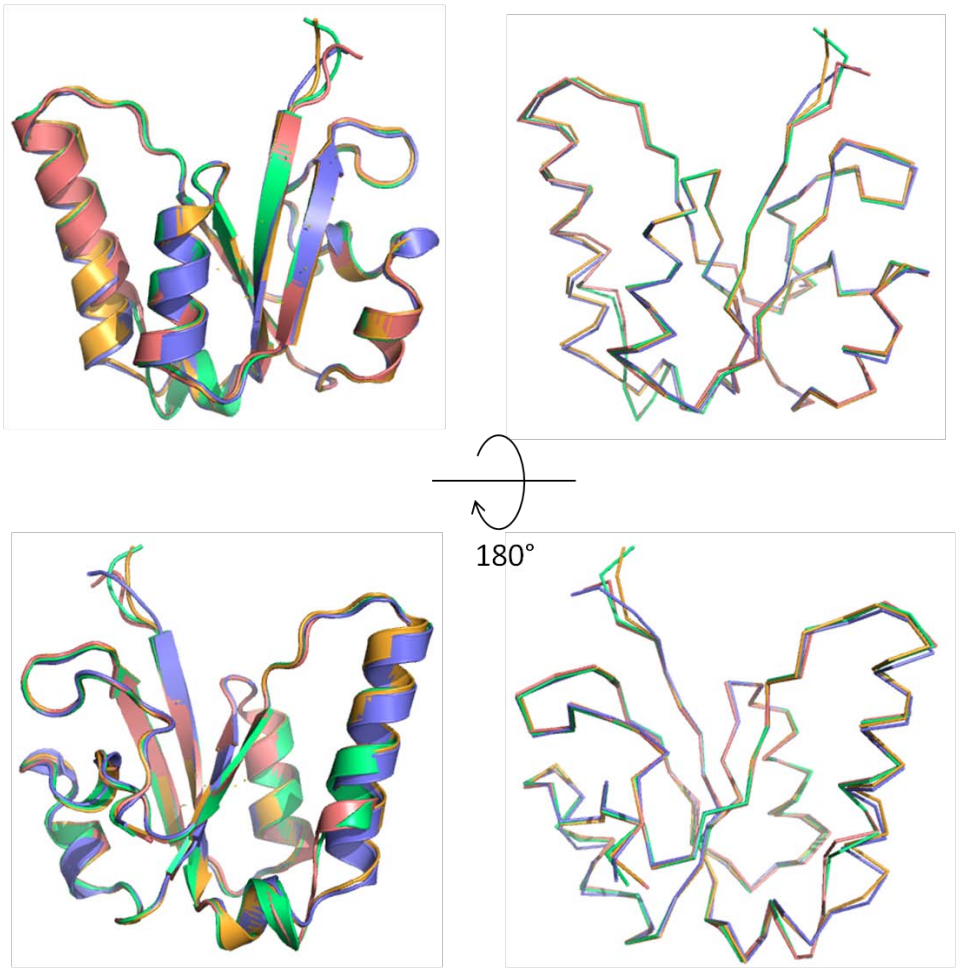

Figure 6.21 Superimposition of the four YCF54 molecules located in each R82A unit cell

Secondary structure alignment and superimposition of the R82A mutant Ycf54 proteins located within each unit cell; Chain A (green), Chain B (orange), Chain C (pink), Chain D (mauve). Diagrams to the left show how the secondary structures align and the diagrams to the right show the alignment of the Cα backbone. Chain A on Chain B RMSD = 0.241 Å, Chain A on Chain C RMSD = 0.254 Å, Chain A on Chain D RMSD = 0.381 Å, RMSD values are inclusive of all atoms.

6.2.10 Comparison of wild type and mutant Ycf54 structures

To investigate the effects caused by the two point mutations A9G and R82A, the crystal structures obtained for these mutant Ycf54 proteins were compared with the wild type Ycf54 crystal structure. Both the wild type and A9G crystal structures were found to crystallise within the same space group (P 2 2 2_1) with a single molecule in the asymmetric unit, whereas R82A was found to crystallise in an alternative space group (P 2_1 2_1 2_1) with four molecules (Chains A, B, C and D) in the asymmetric unit. As displayed in **Figure 6.21** the four chains of R82A crystallise in two different conformations, whose secondary structures are very similar aside from the difference that occurs at the N-terminal end of helix α2. Therefore in all structural comparisons of R82A with wild type Ycf54, the two chains with the best density (Chain A and Chain B) representing both crystallisation conformations were used.

Figure 6.22 displays the superimposition of the secondary structures of the wild type, A9G, R82A Chain A and R82A Chain B Ycf54 molecules. These images and the RMSD of the Cα backbone calculated show all the structures to have a highly conserved secondary structure with no major abnormalities. The only region at which the structures visibly deviate is at the N-terminal of helix α2 (highlighted in **Figure 6.22**), where wild type and R82A Chain B have the same conformation and A9G and R82A Chain A follow the same conformation. This suggests that the Ycf54 structure may be flexible in this region, which is why two conformations, conserved between two different crystal structures, are observed. Additionally, the crystal structures of the Ycf54 homologues from *Thermosynechococcus elongatus* and *Nostoc* sp were observed to crystallise in the same conformation as A9G and R82A Chain A (**Figure 6.13**), lending validity to this hypothesis.

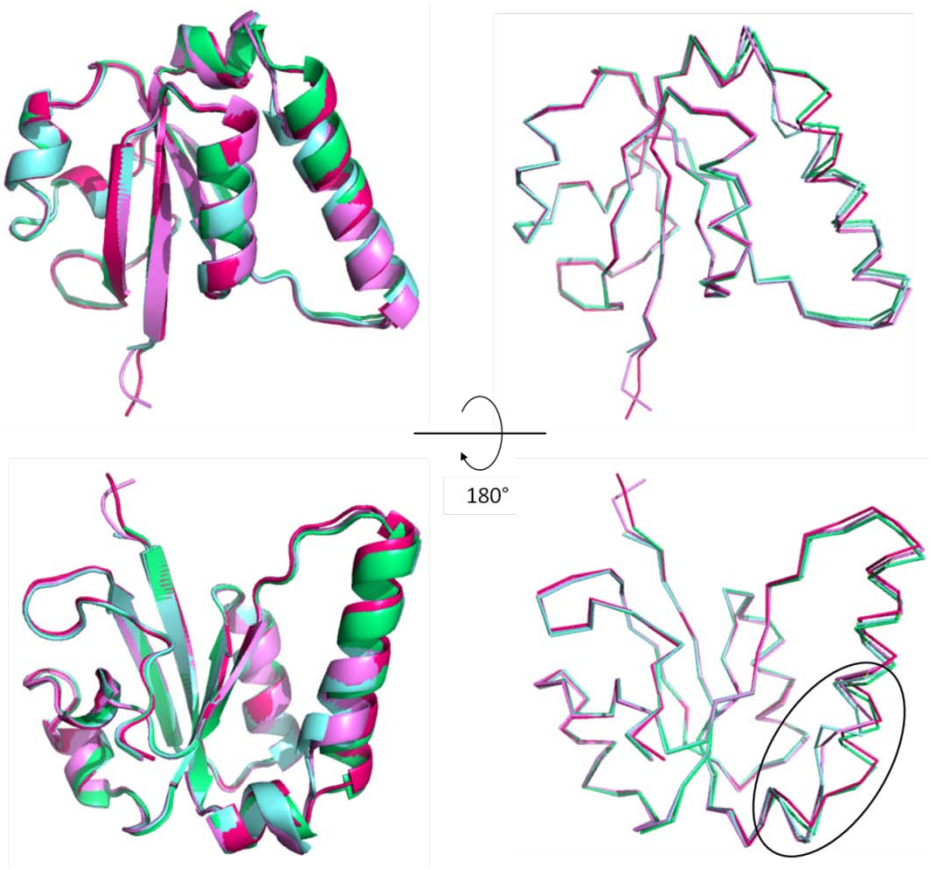

Figure 6.22 Superimposition of wild type A9G and R82A Ycf54 structures

Secondary structure alignment and superimposition of Ycf54 proteins wild type (green), A9G (blue), R82A Chain A (violet) and R82A Chain B (pink) generated using PyMol. Diagrams to the left show how the secondary structures align and the diagrams to the right show the alignment of the Cα backbone. A9G on wild type RMSD = 0.193 Å, R82A Chain A on wild type RMSD = 0.282 Å and R82A Chain B on wild type RMSD = 0.335 Å. Highlighted by the circle is the region on greatest disparity between the structures.

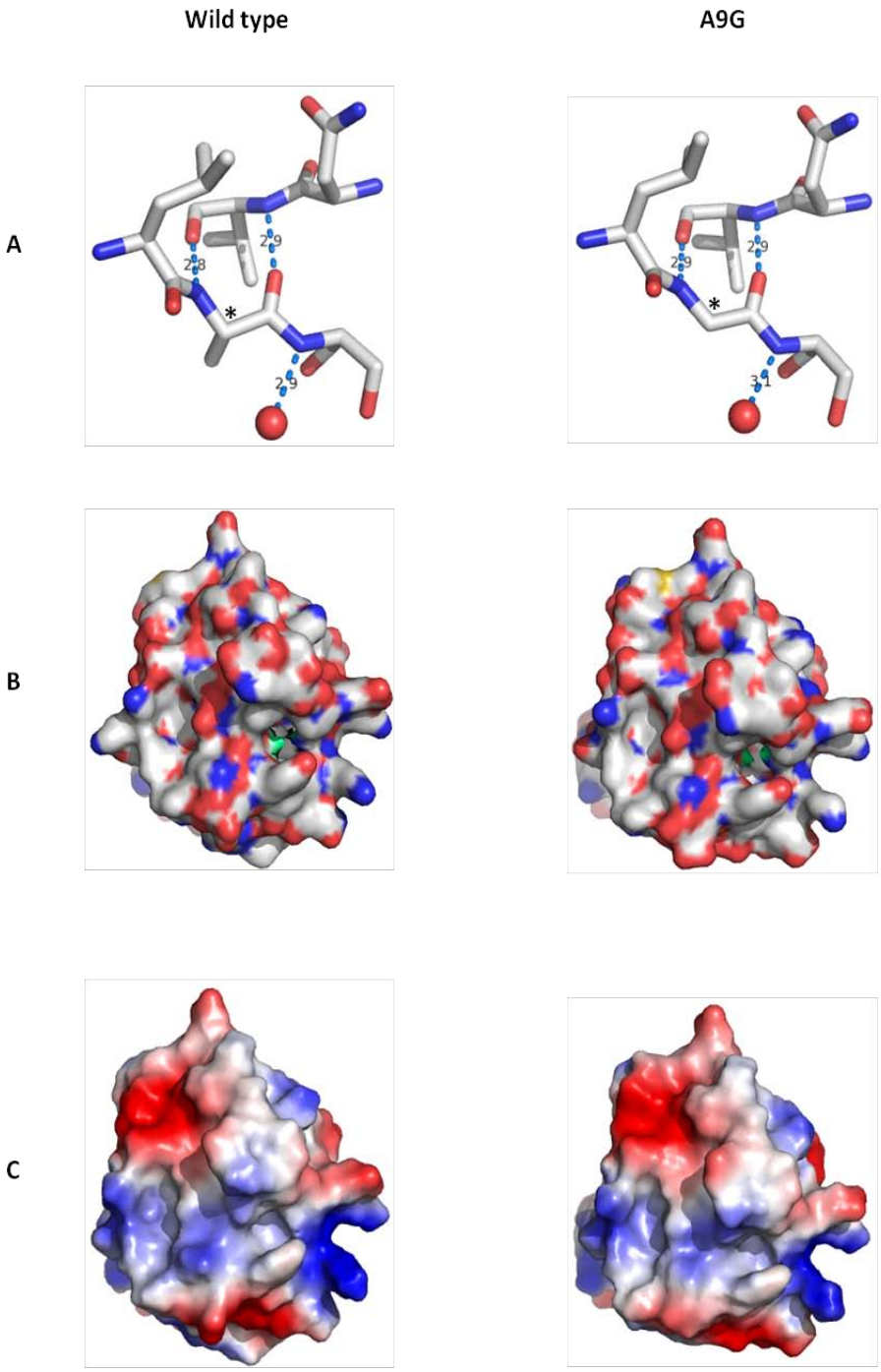

Figure 6.23 Analysis of the hydrogen bonds and electrostatic surface charges in the vicinity of the A9G base substitution

Analysis of the hydrogen bonding between residue 9 and the surrounding residues in the wild type Ycf54 structure and A9G mutant structure is shown in (**A**); residue 9 is highlighted by an asterisk (*****), the waters are shown as red spheres and the hydrogen bonds are marked by blue dashes. The location of residue 9 on the surface of the wild type and A9G Ycf54 proteins is highlighted in green (**B**) and representations of the surface electrostatics on this face of the protein are shown in (**C**). Surface electrostatics were calculated using the Adaptive Poisson-Boltzman Solver (APBS) in PyMOL. Blue: positive charge, Red: negative charge, White: neutral.

As no major alterations in the secondary structure are observed for either point mutation, the effects of A9G and R82A base substitutions on the conformation of the local structure and surface charges was examined. **Figure 6.23 A** shows the hydrogen bonds in the local region surrounding residue 9 are identical between the wild type and A9G structures. Similarly the surface exposed area of residue 9 and the surface electrostatics observed are identical between the two structures (**Figure 6.23 B** and **C**), demonstrating the loss of a CH_3 R-group from residue 9 has no discernible effects on the local structure of the surrounding area.

Unlike the A9G point mutation, closer examination of the structure surrounding residue 82 shows alterations in the hydrogen bonding network (**Figure 6.24**) and the surface electrostatics (**Figure 6.25**), despite there being no obvious alterations in the conformation of the residues surrounding residue 82 (**Figure 6.26**).

In the wild type structure R82 crystallises in two conformations (**Figure 6.4**). In one, R82 forms two direct hydrogen bonds with the side chain of E17 as well as the backbone carbonyl of L83, and two indirect hydrogen bonds mediated by a water are formed between the backbone carbonyl of F20 and the side chain of E17. In the second conformation, R82 forms a water-mediated hydrogen bond with the side chain of W78 (**Figure 6.24**). However, upon substitution of R82 for an alanine, none of the hydrogen bonds between the arginine side chain and the surrounding residues may form, thus removing this network of interactions. Instead an "empty" space is observed where R82 should reside and no hydrogen bonding interactions occur between the regions encompassing residues E17-F20 of α2 and residues W78-L83 of α5. Furthermore, by substituting the arginine with an alanine, residue 82 can no longer switch between the two conformations observed. Unfortunately, as the precise role of R82 in cyclase function is not known, it is impossible to speculate how the dual conformations of this residue and its local hydrogen bonding network are involved in interacting with the catalytic component of the cyclase, Sll1214/Sll1874, or cyclase activity.

The surface occupancy of residue 82 (viewed **Figure 6.25 A** where R82 or A82 is highlighted in green in the wild type and mutant structures respectively) is vastly reduced in the mutant compared to the wild type, altering the topography of the surface area in this region. Additionally, the surface electrostatics in this region are also altered (**Figure 6.25 B** and **C**); the base substitution R→A changes the surface electrostatics from predominately positive to predominately negative. Thus, it is possible to imagine an electrostatic interaction mediated between the positive region in the vicinity of R82 and a negatively charged region on an as yet unidentified interaction partner (perhaps Sll1214/Sll1874, given R82A can no longer

interact with these proteins), would be disrupted by the altered surface charges on the R82A Ycf54 mutant.

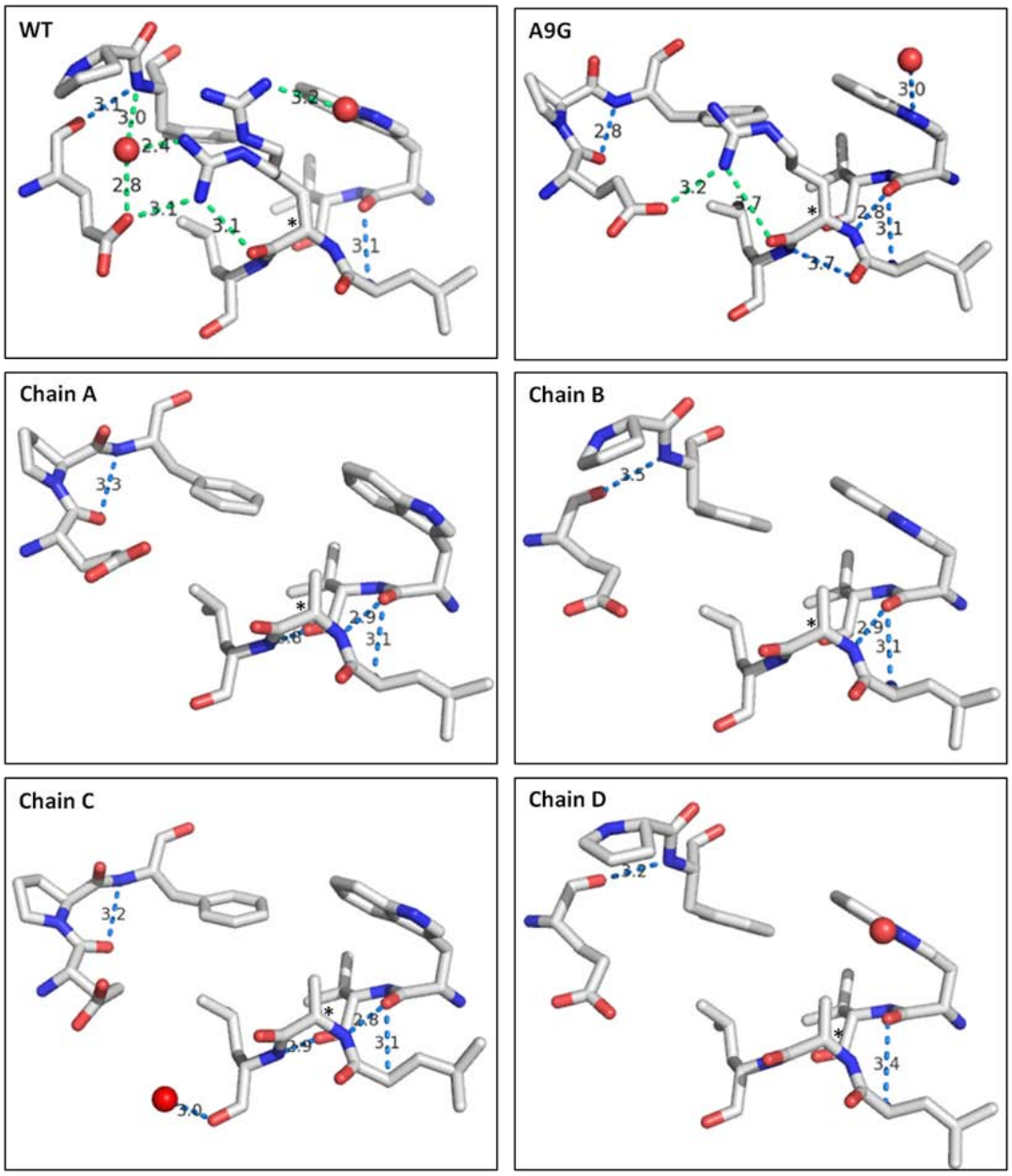

Figure 6.24 Analysis of the hydrogen bonds around residue 82 in wild type Ycf54 and mutant Ycf54 proteins

Analysis of the hydrogen bonding between residue 82 and the surrounding residues in the wild type Ycf54 structure, A9G mutant structure and all four chains of the R82A crystal structure. Residue 82 is highlighted by an asterisk (*), the waters are shown as red spheres, the hydrogen bonds present in all structures are marked by blue dashes, the hydrogen bonds present only in the wild type and A9G structures are marked by green dashes.

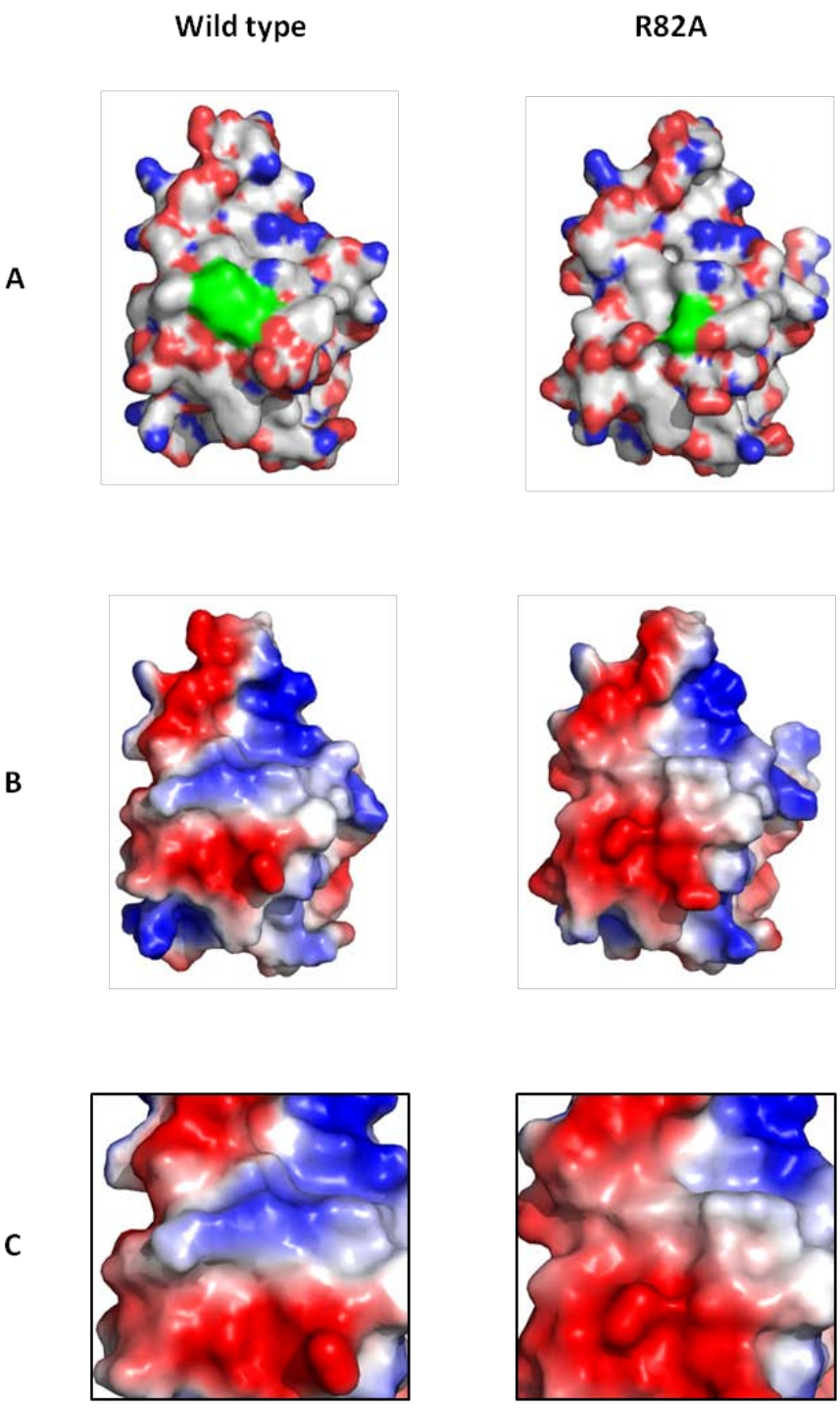

Figure 6.25 Comparison of the surface electrostatics between wild type and R82A structures

The location of the R82A residue on the surface of the wild type and R82A Ycf54 proteins is highlighted in green (**A**). Representations of the surface electrostatics of this face of the protein are shown in (**B**) and a close up of the region encompassing residue 82 in the wild type and R82A Ycf54 structures is shown in (**C**). Surface electrostatics were calculated using the Adaptive Poisson-Boltzman Solver (APBS) in PyMOL. Blue: positive charge, Red: negative charge, White: neutral.

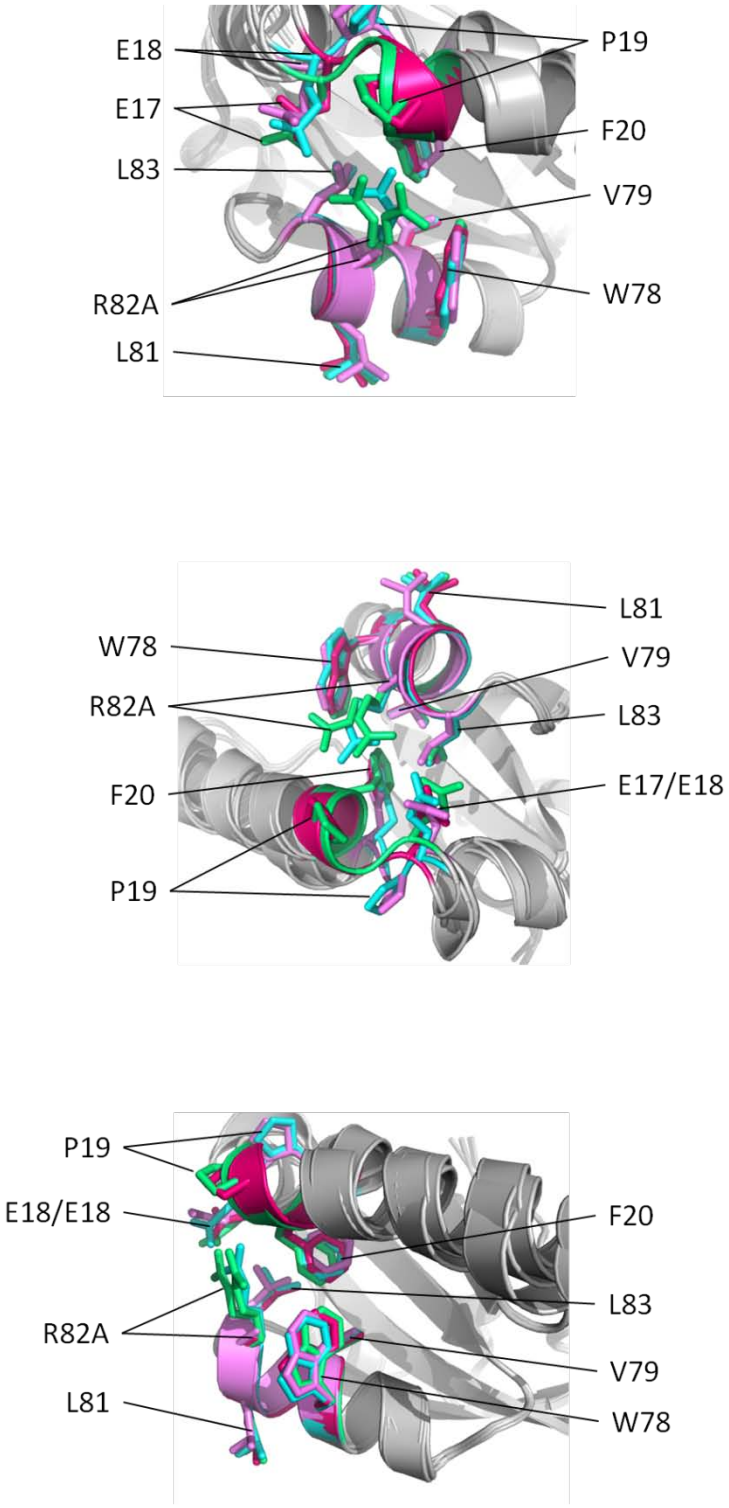

Figure 6.26 Superimposition of the residues surrounding residue 82 in the wild type A9G and R82A Ycf54 structures

Secondary structure alignment and superimposition of the residues surrounding R82A in the Ycf54 proteins wild type (green), A9G (blue), R82A Chain A (violet) and R82A Chain B (pink) generated using PyMOL. In the wild type structure, R82 adopts a dual conformation, which is shown in the diagrams above.

6.3 Discussion

It is commonly accepted that the ancestor of the plastids found in plants and algae was once a free-living cyanobacterium that became enslaved by a host eukaryote cell. Over the course of evolution, many essential genes encoded within the symbiont were transferred to the hosts' nucleus, whilst others became redundant and were lost (Race et al., 1999). However, a small number of proteins (~10 %) essential for fully functional chloroplasts have been retained in the plastid genome. Analyses of the genetic information retained in the plastid genomes of various plant and algal species showed that they primarily encode components essential for the biochemistry of the plastid, specifically photosynthetic proteins, integral membrane proteins, fatty acid biosynthesis proteins and the associated machinery required for their expression (Leister, 2003). Of these genes, ~80-90 open reading frames encode proteins not involved in the gene-expression/ translation machinery and the functions of a large proportion of these have been elucidated (Kleine et al., 2009). The remaining genes, with no known function were given the prefix *ycf* (hypothetical chloroplast open reading frame). Ycf54 is one of these unknown hypothetical genes and its function has recently started to be elucidated through the work presented in this thesis, in addition to the manuscripts of Hollingshead et al., (2012) and Albus et al., (2012) and the doctoral thesis of Wallner T (2012).

Ycf54 is a highly conserved protein with identifiable homologues in all sequenced oxygenic photosynthetic organisms. All homologues share the DUF (domain of unknown function) domain DUF2488 also known as the Ycf54 domain. Examination of the evolutionary relationships between the Ycf54 homologues using molecular phylogenetic analyses (**Figure 6.13**) show the plant Ycf54 homologues to be closely clustered, indicating tight evolutionary constraints, whilst the Ycf54 homologues from stramenopiles, mosses, green algae, red algae and cyanobacteria are interspersed with each other, a possible outcome of horizontal gene transfer in addition to adaptation to specific ecological niches.

