Characterisation of Chlorophyll Synthases from Cyanobacteria and Plants

Matthew Stephen Proctor

A thesis submitted for the degree of Doctor of Philosophy

Faculty of Science

Department of Molecular Biology and Biotechnology

The University of Sheffield

September, 2018

Summary

During the process of photosynthesis, oxygenic photosynthetic organisms utilise chlorophyll (Chl) molecules, spatially organised within membrane-associated protein complexes called photosystems, to capture light from the sun and convert it into chemical energy. Chls are tetrapyrrole molecules featuring a fifth ring, a central Mg^{2+} ion and a hydrophobic phytyl tail. The integral membrane protein chlorophyll synthase (ChlG) catalyses the addition of the tail to the chlorin ring. In the model photosynthetic cyanobacterium *Synechocystis*, ChlG forms a protein-pigment complex with high-light inducible proteins C and D (HliC/HliD), photosystem II assembly factor Ycf39, the YidC/Alb3 insertase and pigments zeaxanthin, myxoxanthophyll, β-carotene and Chl. This complex is postulated to act at the interface between Chl biosynthesis and photosystem assembly, coordinating co-translational insertion of *de novo* Chl molecules into Chl-binding proteins in a poorly understood process.

To gain an insight into the ubiquity of the ChlG complex in higher photosynthetic organisms, ChlG genes from a plant and algae were FLAG-tagged and heterologously expressed in *Synechocystis*. The eukaryotic ChlG homologs could complement the function of the native bacterial protein but the isolated enzyme did not associate with HliD or Ycf39, maintaining an association only with YidC. This indicates that the ChlG-YidC/Alb3 association may be evolutionarily conserved in algae and higher plants.

Abolishing the synthesis of zeaxanthin and myxoxanthophyll in *Synechocystis* prevented association of HliD and Ycf39 with ChlG, indicating that these carotenoids mediate formation of the ChlG complex. Selective abolishment of myxoxanthophyll restored binding of these proteins, suggesting that zeaxanthin alone can facilitate the ChlG-HliD-Ycf39 interaction.

Structural investigation by chemical cross-linking revealed sites of interaction between members of the ChlG complex. These were found to be confined to the cytoplasm. The N-terminal domain of ChlG was the only region of the enzyme found to interact with its partner proteins. The results enabled the generation of a model of the ChlG complex.

The N-terminus of ChlG was sequentially truncated to investigate the importance of this domain to formation of the ChlG complex. Four truncations were made, removing 11, 23, 32 and 39 residues from the N-terminus, up until the start of the first predicted transmembrane helix. While the binding of YidC, HliD and Ycf39 was not impeded in any case, the enzyme activity of ChlG reduced as the truncations became larger. ChlGs lacking 32 or more residues were unable to complement the function of the native enzyme *in vivo* and showed significantly reduced activity *in vitro*. The results indicate that the N-terminal domain of ChlG is important for facilitating its catalytic activity.

A method for the rapid generation and *in vitro* testing of point mutations to the function of *Arabidopsis* ChlG was developed. This involved optimising production of ChlG in *E. coli* in addition to developing a method for producing the enzyme's substrate, chlorophyllide. A ChlG model was generated using the crystal structure of a related protein, UbiA, as a template. The optimised methods were used to generate six point mutations, predicted from the model and sequence alignments of ChlG with UbiA to be important for enzyme activity or substrate binding. Three of the mutants were devoid of activity, demonstrating the importance of these residues to enzyme function.

Acknowledgements

Firstly, I would like to thank my supervisor Professor Neil Hunter for giving me the chance to undertake a PhD and for his constant support, encouragement and enthusiasm throughout the four years I've been with the lab.

I owe a great deal to Dr. Andrew Hitchcock who in many ways has acted as a second supervisor throughout the course of my PhD and has been unfailingly generous with his help and support, however stressed and busy he may have been! His dedication to my success has made my time in the lab so much easier and more enjoyable. I could not have done it without him.

Thanks to Dr. Jack Chidgey for all of his help and advice with my various ChlG projects ("just chuck it all in the bin mate"), for showing me how to use numerous pieces of lab equipment, ensuring that Shark *et al*. was published and for taking the time to read my PhD thesis ("bit dry this Shark"). Huge thanks must go to Dr. Phil Jackson without whom Chapter 5 in this thesis would not have been possible. His expertise in mass spectrometry has been invaluable. Thank you also to Dave Farmer for contributing two protein models to this thesis, and to Dr. Dan (Ken) Canniffe for his contributions to Chapter 4.

Thanks to all members of the Hunter Lab, past and present, for making life in the lab a pleasure, in particular to George (J-dog) who has been a constant source of entertainment and irritation. Thanks to Dave Swaino for his protein purification and Star Trek knowledge, Lizzy, Mo and Paul for all of their help with the molecular biology aspects of my PhD. Thanks also to my friends outside the Hunter lab for some great holidays, many trips to the pub and for keeping me sane over the last four years.

Thank you to Dr. Roman Sobotka for warmly welcoming me into his laboratory during my stay in the Czech Republic and his invaluable advice throughout the duration of my PhD. Thanks to members of Roman's lab, in particular Agnes for keeping me busy at the weekends and for introducing me to Czech lager.

Thanks to Mum, Dad and Jon for putting up with my science related chatter and complaints over the years and for keeping me positive when the going got tough.

Finally, I'd like to thank Alexandra for her constant support throughout my PhD, tolerating my working at weekends and for staying awake in the early hours of the morning (twice) to help me to format my thesis with her usual (aggressiveness) enthusiasm.

Table of Contents

List of Figures

Chapter 3:

Chapter 4:

Chapter 5:

Chapter 6:

Chapter 7:

List of Tables

Chapter 2:

Chapter 3:

Chapter 5:

List of Abbreviations

3-vinyl-bacteriochlorophyllide	3V-BChlide
Å	Angström
a.u	Arbitrary units
A225M	AtChlG with alanine 225 substituted by methionine
ADP	Adenosine diphosphate
ALA	Aminolevulinic acid
ALB	Autoinduction LB media
Amp^R	Ampicillin resistance cassette
AP	Allophycocyanin
A. thaliana	*Arabidopsis thaliana*
AtChlG	*Arabidopsis thaliana* chlorophyll synthase
ATP	Adenosine triphosphate
AU	Absorbance units
*bc*1	Cytochrome *bc*1
BChl	Bacteriochlorophyll
BChlide	Bacteriochlorophyllide
BS3	bis-sulfosuccinimidyl-suberate
CAB	Chl *a* binding domain
Chl	Chlorophyll
Chl a_{GG}	Geranylgeranyl-Chlorophyll *a*

ChlG complex	ChlG-YidC-HliD-HliC-Ycf39
chlG-FLAG	*psbAII:: Syn* 6803_*chlG*-3xFLAG
Chlide	Chlorophyllide *a*
CID	Collision induced dissociation
Copro	Coproporphyrin IX;
COR	Chlide *a* reductase
CPO	Coproporphyrinogen III oxidase
Cyclase	Mg-protoporphyrin monomethylester cyclase
cyt b_{559}	Cytochrome b_{559}
Cyt b_6f	Cytochrome b_6f
cyt c_2	Cytochrome c_2
DGDG	Digalactosyldiacylglycerol
DMAPP	Dimethylallyl diphosphate
D-Myx	Deoxy-myxoxanthophyll
DNA	Deoxyribonucleic acid
DSP	Dithiobis[succinimidylpropionate]
DSS	Disuccinimidyl suberate
DSSO	Disuccinimidyl sulfoxide
DTT	Dithiothreitol
DV-Pchlide	Divinyl-protochlorophyllide
E. coli	*Escherichia coli*

Ech	Echinenone
EDTA	Ethylenediaminetetraacetic acid
ELIP	Early light-induced proteins
Ery^R	Erythromycin resistance cassette
ETC	Electron transport chain
FAD	Flavin-adenine dinucleotide
Fd	Ferredoxin
FeCH	Ferrochelatase
FeS	Rieske iron-sulphur cluster
FLAG-*chlG*	*psbAII::*3xFLAG-*Syn* 6803_*chlG*/Δ*chlG*
FLAG-6803	*psbAII::*3xFLAG-*Syn* 6803_*chlG*/Δ*chlG* (Chapter 3)
FLAG-7002	*psbAII::*3xFLAG-*Syn* 7002_*chlG*/Δ*chlG*
FLAG-At	*psbAII::* 3xFLAG-*A. thaliana_chlG*/Δ*chlG*
FLAG-Cr	*psbAII::* 3xFLAG-*C. reinhardtii_chlG*/Δ*chlG*
FMN	Flavin mononucleotide
FNR	Ferredoxin-$NADP^+$ reductase
FPP	Farnesyl pyrophosphate
GG	Geranylgeraniol
GGPPS	Geranylgeranyl pyrophosphate synthase
GGPP	Geranylgeranyl pyrophosphate
GGR	Geranylgeranyl reductase