The primary goal of this chapter was to determine the crystal structure of wild type Ycf54 and the structures of the Ycf54 mutants (A9G, D39A, F40A and R82A) found to have an *in vivo* phenotype in the experiments conducted in Chapter 5. Of these four point mutations, three (D39A, F40A and R82A) were found to affect the ability of the cell to convert MgPME to Pchlide and one (A9G) was found to affect the accumulation of Ycf54 and its interaction partners Sll1214 and Sll1874, but not the ability of the cell to produce Pchlide. To further understand how these mutations affect the inherent structure of Ycf54 attempts were made to solve the crystal structures of each of these proteins. As a result, the crystal structures of

three proteins were successfully solved, wild type, A9G and R82A. Unfortunately, D39A did not crystallise under any of the conditions trialled and although F40A did form crystals, none of these diffracted to a resolution better than 4 Å.

6.3.1 Summary of wild type Ycf54 structure

The structure of wild type Ycf54 was successfully solved to a resolution of 1.3 Å. The final refinement model reveals *Synechocystis* Ycf54 to consist of a single domain, in which a four stranded anti-parallel β-sheet is situated between two short sections of α-helices (**Figure 6.7**). Comparison of this structure to the two Ycf54 structures from *Thermosynechococcus elongatus* and *Nostoc* sp existing in the PDB (**Figure 6.13**) shows that as well as sharing a high sequence identity, these three structures also share high structural similarity. Given the Ycf54 domain is highly conserved throughout all oxygenic photosynthetic organisms (**Figure 5.1** and **Figure 6.12**) the level of sequence and structural homology observed further underlines the importance of this protein.

From the work presented in Chapters 4 and 5 of this thesis and by Albus et al., (2012) and Wallner (2012), it can be concluded that Ycf54 is required for normal procession of the cyclase reaction in *Synechocystis* and *Nicotiana tabacum*; thus, Ycf54 can be regarded as a component of the cyclase complex. Previous biochemical studies using fractionated extracts from cucumber cotelydons (Walker et al., 1991), barley plastids (Rzeznicka et al., 2005), *Synechocystis* cells and *Chlamydomonas rehinhardtii* cells (Bollivar and Beale, 1996a) all found that the cyclase complex consists of several subunits, one which is located in the soluble fraction and two that are membrane bound. The identity of only one cyclase subunit has been discovered, which is the membrane bound AcsF protein (Pinta et al., 2002; Tottey et al., 2003a; Rzeznicka et al., 2005; Minamizaki et al., 2008a). Western blots of fractionated *Synechocystis* cell lysates probed with anti-Ycf54 show that Ycf54 is located predominantly in the soluble fraction. In addition the recombinant Ycf54 protein was found to be highly soluble. Taken together these data suggest Ycf54 may be the unidentified soluble component of the cyclase. However, biochemical assays using fractionated *Synechocystis* membranes and recombinant Ycf54 should be conducted to obtain conclusive evidence.

Often referred to as the catalytic subunit, AcsF (known as Sll1214/Sll1874 in *Synechocystis*) belongs to a family of di-iron carboxylate proteins, a group of non-heme and non-transmembrane iron proteins that co-ordinate a binuclear iron centre via six conserved amino acids (four carboxylate, two histidine). Unlike other mono-oxygenase enzymes, the family of di-iron carboxylate proteins are not soluble and do not traverse the membrane, instead they are bound interfacially to one leaflet of the lipid bilayer (Berthold and Stenmark,

2003). In addition to the cyclase, other members of this family include the alternative oxidase, the plastid terminal oxidase and the di-iron 5-dethoxyquinone hydroxylase. Biochemical studies using fractionated lysates all report that a reducing agent (NADPH) was required for cyclase activity, however analysis of the AcsF and Ycf54 sequences shows there no Rossmann fold in either protein, meaning an additional subunit, with the ability to bind and reduce NADPH remains to be discovered.

To further elucidate the function of Ycf54, the Dali server was employed to identify structural homologues. Two of the top hits, D-lactate dehydrogenase and the nitrogen regulatory protein P_{II}, were of interest because of their respective roles in electron transfer and orchestration of balancing nitrogen fixation and photosynthesis (**Figure 6.14**) (Dym et al., 2000; Fokina et al., 2010). However, without further information on how Ycf54 interacts with the other members of the cyclase complex and the mechanism of the cyclase reaction, it is difficult to draw conclusions from these structural homologues and how they may relate to the function of Ycf54.

6.3.2 Summary of mutant Ycf54 structures

In addition to the structure of wild type Ycf54, the structures of two Ycf54 mutants, A9G and R82A, were solved to resolutions of 1.5 Å and 2.2 Å respectively. These two mutants were both found to have a detectable *in vivo* phenotype in Chapter 5 and in order to understand how these mutations may alter the function of Ycf54 their crystal structures were determined.

Neither secondary structure was significantly altered (**Figure 6.22**) as an overlay of the mutant and wild type Ycf54 structures showed all the Cα atoms to have an RMSD of 0.2 Å for A9G and 0.3 Å for R82A. Thus, the phenotypes observed in A9G and R82A are not a result of the mutant proteins adopting a different conformation or failing to fold. In addition, analyses of the electrostatic surface potential and local hydrogen bonds surrounding residue 9 in A9G, showed no differences between wild type Ycf54 and the A9G mutant (**Figure 6.23**). In conclusion there is no obvious structural basis for the reduced accumulation of Ycf54 observed in the *Synechocystis* strain A9G. Nevertheless, this mutant has been important in determining that the cellular accumulation of Ycf54 and Sll1214 is likely to be in excess, as the *Synechocystis* strain A9G is able to synthesis chlorophyll to the same levels observed in wild type *Synechocystis* (**Table 5.1**).

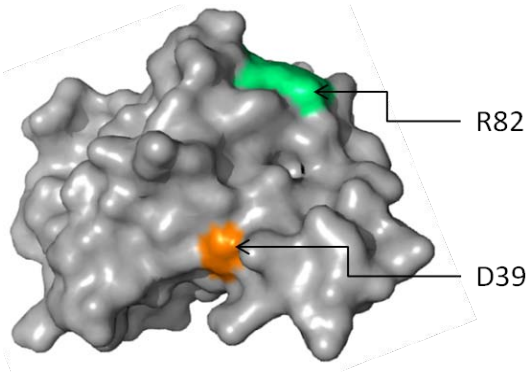

Figure 6.27 Surface availability of conserved residues D39, F40 and R82

Image of the surface area occupied by the conserved residues D39 (orange) and R82 (green) generated using PyMOL.

Upon closer comparison of the R82A and wild type Ycf54 structures, two disparities could be observed. Firstly, there is an alteration in the electrostatic surface potentials at the site of residue 82. In the wild type structure, R82 gives the local surface a positive charge; replacement of this residue with an alanine results in the local surface adopting an overall negative charge (**Figure 6.25**). Secondly, there are several alterations in the network of hydrogen bonds surrounding residue 82. In the wild type, R82 forms several stabilising hydrogen bonds with both "in-space" and "in-sequence" neighbouring residues, directly and through a hydrogen-bonded water molecule (**Figure 6.24**). When replaced with an alanine, residue 82 can no longer form these interactions, which may alter the stability of this region. Lastly, R82 was observed to crystallise in two distinct conformations (**Figure 6.4**), which is no longer an option in the R82A mutant. It may be that the flexibility of R82, along with its associated positive surface potential is required for docking of Ycf54 onto Sll1214/Sll1874, as pulldown assays using FLAG-R82A as bait show this mutant is no longer able to interact with the catalytic subunit of the cyclase (**Figure 5.10**).

6.3.3 Conclusions and future work

Overall the data gathered within this chapter show that the both the sequence and known structures of Ycf54 are highly conserved, which highlights the importance of this protein in chlorophyll biosynthesis. However, the specific role of Ycf54 along with its purpose in the cyclase complex has not yet been fully elucidated; several factors remain to be investigated, some of which are discussed below.

The site at which Ycf54 docks onto the catalytic subunit(s) of the cyclase Sll1214/Sll1874 is unknown. In Chapter 5, *in vivo* pulldown assays using the FLAG-tagged Ycf54 mutants D39A, F40A and R82A showed that none of these mutants were able to interact with

Sll1214/Sll1874. However, when the surface occupancy of each of these molecules is viewed (**Figure 6.27**), it becomes apparent that F40 is buried and that D39 and R82 could be considered to be on different faces of Ycf54, separated by a "over-surface" distance of ~25 Å. Without a structure of D39A or F40A, the effects of these mutations on the folding of Ycf54 is not known, so conclusions cannot be drawn as to how these point mutations affect the interaction between Ycf54 and Sll2114/Sll1874. Nevertheless, it is unlikely that D39A, F40A and R82A form a docking area for Sll1214/Sll1874 on the surface of Ycf54. The most ideal method to map the regions of Sll1214/Sll1874 that interact would be to co-crystallise Sll1214/Sll1874 and Ycf54 to obtain a structure of the complex. However, there are two reasons why this is unlikely to be successful; firstly, Sll1214/Sll1874 is a membrane-associated protein and requires solubilising with detergents in order to be purified from *Synechocystis* (where it only accumulates at low levels, even when expressed with a FLAG-tag from the strong light driven promoter *psbAII*), secondly, numerous expression constructs and conditions have been trialled to produce recombinant Sll1214 or Sll1874, none of which have yielded significant protein.

Although Ycf54 is a soluble protein, located in the cytoplasmic fraction of *Synechocystis* (**Figure 4.7**), it is not known whether it is the soluble protein required for reconstitution of cyclase activity in assays using cell lysates. As discussed in Chapter 4, biochemical assays using the membrane fraction of *Synechocystis* and recombinant Ycf54 should be conducted, using the method of Bollivar and Beale (1996) to determine if Ycf54 is the elusive soluble protein described by Walker et al., (1991).

Upon analysis of the surface electrostatic charges of *Synechocystis* Ycf54, a negatively charged ridge, located on α2, was discovered. Closer inspection of the residues responsible for this area of negative charge, revealed two of the glutamate residues, E22 and E26, to be two of the seven residues that are absolutely conserved across all known Ycf54 homologues (**Figure 5.1**). Comparison of the surface electrostatic charges on the corresponding helices in the *Thermosynechococcus elongatus* and *Nostoc sp* Ycf54 proteins (**Figure 6.28**), showed that these two proteins also have a negatively charged ridge, a strong indication that this feature is structurally conserved. In Chapter 5, E22 and E26 were individually point mutated to an alanine *in vivo*, however no detectable phenotype was observed in either of the *Synechocystis* strains Ycf54-E22A or Ycf54-E26A. As discussed earlier, the observable lack of phenotype may be due a single point mutation not having a significant impact on the overall negative charge of this area; thus any electrostatic interactions in which this ridge may partake may not be perturbed. Further investigation into the role of this interesting structural feature could involve the insertion of several lysine resides (so as to reverse the

charge of the ridge) into the native Ycf54 gene *in vivo*, followed by investigation of the strain(s) generated to determine if they have an altered phenotype.

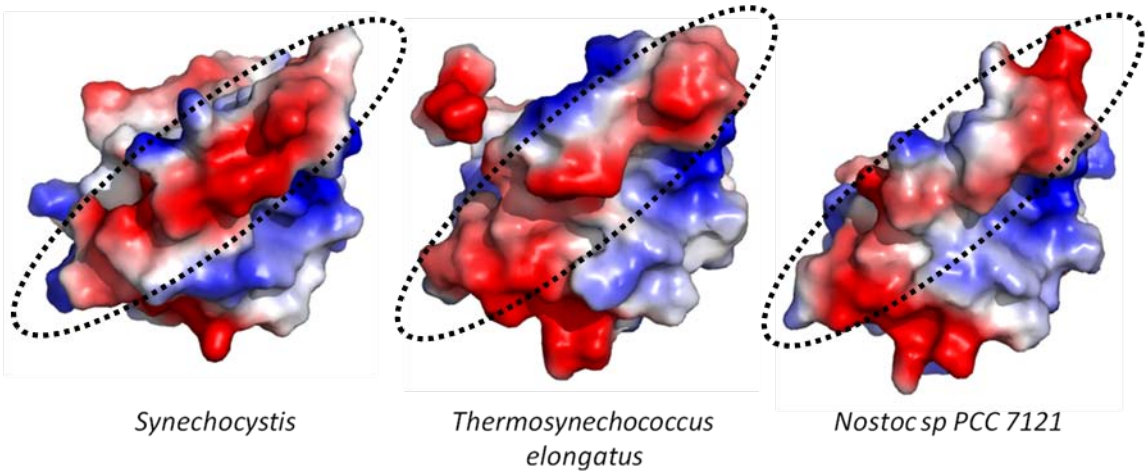

Figure 6.28 Acidic ridge is conserved between *Synechocystis*, *Thermosynechococcus elongatus* and *Nostoc sp* Ycf54 proteins

The surface electrostatic charges of the electronegative ridge, located on helix 2 of *Synechocystis* Ycf54, are displayed for the *Synechocystis*, *Thermosynechococcus elongatus* and *Nostoc sp* Ycf54 proteins. Surface electrostatics were calculated using the Adaptive Poisson-Boltzman Solver (APBS) in PyMOL. Blue: positive charge, Red: negative charge, White: neutral.

Table 6.1 Data collection statistics for Ycf54 crystals

Dataset	**Ycf54**	**A9G**	**R82A**
Spacegroup	C 2 2 2_1	C 2 2 2_1	P 2_1 2_1 2_1
Unit cell parameters			
A	42.73	42.20	55.52
B	45.93	46.13	91.15
C	119.29	119.66	120.30
A	120	90	90
B	120	90	90
Γ	120	90	90
Matthews coefficient $Å^3$ Da	2.03	2.17	2.99
Molecules per asymmetric unit	1	1	4
Temperature (K)	100	100	100
X-ray source	Diamond IO2	Diamond IO2	Diamond IO2
Detector			
Resolution (Å)	1.31	1.50	2.20
Energy (keV)	0.979	0.979	0.979
Mosaicity (°)	0.332	0.165	0.09
Unique observations	28342	18732	31210
R_{merge}	0.04	0.08	0.05
R_{prim}	0.03	0.04	0.02
Completeness (%)	98.9 %	98.5	98.7
Multiplicity	4.4	6.1	6.6
Anomalous multiplicity	2.2	3.1	3.4
Mean $\left(\frac{[dI]}{s[dI]}\right)$	_	1.019	0.913

Table 6.2 Refinement statistics for Ycf54 structures

Parameter	Ycf54	A9G	R82A
Resolution (Å)	1.31	1.50	2.20
Number of reflections	28342	18732	31210
Used reflections	26909	17771	29636
Number of atoms	1029	959	3637
Number of waters	120	83	149
Number of ions	2	1	0
Un-modelled residues	0	0	0
Residues truncated to cα	0	0	0
Ramachandran favoured (%)	98.10	98.10	94.8
Ramachandran outliers (%)	0	0	1.7
Poor rotamers	0	0	
RMSD bond length (Å)	0.03	0.02	0.02
RMSD bond angle (°)	2.22	2.01	1.67
RMS chir volume	0.16	0.14	0.10
Average B factors ($Å^2$)			
Main chain	19.8	-	-
Waters	18.2	-	-
R-factor (%)	18.86	20.60	18.90
R_{free} (%)	21.68	26.00	26.00

Table 6.3 Top Dali hits with *Thermosynechococcus elongatus* PDB 3HZE Chain A used as the search model

PDB code	Chain	Species	Description	Z-score	RMSD (Å)	Sequence Identity (%)
3JSR	A	*Nostoc* Sp. PCC 7120	Ycf54	19.4	0.9	73
1Q8I	A	*Escherichia coli*	DNA Polymerase II	4.7	3.0	14
1F0X	A	*Escherichia coli*	D-Lactate dehydrogenase	4.5	3.4	10
1MW7	A	*Helicobacter pylori*	Transcriptional regulatory protein (HP0162)	4.3	2.8	9
2DCL	B	Pyrococcus horikoshii	Hypothetical protein (PH1503)	4.3	3.2	7
2BOA	B	Human	Carboxypeptidase A4	4.1	3.2	4
3L7P	D	*Streptococcus mutans*	Nitrogen regulatory protein PII	4.1	3.3	6
2QFM	A	Human	Spermine synthase	4.1	3.7	10

Z-score is a measure of the statistical significance of a match in terms of Gaussian statistics. The higher the Z score, the higher the statistical significance of the match.
RMSD is the Root Mean Square Deviation calculated between Cα atoms of matched residues at the best 3D superposition of the query and target structures. Generally the larger the RMSD, the more distant the matched structures are.

Table 6.4 Top Dali hits with *Nostoc* sp PCC 7120 PDB 3JSR used as the search model

PDB code	Chain	Species	Description	Z-score	RMSD (Å)	Sequence Identity (%)
3HZE	A	*Thermosynechococcus elongatus*	Ycf54	19.5	0.8	73
3K59	A	*E. coli*	DNA Polymerase II	5.0	4.3	12
3MHY	C	*Azospirillum brasilense*	PII like protein	4.4	3.6	7
2J9D	F	*Methanocaldococcus jannaschii*	PII like protein	4.4	4.0	8
3DFE	D	*Anabaena variabilis ATCC 29413*	PII like protein	4.4	3.4	7
2DCL	B	*Pyrococcus horikoshii*	Transcriptional regulatory protein (HP0162)	4.3	2.8	9
3MO5	B	Humans	Uncharacterised protein PEPE_1480	4.3	4.3	1

Z-score is a measure of the statistical significance of a match in terms of Gaussian statistics. The higher the Z score, the higher the statistical significance of the match.
RMSD is the Root Mean Square Deviation calculated between Cα atoms of matched residues at the best 3D superposition of the query and target structures. Generally the larger the RMSD, the more distant the matched structures are.

7. Investigating the role of Δ*slr0483* in Synechocystis

7.1 Introduction

In a recent study by Formighieri *et al* (2012), quantitative transcription profiling of a *Chlamydomonas reinhardtii gun4* insertion mutant was performed on cells grown in the dark. The authors found 69 nuclear genes that encoded predicted chloroplast proteins that were differentially expressed in *gun4*$^{-}$ by more than three-fold when compared to the wild type. These proteins were found to be involved in a variety of cellular processes including signalling, photosystem/ light-harvesting complexes, tetrapyrrole metabolism, gene expression, chromatin reorganisation and nitrogen metabolism. Of interest are the up-regulated tetrapyrrole metabolism genes: GSA, UROIII, CPOIII, the magnesium chelatase subunits ChlH and ChlI (but not ChlD) and the chlorophyll biosynthesis enzymes ChlM, POR1, ChlP, and the up-regulated photosystem/ light harvesting genes: LHCBM4, LHCBM5, CP26, LHCSR1 and LHCA1-9. Such global changes in the expression of the genes associated with photosynthesis in the absence of GUN4, implies that GUN4 may have a wider role in regulating the tetrapyrrole biosynthetic pathway, as well as it's well documented role as part of the MgCH complex (Larkin et al., 2003b; Davison et al., 2005b). In addition to the known photosynthetic genes, 11 genes with no known functional annotation were also found to be differentially expressed. Given the number of tetrapyrrole metabolism and light-harvesting genes whose expression is affected in the absence of GUN4, it is plausible some of these may also have an as yet unidentified role in photosynthesis. Therefore, each of the 11 unknown genes was subjected to BLAST searching, from which one gene was identified as potential photosynthesis candidate, as it is highly conserved across the lineage of photosynthetic organisms. This gene corresponds to *slr0483* in *Synechocystis*.

In Chapter 3, Slr0483 was found in the eluates from FLAG pulldown experiments using the bait proteins CT-FLAG-ChlM, CT-FLAG-Sll1214 and CT-FLAG-ChlP and has also been identified in the eluates from pulldown experiments using NT-FLAG-Protoporphryinogen oxidase, NT-FLAG-Sll1214 and NT-FLAG-FeCH (Dr. Roman Sobotka, personal communication). As the Slr0483 protein represents a highly conserved protein of unknown function, which is conserved across the oxygenic photosynthetic lineages, it was decided to investigate the role of Slr0483 further.

Slr0483 is a member of the PFAM family CAAD/PF14159 and contains a domain of unknown function, DUF4308, which can found in amino-acyl tRNA enzymes as well as in isolation in

single domain proteins. The CAAD domain stands for Cyanobacterial Aminoacyl-tRNA synthetases Appended Domain, and is found in protein domains that are localised in the thylakoid membrane. The domain consists of two putative transmembrane helices, and is present in glutamyl-, isoleucyl- and valyl-tRNA synthetases, as well as several proteins unrelated to the aminoacyl tRNA synthetases (InterPro 2014).

There are five known domain organisations that contain the CAAD domain (according to the PFAM database on 30 December 2013), which can be viewed in **Figure 7.1**. The most common domain architecture, accounting for 267 or 93 % of the known CAAD containing proteins (287 sequences in total), is when CAAD is found in isolation in single domain proteins and is usually situated at the C-terminus; Slr0483 has this domain architecture. So far the CAAD domain has been found in 100 species, including 69 cyanobacteria, 2 microscopic algae, 1 amoeboid, 8 green algae, 1 moss and 19 vascular plants; interestingly CAAD domains have not been reported in any species of red algae or stramenopiles.

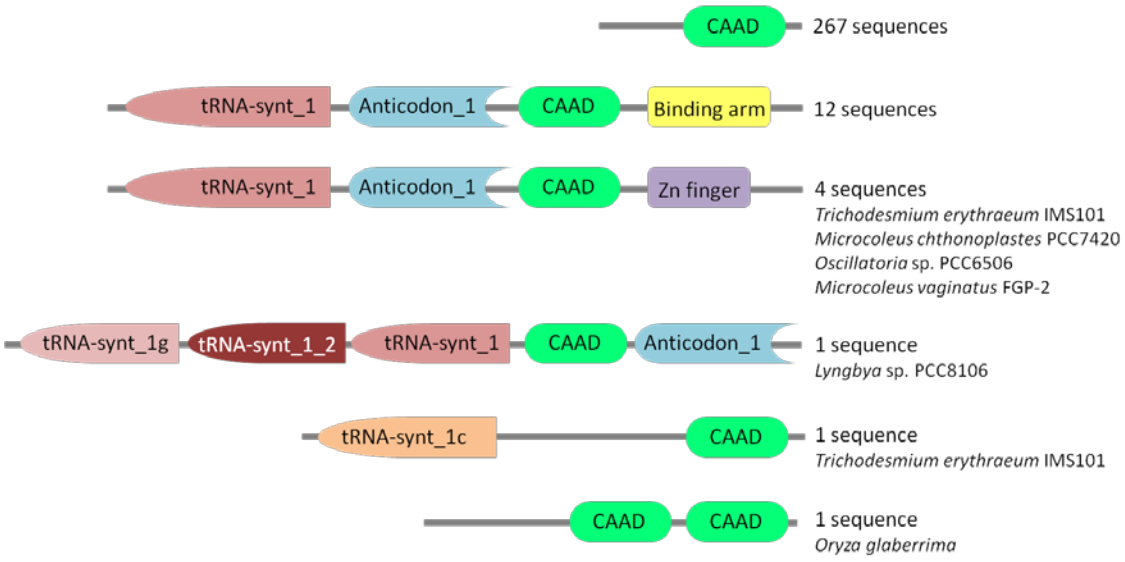

Figure 7.1 Domain architecture in proteins containing the CAAD domain

The six-domain architectures found in proteins containing the CAAD domain, of which Slr0483 is a member.

7.2 Results

7.2.1 Bioinformatics analysis of Slr0483

The Slr0483 primary sequence was subjected to bioinformatics analyses to determine if the protein contained the two predicted transmembrane domains, which constitute the conserved CAAD domain. Analyses performed using the PSIpred program suggests Slr0483

has two transmembrane helices, located towards the C-terminus of the protein (**Figure 7.3 B**), as well as three non-transmembrane helices. The FLAG-tagged Slr0483 protein was found to be associated with the insoluble fraction in FLAG-pull down experiments (**Figure 7.4**), which suggests that, as expected, this protein is membrane associated.

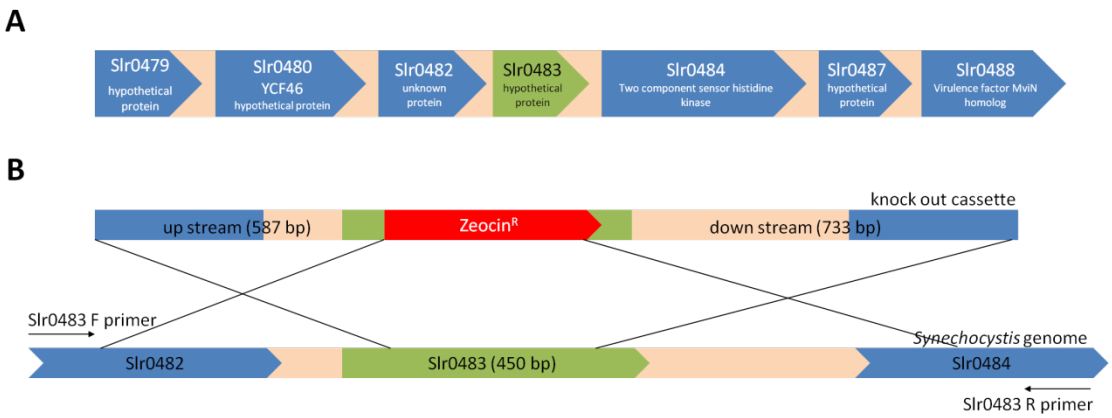

Figure 7.2 Slr0483 locus in the *Synechocystis* genome and knock out cassette used to generate Δ*slr0483*

Location of slr0483 in *Synechocystis* and the flanking genes (**A**); knock out cassette used to generate Δ*slr0483* and the region of the *Synechocystis* genome into which it integrates (**B**).

7.2.2 Creation of a Δ*slr0483* deletion mutant

To investigate the function of the unknown conserved protein Slr0483, a *Synechocystis* strain was constructed in which the *slr0483* gene is interrupted with a zeocin resistance cassette. The strain was constructed using the mega primer method adapted from Ke and Madison (1997). A length of DNA consisting of a zeocin resistance cassette flanked by ~500 bp of the upstream and downstream regions of the slr0483 gene locus was generated though sequential PCR reactions (**Figure 4.3**).

A fully segregated Δ*slr0483 Synechocystis* strain was created by inserting a zeocin resistance cassette into *slr0483* (**Figure 7.2 B**) and sequentially selecting on increasing concentrations of zeocin until full segregation was achieved (Section 2.3.1). Confirmation of full segregation by amplification of the *slr0483* region is shown (**Figure 7.3 A**). During segregation it was noted that although this mutant could grow photoautotrophically and photomixotrophically, it had a significantly longer doubling time than wild type.

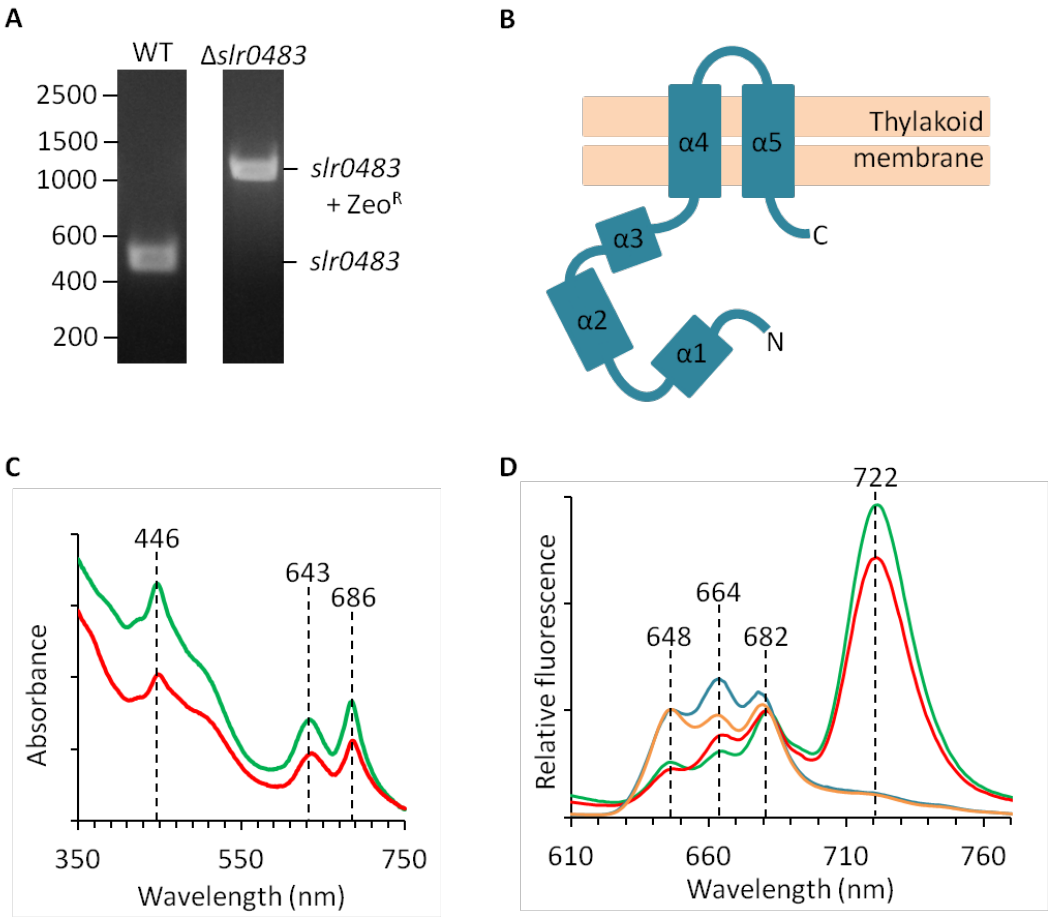

Figure 7.3 Production of Δ*slr0483* and spectroscopic analysis of the resulting phenotype

PCR analysis to show complete segregation of *slr0483* disrupted with a zeocin resistance cassette (A). Full segregation is shown by the absence of PCR product in the WT strain and the presence of a new, higher molecular weight product in the mutant strain. PCR reactions were resolved on a 1 % agarose gel (w/v), visualised with ethidium bromide. The *slr0483* primary sequence was analysed using the PSIPred program. *Slr0483* is predicted to form five α-helices, two of which (α4 and α5) are transmembrane helices (B). Whole cell absorbance spectra of wild type (green) and Δ*slr0483* mutant (red) *Synechocystis* cells (C) were recorded for cells were grown under normal light conditions in BG-11 medium supplemented with 5 mM glucose and 10 mM TES pH 8.2. Spectra are measured for cells in log phase normalised for cell density at OD_{750}. For the 77K emission spectra, all samples were re-suspended in 80 % glycerol to an OD_{750} = 0.1. 77K whole cell emission spectra (D) were recorded for excitation at 435 nm (wild type, green line; Δ*slr0483* mutant, red line) and 580 nm (wild type, blue line; Δ*slr0483* mutant, orange line). For further comparability, the 435 nm emission spectra were normalised to the PSII emission peak at 682 nm and the 580 nm emission spectra were normalised to the phycocyanin fluorescence at 648 nm.