GSPP	Geranyl thiolopyrophosphate
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HliA/B/C/D	High-light inducible protein A/B/C/D
Hlip	High-light inducible protein
HPLC	High Performance Liquid Chromatography
IBAQ	Intensity based absolute quantification
IMAC	Immobilised metal affinity chromatography
IPP	Isopentenyl diphosphate
IPP	Isoprenylpyrophosphate
IUPAC	International Union of Pure and Applied Chemistry
Kan^R	Kanamycin resistance cassette
Kb	Kilobase
kDa	Kilodalton
L56P	AtChlG leucine 56 substituted by proline
LB	Luria-Bertani
LC-SDA	Succinimidyl 6-(4,4'-azipentanamido)hexanoate
LHCI/LHCII	Light-harvesting complex I/II
LIL	Light-harvesting-like
LPOR/DPOR	Light/dark-NADPH-protochlorophyllide oxidoreductase

MeOH	Methanol
MEP	2C-methyl-D-erythritol 4-phosphate
MgCH	Magnesium chelatase
MGDG	Monogalactosyldiacylglycerol
MgP	Mg-protoporphyrin IX
MgPME	Mg-protoporphyrin IX monomethyl ester
MgPME	Mg-protoporphyrin monomethylester
mRNA	Messenger Ribonucleic acid
MS/MS	Tandem mass spectrometry
MS3	Tandem mass spectrometry in time
MT	Methyl transferase
MV-Chlide	Monovinyl-chlorophyllide
MV-PChlide	Monovinyl protochlorophyllide *a*
Myx	Myxoxanthophyll
N99A	AtChlG with asparagine 99 substituted by alanine
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NanoLC-MS/MS	Nanoscale liquid chromatography coupled to tandem mass spectrometry
NHS	N-hydroxysuccinimide
NTRC	NADPH-dependent thioredoxin reductase C
OCP	Orange carotenoid protein
OEC	Oxygen evolving complex

OHP	One-Helix Proteins
p.s.i	Pounds per square inch
P54F	AtChlG with proline 54 by phenylalanine
PC	Phycocyanin
Pchlide	Protochlorophyllide
PCR	Polymerase Chain Reaction
Pcy	Plastocyanin
PDM	PratA-defined membrane
PE	Phycoerythrin
PG	Phosphatidylglycerol
PHB	*p*-hydroxybenzoate
PLB	Prolamellar body
PM	Cytoplasmic membrane
PMF	Proton motive force
POR-interacting TPR protein	Pitt
PPOX	Protoporphyrinogen oxidase
PPP	Pytyl pyrophosphate
PQ	Plastoquinone
Proto	Protoporphyrin IX;
PSI/PSII	Photosystem I/II
PSM	Peptide spectrum match

Q46E	AtChlG with glutamine 46 substituted by glutamic acid
Q_A/Q_B	Quinone (primary/terminal electron acceptor of PSII)
Q_BH_2	Hydroquinol (reduced plastoquinone)
Rba	*Rhodobacter*
RC	Reaction centre
ROS	Reactive oxygen species
rpm	Revolutions per minute
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosylmethionine
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SQ	Semiquinone
SQDG	Sulfoquinovosyldiacylglycerol
sulfo-LC-SDA	Sulfosuccinimidyl-6-(4,4'-azipentanamido)hexanoate
Syn 7002	*Synechococcus* sp. PCC 7002
Synechocystis	*Synechocystis* sp. PCC 6803
TFA	Trifluoroacetic acid
TM	Thylakoid membrane
TPR	Tetratricopeptide repeat
UbiA	Ubiquinone synthase

UROD	Uroporphyrinogen III decarboxylase
V60Y	AtChlG valine 60 substituted by tyrosine
Vipp1	Vesicle inducing protein in plastids 1
WT	Wild-type
Y3IP1	Ycf3 interacting protein
Zea	Zeaxanthin
Zeo^R	Zeocin-resistance cassette
Z-ISO	ζ-carotene isomerase
β-car	β-carotene
β-DDM `	n-Dodecyl-β-D-maltoside
Δ1-11	*psbAII::*3xFLAG-*Syn* 6803_*chlG*_ Δ1-11
Δ1-11/Δ*chlG*	*psbAII::*3xFLAG-*Syn* 6803_*chlG* _ Δ1-11/ Δ*chlG*
Δ1-23	*psbAII::*3xFLAG-*Syn* 6803_*chlG*_ Δ1-23
Δ1-23/Δ*chlG*	*psbAII::*3xFLAG-*Syn* 6803_*chlG* _ Δ1-23/ Δ*chlG*
Δ1-32	*psbAII::*3xFLAG-*Syn* 6803_*chlG*_ Δ1-32
Δ1-39	*psbAII::*3xFLAG-*Syn* 6803_*chlG*_ Δ1-39
Δ1-45	*psbAII::*3xFLAG-*Syn* 6803_*chlG*_ Δ1-45
Δ1-51	*psbAII::*3xFLAG-*Syn* 6803_*chlG*_ Δ1-51
Δ*crtR*	*psbAII::*3xFLAG-*Syn* 6803_*chlG*/Δ*chlG*/Δ*crtR*
Δ*crtR*/Δ*cruF*	*psbAII::*3xFLAG-*Syn* 6803_*chlG*/Δ*chlG*/Δ*crtR*/Δ*cruF*
Δ*cruF*	*psbAII::*3xFLAG-*Syn* 6803_*chlG*/Δ*chlG*/Δ*cruF*

Chapter 1: Introduction

1.1 Overview of photosynthesis

The sun is the primary energy source for nearly all life on Earth. Photosynthesis is the process by which organisms collectively known as phototrophs, such as plants, algae and photosynthetic bacteria, collect solar radiation and use the energy to drive the biosynthesis of carbohydrates from CO_2. At its simplest, photosynthesis can be represented by the following equation (Van Niel, 1962):

$$6CO_2 + 12H_2O \xrightarrow{hV} C_6H_{12}O_6 + 6O_2 + 6H_2O$$

In the first step, known as the light reactions; H_2A in the equation above is the reducing agent that provides electrons for photosynthetic fixation of CO_2. The photosynthetic apparatus absorbs sunlight to increase the energy state of the electrons which then flow along an electron transport chain (ETC) yielding energy that is harnessed to produce high energy compounds such as ATP and NADPH. In the second step, referred to as the dark reactions, these compounds are used as sources of chemical energy to fix CO_2 into carbohydrates, energy storage molecules, in a process known as the Calvin-Benson cycle.

Photosynthesis consists of a series of complex physical and chemical reactions which must occur in a highly coordinated manner, requiring many protein complexes operating within specialised membrane structures called thylakoid membranes (TM) in cyanobacteria, algae and plants. The mechanisms of photosynthesis vary between organisms but can, in general, be divided into two classes, oxygenic and anoxygenic photosynthesis. Oxygenic photosynthesis occurs in plants, algae and cyanobacteria and utilises H_2O as a reducing agent. The oxidation of H_2O by a photosynthetic protein complex splits the compound into its component molecules, hydrogen and oxygen, and liberates electrons. The electrons flow along an ETC consisting of three protein complexes; photosystem II (PSII), cytochrome b_6f (cyt b_6f) and photosystem I (PSI). The energy of the electrons is used to pump protons (H^+) across the TM and establish a proton motive force (PMF). The PMF is subsequently used by a 4th complex, ATP-synthase, to drive production of ATP. The electrons are used to reduce $NADP^+$ to NADPH (Blankenship, 2014).

Anoxygenic photosynthesis utilises compounds other than H_2O, such as H_2S and nitrites, as the reducing group (Griffin *et al*., 2007). The ETC of photosynthetic bacteria, including purple bacteria, green non-sulphur bacteria, green sulphur bacteria and heliobacteria consists of a RC (RC) complex, cytochrome bc_1 (bc_1) and cytochrome c_2 (cyt c_2). These complexes facilitate light driven cyclic electron transfer in which the electrons are recycled within the system, in contrast to oxygenic photosynthesis where the electrons are used to reduce $NADP^+$. As such, no NAD(P)H is produced directly in anoxygenic photosynthesis; however the energy released by cyclic electron transfer is similarly used to generate a PMF across the photosynthetic membrane which is in turn used to produce ATP, and indirectly NAD(P)H via complex I (NAD(P)H dehydrogenase) by reverse electron flow. Both modes of photosynthesis will be discussed in further detail.