7.2.3 Spectral properties of Δ*slr0483*

The whole cell absorbance spectra of wild type and Δ*slr0483* cells normalised for an OD_{750} ~0.2 were recorded using a UV-VIS spectrophotometer (**Figure 7.3 C**). The Δ*slr0483* spectrum has markedly lower absorbance maxima at 446 nm and 686 nm, which represents the

absorbance maxima of chlorophyll-containing complexes, indicating that a population of chlorophyll binding proteins is reduced, either though retarded synthesis or enhanced degradation, or that chlorophyll biosynthesis is perturbed in Δ*slr0483*. Quantification of the amount of chlorophyll *a* in wild type and Δ*slr0483* confirms that chlorophyll levels are decreased by ~50 % in the mutant (**Table 7.1**). The absorbance spectra also show that the phycobilisome absorbance peak at 643 nm is reduced in Δ*slr0483*, suggesting that the accumulation of phycobilisomes may also be reduced in the mutant.

Table 7.1 Chlorophyll *a* content of wild type and Δ*slr0483* cells grown under photomixotrophic conditions

Strain	Chlorophyll *a* mg · L^{-1} · OD_{750}^{-1}	% reduction compared to WT
WT	3.42 (± 0.13)	_
Δ*slr0483*	1.84 (± 0.12)	54

Given the observed ~50 % decrease in chlorophyll in Δ*slr0483* compared to wild type, the photosynthetic pigments were extracted from an equal quantity of wild type and Δ*slr0483* cells. HPLC analyses of these showed there to be no build-up of precursor pigments or partially reduced geranylgeranyl-chlide in Δ*slr0483*, which suggests Slr0483 may not be directly involved in chlorophyll biosynthesis (data not shown).

Low temperature fluorescence spectroscopy was used to investigate differences in the spectral properties of the photosystems and phycobilisome light-harvesting antennae in Δ*slr0483*. The emission spectra were recorded after excitation of chlorophylls at 435 nm and phycobilisomes at 580 nm for wild type and Δ*slr0483* cells re-suspended in 80 % glycerol to an OD_{750} of 0.1, so as to minimise the effects of cell scattering (**Figure 7.3 D**). For ease of comparison, the 435 nm spectra were normalised for PSII fluorescence at 682 nm and the 580 nm spectra were normalised to phycocyanin fluorescence at 646 nm.

After excitation at 435 nm (**Figure 7.3 D**, wild-type is in green and Δ*slr0483* is in red), four clearly defined maxima are observed at 648 nm, 664 nm, 682 nm and 722 nm, which arise from phycocyanin, allophycocyanin, PSII (CP43) and PSI respectively. A small shoulder can also be observed to the right of the PSII (CP43) maxima at 693 nm, which represents fluorescence from PSII (CP47) (Andrizhiyevskaya et al., 2005). Comparison of the wild type and Δ*slr0483* 435 nm emission spectra shows an observable decrease in allophycocyanin fluorescence and a decreased PSI to PSII ratio (PSI is decreased in relation to PSII) in Δ*slr0483*.

Three highly defined maxima at 646 nm, 664 nm and 681 nm are observed after excitation at 580 nm (**Figure 7.3 D**, wild type is blue and Δ*slr0483* is orange). These represent phycocyanin, allophycocyanin and terminal phycobiliprotein fluorescence respectively (Andrizhiyevskaya et al., 2005). With respect to the level of phycocyanin the levels of both allophycocyanin and terminal phycobilisomes fluorescence are reduced in Δ*slr0483* compared to the wild type.

7.2.3 Slr0483 interacts with enzymes in the tetrapyrrole biosynthesis pathway

In Chapter 3 of this thesis, mass spectrometry analysis of the eluates from pulldown experiments with C-terminal FLAG-ChlM, C-terminal FLAG-Sll1214 and C-terminal FLAG-ChlP identified Slr0483 as a potential interaction partner of each protein. In addition Slr0483 has been identified as interacting with NT-FLAG-Protoporphyrinogen oxidase, NT-FLAG-FeCH and NT-FLAG-Sll1214 and (Dr Roman Sobotka, personal communication). To further strengthen the hypothesis that these proteins interact, reciprocal pulldown experiments with N-terminal FLAG-Slr0483 as bait were performed. The construction of N-terminal FLAG-Slr0483 is described earlier in this chapter in Section 3.2.1; a PCR confirming full segregation of the FLAG-Slr0483 *Synechocystis* mutant is shown in **Figure 3.2**.

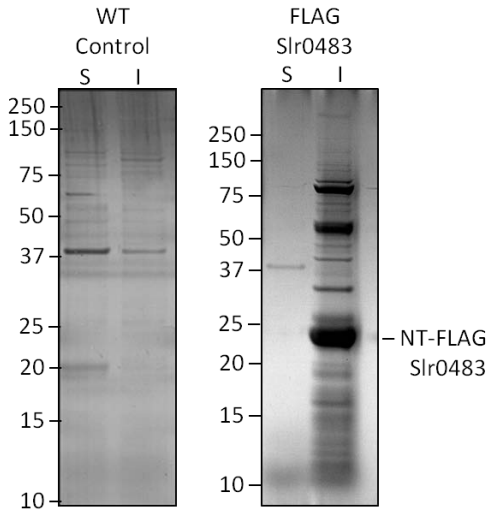

Figure 7.4 SDS-PAGE analysis of FLAG-Slr0483 eluate

FLAG-tagged proteins were purified from *Synechocystis* soluble cytoplasmic (S) and β-DDM solubilised thylakoid (I) fractions using anti-FLAG M2 affinity gel and eluted using the anti-FLAG peptide. The eluates from the *Synechocystis* wild type control and N-terminal FLAG-tagged Slr0483 pulldown experiments were 4x concentrated, separated by SDS-PAGE and silver strained.

Three sets of FLAG-affinity pulldown experiments were performed with the N-terminal FLAG-Slr0483 mutant (described in Section 2.6.9). The elution fractions from each of the pulldown experiments were analysed by silver stained SDS-PAGE (**Figure 7.4**). N-terminal FLAG-Slr0483

has a predicted molecular weight of ~20 kDa, but appears to migrate at a molecular weight of ~ 23 kDa on the silver stained SDS-PAGE. FLAG peptides have a high content of acidic residues such as aspartic acid, which could influence the binding of SDS affecting the relative mobility, thus giving the protein a larger apparent molecular weight. This may be the cause for N-terminal FLAG-Slr0483 having a larger than expected molecular weight. Bioinformatics analysis of the Slr0483 primary sequence suggested that Slr0483 contains two transmembrane helices (**Figure 7.2 B**); SDS-page analysis of the N-terminal FLAG-Slr0483 pulldown eluate show that Slr0483 and its associated interaction partners are, as would be expected, located in the insoluble fraction.

7.2.4 Abundance of chlorophyll biosynthesis enzymes in Δ*slr0483*

Western blot and mass spectrometry analyses indicate that the Δ*slr0483* strain may interact with protoporphyrinogen oxidase and FeCH, as well as the chlorophyll biosynthesis enzymes Sll1214 and ChlP. Additionally, a Slr0483 homologue was identified as being up-regulated in a *Chlamydomonas reinhardtii* strain deficient in GUN4. Therefore, it was decided to investigate whether the absence of Slr0483 affected the cellular accumulation of any chlorophyll biosynthesis enzymes (**Figure 7.5**), with specific focus on GUN4, the magnesium chelatase subunits (ChlH, ChlI and ChlD), the cyclase subunits Sll1214/Sll1874, ChlM and ChlP. Whole cell lysates were obtained from wild type and Δ*slr0483* cells grown under photomixotrophic conditions to an OD_{750} of 0.6-0.8, using the method described in Section 2.6.7. These were then quantified by protein concentration (using the method of Kalb and Bernlohr, 1977), rather than chlorophyll, as Δ*slr0483* accumulates ~50 % less chlorophyll than wild type *Synechocystis* (**Table 3.1**). Equal quantities of protein were then separated by SDS-PAGE and transferred to a nitrocellulose membrane, which was probed with antibodies to FeCH (AbCam), ChlH, ChlI, ChlD, GUN4, ChlM, CHL27 (Agrisera), POR, DVR, ChlP and D1 (Agrisera). Of the proteins probed, only the cellular accumulation of POR is altered, which is more abundant in Δ*slr0483* than in wild type cells. The reason for the increase in POR is unknown and requires further investigation.

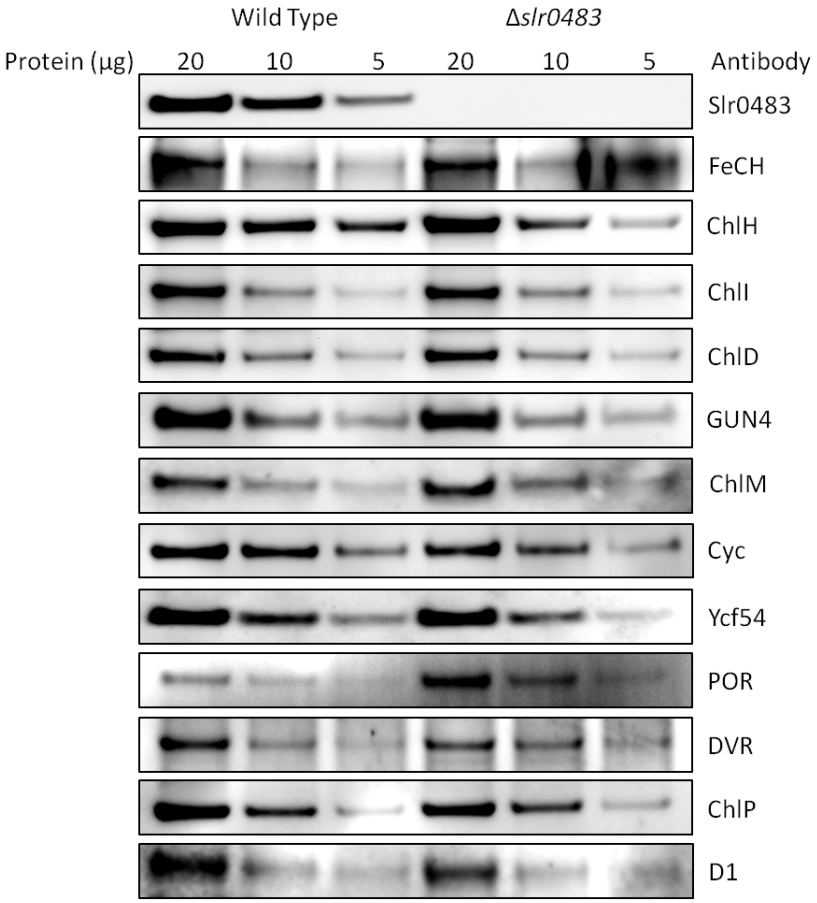

Figure 7.5 Analysis of chlorophyll biosynthesis enzyme accumulation in Δ*slr0483* mutant cells

Western blot analysis of the chlorophyll biosynthesis enzymes in wild type and Δ*slr0483* cells (**A**). Whole cell lysates from wild type and Δ*slr0483* cells, grown under photomixotrophic conditions, were normalised for protein concentration (5 µg, 10 µg or 20 µg), separated by SDS-PAGE and transferred to a nitrocellulose membrane. The blots were probed with antibodies to Slr0483 (kindly gifted by Prof. Dario Lester), the porphyrin biosynthesis enzymes, FeCH, ChlH, ChlI, ChlD, GUN4, ChlM, Cyc (CHL27,*Arabidopsis* homologue to Sll1214/Sll1874), Ycf54, POR, DVR and ChlP, and the PSII subunit D1.

In Chapter 3, it was suggested that Slr0483 could form a docking site, to which the interfacial-membrane localised chlorophyll biosynthesis enzymes could associate, facilitating the formation of a chlorophyll biosynthesis super-comeplex. Western blot analyses were employed to investigate whether the membrane localisation of the chlorophyll biosynthesis enzymes in Δ*slr0483* is altered in comparison to wild type. Following breakage of the Δ*slr0483* and wild type cells, the cell lysates were separated into their respective soluble and insoluble fractions by ultra-centrifugation (as described in Section 2.6.8). After separation of the soluble and insoluble fractions, remaining insoluble contaminants in the soluble fraction were removed by a further round of ultra-centrifugation and remaining soluble contaminants in the insoluble fraction were removed by washing the insoluble pellet in FLAG-buffer. The membrane associated proteins in the insoluble pellet were then solubilised with β-DDM at a

final concentration of 2 % (v/v) (as described in Section 2.6.8). Equal quantities of protein (calculated using the method of Kalb and Bernlohr, 1977) were then separated by SDS-PAGE and transferred to a nitrocellulose membrane, which was probed with antibodies to FeCH (AbCam), ChlH, ChlI, ChlD, GUN4, ChlM, CHL27 (Agrisera), POR and ChlP (**Figure 7.6**). These analyses show that Δ*slr0483* accumulates more of the MgCH subunits ChlH, ChlI, ChlD and GUN4 in the soluble fraction than wild type and that CHL27 accumulation is reduced in the insoluble fraction of Δ*slr0483* in comparison to wild type (**Figure 7.6**). The reason for the increased accumulation of the MgCH subunits in the soluble fraction of Δ*slr0483* is unknown and requires further investigation.

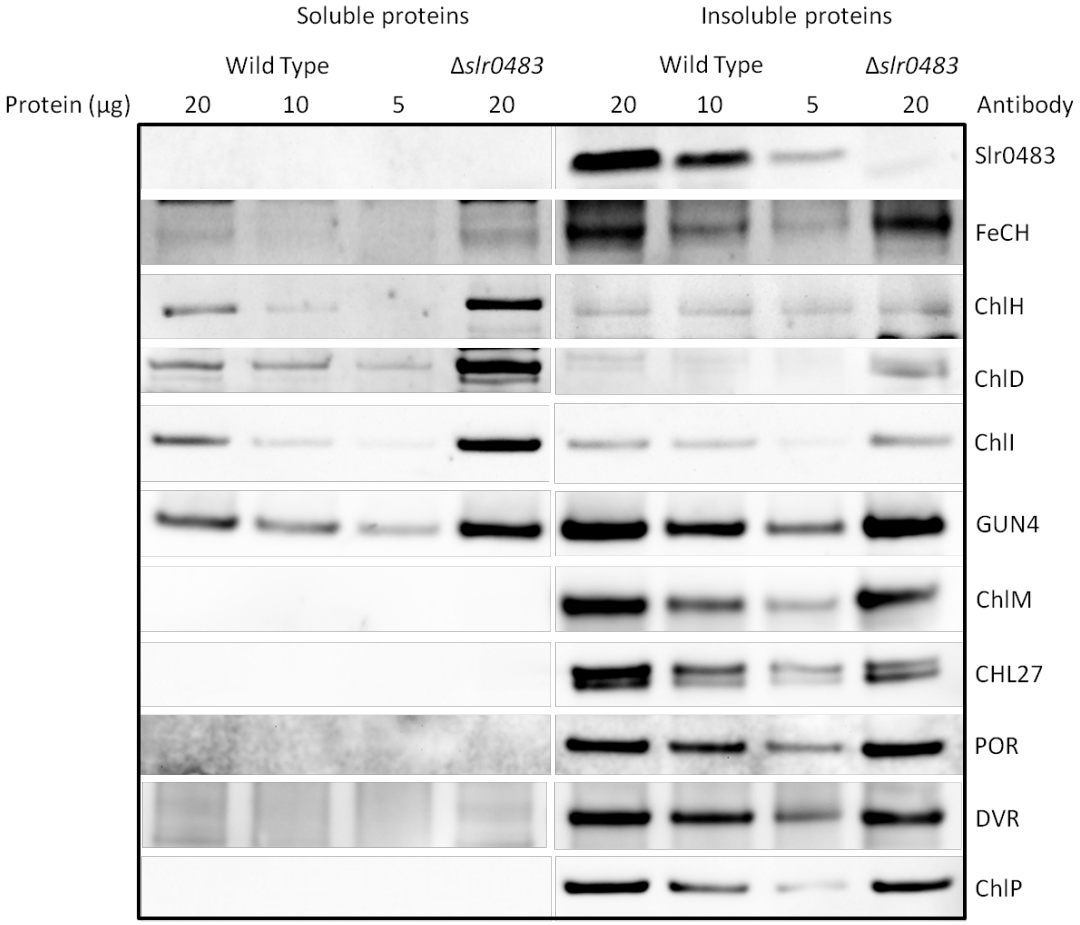

Figure 7.6 Analysis of chlorophyll biosynthesis enzyme location in wild type and Δ*slr0483* mutant cells

Western blot analysis of the localisation of the chlorophyll biosynthesis enzymes in wild type and Δ*slr0483* cells. Whole cell lysates from wild type and Δ*slr0483* cells, grown under photomixotrophic conditions, were separated into their respective soluble and insoluble membrane fractions (the membrane fractions were solubilised with 2% (v/v) β-DDM), normalised for protein concentration (5 µg, 10 µg or 20 µg), separated by SDS-PAGE and transferred to a nitrocellulose membrane. The blots were probed with antibodies to Slr0483 (kindly gifted by Prof. Dario Lester), and the porphyrin

biosynthesis enzymes, FeCH, ChlH, ChlI, ChlD GUN4, ChlM, Sll1214/Sll1874 (CHL27), POR, DVR and ChlP.

7.2.5 Abundance of chlorophyll containing complexes in Δslr0483

To further investigate the relative levels of phycobilisomes and photosystem complexes, the proteins associated with the membranes were separated by sucrose density gradient ultra-centrifugation. Cultures of wild type and Δ*slr0483* cells were harvested at mid-log phase (OD_{750} 0.6-0.8), the cells broken and the membranes isolated by ultra-centrifugation. The membrane fraction was re-suspended and normalised to a protein concentration of 10 mg ml^{-1}, then solubilised with β-DDM at a final concentration of 1 % (v/v) in a homogeniser. The solubilised material was diluted to a final protein concentration of 1 mg ml^{-1} and 1 mg of protein was separated by ultracentrifugation on a continuous 10 – 30 % (w/v) sucrose gradient (experiments were adapted from Dühring et al (2006) and performed as described in Section 2.5.7).

Figure 7.7 A shows the results of the detergent fractionation and sucrose density gradient ultracentrifugation. Both gradients have three clearly defined coloured bands which contain the free pigments, the PSI/PSII monomers and the PSI trimers respectively. 0.5 ml fractions were harvested from the gradients (indicated in **Figure 7.7 A**) and the absorbance spectra recorded for each.

Plotted on the graph in **Figure 7.7 B** are the absorbance maxima for chlorophyll *a* (678 nm), phycobiliproteins (620 nm) and β-carotene (487 nm) for each of the 16 harvested gradient fractions. This traces the presence of chlorophyll complexes (green line), phycobiliproteins (blue line) and carotenoids (orange line) through the gradients, giving an indication as to their relative abundances in wild type (solid lines) and Δ*slr0483* (dashed lines) membranes.

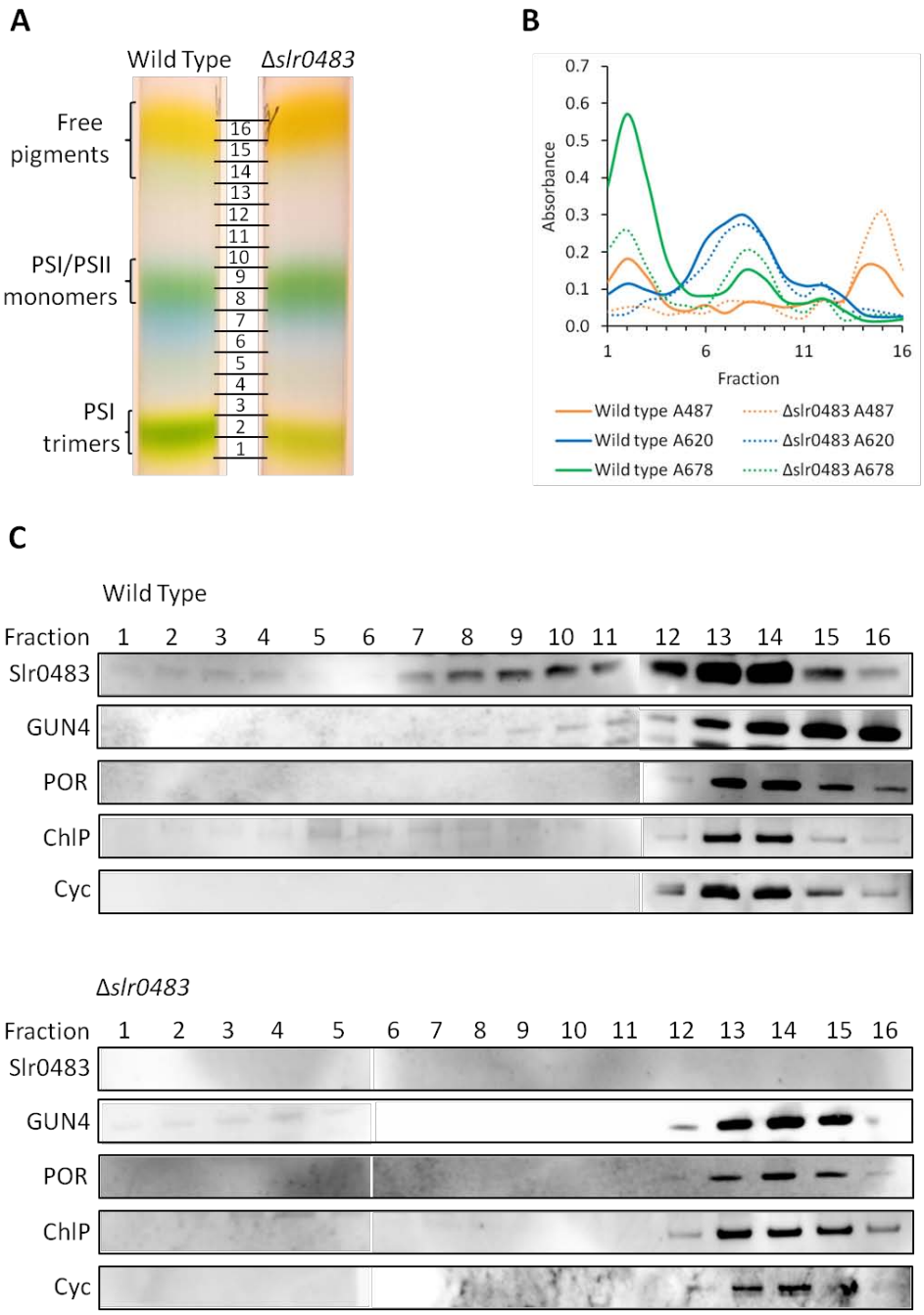

Figure 7.7 Separation of membrane protein complexes in wild type and Δslr1780 Synechocystis cells via sucrose density gradients

10-30 % (w/v) continuous sucrose gradients were used to separate membrane complexes, at a protein concentration of 1 mg ml^{-1}, from β-DDM solubilised membranes isolated from wild type and *Δslr1780 Synechocystis* cells. Three clearly coloured bands were resolved, which represent free pigments, PSI/PSII monomers and PSI trimers (**A**). Sixteen 0.5 ml fractions were harvested from the gradient and their absorbance spectra recorded. Plotted are the absorption maxima for β-carotene (487 nm), phycobilisomes (620 nm) and chlorophyll (678 nm) (**B**). The fractions were also separated by SDS-PAGE, transferred to a nitrocellulose membrane and probed with antibodies to Slr0483 (kindly gifted by Prof. D. Lester), GUN4, POR, CHL27 (cyc) and ChlP (**C**).

The PSI band (located in fractions 1 and 2) is visibly greener in the wild type gradient than in the Δ*slr0483* gradient, indicating the wild type membranes have more PSI trimers than Δ*slr0483* membranes. The graph also shows that wild type fractions 1 and 2 have almost double the absorbance for chlorophyll and β-carotene than the corresponding fractions in Δ*slr0483*. Inspection of fractions 1 and 2 (**Figure 7.7 B**) show that there are ~50 % fewer PSI trimers in Δ*slr0483* membranes. Separation of the membrane-associated complexes on a clear native gel, (**Figure 7.8**, data courtesy of Dr Roman Sobotka and Dr Jana Kopecna), shows a significant reduction in the abundance of both the high molecular weight green super complex and PSI trimers in Δ*slr0483* in comparison to wild type. Additionally, the clear native of Δ*slr0483* shows the appearance of two unusual phycobiliprotein complexes below the PSI monomer (**Figure 7.8 A**), which are not present in the wild type sample.

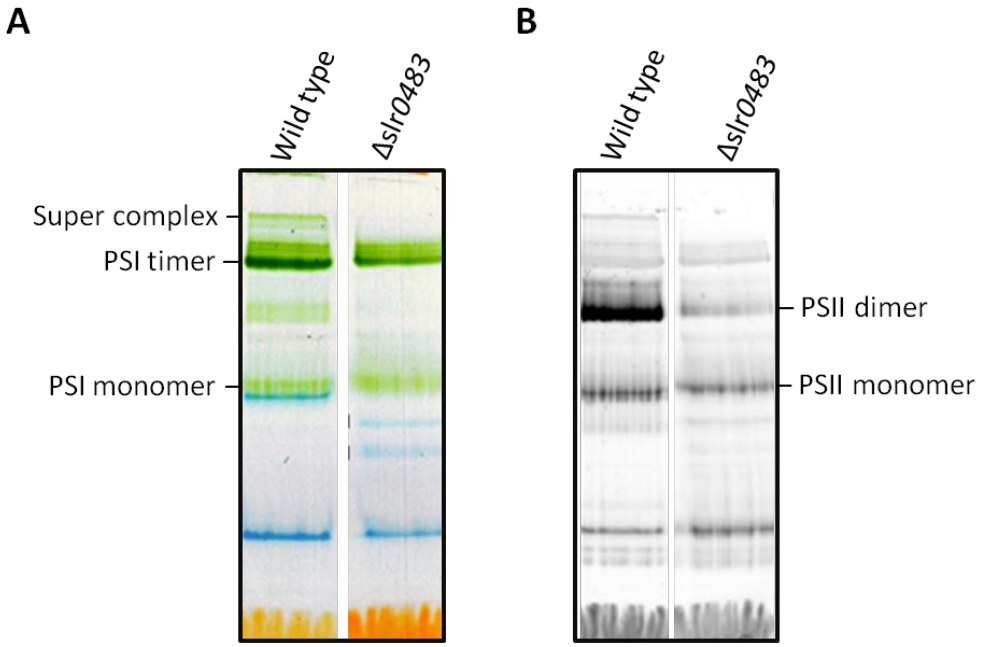

Figure 7.8 Accumulation of chlorophyll containing complexes in photomixotrophically grown wild type and Δ*slr0483* cells

Membranes associated photosynthetic complexes were solubilised from the membrane fraction using 2% β-DDM and separated by clear-native-PAGE (**A**). Coloured complexes can be viewed in (**A**) and the fluorescence emission of chlorophyll containing complexes, visualised by excitation of chlorophylls at 435 nm, can be viewed in (**B**). The clear-native gel was kindly produced by R Sobotka and J Kopecna.

The chlorophyll fluorescence recorded from the clear-native gel (**Figure 7.8**) shows the accumulation of PSII dimers in Δ*slr0483* membranes is significantly reduced. Second dimension separation of these complexes by SDS-PAGE (**Figure 7.9**, data courtesy of Dr R Sobotka and Dr J Kopecna) shows that the levels of PSII subunits, CP43, CP47, D1 and D2 are greatly reduced in Δ*slr0483* compared to wild type.

Phycobilisome stability also appears to be compromised in Δ*slr0483* compared to wild type, as Δ*slr0483* accumulated two lower molecular weight phycobiliprotein complexes (indicated by the arrows), which are not observed in the wild type. Also observable in both the sucrose gradients and clear-native gels is the increased levels of free carotenoid pigments in Δ*slr0483* than wild type, which is indicated by the intensity of the free pigment band (**Figure 7.7 A and Figure 7.8**) and the greater absorbance at 487 nm in Δ*slr0483* (**Figure 7.7 B**).