1.2 Model photosynthetic organisms used in this study

1.2.1 Cyanobacteria

The phylum Cyanobacteria, formally blue-green algae, is believed to have evolved approximately 3 to 4 billion years ago and are the first organisms capable of oxygenic photosynthesis (Shih, 2015). Cyanobacteria contributed significantly to the oxygen content of earth’s atmosphere, paving the way for the evolution of large complex organisms that are reliant on aerobic respiration.

The photosynthetic apparatus in cyanobacteria is analogous to that found in plants and algae due to the evolutionary relationship shared between these organisms. The chloroplasts, organelles that house the photosynthetic apparatus in plants and algae, are thought to have originated from cyanobacteria through endosymbiosis between 1 and 2 billion years ago. It is thought that a eukaryotic cell internalised a cyanobacterium which provided its host with food until, over time, it was assimilated by the larger cell (McFadden, 2001). Being the progenitor of the chloroplast, cyanobacteria do not contain chloroplasts of their own. Instead the TM form pairs layered into sheets that follow the periphery of the cell and converge at points near the cytoplasmic membranes as (discussed further in Section 1.5) (Van De Meene *et al*., 2006).

Today the cyanobacteria are responsible for a large part of the photosynthetic productivity of the world's oceans and can be found in almost all environments, including under extreme conditions such as those of Antarctica and in hot springs (Chorus and Bartram, 1999). Cyanobacteria account for 20–30% of Earth's primary photosynthetic productivity and convert solar energy into biomass-stored chemical energy at the rate of 450 TW (Pisciotta *et al*., 2010) which is 0.2% to 0.3% the total energy reaching the earth from the Sun (178,000 TW in total) (Kruse *et al*., 2005).

1.2.2 *Synechocystis* sp. PCC 6803

Synechocystis sp. PCC 6803 (hereafter *Synechocystis*) is a spherical fresh water cyanobacterium 1.5-2 µm in size (Figure 1.1A). S*ynechocystis* is capable of oxygenic photosynthesis, allowing it to grow photoautotrophically, as well as heterotrophic growth in dark conditions when a suitable carbon source, such as sugar, is available. The first photosynthetic organism to have its entire genome sequenced (Ikeuchi, 1996), *Synechocystis* is readily transformable; able to acquire new genes through homologous recombination (Zang *et al*., 2007). This, combined with its ability to grow heterotrophically, makes *Synechocystis* the ideal candidate for studying photosynthesis, pigment synthesis, tocopherol synthesis, carbon metabolism and respiration, among other processes. It is also being researched as a potential phototrophic "cell factory" for the production of renewable biofuels and other valuable chemicals (Yu *et al*., 2013).

1.2.3 Purple photosynthetic bacteria

Purple bacteria are Gram negative anoxygenic phototrophic microorganisms that inhabit both terrestrial and aquatic environments. There are three classes of purple bacteria called aerobic anoxygenic phototrophs, purple sulphur bacteria and purple non-sulphur bacteria. Anoxygenic phototrophs are unable to utilise CO_2 as a carbon source and require oxygen to grow. The latter two classes are able to grow photoautotrophically (fixing CO_2), photoheterotrophically (utilising carbon sources other than CO_2) and heterotrophically (either by fermentation or aerobic respiration). Reductants other than H_2O are required for anoxygenic photosynthesis. Purple sulphur bacteria are able to use H_2S as a reductant whereas as purple non-sulphur bacteria are not, instead relying on H_2 as an electron donor. Purple bacteria synthesise pigments called bacteriochlorophylls (BChl) as their main light absorbing molecules. The synthesis of these pigments is repressed by the abundance of O_2 and so low oxygen conditions are required for photosynthesis (Cohen-Bazire *et al.*, 1957). These conditions are often found in still aquatic environments, such as lakes and ponds. In response to low oxygen conditions, the cytoplasmic membrane of purple bacteria will develop intracytoplasmic membranes, vesicles formed by the invagination of the cytoplasmic membrane at multiple points (Tucker *et al.*, 2010; Zeilstra-Ryalls *et al.*, 1998). The photosynthetic apparatus is assembled within these membranes (Kiley and Kaplan, 1988; Tucker *et al.*, 2010).

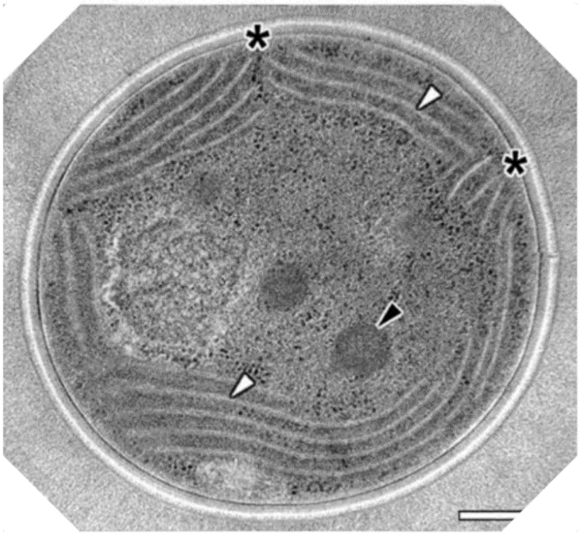
1.2.4 *Rhodobacter sphaeroides*

Rhodobacter (Rba.) sphaeroides is a non-sulphur purple bacterium found in numerous environments including soil, sewage lagoons and anoxic water (Cooper *et al.*, 1975). The bacterium is rod shaped and motile by means of a single sub-polar flagellum (Mackenzie *et al.*, 2007) (Figure 1.1B). *Rba. sphaeroides* grows optimally in anaerobic photoheterotrophic and aerobic chemoheterotrophic conditions. The organic compounds required for either condition act as both a source of carbon and as a reductant for photoheterotrophic and chemoheterotrophic growth (Pfennig, 1978).

CO_2 is used as the sole carbon source under photoautotrophic growth conditions, with H_2 utilised as a reductant (Woese *et al.*, 1984). Decreasing oxygen availability stimulates the switch to photosynthetic growth and promotes induction of photosystem synthesis and assembly. The variations in light intensity found in nature determine the cellular level of intracytoplasmic membrane formation (Adams and Hunter, 2012).

Rba. sphaeroides is an ideal model organism for photosynthesis research. It is able to grow aerobically and is thus not reliant on photosynthesis for survival. Therefore, genes involved in phototrophic growth can be manipulated and mutated without affecting the integrity of the cell, whilst non-photosynthetic growth is maintained. Most of the photosynthesis genes are located within a single gene cluster (Coomber and Hunter, 1989), simplifying discovery of novel photosynthetic genes.

A



B

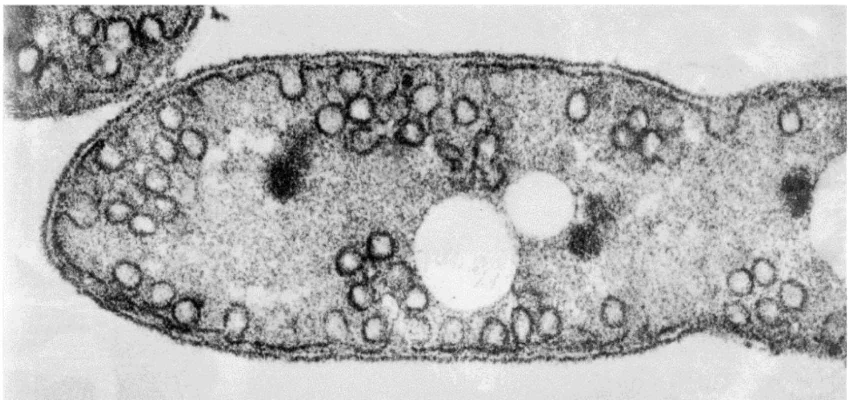


Figure 1.1: Transmission electron micrographs of phototrophic bacteria. Transmission electron micrograph of an ultra-thin section of a *Synechocystis* cell (A) showing TM (white arrows) cytoplasmic inclusions (black arrow) and sites at which the TM and PM converge (asterisk), modified from Allison *et al* (2005). Transmission electron micrograph (B) of a *Rba. sphaeroides* cell showing the invagination of the membrane to from ICM vesicles, taken from: http://www.news.wisc.edu/newsphotos/images/bacterium_microscopic02.jpg.

1.3 Anoxygenic photosynthesis

Anoxygenic photosynthesis requires only one photosystem, which limits the electron transport chain to cyclic electron transfer where the electrons flow between the RC and cytochrome bc_1 (cyt bc_1) via a quinones and cytochrome c_2 (Figure 1.2). Anoxygenic phototophs lack the ability to oxidise H_2O and so the electrons for NADPH and anaerobic respiration are provided by reducing compounds such as H_2S and H_2. Energy from the antenna system passes to a pair of bacteriochlorophyll molecules within the RC, eliciting a charge separation. A nearby pheophytin molecule receives the electron, which is passed to a quinone (Q_A) which becomes a semiquinone radical. A second, loosely bound quinone molecule (Q_B) accepts the electron from Q_A and following a second photochemical event and another series of electron transfers, Q_B binds $2H^+$, becoming a hydroquinol (Q_BH_2). Q_BH_2 diffuses into the TM and docks with cyt bc_1. This complex oxidises Q_BH_2 and the electron released is accepted by cytochrome c_2 (cyt c_2) which can return it to the RC where the cycle repeats. The oxidation of Q_BH_2 also releases $2H^+$, contributing to the acidification of the periplasm and generating a PMF across the photosynthetic membrane. The PMF is utilised by ATP-synthase to produce ATP, analogous to oxygenic photosynthesis. All events described have now been modelled structurally and kinetically (Cartron *et al*., 2014; Sener *et al*., 2016).

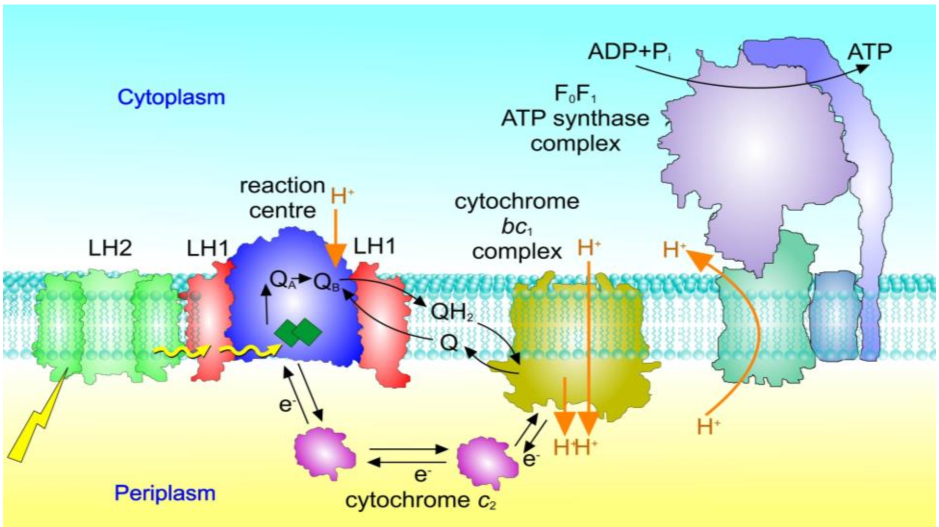


Figure 1.2: Anoxygenic photosynthesis. Schematic diagram showing the flow of electrons (black arrows) around the photosynthetic apparatus of anoxygenic phototrophs. Protons (orange arrows) are transported across the chromatophore membrane from the cytoplasm into the chromatophore lumen.

1.4 Oxygenic photosynthesis

In plants, oxygenic photosynthesis takes place within organelles called chloroplasts, which contain stacks of TM resembling flat hollow discs called grana that house the protein complexes required for photosynthesis. During photosynthesis, protons are pumped into the thylakoid lumen from the thylakoid stroma to establish a pH gradient and generate a PMF across the TM. Cyanobacteria are oxygenic photosynthetic bacteria that contain TM pairs, which form layered sheets parallel to the periphery of the cell and contiguous with the cytoplasmic membrane (PM) (Van De Meene *et al.*, 2006).

The main stages of oxygenic photosynthesis are:

1. Absorption of light by photosystem antennae complexes
2. Reduction of quinone by PSII
3. Reduction of plastocyanin (or cytochrome c_6) by cyt b_6f
4. Oxidation of plastocyanin/cytochrome c_6 by PSI and reduction of $NADP^+$
5. Production of ATP by ATP-synthase

The protein complexes that form the oxygenic photosynthetic apparatus are presented in Figure 1.3. Figure 1.7C shows a Z-scheme, detailing the movement of electrons through these complexes. The details of each of these five steps are described in the following sections.

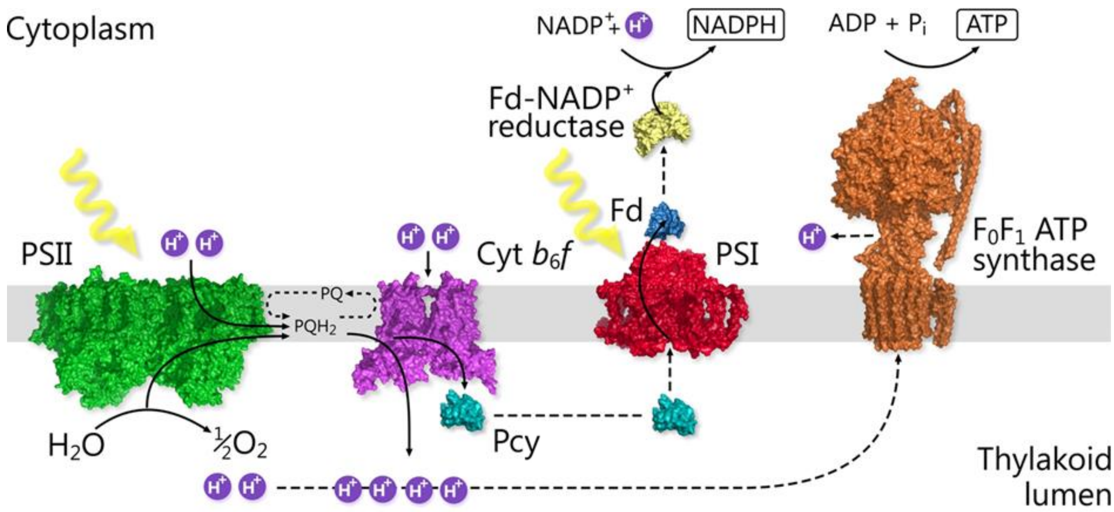


Figure 1.3: Oxygenic photosynthesis. Arrangement of the oxygenic photosynthetic apparatus in the thylakoid membrane. Black arrows show the flow of protons during the light driven chemical reactions of photosynthesis. Taken from MacGregor-Chatwin *et al.* (2017).

1.4.1 Photosystem antenna complexes

The primary process in photosynthesis is the harvesting of light. This task is performed by various light-harvesting complexes that absorb light via bound pigment molecules, the most common of which is chlorophyll (Chl), and channel the energy to the photosystems where it is used to drive the ETC. If every photosystem reaction centre relied on its own photochemically active pigment molecules, photosynthesis would be very inefficient. Due to the low availability of light of the appropriate wavelengths, at any given moment most of the photosystems would be sitting idly, wasting the energy used to synthesise and assemble them to the detriment of the organism (Blankenship, 2014). Structural studies have revealed that there are between 150 and 250 extra Chl antenna molecules associated with each PSII RC (Barros and Kühlbrandt, 2009).

In plants, PSI is associated with the Light-Harvesting complex I (LHC-I) antenna and PSII is associated with Light Harvesting Complex II (LHC-II) (Figure 1.4B), although LHCII is also found with PSI (Pan *et al*., 2018). The proteins that constitute the LHCs come from a superfamily collectively known as Chl *a*/*b*-binding (Cab) proteins (Jansson, 1999; Peter and Thornber, 1991). The polypeptide constituents of LHC-I are Lhca1-6

(Jansson, 1999). LHC-II has been more widely studied and consists mainly of Lhcb1, Lhcb2 and Lhcb3 that bind Chl *a*/*b* and carotenoid pigment molecules. Among the most abundant membrane proteins on earth, the LHC polypeptides are able to spontaneously form trimers of all possible configurations within the thylakoid membrane (Standfuss and Kühlbrandt, 2004). It is estimated that up to eight LHC-II trimers serve each PSII core complex which equates to between 130 and 250 accessory Chl molecules (Wei *et al*., 2016). The LHC-II trimers and PSII core complex form a number of different supercomplexes of varying stoichiometries, the structures of which have been determined by cryo-electron microscopy (Section 1.4.2).