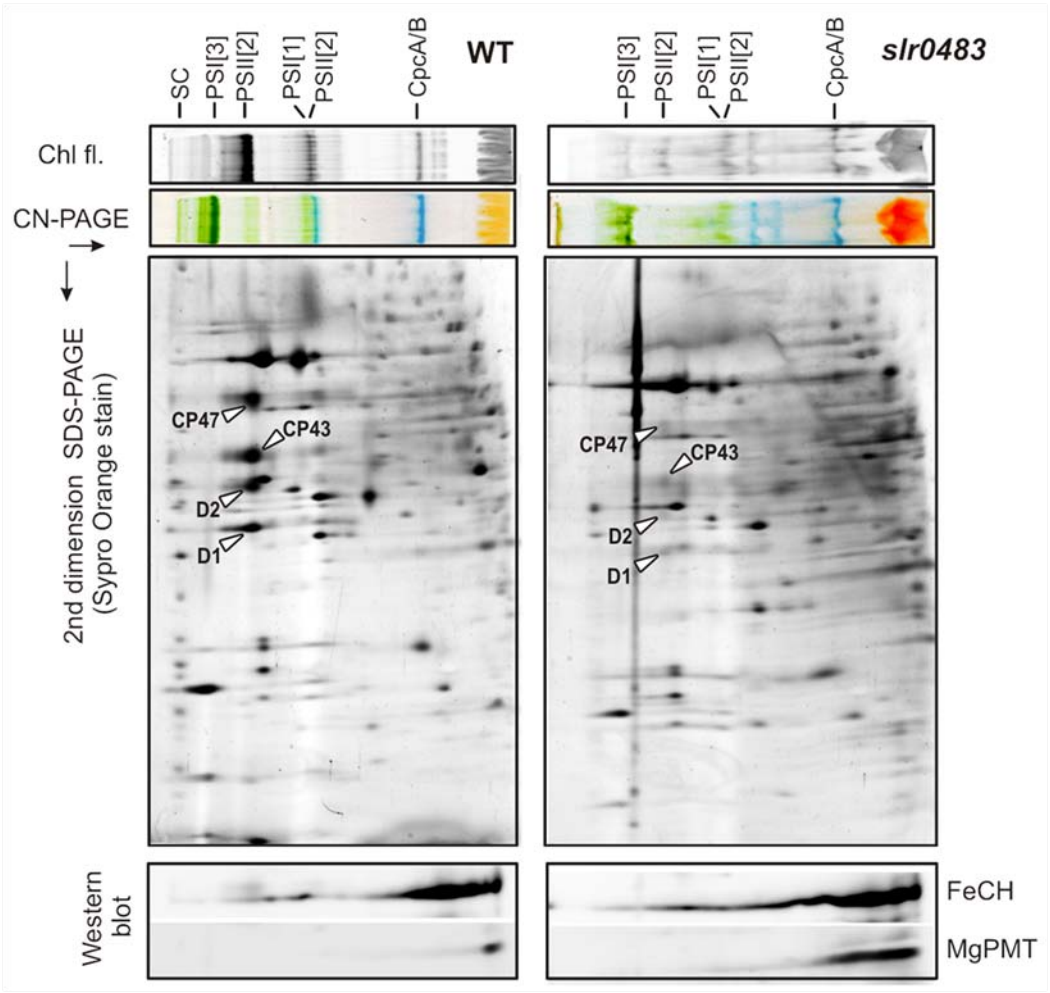

Figure 7.9 2D-gel analysis of photosynthetic complexes in photoautotrophically grown wild type and Δ*slr0483* cells

A second dimension of the clear-native gels depicted in **Figure 7.7** was resolved by SDS-PAGE and stained with Sypro Orange. Proteins were then transferred to a nitrocellulose membrane, which was probed with antibodies to FeCH and ChlM (MgPMT). 2D-gel was kindly produced by R Sobotka and J Kopecna.

To investigate the distribution of Slr0483, the wild type and Δ*slr0483* sucrose gradient fractions were separated by SDS-PAGE and transferred to a nitrocellulose membrane for immuno-blot analysis. Antibodies raised against Slr0483 (kindly gifted by Prof. Dario Lester), GUN4, POR and ChlP were used to probe the blots (**Figure 7.7 B**). The blot, probed with anti-

Slr0483, shows Slr0483 to be mainly concentrated in the slower sedimenting fraction of the gradient, but small levels of this protein are found distributed throughout the gradient, in the regions where PSI/PSII monomers and PSI trimers are found. Importantly, the Δ*slr0483* fractions contained no detectable Slr0483, which demonstrates that the Δ*slr0483* mutant is fully segregated. The blots show that GUN4 and POR association with the membrane is reduced in Δ*slr0483* compared to wild type, but that the amount of ChlP associated with the membrane is increased in Δ*slr0483*. These results are interesting, as the abundance of GUN4 and ChlP is unchanged in Δ*slr0483* compared to wild type (**Figure 7.5**), but the abundance of POR is increased in Δ*slr0483*.

7.3 Discussion

7.3.1 Slr0483 is related to the *Arabidopsis* CURVATURE THYLAKOID1 proteins

Although designated as an unknown hypothetical gene in the *Synechocystis* genome, Slr0483 is homologous (28 % sequence identity) with the *Arabidopsis* TMP14 protein, which was first identified in a screen to find phosphorylated proteins associated with the thylakoid membrane (Hansson and Vener, 2003). This protein is exclusively located in the lamella and was found to co-localise with PSI when the lamella proteins were separated via 2D-BN-PAGE/SDS-PAGE or sucrose density gradients (Khrouchtchova et al., 2005). Analysis of TMP14 accumulation in various PSI mutants showed that TMP14 is up-regulated in the absence of PsaO, and virtually undetectable in the absence of PsaF, PsaG and PsaL (Khrouchtchova et al., 2005). Later, a yeast-two hybrid experiment found TPM14 to interact with PsaL (Yu et al., 2008). Taken together these data suggest that TMP14 interacts with PSI in *Arabidopsis*.

However, in a recent publication by Armbruster *et al* (2013) TMP14 was described as a member of the CURVATURE THYLAKOID1 (CURT1) family, of which there are four reported members, CURT1A, CURT1B (TMP14), CURT1C and CURT1D. Here, they presented evidence that the CURT1 proteins were enriched at the grana margins and found that the grana architecture was dependent on the levels of CURT1 proteins; fewer grana margins were found in plants without CURT1 and an increased number of grana membrane layers, with a smaller diameter, were seen in plants over-expressing CURT1A (Armbruster et al., 2013). Overall, the authors concluded the CURT1 genes were essential for facilitating grana curvature and the formation of wild type grana architecture.

CURT1 proteins are evolutionarily conserved and homologues are found in cyanobacteria, algae and plants. However, CURT1 homologues are absent in the early diverging cyanobacteria lineages (including *Gloeobacter violaceus* and *Synechococcus* PCC7336), which suggests the CURT1 protein originated after the initial cyanobacteria divergence (Luque and Ochoa de Alda, 2013). A sequence alignment of the four *Arabidopsis* CURT1 proteins and the *Synechocystis* CURT1 homologue (Slr0483) is shown in **Figure 7.10**.

```
CURT1A    MAISVAASSSMAVMVPRVP--------AVSTRCSAVPYLP-PRSFGRSS-FTVPLKLVSG 50
CURT1B    -MASLSVSSSSTIIDSRAPPSRLASASASSPSCISLPTLP-IQSHTRAAKATAYCRKIVR 58
CURT1C    -MASISATLPSPLLLTQRK----------SNLTSIQKLP--FSLTRGTNDLSPLSLTRN 46
CURT1D    --MELCTRSTTIITHLPASFNGHGYLAGKSVDRISLPLQRNVASLVLQSRTLRCSRKFPG 58
Slr0483   ----MGRKHSIRIIKDWLK---------MEEQKTATAGIK--TDVGPITTPNPQKSPITD 45

CURT1A    NGLQ--KVELLKTR----------ASSEETSSIDTNELITDLKEKWDGLENK-------S 91
CURT1B    NVVTRATTEVGEAP----------ATTTEAETTELPEIVKTAQEAWEKVDDK-------Y 101
CURT1C    PSSISLMVKASGES----------------SDSSTDLDVVSTIQNVWDKSEDR-------L 84
CURT1D    ETVTEETSTGVNEFGVEDRDGVVVAAEEKNSNSEAPQAEDEETQALEFLNDIKLDSDKTY 118
Slr0483   QAWQEWLQPVWEVLG---------------------KIPQMTGEFFEDNKQP-------L 77

CURT1A    TVLIYGGGAIVAVWLSSIVVGAINSVPLLPKVMELVGLGYTGWFVYRYLLFKSSRKELAE 151
CURT1B    AIGSLAFAGVVALWGSAGMISAIDRLPLVPGVLELVGIGYTGWFTYKNLVFKPDREALFE 161
CURT1C    GLIGLGFAGIVALWASLNLITAIDKLPVISSGFELVGILFSTWFTYRYLLFKPDRQELSK 144
CURT1D    SILLYGSGAIVALYLTSAIVSSLEAIPLFPKLMEVVGLGYTLWFTTRYLLFKRNREELKT 178
Slr0483   ISLGIILLGIISVKILIAVLDAINDIPLLAPTLQLIGMGYTAWFIWRYLWKAEKRQELAS 137

CURT1A    DIESLKKKITAGSE-- 164
CURT1B    KVKSTYKDILGSS-- 174
CURT1C    IVKKSVADILGQ--- 156
CURT1D    KVSEIKKQVLGSDSE 193
Slr0483   EFGALKEQIFGG--- 149
```

Figure 7.10 Sequence alignment showing the evolutionarily conserved residues of the Ycf54 domain
Primary sequence alignment of the *Arabidopsis* CURT1 proteins (CURT1A, CURT1B, CURT1C and CURT1D) and the *Synechocystis* Slr0483 protein, generated by CLUSTALW2. Conserved, highly similar and similar residues are highlighted in blue, dark grey and black respectively.

CURT1 could play an analogous role in curving thylakoid membranes. Although Armbruster *et al* (2013) were not successful in generating a fully segregated Slr0483 *Synechocystis* knock-out mutant, they did complement the *slr0483* with *Arabidopsis curt1A*. They found CURT1A could only partially complement Slr0483 and suggested from electron micrographs of the cellular architecture that the phycobilisomes were detached from the thylakoid membrane in the mutant. However, what the authors indicate as detached phycobilisomes closely resemble glycogen granules, which can be viewed in the electron micrographs of the Ycf54 mutant (**Figure 4.10**) and (Kopecna et al., 2012).

Although Slr0483 has been identified as a homologue with the CURT proteins, this homology is only shared by the C-terminal CAAD domain, which corresponds to the two putative transmembrane helices, responsible for anchoring these proteins to the thylakoid membranes (**Figure 7.10**). These proteins share little homology at the N-terminus, which is

the region likely to confer a function and potentially interact with other proteins in the cell; therefore it is plausible that although these proteins potentially evolved from a membrane-anchored common ancestor, the variable N-terminus may have adapted to perform different functions within the cell. This theory potentially explains the discrepancies observed between the *Arabidopsis* CURT mutants, in which PSI and PSII accumulation is unaffected, and the *Synechocystis* Δ*slr0483* mutant, in which PSI and PSII accumulation is significantly affected.

7.3.2 Photosystem accumulation is detrimentally affected in Δ*slr0483*

Analysis of the PSI and PSII subunits in the CURT1 deficient *Arabidopsis* plants revealed that their accumulation is akin to wild type (Armbruster et al., 2013). However, analysis of PSI and PSII accumulation in Δ*slr0483* showed that both complexes are significantly reduced (**Figure 7.8** and **Figure 7.9**), indicating that in *Synechocystis* Slr0483 either plays an important role in PSI/PSII stability or assembly and in the absence of this protein, accumulation of PSI and PSII is compromised. It is likely that Slr0483 may be involved in PSI and PSII assembly, for example by facilitating the transfer of chlorophyll molecules to PSI/PSII from the terminal chlorophyll biosynthesis enzymes, as Slr0483 has been identified in the pulldown experiments as an interaction partner of PPO, FeCH, Sll1214 and ChlP (personal correspondence, Roman Sobotka). Additionally, CURT1B (TMP14) has been identified as interacting with PSI in *Arabidopsis* (Khrouchtchova et al., 2005); it would be interesting to see if any PSI subunits are located in the FLAG-Slr0483 pulldowns.

7.3.3 Further work

Experimental characterisation of Δ*slr0483* is in its initial stages and further work is required to fully elucidate the role of this protein in *Synechocystis*. Several avenues should be investigated to characterise the role of this protein, which are detailed below.

Initially, the *Arabidopsis* CURT1B (TMP14) of Slr0483 was identified as a phosphorylated protein associated with PSI in the chloroplast lamella (Hansson and Vener, 2003). Early work on determining the phosphorylation state of Slr0483 suggests that this protein is not phosphorylated in *Synechocystis* (Dr Phil Jackson, personal communication).

It should be established whether the reduced chlorophyll accumulation in Δ*slr0483* (**Table 3.**2) is due to a reduced rate of chlorophyll biosynthesis or to an overall reduction in the chlorophyll binding proteins, such as PSI/PSII complexes. **Figure 7.5** shows the cellular levels of all the chlorophyll biosynthesis enzymes (with the exception of POR, which is up-regulated in Δ*slr0483*) are unchanged between wild-type and Δ*slr0483* mutant cells, although the association of some of these enzymes with the thylakoid membrane is altered in Δ*slr0483*

compared to wild type (**Figure 7.7**). Additionally, the rate of chlorophyll biosynthesis in wild-type and Δ*slr0483* should be compared. This could be achieved by monitoring the incorporation of [^{14}C]Glu into chlorophyll using a pulse-chase experiment as described in Kopecna *et al* (2012).

Thin section electron micrographs of the cellular architecture of Δ*slr0483* will determine the number and morphology of thylakoid membranes in comparison with the wild-type. The localisation of Slr0483 in the cell should also be determined, to ascertain whether the protein accumulates at the curved edges of the thylakoids, as suggested in Armbruster *et al* (2013). Work has already commenced to GFP and YFP label Slr0483, for imaging using confocal microscopy (Spence et al., 2003) and STORM (Huang et al., 2010).

8. Concluding remarks

The work reported in Chapter 3 of this thesis has lead to the first hypothetical models of interactions between the chlorophyll biosynthesis enzymes, and suggests that, as hypothesised, the enzymes of the chlorophyll biosynthesis pathway may indeed form a substrate channelling super-complex. In addition, the *in vivo* FLAG-pulldown studies have lead to the identification of two previously unknown proteins Slr1780 and Slr0483 as having roles in chlorophyll biosynthesis and photosystem stability respectively.

Chapter 4, Chapter 5 and Chapter 6 of this thesis has focused upon investigating the previously unknown hypothetical protein Ycf54 (Slr1780), which has been identified as a component of the Mg-protoporphyrin IX monomethylester cyclase complex. The data in these chapters represents the first studies on characterising the Ycf54 protein and validation of its role as a component of the cyclase. In addition, the work carried out in Chapter 5 of this thesis identified that three of the seven absolutely conserved residues in the Ycf54 protein are required for Pchlide formation in *Synechocystis*. The work in this Chapter also lead to the finding that Ycf54 proteins in which any of the three absolutely conserved residues (D39, F40 and R82) are mutated to an alanine cannot interact with the di-iron catalytic component of the cyclase; thus establishing that Ycf54 interaction with Sll1214/Sll1874 is required for cyclase activity. This knowledge was expanded upon in Chapter 6, in which the first crystal structure of *Synechocystis* wild Ycf54 and two Ycf54 mutants were reported. These crystal structures lead to the conclusion that the residue R82 plays a functional role in the interaction of Ycf54 with Sll1214/Sll1874.

Chapter 7 of this thesis represents the first studies into the function of the highly conserved, but previously unknown protein Slr0483. Although there is much investigation required to fully elucidate the exact role of this protein, the work in this chapter suggests that Slr0483 is required for the stability of PSI trimers and PSII.

Overall, the results from this work show that *in vivo* FLAG-pulldown experiments are a successful approach in fishing expeditions to identify as yet unknown components in a partially characterised biological pathway.

Appendix

Mass spectrometry data from FLAG-Pulldown experiments reported in Chapter 3

Table 1. Hits from *Synechocystis* wild type control pulldown experiments

Synechocystis Gene ID	*Synechocystis* wild type control pulldown experiments																Hits identified in >1 pulldown with a Mascot score >30
	Individual pulldown experiments using the soluble fraction								Individual pulldown experiments using the insoluble fraction								
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	
Sll1577	78			104				29	69				40	59	46	45	Sll1577
Slr1884	212	29	54	254	114	84	120	122	138	209	85	105	291	252	174	50	Slr1884
Sll1568				123					105				175	158	75		Sll1568
Slr2067				148				34	83				146	157	54		Slr2067
Sll1746											35	72	104	30			Sll1746
Sll1099													61				
Sll1043	40												328	123			Sll1043
Sll1818													49	96			Sll1818
Slr0925													49	44			Slr0925
Sll0839		35	47						49							47	Sll0839
Slr1102		157		58			160	150			172	194	483	160	203		Slr1102
Sll0033			32													32	Sll0033
Sll1204				28											26		Sll1204
Sll1712				62									107	114	61		Sll1712
Slr0455				104											58		Slr0455
Sll1687				116	118		55						907	710	558		Sll1687
Sll1681							26										
Slr0162							35										
Slr2141								30									
Slr1270									27								
Sll1665										35	131	36					Sll1665
Slr6047											27						
Sll1166												27					

Table 2. Hits from *Synechocystis* CT-FLAG Slr0525 (ChlM) pulldown experiments

Synechocystis Gene ID	*Synechocystis* CT-FLAG ChlM pulldown experiments						Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction			Individual pulldown experiments using the insoluble fraction					
	1	2	3	1	2	3			
Slr0525	621	657	796	592	574	868	Slr0525		Slr0525
Sll1945	323			203	183		Sll1945		Sll1945
Slr0447	247			104	115		Slr0447		Slr0447
Sll1578	181		46	163	101		Sll1578	Sll1578	
Sll1450	145			124	51		Sll1450		Sll1450
Sll1184	144			45			Sll1184		Sll1184
Slr0384	121								
Slr0506	114			173	133		Slr0506		Slr0506
Sll0020	108			69	63		Sll0020		Sll0020
Slr1835	96			56	43		Slr1835		Slr1835
Slr0737	94			102	69		Slr0737		Slr0737
Sll0819	94			74	83		Sll0819		Sll0819
Sll1769	89			92			Sll1769		Sll1769
Ssl0294	86			110			Ssl0294		Ssl0294
Sll0982	85			29	46		Sll0982		Sll0982
Sll1746	83			74	64		Sll1746	Sll1746	
Slr0483	79			74	59		Slr0483		Slr0483
Slr1254	78			56			Slr1254		Slr1254
Sll1577	76			76	150		Sll1577	Sll1577	
Sll1147	75			75	76		Sll1147		Sll1147
Sll1260	73			59	145		Sll1260		Sll1260
Sll1326	67			100	41		Sll1326		Sll1326
Sll1665	65			84	66		Sll1665	Sll1665	
Slr1220	61				41		Slr1220		Slr1220
Slr1311	60			42	57		Slr1311		Slr1311
Slr1091	60			68	59		Slr1091		Slr1091
Sll0237	59			31	59		Sll0237		Sll0237

Slr1963	58				29		Slr1963		Slr1963
Sll0378	54								
Sll1528	54				98		Sll1528		Sll1528
Sll0947	53			34			Sll0947		Sll0947
Slr1128	50			103			Slr1128		Slr1128
Sll1218	46			38	44		Sll1218		Sll1218
Sll1867	45			30	79		Sll1867		Sll1867
Slr1665	41			47	46		Slr1665		Slr1665
Sll1530	40								
Sll0839	39			50	40		Sll0839	Sll0839	
Slr2067	38								
Slr1226	38			37			Slr1226		Slr1226
Ssr2831	37			31	46		Ssr2831		Ssr2831
Ssl0707	36				28		Ssl0707		Ssl0707
Sll1694	35			39	43		Sll1694		Sll1694
Sll0519	35			87			Sll0519		Sll0519
Ssl0563	33			31			Ssl0563		Ssl0563
Sll1561	33								
Slr1329	33			29	44		Slr1329		Slr1329
Slr2067	32			28	53		Slr2067		Slr2067
Slr0161	30			29	37		Slr0161		Slr0161
Sll1029	29								
Sll5097	27								
Slr1780		77				57	Slr1780		Slr1780
Slr0757		35	28			30	Slr0757		Slr0757
Slr0536		31							
Sll0837			85	44		77	Sll0837		Sll0837
Sll0569				55					
Ssl2598				54	27		Ssl2598		Ssl2598
Sll0058				50					

Table 3. Hits from *Synechocystis* NT-FLAG Sll1214 pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG Sll1214 pulldown experiments												Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction						Individual pulldown experiments using the insoluble fraction								
	1	2	3	4	5	6	1	2	3	4	5	6			
Sll1214	1082	648	630	714	341	359	171	641	927	469	488	955	Sll1214		Sll1214
Sll1578		88	113					256	139	187			Sll1578	Sll1578	
Sll1874		66													
Sll0416		40			64			539	268	925			Sll0416		Sll0416
Sll1080		33													
Slr1918		33													
Slr2141		27													
Slr0506			264	375									Slr0506		Slr0506
Sll1099			66					1025	131	36			Sll1099		Sll1099
Sll1577			51	147				344	349	366			Sll1099	Sll1099	
Sll0296			31								33		Sll0296		Sll0296
Slr1986			27					147	110	80			Slr1986		Slr1986
Sll1106				62					100				Sll1106		Sll1106
Sll0058				52											
Sll0819				50					71				Sll0819		Sll0819
Slr1835				43					47				Slr1835		Slr1835
Slr1655				37					36				Slr1655		Slr1655
Ssl7039				28											
Slr7088				28											
Slr1780					33	40		73	31				Slr1780		Slr1780
Slr2076						35		220	220	184			Slr2076	Slr2076	
Slr1884									35	274			Slr1884	Slr1884	
Slr0288															
Sll1130								41	56	72			Sll1130		Sll1130
Slr1923							125					75	Slr1923		Slr1923
Sll1091							79	138	30			152	Sll1091		Sll1091
Sll1147							31								
Sll1908								208	87				Sll1908		Sll1908

Slr1463								184	30				Slr1463		Slr1463
Slr1329								162	346				Slr1329		Slr1329
Slr0844								103							
Sll1580								100	166	90			Sll1580		Sll1580
Sll1604								90	128				Sll1604		Sll1604
Ssl0707								74	60	59			Ssl0707		Ssl0707
Sll1097								74	107	44			Sll1097		Sll1097
Slr1934								72	44				Slr1934		Slr1934
Sll1744								67	133	118			Sll1744		Sll1744
Sll1808								63	148	97			Sll1808		Sll1808
Sll1326								62	306	39			Sll1326		Sll1326
Slr0394								58	63				Slr0394		Slr0394
Sll1316								53							
Sll1687								52	51	172			Sll1687	Sll1688	
Sll1317								51							
Ssr1399								48							
Slr0743								44							
Slr0552								39	160				Slr0552		Slr0552
Slr0039								38							
Sll1746								38	164	109			Sll1746	Sll1746	
Ssr1156								35	98	95			Ssr1156		Ssr1156
Sll0359								35							
Slr1198								34	83	34			Slr1198		Slr1198
Slr0947								33	45				Slr0947		Slr0947
Slr7041								33							
Ssl3093								29	56	88			Ssl3093		Ssl3093
Slr666								28							
Ssl3432								28		30			Ssl3432		Ssl3432
Slr1963									284						
Sll0020									261						
Slr0335									204	46			Slr0335		Slr0335
Slr1102									186						
Sll0058									172	57			Sll0058		Sll0058
Slr1834									149						
Sll0947									143	65			Sll0947		Sll0947

Slr0737									139						
Sll1261									136						
Slr1128									117						
Sll1274									112	153			Sll1274		Sll1274
Sll1812									102	57			Sll1812		Sll1812
Dlr1678									81	81			Dlr1678		Dlr1678
Sll1821									61	34			Sll1821		Sll1821
Slr1329									60						
Sll1819									58						
Slr1854									56						
Sll1810									55	38			Sll1810		Sll1810
Sll1801									54	59			Sll1801		Sll1801
Slr0924									54						
Sll1811									53						
Sll1818									52	36			Sll1818	Sll1818	
Sll0569									50						
Sll1342									50	82			Sll1342		Sll1342
Sll1043									50	55			Sll1043	Sll1043	
Sll1621									47	34			Sll1621		Sll1621
Sll1816									46	31			Sll1816		Sll1816
Sll1789									45	105			Sll1789		Sll1789
Sll1471									34	46			Sll1471		Sll1471
Ssr3383									32	28			Ssr3383		Ssr3383
Slr1223									31						
Sll1324									31						
Sll1185									31	56			Sll1185		Sll1185
Ssl3436									31	109			Ssl3436		Ssl3436
Slr2075										330		56	Slr2075		Slr2075

Table 4. Hits from *Synechocystis* CT-FLAG Sll1214 pulldown experiments

Synechocystis Gene ID	*Synechocystis* CT-FLAG Sll1214 pulldown experiments								Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction				Individual pulldown experiments using the insoluble fraction						
	1	2	3	4	1	2	3	4			
Sll1867	327	47	323	68		397			PsaA		PsaA
Sll0416	235		90						GroEL-2		GroEL-2
Sll1214	225		515	540	942	513	724	1027	Sll1214		Sll1214
Slr2076	225		201			91			GroEL		GroEL
Sll1091	205		148			221			ChlP		ChlP
Slr0737	198	73	221			224			PsaD		PsaD
Sll1604	193										
Sll0819	184	43	258			258			PsaF		PsaF
Slr1330	155					96			AtpC		AtpC
Slr2067	150	41	181	32		198	40		ApcA	ApcA	
Sll1694	143		112			172			HofG		HofG
Ssl0563	128		269			261			PsaC		PsaC
Sll0617	126		105			227			Im30		Im30
Sll182	117		38			35			RplB		RplB
Sll1326	115		349	67		402			AtpA		AtpA
Sll1324	114		50			85			AtpF		AtpF
Slr2075	98										
Sll1807	92										
Slr1329	88	39	211			522			AtpB		AtpB
Slr1986	87	82	52			175			ApcB		ApcB
Slr0906	84		111			251			PsbB		PsbB
Sll1799	82										
Sll1577	79	124	253			354			CpcB	CpcB	
Sll1746	77	68	79			84			RplL	RplL	
Sll1274	77		71						RplI		RplI

Slr1835	72		297	28		287			PsaB		PsaB
Slr1884	71	132	229	38		138			TrpS	TrpS	
Slr0342	71										
Sll1530	68		61						Sll1530		Sll1530
Sll0226	67					46			Ycf4		Ycf4
Slr1329	63					84			AtpD		AtpD
Ssr2799	63					50			RpmA		RpmA
Sll0851	63		197			202			PsbC		PsbC
Slr1311	61		283			225			PsbA		PsbA
Slr0156	60										
Sll1184	60										
Sll1194	60		56			152			PsbU		PsbU
Sll1218	58										
Ssl1046	57										
Slr0483	56	42	65			60			Slr0483		Slr0483
Slr0909	34		56						Slr0909		Slr0909
Sll1578		123	139			266	42		CpcA	CpcA	
Slr0009		109	39						RbcL		RbcL
Sll1712		81						88	DBH	DBH	
Ssl3093		59									
Slr1655		49	91			56			PsaL		PsaL
Sll0839		45	47			52	51	53	Sll0839	Sll0839	
Ssl0707		40				51			GlnB		GlnB
Sll0033		33						33	Sll0033	Sll0033	
Slr0906			159			187			PsbD		PsbD
Ssr2831			137			163			PsaE		PsaE
Sll1744			119			76			RplA		RplA
Slr0447			104			66			AmiC		AmiC
Sll1808			101			78			RplE		RplE
Sll1317			101			101			PetA		PetA
Sll1528			98			29			Sll1528		Sll1528
Slr1185			96			98			PetC		PetC
Slr1102			96	59		62			Slr1102	Slr1102	
Sll1463			90			44			FtsH		FtsH
Sll1665			82			137			Sll1665	Sll1665	

Ssl2598			63			32			PsbH		PsbH
Sll1810			52			52			RplF		RplF
Slr1506			48			37			Slr1506		Slr1506
Sll0629			43			54			PsaK2		PsaK2
Sll0533			41								
Sll1450			40			76			NrtA		NrtA
Smr0006			37			49			PsbF		PsbF
Sll1327			29			39			AtpG		AtpG
Slr0940			29								
Sll1106				37		46			Sll1106		Sll1106

Table 5. Hits from *Synechocystis* NT-FLAG Sll1874 pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG Sll1874 pulldown experiments								Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction				Individual pulldown experiments using the insoluble fraction						
	1	2	3	4	1	2	3	4			
Sll1874	326	41	41	40	88	569	207	53	Sll1874		
Sll1577	32				226	274	35	91	Sll1577	Sll1577	
Sll1214			28	921		120			Sll1214		Sll1214
Slr0937				108							
Slr2075				85	220			144	Slr2075		Slr2075
Sll1147				50							
Slr2076				47	152	78	28	250	Slr2076		Slr2076
Sll1580					255						
Sll2067					206	176		130	Sll2067	Sll2067	
Sll1578					178	217	36	65	Sll1578	Sll1578	
Sll0416					168			157	Sll0416		Sll0416
Slr1884					162	185	99	176	Slr1884	Slr1884	
Sll1687					108		102	323	Sll1687	Sll1687	
Sll1746					95	214	144	154	Sll1746	Sll1746	
Ssr3383					62						
Ssl3093					56						
Sll1694					40	44			Sll1694		Sll1694
Ssl0707					35						
Slr1986					33	78			Slr1986		Slr1986
Sll1471					30						
Slr1102						939	175		Slr1102	Slr1102	
Slr1780						234					
Slr0244						212					
Sll1326						211					
Slr0906						198					

Sll1808						183		76	Sll1808		Sll1808
Sll1867						171					
Slr1311						154					
Slr1128						152					
Sll1812						151					
Sll1099						148					
Sll1106						137					
Sll0849						136					
Sll1322						128					
Sll1665						127	43		Sll1665	Sll1665	
Slr1835						124					
Slr1963						85	39		Slr1963		Slr1963