Cyanobacteria do not contain an LHC antenna to harvest light; instead this function is performed by structures called phycobilisomes that form tight rows on the cytoplasmic surface of the TM. Unlike LHCs, the major light-harvesting proteins of cyanobacteria belong to a family of polypeptides called phycobiliproteins, of which three classes form the main component of phycobilisomes. These are phycoerythrin (PE), phycocyanin (PC) and allophycocyanin (AP). The pigment molecules that invariantly bind to these proteins are phycocyanobilin, phycoerythrobilin, phycouroblin and phycobiliviolin. Three (two in some cyanobacteria) hexamers of phycobiliproteins form a cylindrical complex that trimerises with a second cylinder to form the core of the phycobilisome. From the core radiates six rod like structures, each consisting of three hexamers of phycobiliproteins, for a total of six peripheral groups per phycobilisome. Some bacteria are able to change the size of the rod structures from three to four hexamers of phycobiliproteins in response to environmental light intensities (Grossman, 1990) in a process called chromatic adaption. The antenna complex sits above the photosystem RC and channels light energy to it (Figure 1.4A).

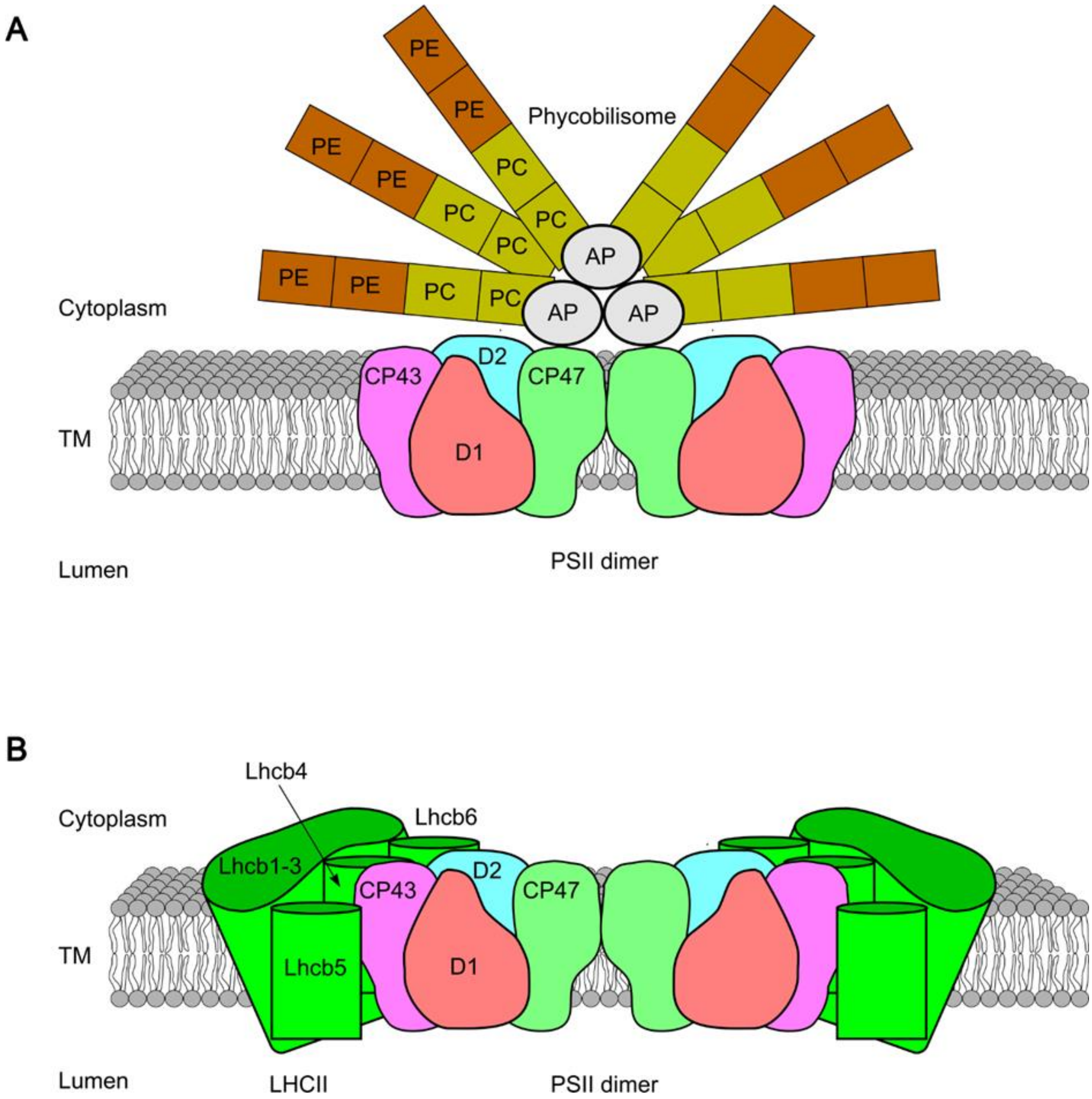


Figure 1.4: Photosystem II antennae complexes in cyanobacteria and plants. PSII dimers associate with phycobilisomes (A) in cyanobacteria which channel light energy to the PSII RC consisting of D1 (red), D2 (blue) and inner light harvesting antennae CP43 (purple) and CP47 (green). The phycobilisome consists of a trimeric core complex containing allophycocyanin (AP) and peripheral rod structures consisting of hexameric phycobiliproteins that contain phycoerythrin (PE) and phycocyanin (PC). In plants, PSII instead associates with light-harvesting complex II (B) consisting of six subunits LHCb1-6 (dark green), which are membrane intrinsic and encompass the PSII dimers, channelling light energy to the RC.

1.4.2 PSII Structure and Function

Photosystem II (PSII) is one of two RC complexes within oxygenic phototrophs and consists of a large multi-subunit enzyme integral to the TM lipid bilayer. PSII can either directly absorb light energy of the appropriate wavelength via its bound Chl molecules, or the antenna complexes discussed above can channel light energy to the PSII RC. Within the RC, a pair of Chl pigments termed P680 becomes electronically excited, making it a strong reducing species (Rappaport *et al*., 2009). Alternatively a Chl situated nearby to P680 called Chl_{D1} can also become electronically excited. Once excited, both species can act as primary electron donors (Acharya *et al*., 2012). To prevent a wasteful recombination reaction occurring, secondary reactions rapidly occur to physically distance the oxidised and reduced species. A nearby molecule of pheophytin (a Chl molecule lacking the central Mg^{2+} ion) accepts the electron from the primary electron donor, physically separating the charge and converting electronic excitation energy into a chemical redox form. Charge separation occurs within a few picoseconds and is the primary reaction of photosynthesis. Electrons lost from the primary Chl electron donor are replaced by the oxidation of H_2O in the oxygen evolving complex (OEC), a sub-complex of PSII. The protons produced by H_2O oxidation are released into the thylakoid lumen. Electrons from primary photochemistry pass from pheophytin to a pair of plastoquinone molecules called Q_A and Q_B respectively. The Q_B molecule is reduced twice by two primary reaction events, binds $2H^+$ (PQH_2) and diffuses into the TM lipid bilayer, allowing an oxidised Q_B to bind to PSII in its place. The PSII ETC is shown in Figure 1.7A.

The crystal structure of PSII complex from cyanobacteria has been solved many times culminating in a structure of PSII from *Thermosynechococcus vulcanus* at 1.9 Å resolution (Umena *et al*., 2011) (Figure 1.5A). Recent advances in single-particle cryo-electron microscopy resulted in the elucidation of the spinach PSII-LHCII complex to 3.2 Å resolution (Wei *et al*., 2016).

PSII exists primarily as a dimer within the TM (Boekema *et al*., 1995; Kouřil *et al*., 2012) although monomeric PSII has been reported (Watanabe *et al*., 2009). In the cyanobacterium *Thermosynechococcus elongatus* it is a large dimeric complex, 105 Å in depth, 205 Å in length, and 110 Å in width (Ferreira *et al*., 2004; Gao *et al*., 2018). Each PSII monomer consists of 20 protein subunits and over 80 cofactors including pigments (35 Chl, 12 β-carotene, 2 haems), pheophytins, plastoquinones, lipids, various ions such as Ca^{2+}, Cl^- and an Mn_4O_5Ca cluster as well as over 1000 bound H_2O molecules (Suga *et al*., 2015; Umena *et al*., 2011).

The PSII protein subunit components differ slightly between plants and cyanobacteria; however the composition and organisation of the core subunits are comparable. Both contain a RC comprising core proteins D1 and D2 and their respective antenna subunits, CP47 and CP43, which facilitate the primary reactions of photosynthesis. D1 and D2 form a heterodimer at the heart of the PSII complex (Nixon *et al*., 2010) and bind six Chl pigments, including the special pair (P680) involved in charge splitting, along with two pheophytins and two plastoquinones (Q_A and Q_B) that facilitate electron transfer through PSII (Barber, 2006; Cardona *et al*., 2012). The D1/D2 core co-ordinates the linear transfer of electrons from P680, through pheophytin, to Q_A and Q_B. CP43 and CP47 bind Chl and β-carotene pigments and act as inner light harvesting complexes, channelling absorbed light energy to P680 as well as connecting the RC with the external light harvesting complexes (Ballottari *et al*., 2012; Barber, 2006).

Surrounding the RC are 14 small subunits: cytochrome b_{559} (cyt b_{559}), PsbE, PsbF, PsbL, PsbM, PsbT, PsbH, PsbI, PsbJ, PsbK, PsbX, PsbY, PsbZ, Psb30. The functions of these are mostly unknown although cyt b_{559}, PsbE and PsbF are known to play a photoprotective role, while PsbL, PsbM and PsbT may facilitate dimerisation of PSII (Pospíšil, 2011; Shi *et al*., 2012).

Oxidation of H_2O is catalysed by the Mn_4O_5Ca cluster which is stabilised and shielded by multiple extrinsic subunits that together form the oxygen evolving complex (OEC). The cyanobacterial extrinsic subunits comprise PsbO, CyanoP, CyanoQ, PsbU and PsbV whereas in terrestrial plants PsbU and PsbV have been lost over the course of evolution (Bricker *et al*., 2012; Thornton *et al*., 2004). Plants and algae have gained two extra subunits, PsbW and PsbR, that are not native to cyanobacteria and contain homologs of CyanoP and CyanoQ named PsbP and PsbQ respectively.

The OEC Mn_4O_5Ca cluster is co-ordinated by ligands provided by the D1 and CP43 subunits (Hwang *et al*., 2007; Service *et al*., 2011). Charge separation in P680 is used to drive oxidation of the Mn_4O_5Ca cluster which, after four sequential oxidation equivalents, is reduced again when it splits two molecules of H_2O into 4 H^+ and O_2 (Grundmeier and Dau, 2012; Renger, 2011; Siegbahn, 2009). The oxidised P680 receives an electron from the redox active Tyrosine Z residue of D1, "resetting" the system (Barber, 2016; Debus *et al*., 1988).

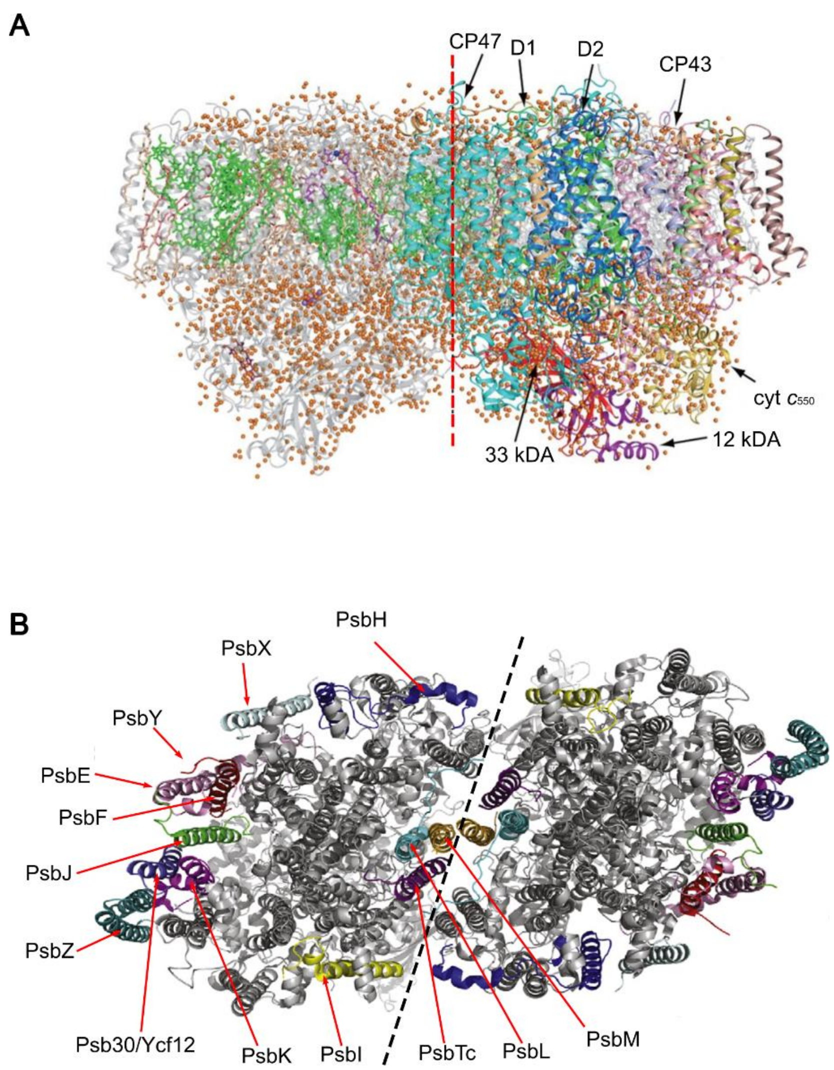


Figure 1.5: Structure of PSII. The structure of PSII. A side view of the complex (A) reveals the core subunits of PSII which are D1 (green), D2 (blue), CP47 (cyan) and CP43 (pink). A top view (B) shows the peripheral subunits which surround the core. Modified from (A) (Umena *et al*., 2011) and (B) (Shi *et al*., 2012).

1.4.3 Cytochrome b_6f

PQH_2 docks with, and donates electrons to, a large dimeric complex known as cytochrome b_6f (cyt b_6f). A complicated series of reactions called the Q cycle unfolds, resulting in the release of $2H^+$ into the thylakoid lumen and reduction of plastocyanin (Pcy) or cytochrome *c* (Baniulis *et al.*, 2011). In brief, PQH_2 binds to the cyt b_6f complex on the luminal side of the complex and donates one electron to a Rieske iron-sulphur cluster (FeS) and a second electron to quinone, reducing it to semiquinone (SQ). The PQH_2 molecule is oxidised and releases $2H^+$ to the thylakoid lumen. The Rieske iron-sulphur cluster reduces Pcy via cytochrome *f* on the luminal side of the TM and the SQ reduces cytochrome *b* on the stromal side. Pcy is released into the thylakoid lumen enabling an oxidised Pcy to bind in its place. This chain of events repeats, reducing cytochrome *b* a second time and enabling it to reduce a PQ which binds $2H^+$ from the stroma and diffuses into the TM where it can join the Q cycle. A complete Q cycle results in the transfer of $4H^+$ to the thylakoid lumen, the oxidation of PQH_2 and reduction of 2Pcy (Figure 1.7D). This is summarised by the following equation:

$$PQH_2 + 2Pcy\ (oxidised) + 2H^+\ (stroma) \rightarrow PQ + 2Pcy\ (reduced) + 4H^+\ (lumen)$$

1.4.4 PSI Structure and Function

PSI is a multimeric pigment protein complex situated within the TM of oxygenic phototrophs. Like PSII, PSI has a RC containing a special pair of Chl molecules denoted P700. Upon absorption of light and channelling to the special pair, charge splitting occurs by a mechanism similar to that in PSII, resulting in the transfer of an electron to a nearby Chl monomer called A_0. From there the electron is transferred to A_1, a phylloquinone in most organisms, and then through three iron sulphur clusters, called $Fe\text{-}S_X$, $Fe\text{-}S_A$ and $Fe\text{-}S_B$, towards the stromal side of the TM where a molecule of ferredoxin (Fd) is bound via electrostatic interaction with PSI. Once Fd has accepted an electron from $Fe\text{-}S_B$, it diffuses from the PSI complex and donates the electron to ferredoxin- $NADP^+$ reductase (FNR), an enzyme which holds the electron via an FAD co-factor until it receives a second electron from another Fd. Upon receiving two electrons, FNR reduces $NADP^+$ to NADPH using the energy it has gained through electron transfer. The Pcy molecules reduced by cyt b_6f bind PSI on the lumenal side of the complex and reduce the PSI RC, essentially "re-setting" the system (Brettel and Leibl, 2001). The PSI ETC is shown in Figure 1.7B.