Table 6. Hits from *Synechocystis* NT-FLAG Slr1780 (Ycf54) pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG Ycf54 pulldown experiments						Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction			Individual pulldown experiments using the insoluble fraction					
	1	2	3	1	2	3			
Slr1666	117		142		119	129	Slr1666		Slr1666
Slr2067	113	291	82				Slr2067	Slr2067	
Sll1577	104	508	103	84	57	72	Sll1577	Sll1577	
Slr1140	101		178		72	38	Slr1140		Slr1140
Sll1746	45	200	80				Sll1746	Sll1746	
Sll1821	40	102					Sll1821		Sll1821
Slr1198	39	79	35				Slr1198		Slr1198
Slr1780	38	103	41	36	35	33	Slr1780		Slr1780
Slr1884	34	242			59	150	Slr1884	Slr1884	
Slr2076	34	335	108				Slr2076		Slr2076
Sll1578	28	266	120		65	45	Sll1578	Sll1578	
Sll1130	28	27					Sll1130		Sll1130
Sll0416		431							
Sll1580		319							
Slr1986		318	43				Slr1986		Slr1986
Slr0335		204							
Sll1099		187	63				Sll1099		Sll1099
Sll1687		172			149	197	Sll1687	Sll1687	
Sll1214		31				35	Sll1214		Sll1214
Sll1712		30							
Sll1098		30							
Slr0288			50						
Slr0376			38		47	51	Slr0376		Slr0376

Table 7. Hits from *Synechocystis* CT-FLAG Slr1780 (Ycf54) pulldown experiments

Synechocystis Gene ID	*Synechocystis* CT-FLAG Ycf54 pulldown experiments							Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction				Individual pulldown experiments using the insoluble fraction					
	1	2	3	4	1	2	3			
Sll1578	135	131		68		63	57	Sll1578	Sll1578	
Slr1780	79	49	374	302	64	90	379	Slr1780		Slr1780
Sll1577	76	46						Sll1577	Sll1577	
Slr0009	58	39						Slr0009		Slr0009
Slr1986	44	44				38		Slr1986		Slr1986
Slr1884	43					35		Slr1884	Slr1884	
Slr0012	40									
Sll1874		504				340		Sll1874		Sll1874
Sll1106		293				228		Sll1106		Sll1106
Slr0737		69								
Ssl0294		46								
Ssr2831		46				33		Ssr2831		Ssr2831
Sll1214		36				33		Sll1214		Sll1214
Sll1867		35								
Slr0483		34								
Slr1655		32				29		Slr1655		Slr1655
Sll1322		30				29		Sll1322		Sll1322
Sll1294		26								
Slr5058				26				Slr5058		Slr5058
Slr0619					28			Slr0619		Slr0619
Slr2067						32	74	Slr2067	Slr2067	
Ssl0707						28				

Table 8. Hits from *Synechocystis* NT-FLAG Slr0506 (POR) pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG POR pulldown experiments									Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction					Individual pulldown experiments using the insoluble fraction						
	1	2	3	4	5	1	2	3	4			
Slr0506	1628	428	492	496	501	1454	655	586	491	Slr0506		Slr0506
Sll1097	121	52	38	49		146	45	53		Sll1097		Sll1097
Ssl3436					68	97		33		Ssl3436		Ssl3436
Sll1746	99	127	96	136	158	96	147	166		Sll1746	Sll1746	
Sll1744	125	38	49	40		92	90	181		Sll1744		Sll1744
Sll1106						91						
Sll1812	52				38	87	42	64		Sll1812		Sll1812
Sll1099	195					80				Sll1099		Sll1099
Sll1802					41	77				Sll1802		Sll1802
Slr0244	82				94	76				Slr0244		Slr0244
Sll1101					28	64				Sll1101		Sll1101
Sll0043						59						
Sll1577	105	194	120	137	64	54	103	134	50	Sll1577	Sll1577	
Slr2067	57	135	155	125	106	53	71	142	71	Slr2067	Slr2067	
Slr1963						53						
sll0947						52						
Sll1801						50						
Slr0335						49						
Ssr1399	32				37	47	46			Ssr1399		Ssr1399
Sll1033	27					44						
Slr1835						42		33		Slr1835		Slr1835
Sll1578	38	125	187	144	113	42	109	88	47	Sll1578	Sll1578	
Sll1810				42		34	46	113		Sll1810		Sll1810
Slr2001						33						
Sll1824	48					33				Sll1824		Sll1824
Sll1180						32						
Sll1799						32	29					
Sll1743						30						

Sll1800						30	29	75		Sll1800		Sll1800
Ssl3432						30			34	Ssl3432		Ssl3432
Sll0587	38					29						
Slr1923	96					30				Slr1923		Slr1923
Sll1740						29						
Sll1561						28						
Slr2121	35					28						
Sll1151						27						
Slr1102		42	44	30			118	122		Slr1102	Slr1102	
Sll1808	174	69	45	82	41		117	103		Sll1808		Sll1808
Slr1844		156	227	236	111		114	121		Slr1844	Slr1844	
Slr1678				32			65	110		Slr1678		Slr1678
Sll1665							63	97		Sll1665		Sll1665
Sll1817							53	60		Sll1817		Sll1817
Sll1274			59	46			52	134		Sll1274		Sll1274
Sll1804							47	39		Sll1804		Sll1804
Sll1803							41					
Sll1821			38		71		31	81		Sll1821		Sll1821
Slr2076	140	71	122	52	81					Slr2076		Slr2076
Sll1109	125				48					Sll1109		Sll1109
Sll0837	84											
Sll1580	74				80					Sll1580	Sll1580	
Sll0819								47				
Sll1805				47	35			35		Sll1805		Sll1805
Sll1043		61	40	44				29		Sll1043	Sll1043	
Sll1634								29				
Sll1687		512	455	508						Sll1687	Sll1687	
Sll0416		47			326					Sll0416		Sll0416
Sll1809		29										
Sll1821			41		40					Sll1821		Sll1821
Slr1986			37	29								

Table 9. Hits from *Synechocystis* NT-FLAG Slr1923 (DVR) pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG Sll1874 pulldown experiments								Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction				Individual pulldown experiments using the insoluble fraction						
	1	2	3	4	1	2	3	4			
Slr1923	319	2039	1030	193	480	2040	771	855	Slr1923		Slr1923
Sll1687	129		104						Sll1687	Sll1687	
Sll1577	127		80	45				65	Sll1577	Sll1577	
Sll1578	109		118	102			66	65	Sll1578	Sll1578	
Slr2067	64		42	43				52	Slr2067	Slr2067	
Slr1844	63		175				124		Slr1844	Slr1844	
Sll0416	53										
Slr2076	28		87								
Slr2121		36				36			Slr2121		Slr2121
Sll1099		35									
Slr2081		29				29					
Slr1102							101	72	Slr1102	Slr1102	
Sll1746			56				81	90	Sll1746	Sll1746	
Sll0819								32			
Slr2080			45								
Sll1043			44								

Table 10. Hits from *Synechocystis* NT-FLAG Sll1091 (ChIP) pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG ChIP pulldown experiments							Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction			Individual pulldown experiments using the insoluble fraction						
	1	2	3	1	2	3	4			
Sll1091	1248	1878	1652	1720	1246	513	465	Sll1091		Sll1091
Slr1884	106			68	109	143	153	Slr1884	Slr1884	
Slr1102	94				69		352	Slr1102	Slr1102	
Sll1746	38				117	40	47	Sll1746	Sll1746	
Slr1923		114	27					Slr1923		Slr1923
Slr2001		38								
Sll0067			102							
Slr0506			31	43				Slr0506		Slr0506
Sll1577				262		137	42	Sll1577	Sll1577	
Sll1751				204			29	Sll1751		Sll1751
Sll1099				177						
Slr0844				168						
Slr2067				144		86		Slr2067	Slr2067	
Sll1580				88						
Ssl2296				77						
Slr1578				72		67		Slr1578	Slr1578	
Sll1808				47						
Ssr1399				38						
Sll1317				33						
Slr1140				31						
Sll1762				28						
Sll1665					43					
Sll1821					41					
Sll0819					36					
Slr1655					32					
Sll0416						214	379	Sll0416		Sll0416
Sll1687						171	228	Sll1687	Sll1688	

Slr2075						67	35	Slr2075		Slr2075
Sll5076						35				
Slr2076						32	66	Slr2076		Slr2076

Table 11. Hits from *Synechocystis* CT-FLAG Sll1091 (ChIP) pulldown experiments

Synechocystis Gene ID	*Synechocystis* CT-FLAG ChIP pulldown experiments									Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction					Individual pulldown experiments using the insoluble fraction						
	1	2	3	4	5	1	2	3	4			
Sll1091	938	936	886	996	899	1318	959	660	398	Sll1091		Sll1091
Slr2076	878	416	512		38	32				Slr2076		Slr2076
Sll0416	388	72	385							Sll0416		Sll0416
Slr2075	298	137	135							Slr2075		Slr2075
Slr1096	219											
Sll1326	172	32	87							Sll1326		Sll1326
Sll1694	168	188								Sll1694		Sll1694
Sll0020	168		41							Sll0020		Sll0020
Sll1577	159							82	40	Sll1577	Sll1577	
Sll1769	147											
Slr1884	139	225	336	44	47	271		75	30	Slr1884	Slr1884	
Sll1687	126	33	84							Sll1687	Sll1687	
Sll1578	112	30					33	126	130	Sll1578	Sll1578	
Slr0737	105											
Sll0982	94											
Sll1184	85											
Sll1867	76		68							Sll1867		Sll1867
Slr0058	76											
Slr0731	73											
Sll1218	61											
Slr0552	55											
Slr0244	50		34			75				Slr0244		Slr0244
Sll0823	48											
Slr1220	45		80							Slr1220		Slr1220
Slr1329	44											

Sll1665	44		69							Sll1665	Sll1665	
Slr1835	42		92							Slr1835		Slr1835
Sll1536	39		58							Sll1536		Sll1536
Ssl0563	39											
Slr2067	39		82					63	64	Slr2067	Slr2067	
Sll0569	38											
Sll1633	36											
Slr1963	36											
Slr0483	34		39							Slr0483		Slr0483
Sll1456	32											
Slr0886	31											
Smr0006	30											
Sll1261	29											
Sll1580	27											
Sll1945			167									
Slr0447			114									
Sll1746			84			81				Sll1746	Sll1746	
Sll0819			79									
Slr1986			69					67	42	Slr1986		Slr1986
Sll1430			60									
Ssr2831			48									
Slr1102			47		126	516				Slr1102	Slr1102	
Sll1450			47									
Sll1029			45									
Slr6021			35			35				Slr6021		Slr6021

Table 12. Hits from *Synechocystis* NT-FLAG Slr1644 (Pitt) pulldown experiments

Synechocystis Gene ID	*Synechocystis* NT-FLAG Slr1644 (Pitt) pulldown experiments					Hits identified in >1 pulldown with a Mascot score >30	Hits identified in control pulldown	Possible interaction partners
	Individual pulldown experiments using the soluble fraction		Individual pulldown experiments using the insoluble fraction					
	1	2	1	2	3			
Slr1044	790							
Slr1644	638	82	288	58	231	Slr1644		Slr1644
Slr1277	287							
Slr1577	196	307	204	267	252	Slr1577	Slr1577	
Slr1834	123							
Slr1884	100	215	96	133	101	Slr1884	Slr1884	
Slr2067	98	133	189	161	206	Slr2067	Slr2067	
Slr1272	94							
Slr0906	92							
Sll1942	91							
Sll1665	61							
Sll1578	60	131	170	143	170	Sll1578	Sll1578	
Sll1106	59							
Sll1326	55							
Slr0552	52							
Sll1837	46							
Sll0818	43							
Ssl3436	42	77				Ssl3436		Ssl3436
Ssr2831	40							
Slr1329	39							
Ssl0563	36							
Ssr3451	35							
Sll1096	32	35				Sll1096		Sll1096
Slr0737	31							
Sll1274	29	110				Sll1274		Sll1274

Slr1102	28							
Slr2076		237			26	Slr2076		Slr2076
Sll1687		212						
Sll0416		124						
Slr1140		105						
Sll1746		98	61		83	Sll1746	Sll1746	
Sll1812		96						
Sll1821		52						
Sll1817		47						
Slr1678		45						
Slr2075		45						
Slr0009		43						
Sll1811		36						
Slr1986		35	86		143	Slr1986		Slr1986
Sll1260		33						
Slr0467		32						
Sll1097		32						
Sll0601		31						
Ssl3093		30	65	109	57	Ssl3093		Ssl3093
Slr0455		27						
Sll1471			68		34	Sll1471		Sll1471
Sll1712			57		49	Sll1712		Sll1712
Slr0394			49					
Sll1789			30	138		Sll1789		Sll1789
Slr1708				29				
Sll1099					72			
Sll1580					36			

References

(1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol Crystallogr.* **50,** 760-763.

Addlesee, H. A.; Gibson, L. C.; Jensen, P. E.; Hunter, C. N. (1996). Cloning, sequencing and functional assignment of the chlorophyll biosynthesis gene, chlP, of Synechocystis sp. PCC 6803. *FEBS Lett* **389,** 126-130.

Addlesee, H. A.; Fiedor, L.; Hunter, C. N. (2000). Physical mapping of bchG, orf427, and orf177 in the photosynthesis gene cluster of Rhodobacter sphaeroides: functional assignment of the bacteriochlorophyll synthetase gene. *J. Bacteriol.* **182,** 3175-3182.

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

Adhikari, N. D.; Froehlich, J. E.; Strand, D. D.; Buck, S. M.; Kramer, D. M.; Larkin, R. M. (2011). GUN4-porphyrin complexes bind the ChlH/GUN5 subunit of Mg-Chelatase and promote chlorophyll biosynthesis in Arabidopsis. *Plant Cell* **23,** 1449-1467.

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

Alawady, A.; Reski, R.; Yaronskaya, E.; Grimm, B. (2005). Cloning and expression of the tobacco CHLM sequence encoding Mg protoporphyrin IX methyltransferase and its interaction with Mg chelatase. *Plant Mol Biol* **57,** 679-691.

Alawady, A. E. and Grimm, B. (2005). Tobacco Mg protoporphyrin IX methyltransferase is involved in inverse activation of Mg porphyrin and protoheme synthesis. *Plant J* **41,** 282-290.

Albus, C. A.; Salinas, A.; Czarnecki, O.; Kahlau, S.; Rothbart, M.; Thiele, W.; Lein, W.; Bock, R.; Grimm, B.; Schottler, M. A. (2012). LCAA, a novel factor required for magnesium protoporphyrin monomethylester cyclase accumulation and feedback control of aminolevulinic acid biosynthesis in tobacco. *Plant Physiol* **160,** 1923-1939.

Alvarez, S.; Zapata, M.; Garrido, J. L.; Vaz, B. (2012). Characterization of [8-ethyl]-chlorophyll c3 from Emiliania huxleyi. *Chem Commun. (Camb.)* **48,** 5500-5502.

Andrizhiyevskaya, E. G.; Chojnicka, A.; Bautista, J. A.; Diner, B. A.; Van, G. R.; Dekker, J. P. (2005). Origin of the F685 and F695 fluorescence in photosystem II. *Photosynth Res* **84,** 173-180.

Ano, A.; Funahashi, H.; Nakao, K.; Nishizawa, Y. (1999). Effect of glycine on 5-aminolevulinic acid biosynthesis in heterotrophic culture of Chlorella regularis YA-603. *J. Biosci. Bioeng.* **88,** 57-60.

Apchelimov, A. A.; Soldatova, O. P.; Ezhova, T. A.; Grimm, B.; Shestakov, S. V. (2007). The analysis of the ChlI 1 and ChlI 2 genes using acifluorfen-resistant mutant of Arabidopsis thaliana. *Planta* **225,** 935-943.

Archibald, J. M. (2005). Jumping genes and shrinking genomes--probing the evolution of eukaryotic photosynthesis with genomics. *IUBMB Life* **57,** 539-547.

Armbruster, U.; Labs, M.; Pribil, M.; Viola, S.; Xu, W.; Scharfenberg, M.; Hertle, A. P.; Rojahn, U.; Jensen, P. E.; Rappaport, F.; Joliot, P.; Dormann, P.; Wanner, G.; Leister, D. (2013). Arabidopsis CURVATURE THYLAKOID1 proteins modify thylakoid architecture by inducing membrane curvature. *Plant Cell* **25,** 2661-2678.

Armstrong, G. A.; Runge, S.; Frick, G.; Sperling, U.; Apel, K. (1995). Identification of NADPH:protochlorophyllide oxidoreductases A and B: a branched pathway for light-dependent chlorophyll biosynthesis in Arabidopsis thaliana. *Plant Physiol* **108,** 1505-1517.

Armstrong, G. A. (1997). Genetics of eubacterial carotenoid biosynthesis: a colorful tale. *Annu Rev Microbiol* **51,** 629-659.

Asao, M. and Madigan, M. T. (2010). Taxonomy, phylogeny, and ecology of the heliobacteria. *Photosynth Res* **104,** 103-111.

Asrar, G.; Fuchs, A.; Kanemasu, E. T.; Hatfield, J. L. (1984). Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agronomy Journal* **76,** 300-306.

Avissar, Y. J. and Beale, S. I. (1989). Biosynthesis of Tetrapyrrole Pigment Precursors : Pyridoxal Requirement of the Aminotransferase Step in the Formation of delta-Aminolevulinate from Glutamate in Extracts of Chlorella vulgaris. *Plant Physiol* **89,** 852-859.

Bang, W. Y.; Jeong, I. S.; Kim, D. W.; Im, C. H.; Ji, C.; Hwang, S. M.; Kim, S. W.; Son, Y. S.; Jeong, J.; Shiina, T.; Bahk, J. D. (2008). Role of Arabidopsis CHL27 protein for photosynthesis, chloroplast development and gene expression profiling. *Plant Cell Physiol* **49,** 1350-1363.

Battersby, A. R.; Fookes, C. J. R.; Hart, G.; Matcham, G. W. J.; Pandey, P. (1983). Biosynthesis of porphyrins and related macrocycles. Part 21. The interaction of deaminase and its product (hydroxymethylbilane) and the relationship between deaminase and cosynthetase. *J. Chem Soc Perkin Trans* 3041-3047.

Beale, S. I. and Castelfranco, P. A. (1974a). The Biosynthesis of delta-Aminolevulinic Acid in Higher Plants: II. Formation of C-delta-Aminolevulinic Acid from Labeled Precursors in Greening Plant Tissues. *Plant Physiol* **53,** 297-303.

Beale, S. I. and Castelfranco, P. A. (1974b). The Biosynthesis of delta-Aminolevulinic Acid in Higher Plants: I. Accumulation of delta-Aminolevulinic Acid in Greening Plant Tissues. *Plant Physiol* **53,** 291-296.

Beale, S. I.; Gough, S. P.; Granick, S. (1975). Biosynthesis of delta-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley. *Proc Natl Acad Sci U S A* **72,** 2719-2723.

Beale, S. I. (1999). Enzymes of chlorophyll biosynthesis. *Photosynthesis Research* **60,** 43-73.

Beale, S. I. (2006). Biosynthesis of 5-Aminolevulinic Acid. In: *Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, FUnctions and Applications* (Grimm, B., Porra, R. J., Rudiger, W., Scheer, H., eds.) Springer, pp. 147-158.

Beatty, J. T. (2013) *Advances in Botanical Research: Genome Evolution of Photosynthetic Bactiera*. (J.Jacquot and P.Gadal, eds.) Elsevier, London .

Begley, T. P. and Young, H. (1989). Protochlorophyllide reductase. 1. Determination of the regiochemistry and the stereochemistry of the reduction of protochlorophyllide to chlorophyllide. *Journal of the American Chemical Society* **111,** 3095-3096.

Benz, J. and Rüdiger, W. (1981). Chlorophyll biosynthesis: Various chlorophyllides as exogenous substrates for chlorophyll synthase. *Z Naturforsch* **36c,** 51-57.

Berthold, D. A. and Stenmark, P. (2003). Membrane-bound diiron carboxylate proteins. *Annu Rev Plant Biol* **54,** 497-517.

Bhattacharya, D.; Yoon, H. S.; Hackett, J. D. (2004). Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *Bioessays* **26,** 50-60.

Blankenship, R. E. (1992). Origin and early evolution of photosynthesis. *Photosynth Res* **33,** 91-111.

Blankenship, R. E. (2010). Early evolution of photosynthesis. *Plant Physiol* **154,** 434-438.

Block, M. A.; Tewari, A. K.; Albrieux, C.; Marechal, E.; 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.; Popovic, R.; Franck, F. (2003). Early reactions of light-induced protochlorophyllide and chlorophyllide transformations analyzed in vivo at room temperature with a diode array spectrofluorometer. *J Photochem Photobiol. B* **69,** 31-39.

Bogorad, L. (1958a). The enzymatic synthesis of porphyrins from porphobilinogen. III. Uroporphyrinogens as intermediates. *J Biol Chem* **233,** 516-519.

Bogorad, L. (1958b). The enzymatic synthesis of porphyrins from porphobilinogen. II. Uroporphyrin III. *J Biol Chem* **233,** 510-515.

Bogorad, L. (1958c). The enzymatic synthesis of porphyrins from porphobilinogen. I. Uroporphyrin I. *J Biol Chem* **233,** 501-509.

Bogorad, L. (1976). Chemistry and Biochemistry of Plant Pigments. In: *Chlorophyll Biosynthesis* (Goodwin, T. W., ed.) Academic Press, New York, pp. 64-148.

Boldareva-Nuianzina, E. N.; Blahova, Z.; Sobotka, R.; Koblizek, M. (2013). Distribution and Origin of Oxygen-Dependent and Oxygen-Independent Forms of Mg-Protoporphyrin Monomethylester Cyclase among Phototrophic Proteobacteria. *Appl Environ Microbiol* **79,** 2596-2604.

Bollivar, D. W.; Jiang, Z. Y.; Bauer, C. E.; Beale, S. I. (1994a). Heterologous expression of the bchM gene product from Rhodobacter capsulatus and demonstration that it encodes S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase. *J. Bacteriol.* **176,** 5290-5296.

Bollivar, D. W.; Wang, S.; Allen, J. P.; Bauer, C. E. (1994b). Molecular genetic analysis of terminal steps in bacteriochlorophyll a biosynthesis: characterization of a Rhodobacter capsulatus strain that synthesizes geranylgeraniol-esterified bacteriochlorophyll a. *Biochemistry* **33,** 12763-12768.

Bollivar, D. W.; Suzuki, J. Y.; Beatty, J. T.; Dobrowolski, J. M.; Bauer, C. E. (1994c). Directed mutational analysis of bacteriochlorophyll *a* biosynthesis in *Rhodobacter capsulatus*. *Journal of Molecular Biology* **237,** 622-640.

Bollivar, D. W. and Beale, S. I. (1996a). The Chlorophyll Biosynthetic Enzyme Mg-Protoporphyrin IX Monomethyl Ester (Oxidative) Cyclase (Characterization and Partial Purification from Chlamydomonas reinhardtii and Synechocystis sp. PCC 6803). *Plant Physiol* **112,** 105-114.

Bollivar, D. W. and Beale, S. I. (1996b). The chlorophyll biosynthetic enzyme Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase: characterization and partial purification from *Chlamydomonas reinhardtii* and *Synechocystis* sp. PCC 6803. *Plant Physiol.* **112,** 105-114.

Bollivar, D. W.; Clauson, C.; Lighthall, R.; Forbes, S.; Kokona, B.; Fairman, R.; Kundrat, L.; Jaffe, E. K. (2004). Rhodobacter capsulatus porphobilinogen synthase, a high activity metal ion independent hexamer. *BMC Biochem* **5,** 17.

Bougri, O. and Grimm, B. (1996). Members of a low-copy number gene family encoding glutamyl-tRNA reductase are differentially expressed in barley. *Plant J* **9,** 867-878.

Boynton, T. O.; Gerdes, S.; Craven, S. H.; Neidle, E. L.; Phillips, J. D.; Dailey, H. A. (2011). Discovery of a gene involved in a third bacterial protoporphyrinogen oxidase activity through comparative genomic analysis and functional complementation. *Appl Environ Microbiol* **77,** 4795-4801.

Brenner, D. A. and Frasier, F. (1991). Cloning of murine ferrochelatase. *Proc Natl Acad Sci U S A* **88,** 849-853.

Brocker, M. J.; Virus, S.; Ganskow, S.; Heathcote, P.; Heinz, D. W.; Schubert, W. D.; Jahn, D.; Moser, J. (2008). ATP-driven reduction by dark-operative protochlorophyllide oxidoreductase from Chlorobium tepidum mechanistically resembles nitrogenase catalysis. *J Biol Chem* **283,** 10559-10567.

Bryant, D. A.; Costas, A. M.; Maresca, J. A.; Chew, A. G.; Klatt, C. G.; Bateson, M. M.; Tallon, L. J.; Hostetler, J.; Nelson, W. C.; Heidelberg, J. F.; Ward, D. M. (2007). Candidatus Chloracidobacterium thermophilum: an aerobic phototrophic Acidobacterium. *Science* **317,** 523-526.

Buhr, F.; El, B. M.; Valdez, O.; Pollmann, S.; Lebedev, N.; Reinbothe, S.; Reinbothe, C. (2008). Photoprotective role of NADPH:protochlorophyllide oxidoreductase A. *Proc Natl Acad Sci U S A* **105,** 12629-12634.

Buick, R. (1992). The antiquity of oxygenic photosynthesis: evidence from stromatolites in sulphate-deficient Archaean lakes. *Science* **255,** 74-77.

Burke, D. H.; Alberti, M.; Hearst, J. E. (1993a). The Rhodobacter capsulatus chlorin reductase-encoding locus, bchA, consists of three genes, bchX, bchY, and bchZ. *J. Bacteriol.* **175,** 2407-2413.

Burke, D. H.; Alberti, M.; Hearst, J. E. (1993b). bchFNBH bacteriochlorophyll synthesis genes of Rhodobacter capsulatus and identification of the third subunit of light-independent protochlorophyllide reductase in bacteria and plants. *J. Bacteriol.* **175,** 2414-2422.

Cahoon, A. B. and Timko, M. P. (2000). yellow-in-the-dark mutants of Chlamydomonas lack the CHLL subunit of light-independent protochlorophyllide reductase. *Plant Cell* **12,** 559-568.

Calvin, M. and Benson, A. (1948) The Path of Carbon in Photosynthesis. *Science* **107**, 476-480.

Camadro, J. M.; Matringe, M.; Scalla, R.; Labbe, P. (1991). Kinetic studies on protoporphyrinogen oxidase inhibition by diphenyl ether herbicides. *Biochem J* **277 (Pt 1),** 17-21.

Canniffe, D. P.; Jackson, P. J.; Hollingshead, S.; Dickman, M. J.; 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.

Castelfranco, C. and Beale, P. (1983) Chlorophyll Biosynthesis: Recent Advances and Areas of Current Interest. *Annual Review of Plant Physiology* **34**, 241-278.

Cavalier-Smith, T. (1999). Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot. Microbiol* **46,** 347-366.

Chamovitz, D.; Sandmann, G.; Hirschberg, J. (1993). Molecular and biochemical characterization of herbicide-resistant mutants of cyanobacteria reveals that phytoene desaturation is a rate-limiting step in carotenoid biosynthesis. *J Biol Chem* **268,** 17348-17353.

Che, F. S.; Watanabe, N.; Iwano, M.; Inokuchi, H.; Takayama, S.; Yoshida, S.; Isogai, A. (2000). Molecular characterization and subcellular localization of protoporphyrinogen oxidase in spinach chloroplasts. *Plant Physiol* **124,** 59-70.

Chen, M. W.; Jahn, D.; O'Neill, G. P.; Soll, D. (1990). Purification of the glutamyl-tRNA reductase from Chlamydomonas reinhardtii involved in delta-aminolevulinic acid formation during chlorophyll biosynthesis. *J Biol Chem* **265,** 4058-4063.

Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol Crystallogr.* **66,** 12-21.

Chereskin, B. M.; Wong, Y. S.; Castelfranco, P. A. (1982). In Vitro Synthesis of the Chlorophyll Isocyclic Ring : Transformation of Magnesium-Protoporphyrin IX and Magnesium-Protoporphyrin IX Monomethyl Ester into Magnesium-2,4-Divinyl Pheoporphyrin A(5). *Plant Physiol* **70,** 987-993.

Chew, A. G. and Bryant, D. A. (2007). Characterization of a plant-like protochlorophyllide a divinyl reductase in green sulfur bacteria. *J Biol Chem* **282,** 2967-2975.

Chow, K. S.; Singh, D. P.; Roper, J. M.; Smith, A. G. (1997). A single precursor protein for ferrochelatase-I from Arabidopsis is imported in vitro into both chloroplasts and mitochondria. *J Biol Chem* **272,** 27565-27571.

Chow, K. S.; Singh, D. P.; Walker, A. R.; Smith, A. G. (1998). Two different genes encode ferrochelatase in Arabidopsis: mapping, expression and subcellular targeting of the precursor proteins. *Plant J* **15,** 531-541.

Coomber, S. A.; Chaudhri, M.; Connor, A.; Britton, G.; Hunter, C. N. (1990). Localized transposon Tn5 mutagenesis of the photosynthetic gene cluster of Rhodobacter sphaeroides. *Mol. Microbiol.* **4,** 977-989.

Cooper, R. (1963). The biosynthesis of coproporphyrinogen, magnesium protoporphyrin monomethyl ester and bacteriochlorophyll by Rhodopseudomonas capsulata. *Biochem J* **89,** 100-108.

Cunningham, F. and Gantt, E. (1998). Genes and Enzymes of Carotenoid Biosynthesis in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **49,** 557-583.

Dahlin, C.; Aronsson, H.; Wilks, H. M.; Lebedev, N.; Sundqvist, C.; Timko, M. P. (1999). The role of protein surface charge in catalytic activity and chloroplast membrane association of the pea NADPH: protochlorophyllide oxidoreductase (POR) as revealed by alanine scanning mutagenesis. *Plant Mol Biol* **39,** 309-323.