Unlike PSII, the differences between plant and cyanobacterial PSI complexes are far more distinct. Aside from subunit composition, plant PSI is monomeric and surrounded by antennae light-harvesting complexes (LHCI) (Ben-Shem *et al*., 2003; Pan *et al*., 2018), whereas the cyanobacterial complex is predominantly trimeric (Jordan *et al*., 2001).

The structure of cyanobacteria PSI has been solved numerous times, starting with *Thermosynechococcus elongatus* PSI at 2.5 Å (Jordan *et al*., 2001), and then most recently the *Synechocystis* trimeric PSI at 2.5 Å (Malavath *et al*., 2018) (Figure 1.6B/C). The two structures were broadly comparable with some differences between the prosthetic groups. The PSI trimer in *Synechocystis* consisted of 33 protein subunits and a multitude of co-factors, including 285 molecules of Chl and 72 carotenoids (Malavath *et al*., 2018) (Figure 1.6A/C). The core of each PSI monomer is composed of subunits PsaA and PsaB which form a heterodimer and coordinate the electron transport components P700, A_0, A_1 and Fe-S_x centre (Chitnis, 1996). Surrounding this are 10 other small subunits designated PsaC, D, E, F, I, J, K, L, M and X (Figure 1.6B). PsaC provides ligands to the Fe-Sx centre (Xu *et al*., 1994) and, along with PsaD and PsaE, is involved in binding Fd at the reducing side of PSI (Ruffle *et al*., 2000). PsaF is dispensable in cyanobacteria but is thought to mediate Pcy binding to PSI (Hippler *et al*., 1998, 1999). PsaL has been shown to be essential to the formation of trimeric PSI (Chitnis *et al*., 1993; Malavath *et al*., 2018).

The plant PSI structure has also been solved multiple times; most recently the crystal structure of the pea PSI-LHCI supercomplex was resolved to 2.6 Å (Mazor *et al*., 2015) (Figure 1.6B-C). The core complex is mostly conserved between prokaryotic and eukaryotic phototrophs. Between the eukaryotic phototrophs there are variations in the size and pigment cofactors of the peripheral LHCI antennae complexes (Qin *et al*., 2015). PsaM and PsaX are missing from plant PSI, although it contains 5 additional protein subunits that are absent in cyanobacteria; these are PsaG, H, N, O and P. PsaG and H, which are involved in binding LHCI and LHCII respectively. PsaN facilitates docking of Pcy (Scheller *et al*., 2001); PsaO is thought to be important for facilitating the process of state transitions, a mechanism that maintains the balance of excitation energy between PSI and PSII by migration of LHCs between the two photosystems which enables the plant to adapt to rapid changes in light intensity (Jensen *et al*., 2004); PsaP is the most recently discovered subunit of PSI but its function within the complex is undetermined (Khrouchtchova *et al*., 2005).

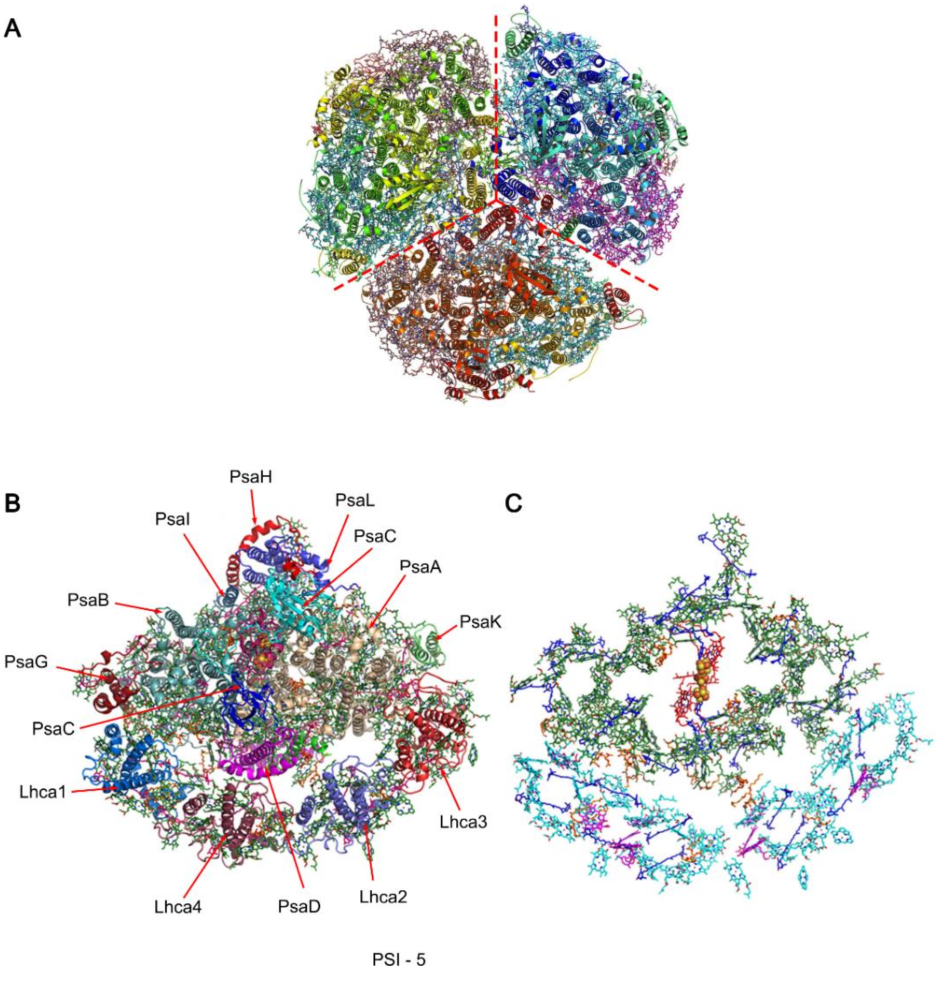


Figure 1.6: Structure of PSI. Structure of a PSI trimer (A) and monomer (B) showing the subunit components of the complex. The Chl content of the complex (C) associated with the PSI core (red), the core antennae (green) and LHCI in (blue). (A) was modified from Malavath *et al* (2018) and (B/C) was modified from Mazor *et al* (2015).

1.4.5 Production of ATP by ATP-synthase

The H^+ released into the thylakoid lumen from the oxidation of H_2O and PQH_2, by PSII and cyt b_6f respectively, results in acidification of the lumen and the generation of a PMF across the TM. ATP-synthase, a large enzyme complex which spans the TM, converts this potential energy into chemical energy that can be used by the cell (Daum *et al*., 2010; Hahn *et al*., 2018). Protons diffuse from the thylakoid lumen to the stroma through the ATP-synthase, which uses the energy to synthesise ATP from ADP (Blankenship, 2014).

The high energy molecules, NADPH generated by the reactions of PSI and ATP by ATP-synthase, feed into the Calvin cycle where they are used to generate energy storing carbohydrates from atmospheric CO_2 (Raines, 2003).