Dailey, H. A. and Karr, S. W. (1987). Purification and characterization of murine protoporphyrinogen oxidase. *Biochemistry* **26,** 2697-2701.

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

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

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

Davison, P. A. and Hunter, C. N. (2011). Abolition of magnesium chelatase activity by the gun5 mutation and reversal by Gun4. *FEBS Lett* **585,** 183-186.

Delfau-Larue, M. H.; Martasek, P.; Grandchamp, B. (1994). Coproporphyrinogen oxidase: gene organization and description of a mutation leading to exon 6 skipping. *Hum. Mol Genet.* **3,** 1325-1330.

Delwiche, C. F. (1999). Tracing the Thread of Plastid Diversity through the Tapestry of Life. *Am Nat* **154,** S164-S177.

Dietmann, S.; Park, J.; Notredame, C.; Heger, A.; Lappe, M.; Holm, L. (2001). A fully automatic evolutionary classification of protein folds: Dali Domain Dictionary version 3. *Nucleic Acids Res* **29,** 55-57.

Duhring, U.; Irrgang, K. D.; Lunser, K.; Kehr, J.; Wilde, A. (2006). Analysis of photosynthetic complexes from a cyanobacterial ycf37 mutant. *Biochim Biophys Acta* **1757,** 3-11.

Dym, O.; Pratt, E. A.; Ho, C.; Eisenberg, D. (2000). The crystal structure of D-lactate dehydrogenase, a peripheral membrane respiratory enzyme. *Proc Natl Acad Sci U S A* **97,** 9413-9418.

Eichacker, L.; Paulsen, H.; Rudiger, W. (1992). Synthesis of chlorophyll a regulates translation of chlorophyll a apoproteins P700, CP47, CP43 and D2 in barley etioplasts. *Eur J Biochem* **205,** 17-24.

Einsle, O.; Tezcan, F. A.; Andrade, S. L.; Schmid, B.; Yoshida, M.; Howard, J. B.; Rees, D. C. (2002). Nitrogenase MoFe-protein at 1.16 A resolution: a central ligand in the FeMo-cofactor. *Science* **297,** 1696-1700.

Elliott, T.; Avissar, Y. J.; Rhie, G. E.; Beale, S. I. (1990). Cloning and sequence of the Salmonella typhimurium hemL gene and identification of the missing enzyme in hemL mutants as glutamate-1-semialdehyde aminotransferase. *J Bacteriol* **172,** 7071-7084.

Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol Crystallogr.* **66,** 486-501.

Erskine, P. T.; Senior, N.; Awan, S.; Lambert, R.; Lewis, G.; Tickle, I. J.; Sarwar, M.; Spencer, P.; Thomas, P.; Warren, M. J.; Shoolingin-Jordan, P. M.; Wood, S. P.; Cooper, J. B. (1997). X-ray structure of 5-aminolaevulinate dehydratase, a hybrid aldolase. *Nat Struct Biol* **4,** 1025-1031.

Erskine, P. T.; Norton, E.; Cooper, J. B.; Lambert, R.; Coker, A.; Lewis, G.; Spencer, P.; Sarwar, M.; Wood, S. P.; Warren, M. J.; Schoolingin-Jordan, P. M. (1999). X-ray structure of 5-aminolevulinic acid dehydratase from *Escherichia coli* complexed with the inhibitor levulinic acid at 2.0 angstrom resolution. *Biochemistry* **38,** 4266-4276.

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. D Biol Crystallogr.* **62,** 72-78.

Falbel, T. G. and Staehelin, L. A. (1994). Characterization of a family of chlorophyll-deficient wheat (Triticum) and barley (Hordeum vulgare) mutants with defects in the magnesium-insertion step of chlorophyll biosynthesis. *Plant Physiol* **104,** 639-648.

Fan, J.; Liu, Q.; Hao, Q.; Teng, M.; Niu, L. (2007). Crystal structure of uroporphyrinogen decarboxylase from Bacillus subtilis. *J Bacteriol* **189,** 3573-3580.

Fodje, M. N.; Hansson, A.; Hansson, M.; Olsen, J. G.; Gough, S.; Willows, R. D.; 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.

Fokina, O.; Chellamuthu, V. R.; Forchhammer, K.; Zeth, K. (2010). Mechanism of 2-oxoglutarate signaling by the Synechococcus elongatus PII signal transduction protein. *Proc Natl Acad Sci U S A* **107,** 19760-19765.

Fookes, C. J. R. and Jeffrey, S. W. (1989). The structure of chlorophyll c3, a novel marine photosynthetic pigment. *Journal of Chemical Society, Chemical Communications* 1827-1828.

Ford, C. and Wang, W. (1980a). Three new yellow loci in Chlamydomonas reinhardtii. *Mol Gen. Genet.* **179,** 259-263.

Ford, C. and Wang, W. Y. (1980b). Temperature sensitive *yellow* mutants of *Chlamydomonas reinhartii*. *Mol Gen. Genet.* **184,** 5-10.

Formighieri, C.; Ceol, M.; Bonente, G.; Rochaix, J. D.; Bassi, R. (2012). Retrograde signaling and photoprotection in a gun4 mutant of Chlamydomonas reinhardtii. *Mol Plant* **5,** 1242-1262.

Frankenberg, N.; Erskine, P. T.; Cooper, J. B.; Shoolingin-Jordan, P. M.; Jahn, D.; Heinz, D. W. (1999). High resolution crystal structure of a Mg2+-dependent porphobilinogen synthase. *J Mol Biol* **289,** 591-602.

Fuesler, T. P.; Wong, Y. S.; Castelfranco, P. A. (1984). Localization of Mg-chelatase and Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase activities within isolated, developing cucumber chloroplasts. *Plant Physiol.* **75,** 662-664.

Fuhrmann, E.; Gathmann, S.; Rupprecht, E.; Golecki, J.; Schneider, D. (2009). Thylakoid membrane reduction affects the photosystem stoichiometry in the cyanobacterium Synechocystis sp. PCC 6803. *Plant Physiol* **149,** 735-744.

Fujita, Y.; Matsumoto, H.; Takahashi, Y.; Matsubara, H. (1993). Identification of a nifDK-like gene (ORF467) involved in the biosynthesis of chlorophyll in the cyanobacterium Plectonema boryanum. *Plant Cell Physiol* **34,** 305-314.

Fujita, Y. and Bauer, C. E. (2000). Reconstitution of light-independent protochlorophyllide reductase from purified bchl and BchN-BchB subunits. In vitro confirmation of nitrogenase-like features of a bacteriochlorophyll biosynthesis enzyme. *J. Biol. Chem.* **275,** 23583-23588.

Fujita, Y. and Bauer, C. E. (2003). The light-indenpendent protochlorophyllide reductase: a nitrogenase-like enzyme catalysing a key reaction for greening in the dark. In: *Porphyrin Handbook* (Kadish, K., ed.) Academic Press, New York , pp. 109-156.

Gabruk, M.; Grzyb, J.; Kruk, J.; Mysliwa-Kurdziel, B. (2012). Light-dependent and light-independent protochlorophyllide oxidoreductases share similar sequence motifs – in silico studies. *Photosynthetica* **50,** 529-540.

Ge, H.; Lv, X.; Fan, J.; Gao, Y.; Teng, M.; Niu, L. (2010). Crystal structure of glutamate-1-semialdehyde aminotransferase from Bacillus subtilis with bound pyridoxamine-5'-phosphate. *Biochem Biophys Res Commun.* **402,** 356-360.

Gest, H. and Favinger, J. (1983). Heliobacterium chlorum, an anoxygenic brownish-green photosynthetic bacterium containing a ''new'' form of bacteriochlrophyll. *Archives of Microbiology* **136,** 11-16.

Gibson, J.; Ludwig, W.; Stackebrandt, E.; Wose, C. R. (1985). The phylogeny of the green photosynthetic bacteria: absence of a close relationship between chlorobium and chloroflexus. *Systematic and applied microbiology* **6,** 152-156.

Gibson, K.; Neuberger, A.; Tait, G. (1963). Studies on the biosynthesis of porphyrin and bacteriochlorophyll by Rhodopseudomonas sphaeroides. 4. S-adenosylmethionine magnesium protoporphyrin IX methyltransferase. *Biochem J* **88,** 325-334.

Gibson, L. C. and Hunter, C. N. (1994). The bacteriochlorophyll biosynthesis gene, bchM, of Rhodobacter sphaeroides encodes S-adenosyl-L-methionine: Mg protoporphyrin IX methyltransferase. *FEBS Lett* **352,** 127-130.

Gibson, L. C.; Willows, R. D.; Kannangara, C. G.; von, W. D.; 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.

Gibson, L. C.; Jensen, P. E.; Hunter, C. N. (1999). Magnesium chelatase from Rhodobacter sphaeroides: initial characterization of the enzyme using purified subunits and evidence for a BchI-BchD complex. *Biochem. J.* **337 (Pt 2),** 243-251.

Giese, K. C. and Vierling, E. (2002). Changes in oligomerization are essential for the chaperone activity of a small heat shock protein in vivo and in vitro. *J Biol Chem* **277,** 46310-46318.

Glazer, A. N.; Williams, R. C.; yamanaka, G.; Schachman, H. K. (1979). Characterization of cyanobacterial phycobilisomes in zwitterionic detergents. *Proc Natl Acad Sci U S A* **76,** 6162-6166.

Gloe, A.; Pfennig, N.; Brockmann, H., Jr.; Trowitzsch, W. (1975). A new bacteriochlorophyll from brown-colored Chlorobiaceae. *Arch Microbiol* **102,** 103-109.

Gomez Maqueo, C. A.; Frigaard, N. U.; Bryant, D. A. (2009). Mutational analysis of three bchH paralogs in (bacterio-)chlorophyll biosynthesis in Chlorobaculum tepidum. *Photosynth Res* **101,** 21-34.

Gora, M.; Chacinska, A.; Rytka, J.; Labbe-Bois, R. (1996). Isolation and functional characterization of mutant ferrochelatases in Saccharomyces cerevisiae. *Biochimie* **78,** 144-152.

Gora, M.; Rytka, J.; Labbe-Bois, R. (1999). Activity and cellular location in Saccharomyces cerevisiae of chimeric mouse/yeast and Bacillus subtilis/yeast ferrochelatases. *Arch Biochem Biophys* **361,** 231-240.

Gorchein, A. (1972). Magnesium protoporphyrin chelatase activity in *Rhodopseudomonas sphaeroides*. Studies with whole cells. *Biochemical Journal* **127,** 97-106.

Goslings, D.; Meskauskiene, R.; Kim, C.; Lee, K. P.; Nater, M.; Apel, K. (2004). Concurrent interactions of heme and FLU with Glu tRNA reductase (HEMA1), the target of metabolic feedback inhibition of tetrapyrrole biosynthesis, in dark- and light-grown Arabidopsis plants. *Plant J* **40,** 957-967.

Goto, T.; Aoki, R.; Minamizaki, K.; Fujita, Y. (2010). Functional differentiation of two analogous coproporphyrinogen III oxidases for heme and chlorophyll biosynthesis pathways in the cyanobacterium Synechocystis sp. PCC 6803. *Plant Cell Physiol* **51,** 650-663.

Gough, S. (1972). Defective synthesis of porphyrins in barley plastids caused by mutation in nuclear genes. *Biochim Biophys Acta* **286,** 36-54.

Gough, S. P.; Kannangara, C. G.; Britton, B. J. (1989). A new method for the synthesis of glutamate 1-semialdehyde. Characterization of its structure in solution by NMR spectroscopy. *Carlsberg. Res Commun.* **54,** 99-108.

Gough, S. P.; Petersen, B. O.; Duus, J. O. (2000). Anaerobic chlorophyll isocyclic ring formation in Rhodobacter capsulatus requires a cobalamin cofactor. *Proc. Natl. Acad. Sci. U. S. A* **97,** 6908-6913.

Grandchamp, B.; Phung, N.; Nordmann, Y. (1978). The mitochondrial localization of coproporphyrinogen III oxidase. *Biochem J* **176,** 97-102.

Granik, S. (1950). The structural and functional relationships between heme and chlorophyll. *Harvey Lectures* 220-245.

Griese, M.; Lange, C.; Soppa, J. (2011). Ploidy in cyanobacteria. *FEMS Microbiol Lett* **323,** 124-131.

Griffiths, W. T. (1975a). Characterization of the terminal stages of chlorophyll (ide) synthesis in etioplast membrane preparations. *Biochem. J.* **152,** 623-635.

Griffiths, W. T. (1975b). Some observations on chlorophyll(ide) synthesis by isolated etioplasts. *Biochem. J.* **146,** 17-24.

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

Grigorieva, G. and Shestakov, S. (1982). Transformation in the cyanobacterium Synechocystis sp. 6803. *FEMS Microbiol Letters* **13,** 367-370.

Guenther, B.; Onrust, R.; Sali, A.; O'Donnell, M.; Kuriyan, J. (1997). Crystal structure of the delta' subunit of the clamp-loader complex of E. coli DNA polymerase III. *Cell* **91,** 335-345.

Gupta, R. S.; Mukhtar, T.; Singh, B. (1999). Evolutionary relationships among photosynthetic prokaryotes (Heliobacterium chlorum, Chloroflexus aurantiacus, cyanobacteria, Chlorobium tepidum and proteobacteria): implications regarding the origin of photosynthesis. *Mol Microbiol* **32,** 893-906.

Hansson, A.; Kannangara, C. G.; von, W. D.; Hansson, M. (1999). Molecular basis for semidominance of missense mutations in the XANTHA-H (42-kDa) subunit of

magnesium chelatase. *Proc Natl Acad Sci U S A* **96,** 1744-1749.

Hansson, M.; Rutberg, L.; Schroder, I.; Hederstedt, L. (1991). The Bacillus subtilis hemAXCDBL gene cluster, which encodes enzymes of the biosynthetic pathway from glutamate to uroporphyrinogen III. *J Bacteriol* **173,** 2590-2599.

Hansson, M. and Hederstedt, L. (1994). Purification and characterisation of a water-soluble ferrochelatase from Bacillus subtilis. *Eur J Biochem* **220,** 201-208.

Hansson, M. and Vener, A. V. (2003). Identification of three previously unknown in vivo protein phosphorylation sites in thylakoid membranes of Arabidopsis thaliana. *Mol Cell Proteomics* **2,** 550-559.

Harada, J.; Wada, K.; Yamaguchi, H.; Oh-Oka, H.; Tamiaki, H.; Fukuyama, K. (2005). Crystallization and preliminary X-ray diffraction study of BchU, a methyltransferase from Chlorobium tepidum involved in bacteriochlorophyll c biosynthesis. *Acta Crystallogr. Sect. F Struct Biol Cryst. Commun.* **61,** 712-714.

Harmer, S. L.; Hogenesch, J. B.; Straume, M.; Chang, H. S.; Han, B.; Zhu, T.; Wang, X.; Kreps, J. A.; Kay, S. A. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* **290,** 2110-2113.

Hart, G. J. and Battersby, A. R. (1985). Purification and properties of uroporphyrinogen III synthase (co-synthetase) from Euglena gracilis. *Biochem J* **232,** 151-160.

Heinemann, I. U.; Jahn, M.; Jahn, D. (2008). The biochemistry of heme biosynthesis. *Arch Biochem Biophys* **474,** 238-251.

Hennig, M.; Grimm, B.; Contestabile, R.; John, R. A.; Jansonius, J. N. (1997a). Crystal structure of glutamate-1-semialdehyde aminomutase: an alpha2- dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity. *Proc. Natl. Acad. Sci. U. S. A* **94,** 4866-4871.

Hennig, M.; Grimm, B.; Contestabile, R.; John, R. A.; Jansonius, J. N. (1997b). Crystal structure of glutamate-1-semialdehyde aminomutase: an alpha2- dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity. *Proc. Natl. Acad. Sci. U. S. A* **94,** 4866-4871.

Heyes, D. J.; Ruban, A. V.; Wilks, H. M.; 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.; Ruban, A. V.; Hunter, C. N. (2003). Protochlorophyllide oxidoreductase: "dark" reactions of a light-driven enzyme. *Biochemistry* **42,** 523-528.

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.; Heathcote, P.; Rigby, S. E.; Palacios, M. A.; Van, G. R.; 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.; Menon, B. R.; Sakuma, M.; Scrutton, N. S. (2008). 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.

Hihara, Y.; Kamei, A.; Kanehisa, M.; Kaplan, A.; Ikeuchi, M. (2001). DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *Plant Cell* **13,** 793-806.

Hihara, Y.; Sonoike, K.; Kanehisa, M.; Ikeuchi, M. (2003). DNA microarray analysis of redox-responsive genes in the genome of the cyanobacterium Synechocystis sp. strain PCC 6803. *J Bacteriol* **185,** 1719-1725.

Hinchigeri, S. B.; Hundle, B.; Richards, W. R. (1997). Demonstration that the BchH protein of Rhodobacter capsulatus activates S-adenosyl-L-methionine:magnesium protoporphyrin IX methyltransferase. *FEBS Lett* **407,** 337-342.

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

Hollingshead, S.; Kopecna, J.; Jackson, P. J.; Canniffe, D. P.; Davison, P. A.; Dickman, M. J.; Sobotka, R.; Hunter, C. N. (2012). Conserved chloroplast open-reading frame ycf54 is required for activity of the magnesium protoporphyrin monomethylester oxidative cyclase in Synechocystis PCC 6803. *J Biol Chem* **287,** 27823-27833.

Holtorf, H.; Reinbothe, S.; Reinbothe, C.; Bereza, B.; Apel, K. (1995). Two routes of chlorophyllide synthesis that are differentially regulated by light in barley (Hordeum vulgare L.). *Proc Natl Acad Sci U S A* **92,** 3254-3258.

Hoober, J. K.; Kahn, A.; Ash, D. E.; Gough, S.; Kannangara, C. G. (1988). Biosynthesis of delta-aminolevulinate in greening barley leaves. IX. Structure of the substrate, mode of gabaculine inhibition, and the catalytic mechanism of glutamate 1-semialdehyde aminotransferase. *Carlsberg. Res Commun.* **53,** 11-25.

Hu, G.; Yalpani, N.; Briggs, S. P.; Johal, G. S. (1998). A porphyrin pathway impairment is responsible for the phenotype of a dominant disease lesion mimic mutant of maize. *Plant Cell* **10,** 1095-1105.

Huang, B.; Babcock, H.; Zhuang, X. (2010). Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* **143,** 1047-1058.

Huang, D. D.; Wang, W. Y.; Gough, S. P.; Kannangara, C. G. (1984). delta-Aminolevulinic acid-synthesizing enzymes need an RNA moiety for activity. *Science* **225,** 1482-1484.

Huang, D. D. and Wang, W. Y. (1986). Chlorophyll biosynthesis in Chlamydomonas starts with the formation of glutamyl-tRNA. *J Biol Chem* **261,** 13451-13455.

Huang, Y. S. and Li, H. M. (2009). Arabidopsis CHLI2 can substitute for CHLI1. *Plant Physiol* **150,** 636-645.

Hunter, C. and Coomber, S. (1988). Cloning and oxygen-regulated expression of the bacteriochlorophyll biosynthesis genes *bch E, B, A* and *C* of *Rhodobacter sphaeroides*. *Journal of General Microbiology* **134,** 1491-1497.

Hunter, C., Daldal, F., Thurnauer, C., Beatty, J. T. (2009) *Advances in Photosynthesis and Respiration: The Purple Photosynthetic Bacteria*. Springer.

Igarashi, R. and Seefeldt, L. (2003). Nitrogen fixation: the mechanism of hte Mo-dependent enzyme. *Crit Rev Biochem Mol Biol* **38,** 351-384.

Invitrogen. (2009) XCell II Blot Module. **MAN0000740**. Invitrogen.

Ishikawa, A.; Okamoto, H.; Iwasaki, Y.; Asahi, T. (2001). A deficiency of coproporphyrinogen III oxidase causes lesion formation in Arabidopsis. *Plant J* **27,** 89-99.

Islam, M. R.; Aikawa, S.; Midorikawa, T.; Kashino, Y.; Satoh, K.; 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.

Ito, H.; Yokono, M.; Tanaka, R.; Tanaka, A. (2008). Identification of a novel vinyl reductase gene essential for the biosynthesis of monovinyl chlorophyll in Synechocystis sp. PCC6803. *J Biol Chem* **283,** 9002-9011.

Jacobs, J. M. and Jacobs, N. J. (1987). Oxidation of protoporphyrinogen to protoporphyrin, a step in chlorophyll and haem biosynthesis. Purification and partial characterization of the enzyme from barley organelles. *Biochem J* **244,** 219-224.

Jacobs, J. M. and Jacobs, N. J. (1993). Porphyrin Accumulation and Export by Isolated Barley (Hordeum vulgare) Plastids (Effect of Diphenyl Ether Herbicides). *Plant Physiol* **101,** 1181-1187.

Jaffe, E. K.; Ali, S.; Mitchell, L. W.; Taylor, K. M.; Volin, M.; Markham, G. D. (1995). Characterization of the role of the stimulatory magnesium of Escherichia coli porphobilinogen synthase. *Biochemistry* **34,** 244-251.

Jaffe, E. K. (1999). The porphobilinogen synthase family of metalloenzymes. *Acta Crystallographica Section D* **D56,** 115-128.

Jaffe, E. K. (2004). The porphobilinogen synthase catalyzed reaction mechanism. *Bioorganic Chemistry* **32,** 316-325.

Jahn, D. (1992). Complex formation between glutamyl-tRNA synthetase and glutamyl-tRNA reductase during the tRNA-dependent synthesis of 5-aminolevulinic acid in Chlamydomonas reinhardtii. *FEBS Lett* **314,** 77-80.

Jahn, D.; Verkamp, E.; Soll, D. (1992). Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. *Trends Biochem. Sci.* **17,** 215-218.

Jensen, P. E.; Gibson, L. C.; Henningsen, K. W.; Hunter, C. N. (1996a). 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. D.; Henningsen, K. W.; Hunter, C. N. (1996b). 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. *Journal of Biological Chemistry* **271,** 16662-16667.

Jensen, P. E.; Willows, R. D.; Petersen, B. L.; Vothknecht, U. C.; Stummann, B. M.; Kannangara, C. G.; von, W. D.; Henningsen, K. W. (1996c). Structural genes for Mg-chelatase subunits in barley: Xantha-f, -g and -h. *Mol Gen. Genet.* **250,** 383-394.

Jensen, P. E.; Gibson, L. C.; 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 (Pt 2),** 335-344.

Jensen, P. E.; Gibson, L. C.; 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 (Pt 1),** 127-134.

Johnson, E. T. and Schmidt-Dannert, C. (2008a). Characterization of three homologs of the large subunit of the magnesium chelatase from Chlorobaculum tepidum and interaction with the magnesium protoporphyrin IX methyltransferase. *J Biol Chem* **283,** 27776-27784.

Johnson, E. T. and Schmidt-Dannert, C. (2008b). Characterization of three homologs of the large subunit of the magnesium chelatase from Chlorobaculum tepidum and interaction with the magnesium protoporphyrin IX methyltransferase. *J Biol Chem* **283,** 27776-27784.

Johnson, M. K. (1998). Iron-sulfur proteins: new roles for old clusters. *Curr. Opin Chem Biol* **2,** 173-181.

Jones, D. T.; Taylor, W. R.; Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl Biosci.* **8,** 275-282.

Jones, O. T. G. (1963a). Magnesium 2,4-divinyl phaeoporphyrin a_5 monomethyl ester, a protochlorophyll-like pigment produced by *Rhodopseudomonas sphaeroides* . *Journal of Biochemistry* **89,** 182-189.

Jones, O. T. G. (1963b). The inhibition of bacteriochlorophyll synthesis in *Rhodopseudomonas sphaeroides* by 8-hydroxyquinoline. *Journal of Biochemistry* **88,** 335-343.

Jordan PM (1991) *Biosynthesis of Tetrapyrroles*. Elsevier.

Jordan, P. M. and Seehra, J. S. (1979). The biosynthesis of uroporphyrinogen III: order of assembly of the four porphobilinogen molecules in the formation of the tetrapyrrole ring. *FEBS Lett* **104,** 364-366.

Jordan, P. M.; Warren, M. J.; Williams, H. J.; Stolowich, N. J.; Roessner, C. A.; Grant, S. K.; Scott, A. I. (1988). Identification of a cysteine residue as the binding site for the dipyrromethane cofactor at the active site of

Escherichia coli porphobilinogen deaminase. *FEBS Lett* **235,** 189-193.

Jordan, P. M. (1991). The biosynthesis of 5-aminoleavelunic acid and its transformation into uroporphyrinogen III. *Biosynthesis of Tetrapyrroles* **19,** 1-86.

Kajiwara, S.; Fraser, P. D.; Kondo, K.; Misawa, N. (1997). Expression of an exogenous isopentenyl diphosphate isomerase gene enhances isoprenoid biosynthesis in Escherichia coli. *Biochem. J.* **324 (Pt 2),** 421-426.

Kalb, V. F., Jr. and Bernlohr, R. W. (1977). A new spectrophotometric assay for protein in cell extracts. *Anal Biochem* **82,** 362-371.

Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirosawa, M.; Sugiura, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3,** 109-136.

Karger, G. A.; Reid, J. D.; 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.

Karr, S. R. and Dailey, H. A. (1988a). The synthesis of murine ferrochelatase in vitro and in vivo. *Biochem J* **254,** 799-803.

Karr, S. R. and Dailey, H. A. (1988b). The synthesis of murine ferrochelatase in vitro and in vivo. *Biochem J* **254,** 799-803.

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

Kauss, D.; Bischof, S.; Steiner, S.; Apel, K.; Meskauskiene, R. (2012a). FLU, a negative feedback regulator of tetrapyrrole biosynthesis, is physically linked to the final steps of the Mg(++)-branch of this pathway. *FEBS Lett* **586,** 211-216.

Kauss, D.; Bischof, S.; Steiner, S.; Apel, K.; Meskauskiene, R. (2012b). FLU, a negative feedback regulator of tetrapyrrole biosynthesis, is physically linked to the final steps of the Mg(++)-branch of this pathway. *FEBS Lett* **586,** 211-216.

Ke, S. H. and Madison, E. L. (1997). Rapid and efficient site-directed mutagenesis by single-tube 'megaprimer' PCR method. *Nucleic Acids Res* **25,** 3371-3372.

Keller, Y.; Bouvier, F.; d'Harlingue, A.; Camara, B. (1998). Metabolic compartmentation of plastid prenyllipid biosynthesis--evidence for the involvement of a multifunctional geranylgeranyl reductase. *Eur. J. Biochem.* **251,** 413-417.

Keppen, O. I.; Tourova, T. P.; Kuznetsov, B. B.; Ivanovsky, R. N.; Gorlenko, V. M. (2000). Proposal of Oscillochloridaceae fam. nov. on the basis of a phylogenetic analysis of the filamentous anoxygenic phototrophic bacteria, and emended description of Oscillochloris and Oscillochloris trichoides in comparison with further new isolates. *Int J Syst Evol Microbiol* **50 Pt 4,** 1529-1537.

Kern, D. and Lapointe, J. (1979a). Glutamyl transfer ribonucleic acid synthetase of Escherichia coli. Effect of alteration of the 5-(methylaminomethyl)-2-thiouridine in the anticodon of glutamic acid transfer ribonucleic acid on the catalytic mechanism. *Biochemistry* **18,** 5819-5826.

Kern, D. and Lapointe, J. (1979b). Glutamyl transfer ribonucleic acid synthetase of Escherichia coli. Study of the interactions with its substrates. *Biochemistry* **18,** 5809-5818.

Khrouchtchova, A.; Hansson, M.; Paakkarinen, V.; Vainonen, J. P.; Zhang, S.; Jensen, P. E.; Scheller, H. V.; Vener, A. V.; Aro, E. M.; Haldrup, A. (2005). A previously found thylakoid membrane protein of 14kDa (TMP14) is a novel subunit of plant photosystem I and is designated PSI-P. *FEBS Lett* **579,** 4808-4812.

Kiang, N. Y.; Siefert, J.; Govindjee; Blankenship, R. E. (2007). Spectral signatures of photosynthesis. I. Review of Earth organisms. *Astrobiology.* **7,** 222-251.

Kim, J.; Eichacker, L. A.; Rudiger, W.; Mullet, J. E. (1994). Chlorophyll regulates accumulation of the plastid-encoded chlorophyll proteins P700 and D1 by increasing apoprotein stability. *Plant Physiol* **104,** 907-916.

Kimble, L. K.; Stevenson, A. K.; Madigan, M. T. (1994). Chemotrophic growth of heliobacteria in darkness. *FEMS Microbiol Lett* **115,** 51-55.

Kimble-Long, L. K. and Madigan, M. T. (2001). Molecular evidence that the capacity for endosporulation is universal among phototrophic heliobacteria. *FEMS Microbiol Lett* **199,** 191-195.

Kirk, J. and Tilney-Basset, R. (1978) *The Plastids: Their Chemistry, Structure, Growth and Inheritance*. Elsevier.

Klappenbach, J. A. and Pierson, B. K. (2004). Phylogenetic and physiological characterization of a filamentous anoxygenic photoautotrophic bacterium ' Candidatus Chlorothrix halophila' gen. nov., sp. nov., recovered from hypersaline microbial mats. *Arch Microbiol* **181,** 17-25.

Kleine, T.; Maier, U. G.; Leister, D. (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **60,** 115-138.

Kobayashi, K.; Kondo, M.; Fukuda, H.; Nishimura, M.; Ohta, H. (2007). Galactolipid synthesis in chloroplast inner envelope is essential for proper thylakoid biogenesis, photosynthesis, and embryogenesis. *Proc Natl Acad Sci U S A* **104,** 17216-17221.

Koch, M.; Breithaupt, C.; Kiefersauer, R.; Freigang, J.; Huber, R.; Messerschmidt, A. (2004). Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. *EMBO J* **23,** 1720-1728.

Kohashi, M.; Clement, R. P.; Tse, J.; Piper, W. N. (1984). Rat hepatic uroporphyrinogen III co-synthase. Purification and evidence for a bound folate coenzyme participating in the biosynthesis of uroporphyrinogen III. *Biochem J* **220,** 755-765.

Kohchi, T.; Shirai, H.; Fukuzawa, H.; Sano, T.; Komano, T.; Umesono, K.; Inokuchi, H.; Ozeki, H.; Ohyama, K. (1988). Structure and organization of Marchantia polymorpha chloroplast genome. IV. Inverted repeat and small single copy regions. *J Mol Biol* **203,** 353-372.

Kopecna, J.; Komenda, J.; Bucinska, L.; Sobotka, R. (2012). Long-term acclimation of the cyanobacterium Synechocystis sp. PCC 6803 to high light is accompanied by an enhanced production of chlorophyll that is preferentially channeled to trimeric photosystem I. *Plant Physiol* **160,** 2239-2250.

Kopecna, J.; Sobotka, R.; Komenda, J. (2013). Inhibition of chlorophyll biosynthesis at the protochlorophyllide reduction step results in the parallel depletion of Photosystem I and Photosystem II in the cyanobacterium Synechocystis PCC 6803. *Planta* **237,** 497-508.

Kovacheva, S.; Ryberg, M.; Sundqvist, C. (2000) ADP/ATP and protein phosphorylation dependence of phototransformable protochlorophyllide in isolated etioplast membranes. *Photosynthesis Research* **64**, 127-136.

Kropat, J.; Oster, U.; Rudiger, W.; Beck, C. F. (2000). Chloroplast signalling in the light induction of nuclear HSP70 genes requires the accumulation of chlorophyll precursors and their accessibility to cytoplasm/nucleus. *Plant J* **24,** 523-531.

Kruse, E.; Mock, H. P.; Grimm, B. (1995). Reduction of coproporphyrinogen oxidase level by antisense RNA synthesis leads to deregulated gene expression of plastid proteins and affects the oxidative defense system. *EMBO J* **14,** 3712-3720.

Kushner, J. P.; Barbuto, A. J.; Lee, G. R. (1976). An inherited enzymatic defect in porphyria cutanea tarda: decreased uroporphyrinogen decarboxylase activity. *J Clin. Invest* **58,** 1089-1097.

Labarre, J.; Chauvat, F.; Thuriaux, P. (1989). Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium Synechocystis strain PCC 6803. *J Bacteriol* **171,** 3449-3457.

Labbe-Bois, R. (1990). The ferrochelatase from Saccharomyces cerevisiae. Sequence, disruption, and expression of its structural gene HEM15. *J Biol Chem* **265,** 7278-7283.

Lange, B. M.; Rujan, T.; Martin, W.; Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A* **97,** 13172-13177.

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

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

Layer, G.; Moser, J.; Heinz, D. W.; Jahn, D.; Schubert, W. D. (2003). Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J.* **22,** 6214-6224.

Layer, G.; Reichelt, J.; Jahn, D.; Heinz, D. W. (2010). Structure and function of enzymes in heme biosynthesis. *Protein Sci.* **19,** 1137-1161.

Lee, D. S.; Flachsova, E.; Bodnarova, M.; Demeler, B.; Martasek, P.; Raman, C. S. (2005). Structural basis of hereditary coproporphyria. *Proc Natl Acad Sci U S A* **102,** 14232-14237.

Leister, D. (2003). Chloroplast research in the genomic age. *Trends Genet.* **19,** 47-56.

Lermontova, I.; Kruse, E.; Mock, H. P.; Grimm, B. (1997). Cloning and characterization of a plastidal and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase. *Proc. Natl. Acad. Sci. U. S. A* **94,** 8895-8900.

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

Liljenberg, C. (1974) Characterisation and properties of a protochlorophyllide ester in leaves of dark grown barley with geranylgeraniol as esterifying alcohol. *Plant Physiol* **32**, 208-213.

Liu, H.; Zhang, H.; Niedzwiedzki, D. M.; Prado, M.; He, G.; Gross, M. L.; Blankenship, R. E. (2013). Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* **342,** 1104-1107.

Lopez, A. B.; Yang, Y.; Thannhauser, T. W.; Li, L. (2008). Phytoene desaturase is present in a large protein complex in the plastid membrane. *Physiol Plant* **133,** 190-198.

Louie, G. V.; Brownlie, P. D.; Lambert, R.; Cooper, J. B.; Blundell, T. L.; Wood, S. P.; Warren, M. J.; Woodcock, S. C.; Jordan, P. M. (1992). Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. *Nature* **359,** 33-39.

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.; ShoolinginJordan, P. M. (1996). The three-dimensional structure of *Escherichia coli* porphobilinogen deaminase at 1.76 Å resolution. *Proteins-Structure Function and Genetics* **25,** 48-78.

Luer, C.; Schauer, S.; Mobius, K.; Schulze, J.; Schubert, W. D.; Heinz, D. W.; Jahn, D.; Moser, J. (2005). Complex formation between glutamyl-tRNA reductase and glutamate-1-semialdehyde 2,1-aminomutase in Escherichia coli during the initial reactions of porphyrin biosynthesis. *J Biol Chem* **280,** 18568-18572.

Luo, J. and Lim, C. K. (1993). Order of uroporphyrinogen III decarboxylation on incubation of porphobilinogen and uroporphyrinogen III with erythrocyte uroporphyrinogen decarboxylase. *The Biochemical journal* **289 (Pt 2),** 529-532.

Luque, I. and Ochoa de Alda, J. A. (2013). CURT1,CAAD-containing aaRSs, thylakoid curvature and gene translation. *Trends Plant Sci.*

Maneli, M. H.; Corrigall, A. V.; Klump, H. H.; Davids, L. M.; Kirsch, R. E.; Meissner, P. N. (2003). Kinetic and physical characterisation of recombinant wild-type and mutant human protoporphyrinogen oxidases. *Biochim Biophys Acta* **1650,** 10-21.

Margulis, L. (1970) *Origin of Eeukaryotic Cells*. Yale University Press, New Haven and London, p. 298.

Martin, W. and Schnarrenberger, C. (1997). The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Curr. Genet.* **32,** 1-18.

Martins, B. M.; Grimm, B.; Mock, H. P.; Huber, R.; Messerschmidt, A. (2001). Crystal structure and substrate binding modeling of the uroporphyrinogen-III decarboxylase from *Nicotiana tabacum*. Implications for the catalytic mechanism. *J Biol Chem* **276,** 44108-44116.

Masephol, B. and Forchhammer, K. (2007). Regulatory cascades to express nitrogenases. In: *Biology of the nitrogen cycle* (Bothe, H., Ferguson, S., Newton, V., eds.) Elsevier, Oxford , pp. 131-145.

Masoumi, A.; Heinemann, I. U.; Rohde, M.; Koch, M.; Jahn, M.; Jahn, D. (2008). Complex formation between protoporphyrinogen IX oxidase and ferrochelatase during haem biosynthesis in Thermosynechococcus elongatus. *Microbiology* **154,** 3707-3714.

Masuda, T.; Suzuki, T.; Takamiya, K. (2000). [Metal chelatases involved in chlorophyll and heme biosynthetic pathways in photosynthetic organisms]. *Tanpakushitsu Kakusan Koso* **45,** 700-709.

Masuda, T.; Suzuki, T.; Shimada, H.; Ohta, H.; Takamiya, K. (2003). Subcellular localization of two types of ferrochelatase in cucumber. *Planta* **217,** 602-609.

Masuda, T. and Takamiya, K. (2004). Novel Insights into the Enzymology, Regulation and Physiological Functions of Light-dependent Protochlorophyllide Oxidoreductase in Angiosperms. *Photosynth Res* **81,** 1-29.

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

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

Matringe, M. and Scalla, R. (1988). Studies on the mode of action of acifluorfen-methyl in nonchlorophyllous soybean cells : accumulation of tetrapyrroles. *Plant Physiol* **86,** 619-622.

Matringe, M.; Camadro, J. M.; Labbe, P.; Scalla, R. (1989a). Protoporphyrinogen oxidase as a molecular target for diphenyl ether herbicides. *Biochem J* **260,** 231-235.

Matringe, M.; Camadro, J. M.; Labbe, P.; Scalla, R. (1989b). Protoporphyrinogen oxidase inhibition by three peroxidizing herbicides: oxadiazon, LS 82-556 and M&B 39279. *FEBS Lett* **245,** 35-38.

Matringe, M.; Camadro, J. M.; Labbe, P.; Scalla, R. (1989c). Protoporphyrinogen oxidase as a molecular target for diphenyl ether herbicides. *Biochem J* **260,** 231-235.

Matringe, M.; Camadro, J. M.; Labbe, P.; Scalla, R. (1989d). Protoporphyrinogen oxidase inhibition by three peroxidizing herbicides: oxadiazon, LS 82-556 and M&B 39279. *FEBS Lett* **245,** 35-38.

Matsumoto, F.; Obayashi, T.; Sasaki-Sekimoto, Y.; Ohta, H.; Takamiya, K.; Masuda, T. (2004). Gene expression profiling of the tetrapyrrole metabolic pathway in Arabidopsis with a mini-array system. *Plant Physiol* **135,** 2379-2391.

Matsuzawa, Y.; Kanbe, T.; Suzuki, J.; Hiraishi, A. (2000). Ultrastructure of the acidophilic aerobic photosynthetic bacterium Acidiphilium rubrum. *Curr. Microbiol* **40,** 398-401.

Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* **33,** 491-497.

Mayer, S. M. and Beale, S. I. (1992). Succinyl-Coenzyme A Synthetase and its Role in delta-Aminolevulinic Acid Biosynthesis in Euglena gracilis. *Plant Physiol* **99,** 482-487.

Mayer, S. M.; Rieble, S.; Beale, S. I. (1994). Metal requirements of the enzymes catalyzing conversion of glutamate to delta-aminolevulinic acid in extracts of *Chlorella vulgaris* and *Synechocystis* sp. PCC 6803. *Archives of Biochemistry and Biophysics* **312,** 203-209.

McGlynn, P. and Hunter, C. N. (1993). Genetic analysis of the bchC and bchA genes of Rhodobacter sphaeroides. *Mol. Gen. Genet.* **236,** 227-234.

McLean, S. and Hunter, C. N. (2009). An enzyme-coupled continuous spectrophotometric assay for magnesium protoporphyrin IX methyltransferases. *Anal Biochem* **394,** 223-228.

Medlock, A. E.; Dailey, T. A.; Ross, T. A.; Dailey, H. A.; Lanzilotta, W. N. (2007). A pi-helix switch selective for porphyrin deprotonation and product release in human ferrochelatase. *J Mol Biol* **373,** 1006-1016.

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

Mehta, P. K. and Christen, P. (1994). Homology of 1-aminocyclopropane-1-carboxylate synthase, 8-amino-7-oxononanoate synthase, 2-amino-6-caprolactam racemase, 2,2-dialkylglycine decarboxylase, glutamate-1-semialdehyde 2,1-aminomutase and isopenicillin-N-epimerase with aminotransferases. *Biochem Biophys Res Commun.* **198,** 138-143.

Mereschkowsky, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Zentralbl.* **25,** 593-604.

Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol. Zentralbl.* **30,** 278-367.

Meskauskiene, R.; Nater, M.; Goslings, D.; Kessler, F.; op den, C. R.; Apel, K. (2001a). FLU: a negative regulator of chlorophyll biosynthesis in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **98,** 12826-12831.

Meskauskiene, R.; Nater, M.; Goslings, D.; Kessler, F.; op den, C. R.; Apel, K. (2001b). FLU: a negative regulator of

chlorophyll biosynthesis in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **98,** 12826-12831.

Meskauskiene, R. and Apel, K. (2002). Interaction of FLU, a negative regulator of tetrapyrrole biosynthesis, with the glutamyl-tRNA reductase requires the tetratricopeptide repeat domain of FLU. *FEBS Lett* **532,** 27-30.

Meurer, J.; Plucken, H.; Kowallik, K. V.; Westhoff, P. (1998). A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in Arabidopsis thaliana. *EMBO J* **17,** 5286-5297.

Mimuro, M. and Fujita, Y. (1977). Estimation of chlorophyll a distribution in the photosynthetic pigment systems I and II of the blue-green alga Anabaena variabilis. *Biochim Biophys Acta* **459,** 376-389.

Minamizaki, K.; Mizoguchi, T.; Goto, T.; Tamiaki, H.; Fujita, Y. (2008a). Identification of two homologous genes, chlAI and chlAII, that are differentially involved in isocyclic ring formation of chlorophyll a in the cyanobacterium Synechocystis sp. PCC 6803. *J. Biol. Chem.* **283,** 2684-2692.

Minamizaki, K.; Mizoguchi, T.; Goto, T.; Tamiaki, H.; Fujita, Y. (2008b). Identification of two homologous genes, chlAI and chlAII, that are differentially involved in isocyclic ring formation of chlorophyll a in the cyanobacterium Synechocystis sp. PCC 6803. *J. Biol. Chem.* **283,** 2684-2692.

Mitchell, L. W. and Jaffe, E. K. (1993). Porphobilinogen synthase from Escherichia coli is a Zn(II) metalloenzyme stimulated by Mg(II). *Arch Biochem Biophys* **300,** 169-177.

Mochizuki, N.; Brusslan, J. A.; Larkin, R.; Nagatani, A.; 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.

Mock, H. P. and Grimm, B. (1997). Reduction of Uroporphyrinogen Decarboxylase by Antisense RNA Expression Affects Activities of Other Enzymes Involved in Tetrapyrrole Biosynthesis and Leads to Light-Dependent Necrosis. *Plant Physiol* **113,** 1101-1112.

Mock, H. P.; Keetman, U.; Kruse, E.; Rank, B.; Grimm, B. (1998). Defense responses to tetrapyrrole-induced oxidative stress in transgenic plants with reduced uroporphyrinogen decarboxylase or coproporphyrinogen oxidase activity. *Plant Phisiology* **113,** 1112.

Mohamed, H. E.; van de Meene, A. M.; Roberson, R. W.; Vermaas, W. F. (2005). Myxoxanthophyll is required for normal cell wall structure and thylakoid organization in the cyanobacterium Synechocystis sp. strain PCC 6803. *J Bacteriol* **187,** 6883-6892.

Monaco, T., Weller, S., Ashton, F. (2001) *Prinicples of Weed Science*. John Wiley & Sons.

Moseley, J.; Quinn, J.; Eriksson, M.; Merchant, S. (2000). The Crd1 gene encodes a putative di-iron enzyme required for photosystem I accumulation in copper deficiency and hypoxia in Chlamydomonas reinhardtii. *EMBO J* **19,** 2139-2151.

Moseley, J. L.; Page, M. D.; Alder, N. P.; Eriksson, M.; Quinn, J.; Soto, F.; Theg, S. M.; Hippler, M.; Merchant, S. (2002). Reciprocal expression of two candidate di-iron enzymes affecting photosystem I and light-harvesting complex accumulation. *Plant Cell* **14,** 673-688.

Moser, J.; Lorenz, S.; Hubschwerlen, C.; Rompf, A.; Jahn, D. (1999). Methanopyrus kandleri glutamyl-tRNA reductase. *J Biol Chem* **274,** 30679-30685.

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

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

Muller, A. H. and Hansson, M. (2009). The barley magnesium chelatase 150-kd subunit is not an abscisic acid receptor. *Plant Physiol* **150,** 157-166.

Muller, P.; Li, X. P.; Niyogi, K. K. (2001). Non-photochemical quenching. A response to excess light energy. *Plant Physiol* **125,** 1558-1566.

Mullet, J. E.; Klein, P. G.; Klein, R. R. (1990). Chlorophyll regulates accumulation of the plastid-encoded chlorophyll apoproteins CP43 and D1 by increasing apoprotein stability. *Proc Natl Acad Sci U S A* **87,** 4038-4042.

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

Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol Crystallogr.* **53,** 240-255.

Nagata, N.; Tanaka, R.; Satoh, S.; Tanaka, A. (2005a). 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.

Nagata, N.; Tanaka, R.; Satoh, S.; Tanaka, A. (2005b). 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.

Nakanishi, H.; Nozue, H.; Suzuki, K.; Kaneko, Y.; Taguchi, G.; Hayashida, N. (2005). Characterization of the Arabidopsis thaliana mutant pcb2 which accumulates divinyl chlorophylls. *Plant Cell Physiol* **46,** 467-473.

Nasrulhaq-Boyce, A.; Griffiths, W. T.; Jones, O. T. (1987a). The use of continuous assays to characterize the oxidative cyclase that synthesizes the chlorophyll isocyclic ring. *Biochem J* **243,** 23-29.

Nasrulhaq-Boyce, A.; Griffiths, W. T.; Jones, O. T. G. (1987b). The use of continuous assays to characterize the oxidative cyclase that synthesizes the chlorophyll isocyclic ring. *Biochem J* **243,** 23-29.

Naylor, G. W.; Addlesee, H. A.; Gibson, L. C. D.; Hunter, C. N. (1999). The photosynthesis gene cluster of *Rhodobacter sphaeroides*. *Photosynthesis Research* **62,** 121-139.

Neuwald, A. F.; Aravind, L.; Spouge, J. L.; 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.

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

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

Nomata, J.; Swem, L. R.; Bauer, C. E.; Fujita, Y. (2005). Overexpression and characterization of dark-operative protochlorophyllide reductase from Rhodobacter capsulatus. *Biochim Biophys Acta* **1708,** 229-237.

Nomata, J.; Kitashima, M.; Inoue, K.; Fujita, Y. (2006). Nitrogenase Fe protein-like Fe-S cluster is conserved in L-protein (BchL) of dark-operative protochlorophyllide reductase from Rhodobacter capsulatus. *FEBS Lett* **580,** 6151-6154.

Nomata, J.; Ogawa, T.; Kitashima, M.; Inoue, K.; Fujita, Y. (2008). NB-protein (BchN-BchB) of dark-operative protochlorophyllide reductase is the catalytic component containing oxygen-tolerant Fe-S clusters. *FEBS Lett* **582,** 1346-1350.

Nugent, J. M. and Palmer, J. D. (1988). Location, identity, amount and serial entry of chloroplast DNA sequences in crucifer mitochondrial DNAs. *Curr. Genet.* **14,** 501-509.

Nureki, O.; Vassylyev, D. G.; Katayanagi, K.; Shimizu, T.; Sekine, S.; Kigawa, T.; Miyazawa, T.; Yokoyama, S.; Morikawa, K. (1995). Architectures of class-defining and specific domains of glutamyl-tRNA synthetase. *Science* **267,** 1958-1965.

Oesterhelt, D. and Stoeckenius, W. (1971). Rhodopsin-like protein from the purple membrane of Halobacterium halobium. *Nat New Biol* **233,** 149-152.

Ogawa, T.; Bovey, F.; Shibata, K. (1975) An intermediate in the phytylation of chlorophyllide a in vivo. *Plant Cell Physiology* **16**, 199-202.

Olson, J. M. and Blankenship, R. E. (2004). Thinking about the evolution of photosynthesis. *Photosynth Res* **80,** 373-386.

Oosawa, N.; Masuda, T.; Awai, K.; Fusada, N.; Shimada, H.; Ohta, H.; Takamiya, K. (2000). Identification and light-induced expression of a novel gene of NADPH-protochlorophyllide oxidoreductase isoform in Arabidopsis thaliana. *FEBS Lett* **474,** 133-136.

Ormerod, J. G.; Kimble, L. K.; Nesbakken, T.; Torgersen, Y. A.; Woese, C. R.; Madigan, M. T. (1996). Heliophilum fasciatum gen. nov. sp. nov. and Heliobacterium gestii sp. nov.: endospore-forming heliobacteria from rice field soils. *Arch Microbiol* **165,** 226-234.

Osanai, T.; Kanesaki, Y.; Nakano, T.; Takahashi, H.; Asayama, M.; Shirai, M.; Kanehisa, M.; Suzuki, I.; Murata, N.; Tanaka, K. (2005). Positive regulation of sugar catabolic pathways in the cyanobacterium Synechocystis sp. PCC 6803 by the group 2 sigma factor sigE. *J Biol Chem* **280,** 30653-30659.

Osanai, T.; Imashimizu, M.; Seki, A.; Sato, S.; Tabata, S.; Imamura, S.; Asayama, M.; Ikeuchi, M.; 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. and Rüdiger, W. (1997). The G4 gene of Arabidopsis thaliana encodes a chlorophyll synthase of etiolated plants. *Bot Acta* **110,** 420-423.

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

Oster, U.; Bauer, C. E.; Rudiger, W. (1997b). Characterization of chlorophyll a and bacteriochlorophyll a synthases by heterologous expression in Escherichia coli. *J. Biol. Chem.* **272,** 9671-9676.

Ouchane, S.; Steunou, A. S.; Picaud, M.; Astier, C. (2004). Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: a strategy adopted to bypass the repressive oxygen control system. *J Biol Chem* **279,** 6385-6394.

Overmann, J. (2001). Diversity and ecology of phototrophic sulfur bacteria. *Microbiology Today* **28,** 116-1.

Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science* **276,** 734-740.

Papenbrock, J.; Grafe, S.; Kruse, E.; Hanel, F.; 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, CHL H and CHL I. *Plant J* **12,** 981-990.

Papenbrock, J.; Mock, H. P.; Tanaka, R.; Kruse, E.; Grimm, B. (2000). Role of magnesium chelatase activity in the early steps of the tetrapyrrole biosynthetic pathway. *Plant Physiol* **122,** 1161-1169.

Pardo, A. D.; Chereskin, B. M.; Castelfranco, P. A.; Franceschi, V. R.; Wezelman, B. E. (1980). ATP requirement for Mg chelatase in developing chloroplasts. *Plant Physiology* **65,** 956-960.

Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551-3567.

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

Peter, E. and Grimm, B. (2009). GUN4 is required for posttranslational control of plant tetrapyrrole biosynthesis. *Mol Plant* **2,** 1198-1210.

Peter, E.; Salinas, A.; Wallner, T.; Jeske, D.; Dienst, D.; Wilde, A.; Grimm, B. (2009). Differential requirement of

two homologous proteins encoded by sll1214 and sll1874 for the reaction of Mg protoporphyrin monomethylester oxidative cyclase under aerobic and micro-oxic growth conditions. *Biochim. Biophys. Acta* **1787,** 1458-1467.

Peters, J.; Fisher, K.; Dean, D. (1995). Nitrogenase Structure and Function: A Biochemical-Genetic Perspective. *Annu Rev Microbiol* **49,** 335-366.

Petricek, M.; Petrickova, K.; Havlicek, L.; Felsberg, J. (2006). Occurrence of two 5-aminolevulinate biosynthetic pathways in Streptomyces nodosus subsp. asukaensis is linked with the production of asukamycin. *J Bacteriol* **188,** 5113-5123.

Pham, P.; Rangarajan, S.; Woodgate, R.; Goodman, M. F. (2001). Roles of DNA polymerases V and II in SOS-induced error-prone and error-free repair in Escherichia coli. *Proc Natl Acad Sci U S A* **98,** 8350-8354.

Phillips, J. D.; Whitby, F. G.; Kushner, J. P.; Hill, C. P. (2003). Structural basis for tetrapyrrole coordination by uroporphyrinogen decarboxylase. *EMBO J* **22,** 6225-6233.

Phillips, J. D.; Whitby, F. G.; Warby, C. A.; Labbe, P.; Yang, C.; Pflugrath, J. W.; Ferrara, J. D.; Robinson, H.; Kushner, J. P.; 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.; Hill, C. P. (2009). Substrate shuttling between active sites of uroporphyrinogen decarboxylase is not required to generate coproporphyrinogen. *J. Mol. Biol.* **389,** 306-314.

Pierson, B. K. and Castenholz, R. W. (1974). A phototrophic gliding filamentous bacterium of hot springs, Chloroflexus aurantiacus, gen. and sp. nov. *Arch Microbiol* **100,** 5-24.

Pinta, V.; Picaud, M.; Reiss-Husson, F.; Astier, C. (2002). Rubrivivax gelatinosus acsF (previously orf358) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. *J. Bacteriol.* **184,** 746-753.

Pontier, D.; Albrieux, C.; Joyard, J.; Lagrange, T.; Block, M. A. (2007). Knock-out of the magnesium protoporphyrin IX methyltransferase gene in Arabidopsis. Effects on chloroplast development and on chloroplast-to-nucleus signaling. *J Biol Chem* **282,** 2297-2304.

Pontoppidan, B. and Kannangara, C. G. (1994). Purification and partial characterisation of barley glutamyl-tRNA(Glu) reductase, the enzyme that directs glutamate to chlorophyll biosynthesis. *Eur J Biochem* **225,** 529-537.

Porra, R. J.; THOMPSON, W. A.; KRIEDEMANN, P. E. (1989). Determination of accurate extinction coefficients and simulataneous equations for assaying chlorophylls *a* and *b* extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta* **975,** 384-394.

Porra, R. J.; Schafer, W.; Gad'on, N.; Katheder, I.; Drews, G.; Scheer, H. (1996). Origin of the two carbonyl oxygens of bacteriochlorophyll a. Demonstration of two different pathways for the formation of ring E in Rhodobacter sphaeroides and Roseobacter denitrificans, and a common hydratase mechanism for 3-acetyl group formation. *Eur. J. Biochem.* **239,** 85-92.

Porra, R. J.; Urzinger, M.; Winkler, J.; Bubenzer, C.; Scheer, H. (1998). Biosynthesis of the 3-acetyl and 13(1)-oxo groups of bacteriochlorophyll a in the facultative aerobic bacterium, Rhodovulum sulfidophilum--the presence of both oxygenase and hydratase pathways for isocyclic ring formation. *Eur J Biochem* **257,** 185-191.

Price, D. C.; Chan, C. X.; Yoon, H. S.; Yang, E. C.; Qiu, H.; Weber, A. P.; Schwacke, R.; Gross, J.; Blouin, N. A.; Lane, C.; Reyes-Prieto, A.; Durnford, D. G.; Neilson, J. A.; Lang, B. F.; Burger, G.; Steiner, J. M.; Loffelhardt, W.; Meuser, J. E.; Posewitz, M. C.; Ball, S.; Arias, M. C.; Henrissat, B.; Coutinho, P. M.; Rensing, S. A.; Symeonidi, A.; Doddapaneni, H.; Green, B. R.; Rajah, V. D.; Boore, J.; Bhattacharya, D. (2012). Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. *Science* **335,** 843-847.

Qian, P.; Marklew, C. J.; Viney, J.; Davison, P. A.; Brindley, A. A.; Soderberg, C.; Al-Karadaghi, S.; Bullough, P. A.; Grossmann, J. G.; Hunter, C. N. (2012a). 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.

Qian, P.; Marklew, C. J.; Viney, J.; Davison, P. A.; Brindley, A. A.; Soderberg, C.; Al-Karadaghi, S.; Bullough, P. A.; Grossmann, J. G.; Hunter, C. N. (2012b). 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.

Race, H. L.; Herrmann, R. G.; Martin, W. (1999). Why have organelles retained genomes? *Trends Genet.* **15,** 364-370.

Radmer, R. J. and Bogorad, L. (1967). (Minus) S-adenosyl-L-methionine-magnesium protoporphyrin methyltransferase, an enzyme in the biosynthetic pathway of chlorophyll in Zea mays. *Plant Physiol* **42,** 463-465.

Raymond, J. (2008). Coloring in the tree of life. *Trends Microbiol* **16,** 41-43.

Rebeiz, C. A.; Wu, S. M.; Kuhadja, M.; Daniell, H.; Perkins, E. J. (1983). Chlorophyll a biosynthetic routes and chlorophyll a chemical heterogeneity in plants. *Mol Cell Biochem* **57,** 97-125.

Rebeiz, C. A.; Parham, R.; Fasoula, D. A.; Ioannides, I. M. (1994). Chlorophyll a biosynthetic heterogeneity. *Ciba Found. Symp.* **180,** 177-189.

Rebeiz, C. A.; Montazer-Zouhoor, A.; Mayasich, J.; Tripathy, B. C.; Wu, S. M.; Rebeiz, C. (2014). Photodyanic herbicides. Recent developements and molecular basis of selectivity. *Critical Reviews in Plant Sciences* **6,** 385-436.

Reid, J. D.; Siebert, C. A.; Bullough, P. A.; 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.

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,** 26893-26899.

Reinbothe, C.; El, B. M.; Buhr, F.; Muraki, N.; Nomata, J.; Kurisu, G.; Fujita, Y.; Reinbothe, S. (2010). Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. *Trends Plant Sci* **15,** 614-624.

Reinbothe, S. and Reinbothe, C. (1996). Regulation of Chlorophyll Biosynthesis in Angiosperms. *Plant Physiol* **111,** 1-7.

Renger, G. and Holzwart, A. R. (2005). The light driven water: plastoquinone oxidoreductase. In: *Primary electron transfer in Photosystem II* (Wydrzynski, T. J. and Satoh, K., eds.) Springer, Dordrecht , pp. 139-175.

Rengstl, B.; Oster, U.; Stengel, A.; Nickelsen, J. (2011). An intermediate membrane subfraction in cyanobacteria is involved in an assembly network for Photosystem II biogenesis. *J Biol Chem* **286,** 21944-21951.

Rhie, G.; Avissar, Y. J.; Beale, S. I. (1996). Structure and expression of the Chlorobium vibrioforme hemB gene and characterization of its encoded enzyme, porphobilinogen synthase. *J Biol Chem* **271,** 8176-8182.

Rimington, C. and MILES, P. A. (1951). A study of the porphyrins excreted in the urine by a case of congenital porphyria. *Biochem J* **50,** 202-206.

Rippka R; Deruelles J; Waterbury JB; Herdman M; Stainier RY (1979). Genetic assignments, strain histories and properties of pure cultures of cyanobacteria. *Journal of General Microbiology* **111,** 1-61.

Rogner, M.; Nixon, P. J.; Diner, B. A. (1990). Purification and characterization of photosystem I and photosystem II core complexes from wild-type and phycocyanin-deficient strains of the cyanobacterium Synechocystis PCC 6803. *J Biol Chem* **265,** 6189-6196.

Romeo, P. H.; Raich, N.; Dubart, A.; Beaupain, D.; Pryor, M.; Kushner, J.; Cohen-Solal, M.; Goossens, M. (1986). Molecular cloning and nucleotide sequence of a complete human uroporphyrinogen decarboxylase cDNA. *J Biol Chem* **261,** 9825-9831.

Rowe, J. D. and Griffiths, W. T. (1995). Protochlorophyllide reductase in photosynthetic prokaryotes and its role in chlorophyll synthesis. *Biochem J* **311 (Pt 2),** 417-424.

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

Rzeznicka, K.; Walker, C. J.; Westergren, T.; Kannangara, C. G.; von, W. D.; Merchant, S.; Gough, S. P.; Hansson, M. (2005). Xantha-l encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. *Proc Natl Acad Sci U S A* **102,** 5886-5891.

Sager, R. (1955). Inheritance in the green alga *Chlamydomonas reinhardtii*. *Genetics* **40,** 476-489.

Sarma, R.; Barney, B. M.; Hamilton, T. L.; Jones, A.; Seefeldt, L. C.; Peters, J. W. (2008). Crystal structure of the L protein of Rhodobacter sphaeroides light-independent protochlorophyllide reductase with MgADP bound: a homologue of the nitrogenase Fe protein. *Biochemistry* **47,** 13004-13015.

Sato, S.; Shimoda, Y.; Muraki, A.; Kohara, M.; Nakamura, Y.; Tabata, S. (2007). A large-scale protein protein interaction analysis in Synechocystis sp. PCC6803. *DNA Res* **14,** 207-216.

Sawers, R.; Viney, J.; Farmer, P.; Bussey, R.; Olsefski, G.; Anufrikova, K.; Hunter, C. N.; Brutnell, T. (2006). The Maize *Oil Yellow1* (*Oy1*) Gene Encodes the I Subunit of Magnesium Chelatase. *Plant Molecular Biology* **60,** 95-106.

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

Schimmel, P. R. (1979). Understanding the recognition of transfer RNAs by aminoacyl transfer RNA synthetases. *Adv. Enzymol Relat Areas Mol Biol* **49,** 187-222.

Schoch, S.; Hehlein, C.; Rudiger, W. (1980). Influence of Anaerobiosis on Chlorophyll Biosynthesis in Greening Oat Seedlings (Avena sativa L.). *Plant Physiol* **66,** 576-579.

Schon, A.; Krupp, G.; Gough, S.; Berry-Lowe, S.; Kannangara, C. G.; Soll, D. (1986). The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. *Nature* **322,** 281-284.

Schopf, J. W. (1993). Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science* **260,** 640-646.

Schopf, J. W.; Kudryavtsev, A. B.; Agresti, D. G.; Wdowiak, T. J.; Czaja, A. D. (2002). Laser--Raman imagery of Earth's earliest fossils. *Nature* **416,** 73-76.

Schottkowski, M.; Ratke, J.; Oster, U.; Nowaczyk, M.; Nickelsen, J. (2009). Pitt, a novel tetratricopeptide repeat protein involved in light-dependent chlorophyll biosynthesis and thylakoid membrane biogenesis in Synechocystis sp. PCC 6803. *Mol Plant* **2,** 1289-1297.

Schroder, I.; Hederstedt, L.; Kannangara, C. G.; Gough, P. (1992). Glutamyl-tRNA reductase activity in Bacillus subtilis is dependent on the hemA gene product. *Biochem J* **281 (Pt 3),** 843-850.

Schulz, R. and Senger, H. (1993). Protochlorophyllide reductase: a key enzyme in the greening process. In: *Pigment-protein complexes in plastids: Synthesis and assembly* (Sundqvist, C. and Ryberg, M., eds.) Academic press, New York , pp. 179-218.

Schulze, J. O.; Masoumi, A.; Nickel, D.; Jahn, M.; Jahn, D.; Schubert, W. D.; Heinz, D. W. (2006). Crystal structure of a non-discriminating glutamyl-tRNA synthetase. *J Mol Biol* **361,** 888-897.

Scott, A. I.; Clemens, K. R.; Stolowich, N. J.; Santander, P. J.; Gonzalez, M. D.; Roessner, C. A. (1989). Reconstitution of apo-porphobilinogen deaminase: structural changes induced by cofactor binding. *FEBS Lett* **242,** 319-324.

Sekine, S.; Nureki, O.; Shimada, A.; Vassylyev, D. G.; Yokoyama, S. (2001). Structural basis for anticodon recognition by discriminating glutamyl-tRNA synthetase. *Nat Struct Biol* **8,** 203-206.

Sekine, S.; Shichiri, M.; Bernier, S.; Chenevert, R.; Lapointe, J.; Yokoyama, S. (2006). Structural bases of transfer RNA-

dependent amino acid recognition and activation by glutamyl-tRNA synthetase. *Structure* **14,** 1791-1799.

Senior, N. M.; Brocklehurst, K.; Cooper, J. B.; Wood, S. P.; Erskine, P.; Shoolingin-Jordan, P. M.; Thomas, P. G.; Warren, M. J. (1996). Comparative studies on the 5-aminolaevulinic acid dehydratases from Pisum sativum, Escherichia coli and Saccharomyces cerevisiae. *Biochem J* **320 (Pt 2),** 401-412.

Shalygo, N. V.; Mock, H. P.; Averina, N. G.; Grimm, B. (1998). Photodynamic action of uroporphyrin and protochlorophyllide in greening barley leaves treated with cesium chloride. *J Photochem Photobiol. B* **42,** 151-158.

Shen, G.; Eaton-Rye, J. J.; Vermaas, W. F. (1993). Mutation of histidine residues in CP47 leads to destabilization of the photosystem II complex and to impairment of light energy transfer. *Biochemistry* **32,** 5109-5115.

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.; Fan, R. C.; Xu, Y. H.; Zhang, D. P. (2006a). The Mg-chelatase H subunit is an abscisic acid receptor. *Nature* **443,** 823-826.

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.; Fan, R. C.; Xu, Y. H.; Zhang, D. P. (2006b). The Mg-chelatase H subunit is an abscisic acid receptor. *Nat* **443,** 823-826.

Shepherd, M.; Reid, J. D.; Hunter, C. N. (2003). Purification and kinetic characterization of the magnesium protoporphyrin IX methyltransferase from Synechocystis PCC6803. *Biochem J* **371,** 351-360.

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.; Hunter, C. N. (2005a). Kinetic basis for linking the first two enzymes of chlorophyll biosynthesis. *FEBS J* **272,** 4532-4539.

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

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

Shimada, Y.; Tsuchiya, T.; Akimoto, S.; Tomo, T.; Fukuya, M.; Tanaka, K.; Mimuro, M. (2008). Spectral properties of the CP43-deletion mutant of Synechocystis sp. PCC 6803. *Photosynth Res* **98,** 303-314.

Shioi, Y. and Takamiya, K. (1992). Monovinyl and divinyl protochlorophyllide pools in etiolated tissues of higher plants. *Plant Physiol* **100,** 1291-1295.

Shoolingin-Jordan, P. M.; Al-Dbass, A.; McNeill, L. A.; Sarwar, M.; Butler, D. (2003a). Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. *Biochem Soc Trans* **31,** 731-735.

Shoolingin-Jordan, P. M.; Al-Dbass, A.; McNeill, L. A.; Sarwar, M.; Butler, D. (2003b). Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. *Biochem Soc Trans* **31,** 731-735.

Shpilyov, A. V.; Zinchenko, V. V.; Shestakov, S. V.; Grimm, B.; Lokstein, H. (2005). Inactivation of the geranylgeranyl reductase (ChlP) gene in the cyanobacterium Synechocystis sp. PCC 6803. *Biochim. Biophys. Acta* **1706,** 195-203.

Shpilyov, A. V.; Zinchenko, V. V.; Grimm, B.; Lokstein, H. (2012). Chlorophyll a phytylation is required for the stability of photosystems I and II in the cyanobacterium Synechocystis sp. PCC 6803. *Plant J.*

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

Smith, A. G.; Marsh, O.; Elder, G. H. (1993). Investigation of the subcellular location of the tetrapyrrole-biosynthesis enzyme coproporphyrinogen oxidase in higher plants. *Biochem J* **292 (Pt 2),** 503-508.

Smith, B. B. and Rebeiz, C. A. (1977). Chloroplast biogenesis: detection of Mg-protoporphyrin chelatase in vitro. *Archives of Biochemistry and Biophysics* **180,** 178-185.

Smith, C. A.; Suzuki, J. Y.; Bauer, C. E. (1996). Cloning and characterization of the chlorophyll biosynthesis gene chlM from Synechocystis PCC 6803 by complementation of a bacteriochlorophyll biosynthesis mutant of Rhodobacter capsulatus. *Plant Mol Biol* **30,** 1307-1314.

Smith, K. M. and Goff, D. A. (1986). Synthese of some proposed biosynthetic precursors to the isocyclic ring in chlorophyll a. *Journal of Organic Chemistry* **51,** 657-666.

Smith, M. A.; Kannangara, C. G.; Grimm, B. (1992). Glutamate 1-semialdehyde aminotransferase: anomalous enantiomeric reaction and enzyme mechanism. *Biochemistry* **31,** 11249-11254.

Sobotka, R.; Komenda, J.; Bumba, L.; Tichy, M. (2005). Photosystem II assembly in CP47 mutant of Synechocystis sp. PCC 6803 is dependent on the level of chlorophyll precursors regulated by ferrochelatase. *J Biol Chem* **280,** 31595-31602.

Sobotka, R.; McLean, S.; Zuberova, M.; Hunter, C. N.; Tichy, M. (2008a). The C-terminal extension of ferrochelatase is critical for enzyme activity and for functioning of the tetrapyrrole pathway in Synechocystis strain PCC 6803. *J Bacteriol* **190,** 2086-2095.

Sobotka, R.; Duhring, U.; Komenda, J.; Peter, E.; Gardian, Z.; Tichy, M.; Grimm, B.; Wilde, A. (2008b). Importance of the cyanobacterial Gun4 protein for chlorophyll metabolism and assembly of photosynthetic complexes. *J Biol Chem* **283,** 25794-25802.

Sobotka, R.; Tichy, M.; Wilde, A.; Hunter, C. N. (2011). Functional assignments for the carboxyl-terminal domains of the ferrochelatase from Synechocystis PCC 6803: the CAB domain plays a regulatory role, and region II is essential for catalysis. *Plant Physiol* **155,** 1735-1747.

Sobotka, R. (2014). Making proteins green; biosynthesis of chlorophyll-binding proteins in cyanobacteria. *Photosynth Res* **119,** 223-232.

Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res* **29,** 1097-1106.

Soll, J.; Schultz, G.; Rudiger, W.; Benz, J. (1983). Hydrogenation of geranylgeraniol : two pathways exist in spinach chloroplasts. *Plant Physiol* **71,** 849-854.

Soper, T. S. and Manning, J. M. (1982). Inactivation of pyridoxal phosphate enzymes by gabaculine. Correlation with enzymic exchange of beta-protons. *J Biol Chem* **257,** 13930-13936.

Spence, E.; Sarcina, M.; Ray, N.; Moller, S. G.; Mullineaux, C. W.; Robinson, C. (2003). Membrane-specific targeting of green fluorescent protein by the Tat pathway in the cyanobacterium Synechocystis PCC6803. *Mol Microbiol* **48,** 1481-1489.

Spiller, S. C.; Castelfranco, A. M.; Castelfranco, P. A. (1982). Effects of Iron and Oxygen on Chlorophyll Biosynthesis : I. IN VIVO OBSERVATIONS ON IRON AND OXYGEN-DEFICIENT PLANTS. *Plant Physiol* **69,** 107-111.

Srivastava, R.; Pisareva, T.; Norling, B. (2005). Proteomic studies of the thylakoid membrane of Synechocystis sp. PCC 6803. *Proteomics* **5,** 4905-4916.

Stanier, R. Y.; Doudoroff, M.; Kunisawa, R.; Contopoulou, R. (1959). THE ROLE OF ORGANIC SUBSTRATES IN BACTERIAL PHOTOSYNTHESIS. *Proc Natl Acad Sci U S A* **45,** 1246-1260.

Stark, W.; Hawjer, C.; Hart, G.; Philippides, A.; Petersen, P.; Lewis, J.; Lepper, F.; Battersby, A. (1993). Biosynthesis of porphyrins and related macrocycles. Synthesis of a spiro-lactam related to the proposed spiro-intermediate for porphyrin biosynthesis - inhibition of cosynthetase. *Journal of Chemical Society -Perkin Transactions 1* **23,** 2875-2892.

Strand, A.; Asami, T.; Alonso, J.; Ecker, J. R.; Chory, J. (2003). Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature* **421,** 79-83.

Susek, R. E.; Ausubel, F. M.; Chory, J. (1993). Signal transduction mutants of Arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell* **74,** 787-799.

Suzuki, J. Y. and Bauer, C. E. (1995). Altered monovinyl and divinyl protochlorophyllide pools in *bchJ* mutants of *Rhodobacter capsulatus* - possible monovinyl substrate discrimination of light-independent protochlorophyllide reductase. *Journal of Biological Chemistry* **270,** 3732-3740.

Suzuki, J. Y.; Bollivar, D. W.; Bauer, C. E. (1997). Genetic analysis of chlorophyll biosynthesis. *Annu Rev Genet.* **31,** 61-89.

Suzuki, T.; Masuda, T.; Inokuchi, H.; Shimada, H.; Ohta, H.; Takamiya, K. (2000). Overexpression, enzymatic properties and tissue localization of a ferrochelatase of cucumber. *Plant Cell Physiol* **41,** 192-199.

Suzuki, T.; Masuda, T.; Singh, D. P.; Tan, F. C.; Tsuchiya, T.; Shimada, H.; Ohta, H.; Smith, A. G.; Takamiya, K. (2002). Two types of ferrochelatase in photosynthetic and nonphotosynthetic tissues of cucumber: their difference in phylogeny, gene expression, and localization. *J Biol Chem* **277,** 4731-4737.

Tait, G. and Gibson, K. (1961). The enzymic formation of magnesium protoporphyrin monomethyl ester. *Biochim Biophys Acta* **52,** 614-616.

Taketani, S.; Nakahashi, Y.; Osumi, T.; Tokunaga, R. (1990). Molecular cloning, sequencing, and expression of mouse ferrochelatase. *J Biol Chem* **265,** 19377-19380.

Taketani, S.; Inazawa, J.; Nakahashi, Y.; Abe, T.; Tokunaga, R. (1992). Structure of the human ferrochelatase gene. Exon/intron gene organization and location of the gene to chromosome 18. *Eur J Biochem* **205,** 217-222.

Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28,** 2731-2739.

Tang, K. H.; Wen, J.; Li, X.; Blankenship, R. E. (2009). Role of the AcsF protein in Chloroflexus aurantiacus. *J Bacteriol* **191,** 3580-3587.

Thomas, H. (1997). Tansley Review No. 92 Chlorophyll: a symptom and a regulator of plastid development. *New Phytologist* **136,** 163-181.

Tottey, S.; Block, M. A.; Allen, M.; Westergren, T.; Albrieux, C.; Scheller, H. V.; Merchant, S.; Jensen, P. E. (2003a). Arabidopsis CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc Natl Acad Sci U S A* **100,** 16119-16124.

Tottey, S.; Block, M. A.; Allen, M.; Westergren, T.; Albrieux, C.; Scheller, H. V.; Merchant, S.; Jensen, P. E. (2003b). *Arabidopsis* CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc. Natl. Acad. Sci. USA* **100,** 16119-16124.

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

Tripathy, B. C. and Rebeiz, C. A. (1986). Chloroplast biogenesis. Demonstration of the monovinyl and divinyl monocarboxylic routes of chlorophyll biosynthesis in higher plants. *J. Biol. Chem.* **261,** 13556-13564.

Tripathy, B. C. and Rebeiz, C. A. (1988a). Chloroplast Biogenesis 60 : Conversion of Divinyl Protochlorophyllide to Monovinyl Protochlorophyllide in Green(ing) Barley, a Dark Monovinyl/Light Divinyl Plant Species. *Plant Physiol* **87,** 89-94.

Tripathy, B. C. and Rebeiz, C. A. (1988b). Chloroplast Biogenesis 60 : Conversion of Divinyl Protochlorophyllide to Monovinyl Protochlorophyllide in Green(ing) Barley, a Dark Monovinyl/Light Divinyl Plant Species. *Plant Physiol* **87,** 89-94.

Tsinoremas, N. F.; Kutach, A. K.; Strayer, C. A.; Golden, S. S. (1994). Efficient gene transfer in Synechococcus sp. strains PCC 7942 and PCC 6301 by interspecies conjugation and chromosomal recombination. *J Bacteriol* **176,** 6764-6768.

Tsukatani, Y.; Wen, J.; Blankenship, R. E.; Bryant, D. A. (2010). Characterization of the FMO protein from the aerobic chlorophototroph, Candidatus Chloracidobacterium thermophilum. *Photosynth Res* **104,** 201-209.

Van Niel, C. (1962) The present status of the comparative study of photosynthesis. *Annual Review of Plant Physiology* **13**, 1-26.

Vavilin, D. V. and Vermaas, W. F. (2002). Regulation of the tetrapyrrole biosynthetic pathway leading to heme and chlorophyll in plants and cyanobacteria. *Physiol Plant* **115,** 9-24.

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

Vermaas, W. F. (1993). Molecular-Biological Approaches to Analyse Photosystem II Structure and Function. *Annu Rev Plant Physiol Plant Mol Biol* **44,** 457-481.

Vinti, G.; Hills, A.; Campbell, S.; Bowyer, J. R.; Mochizuki, N.; Chory, J.; Lopez-Juez, E. (2000). Interactions between hy1 and gun mutants of Arabidopsis, and their implications for plastid/nuclear signalling. *Plant J* **24,** 883-894.

Voelker, R. and Barkan, A. (1995). Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J* **14,** 3905-3914.

Vothknecht, U. C.; Kannangara, C. G.; von, W. D. (1998). Barley glutamyl tRNAGlu reductase: mutations affecting haem inhibition and enzyme activity. *Phytochemistry* **47,** 513-519.

Wada, K.; Yamaguchi, H.; Harada, J.; Niimi, K.; Osumi, S.; Saga, Y.; Oh-Oka, H.; Tamiaki, H.; Fukuyama, K. (2006). Crystal structures of BchU, a methyltransferase involved in bacteriochlorophyll c biosynthesis, and its complex with S-adenosylhomocysteine: implications for reaction mechanism. *J Mol Biol* **360,** 839-849.

Walker, C. J.; Mansfield, K. E.; Rezzano, I. N.; Hanamoto, C. M.; Smith, K. M.; Castelfranco, P. A. (1988). The magnesium-protoporphyrin IX (oxidative) cyclase system. Studies on the mechanism and specificity of the reaction sequence. *Biochem. J.* **255,** 685-692.

Walker, C. J.; Mansfield, K. E.; Smith, K. M.; Castelfranco, P. A. (1989). Incorporation of atmospheric oxygen into the carbonyl functionality of the protochlorophyllide isocyclic ring. *Biochem. J.* **257,** 599-602.

Walker, C. J. and Weinstein, J. D. (1991a). Further characterization of the magnesium chelatase in isolated developing cucumber chloroplasts. *Plant Physiology* **95,** 1189-1196.

Walker, C. J. and Weinstein, J. D. (1991b). 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.; Castelfranco, P. A.; Whyte, B. J. (1991). Synthesis of divinyl protochlorophyllide. Enzymological properties of the Mg-protoporphyrin IX monomethyl ester oxidative cyclase system. *Biochem. J.* **276 (Pt 3),** 691-697.

Walker, C. J. and Willows, R. D. (1997). Mechanism and regulation of Mg-chelatase. *Biochemical Journal* **327,** 321-333.

Wallner, T. (2012) Funktionelle analyse konservierter gene aus cyanobakterien und chloroplasten. Fachbereich Biologie und Chemie, Universitat Giessen.

Warren, M. J. and Jordan, P. M. (1988a). Investigation into the nature of substrate binding to the dipyrromethane cofactor of Escherichia coli porphobilinogen deaminase. *Biochemistry* **27,** 9020-9030.

Warren, M. J. and Jordan, P. M. (1988b). Investigation into the nature of substrate binding to the dipyrromethane cofactor of Escherichia coli porphobilinogen deaminase. *Biochemistry* **27,** 9020-9030.

Warren, M. J.; Smith, A. G.; Heyes, D. J.; Neil Hunter, C. (2009). Biosynthesis of Chlorophyll and Barteriochlorophyll. In: *Tetrapyrroles* Springer New York, pp. 235-249.

Watanabe, N.; Che, F. S.; Iwano, M.; Takayama, S.; Yoshida, S.; Isogai, A. (2001). Dual targeting of spinach protoporphyrinogen oxidase II to mitochondria and chloroplasts by alternative use of two in-frame initiation codons. *J Biol Chem* **276,** 20474-20481.

Weisblum, B. (1995). Insights into erythromycin action from studies of its activity as inducer of resistance. *Antimicrobial Agents and Chemotherapy* **39,** 797-805.

Wellington, C. L. and Beatty, J. T. (1989). Promoter mapping and nucleotide sequence of the bchC bacteriochlorophyll biosynthesis gene from Rhodobacter capsulatus. *Gene* **83,** 251-261.

Wen, J.; Tsukatani, Y.; Cui, W.; Zhang, H.; Gross, M. L.; Bryant, D. A.; Blankenship, R. E. (2011). Structural model and spectroscopic characteristics of the FMO antenna protein from the aerobic chlorophototroph, Candidatus Chloracidobacterium thermophilum. *Biochim Biophys Acta* **1807,** 157-164.

Wendt, K. U.; Lenhart, A.; Schulz, G. E. (1999). The structure of the membrane protein squalene-hopene cyclase at 2.0 A resolution. *J Mol Biol* **286,** 175-187.

Whitby, F. G.; Phillips, J. D.; Kushner, J. P.; Hill, C. P. (1998). Crystal structure of human uroporphyrinogen decarboxylase. *EMBO J* **17,** 2463-2471.

Whittaker, P. and Danks, S. (1978) *Mitochondria: Structure, Function and Assembly*.

Whyte, B. J. and Castelfranco, P. A. (1993). Further observations on the Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase system. *Biochem. J.* **290 (Pt 2),** 355-359.

Whyte, B. J. and Griffiths, W. T. (1993). 8-vinyl reduction and chlorophyll a biosynthesis in higher plants. *Biochem. J.* **291 (Pt 3),** 939-944.

Wiktorsson, B.; Ryberg, M.; Sundqvist, C. (1996) Aggregation of NADPH-protochlorophyllide oxidoreductase-pigment complexes is favoured by protein phosphorylation. *Plant Physiology and Biochemistry* **34**, 23-34.

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.

Williams, J. (1988). Construction of specific mutations in the photosystem II photosynthetic reaction centre by genetic engineering methods in the cyanobacterium *Synechocystis* 6803. In: *The Cyanobacteria* (Glazer, A. and Packer, L., eds.) , pp. 809-812.

Williams, R. C.; Gingrich, J. C.; Glazer, A. N. (1980). Cyanobacterial phycobilisomes. Particles from Synechocystis 6701 and two pigment mutants. *J Cell Biol* **85,** 558-566.

Willows, R. D.; Gibson, L. C.; Kanangara, C. G.; Hunter, C. N.; von, W. 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.; 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.

Winter, G. (2009). Xia2: an expert system for macromolecular crystallography data reduction. *J Appl Crystallogr.* **43,** 186-190.

Witty, M.; Wallace-Cook, A. D.; Albrecht, H.; Spano, A. J.; Michel, H.; Shabanowitz, J.; Hunt, D. F.; Timko, M. P.; Smith, A. G. (1993). Structure and expression of chloroplast-localized porphobilinogen deaminase from pea (Pisum sativum L.) isolated by redundant polymerase chain reaction. *Plant Physiol* **103,** 139-147.

Woese, C. R.; Debrunner-Vossbrinck, B. A.; Oyaizu, H.; Stackebrandt, E.; Ludwig, W. (1985). Gram-positive bacteria: possible photosynthetic ancestry. *Science* **229,** 762-765.

Wong, Y. S. and Castelfranco, P. A. (1984a). Resolution and Reconstitution of Mg-Protoporphyrin IX Monomethyl Ester (Oxidative) Cyclase, the Enzyme System Responsible for the Formation of the Chlorophyll Isocyclic Ring. *Plant Physiol* **75,** 658-661.

Wong, Y. S. and Castelfranco, P. A. (1984b). Resolution and reconstitution of Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase, the enzyme system responsible for the formation of the chlorophyll isocyclic ring. *Plant Physiol.* **75,** 658-661.

Wong, Y. S. and Castelfranco, P. A. (1985). Properties of the Mg-Protoporphyrin IX Monomethyl Ester (Oxidative) Cyclase System. *Plant Physiol* **79,** 730-733.

Wong, Y. S.; Castelfranco, P. A.; Goff, D. A.; Smith, K. M. (1985). Intermediates in the formation of the chlorophyll isocyclic ring. *Plant Physiol* **79,** 725-729.

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

Wu, F. Q.; Xin, Q.; Cao, Z.; Liu, Z. Q.; Du, S. Y.; Mei, C.; Zhao, C. X.; Wang, X. F.; Shang, Y.; Jiang, T.; Zhang, X. F.; Yan, L.; Zhao, R.; Cui, Z. N.; Liu, R.; Sun, H. L.; Yang, X. L.; Su, Z.; Zhang, D. P. (2009a). The magnesium-chelatase H subunit binds abscisic acid and functions in abscisic acid signaling: new evidence in Arabidopsis. *Plant Physiol* **150,** 1940-1954.

Wu, F. Q.; Xin, Q.; Cao, Z.; Liu, Z. Q.; Du, S. Y.; Mei, C.; Zhao, C. X.; Wang, X. F.; Shang, Y.; Jiang, T.; Zhang, X. F.; Yan, L.; Zhao, R.; Cui, Z. N.; Liu, R.; Sun, H. L.; Yang, X. L.; Su, Z.; Zhang, D. P. (2009b). The Magnesium-Chelatase H Subunit Binds Abscisic Acid and Functions in Abscisic Acid Signaling: New Evidence in Arabidopsis. *Plant Physiology* **150,** 1940-1954.

Xiliang, G.; Shaomin, S.; Chuan, D.; Feng, F.; Wong, M. S. (2005). Comparative study on the inclusion behavior between meso-tetrakis(4-N-ethylpyridiniurmyl)porphyrin and beta-cyclodextrin derivatives. *Spectrochim. Acta A Mol Biomol. Spectrosc.* **61,** 413-418.

Xiong, J.; Fischer, W. M.; Inoue, K.; Nakahara, M.; Bauer, C. E. (2000). Molecular evidence for the early evolution of photosynthesis. *Science* **289,** 1724-1730.

Xiong, J. and Bauer, C. E. (2002). Complex evolution of photosynthesis. *Annu Rev Plant Biol* **53,** 503-521.

Yamamoto, H.; Kurumiya, S.; Ohashi, R.; Fujita, Y. (2009). Oxygen sensitivity of a nitrogenase-like protochlorophyllide reductase from the cyanobacterium Leptolyngbya boryana. *Plant Cell Physiol* **50,** 1663-1673.

Yamanaka, G. and Glazer, D. A. (1980). Dynamic aspects of phycobilisome structure: phycobilisome turnover during nitrogen starvation in Synechococcus sp. *Arch Microbiol* **124,** 39-47.

Yamazaki, S.; Nomata, J.; 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.

Yang, Z. M. and Bauer, C. E. (1990). Rhodobacter capsulatus genes involved in early steps of the bacteriochlorophyll biosynthetic pathway. *J. Bacteriol.* **172,** 5001-5010.

Yaronskaya, E.; Ziemann, V.; Walter, G.; Averina, N.; Borner, T.; Grimm, B. (2003). Metabolic control of the tetrapyrrole biosynthetic pathway for porphyrin distribution in the barley mutant albostrians. *Plant J* **35,** 512-522.

Yoon, H. S.; Hackett, J. D.; Ciniglia, C.; Pinto, G.; Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* **21,** 809-818.

Yu, J.; Wu, Q.; Mao, H.; Zhao, N.; Vermaas, W. F. (1999). Effects of chlorophyll availability on phycobilisomes in Synechocystis sp. PCC 6803. *IUBMB Life* **48,** 625-630.

Yu, Q. B.; Li, G.; Wang, G.; Sun, J. C.; Wang, P. C.; Wang, C.; Mi, H. L.; Ma, W. M.; Cui, J.; Cui, Y. L.; Chong, K.; Li, Y. X.; Li, Y. H.; Zhao, Z.; Shi, T. L.; Yang, Z. N. (2008). Construction of a chloroplast protein interaction network and functional mining of photosynthetic proteins in Arabidopsis thaliana. *Cell Res* **18,** 1007-1019.

Zhou, S.; Sawicki, A.; Willows, R. D.; Luo, M. (2012). C-terminal residues of oryza sativa GUN4 are required for the activation of the ChlH subunit of magnesium chelatase in chlorophyll synthesis. *FEBS Lett* **586,** 205-210.

Zscheile, F. P.; Comar, C. L.; Harris, D. G. (1944). Spectroscopic stability of chlorophylls *a* and *b* and certain analytical considerations. *Plant Physiol* **19,** 627-637.

Zsebo, K. M. and Hearst, J. E. (1984). Genetic-physical mapping of a photosynthetic gene cluster from R. capsulata. *Cell* **37,** 937-947.