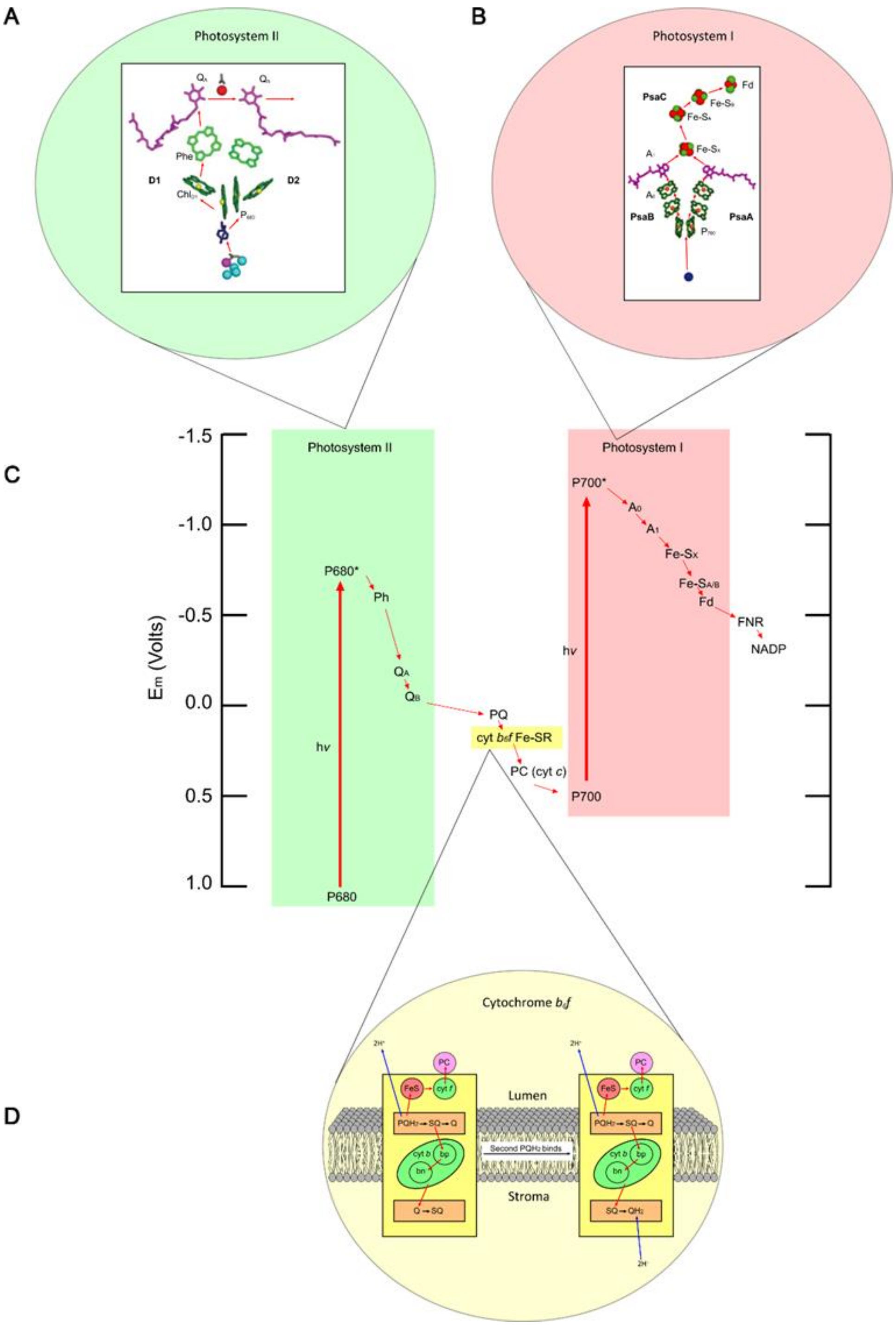


Figure 1.7: Electron transport chain of oxygenic photosynthesis. Light energy absorbed by PSII is channelled to the Chl pair P680 (A) and used for charge splitting. A similar process occurs in PSI whereby the special Chl pair P700 (B) is oxidised upon absorption of light energy, enabling the electron to pass along an ETC to reduce ferredoxin (Fd). The Z-scheme (B) shows the electron transport chain of oxygenic photosynthesis from PSII, through cytochrome *b₆f*, to PSI. Cytochrome *b₆f* (D) facilitates a quinone cycle, resulting in the transport of 4 protons (H^+) into the thylakoid lumen.

1.5 Thylakoid membrane structure and biogenesis

In almost all oxygenic phototrophs the thylakoid membrane (TM) houses the protein-pigment complexes involved in photosynthesis, namely PSI, PSII, cyt b_6f and ATP-synthase as well as the accessory light harvesting antennae mentioned previously. The lipid bilayer of the TM allows the diffusion of the electron carrying proteins of the ETC, quinone and Pcy, and provides the impermeable barrier across which a pH gradient can be established and an electromotive force produced. In cyanobacteria and plants, the lipid content of the TM consists primarily of the galactolipids monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG), as well as glycolipids sulfoquinovosyldiacylglycerol (SQDG) and phospholipids phosphatidylglycerol (PG) (Kobayashi, 2016). MGDG and DGDG make up the bulk of the TM, together accounting for 75% of the total lipid content, and establish a bilayer providing stability to the photosynthetic complexes (Dorne *et al*., 1990). SQDG and PG are acidic lipids and can interact with proteins to aid TM organisation. PG in particular is critical to many photosynthetic processes (Kobayashi, 2016). The ratios of these lipids affect the physical behaviour of the TM under various conditions and so their synthesis is tightly controlled to prevent destabilisation of the photosynthetic apparatus (Demé *et al*., 2014; Moellering and Benning, 2011).

TM biogenesis is a highly complex process, requiring the step-wise assembly of proteins, pigments, lipids, quinones and metal ions in a tightly regulated manner to ensure successful assembly of the photosynthetic unit. Mediation of the various steps requires a large number of dedicated assembly factors (Komenda *et al*., 2012; Schöttler *et al*., 2011). Various models of TM biogenesis in cyanobacteria have been proposed, based on the limited data available (Figure 1.8A-C). The process is thought to originate at specific sites called TM biogenesis centres, defined as the points at which the TM converges with the cytoplasmic membrane (PM) (Van De Meene *et al*., 2006). These resemble cylindrical structures approximately 30 nm in width and 320 nm in length (Kunkel, 1982) and serve as nucleation sites for the integration of the various metabolic pathways involved in photosynthesis, including Chl biosynthesis, protein synthesis and PSII assembly (Nickelsen and Zerges, 2013).

The PSII assembly factor PratA accumulates in semicircular membrane structures that surround the TM biogenesis centres and seem to link both the PM and TM (Stengel *et al*., 2012). These structures are termed PratA-defined membrane (PDM) (Schottkowski *et al*., 2009). PratA features nine tetratricopeptide repeat (TPR) units that have been reported to facilitate interactions between proteins and could serve as a scaffold, aiding the assembly of photosynthetic complexes (Schottkowski *et al*., 2009). PratA is also capable of binding to the core PSII subunit D1 as evidenced by *pratA* knockout studies showing the significant accumulation of pD1 precursor protein and defective PSII assembly. It was postulated that PratA is required for the translocation of PSII proteins from the PDM to the TM where they are further assembled into functional PSII supercomplexes (Schottkowski *et al*., 2009). It has been speculated that, in the early steps of PSII assembly, PratA is also involved in the C-terminal processing of pD1 (Nickelsen *et al*., 2011) and the delivery of Mn^{2+} ions to early PSII intermediates (Nickelsen and Zerges, 2013; Rast *et al*., 2015). Furthermore, the Chl biosynthesis protein NADPH-protochlorophyllide oxidoreductase (POR), the POR-interacting TPR protein (Pitt) and pD1 all localize in higher concentrations within PDMs when PratA expression is perturbed (Nickelsen *et al*., 2011). This highlights the essential role of PratA in the organization of TM biogenesis centres and migration of these proteins from the PDM to the TM.

Vesicle inducing protein in plastids 1 (Vipp1) is another TM biogenesis factor that has been found to be indispensible for TM assembly in some cyanobacteria and plants (Li *et al*., 1994). The protein is capable of assembling into the cylindrical structures mentioned previously (Frain *et al*., 2016) and has been found to be important for the correct assembly of PSI (Stengel *et al*., 2012; Zhang *et al*., 2014). This protein facilitates the budding of vesicles from the inner chloroplast envelope which is an essential process for maintaining the structural and functional integrity of TM (Hennig *et al*., 2015; Kroll *et al*., 2001). Deletion of the *vipp1* homologue in *Synechocystis* resulted in a complete loss of TM formation (Westphal *et al*., 2001); however this was not found to be the case in *Synechococcus* (Zhang *et al*., 2014). In *Synechocystis* Vipp1 exchanges between two TM fractions, one that is concentrated in the high curvature regions of the TM near the periphery of the cell, and one that is evenly distributed. Perturbing this distribution in constant light conditions has no effect on TM biogenesis but if light conditions change the TM does not assemble correctly, indicating a role for Vipp1 in assembly of the complexes involved in the light dependent reactions of photosynthesis (Gutu *et al*., 2018). Vipp1 has also been implicated in the transport of lipids, mediating the transfer of lipids from the PM to the TM (Heidrich *et al*., 2017). In summary, there are many contrasting hypotheses regarding the exact role of Vipp1 in TM biogenesis, and its function(s) remain ambiguous. However, the general consensus is that Vipp1 has a membrane protecting/stabilising function and may have further roles in other photosynthetic processes and/or TM biogenesis (Junglas and Schneider, 2018).

Studies into the composition of the PM in cyanobacteria have shown that Chl-containing photosystem RC intermediates, that are already capable of charge separation, are assembled in the PM. Furthermore PSII and PSI assembly factors have been found to accumulate in the PM (Zak *et al*., 2001). It is therefore feasible that the photosystem RCs are assembled in the PM and then translocated to the TM (Keren *et al*., 2005; Srivastava *et al*., 2006). This translocation possibly occurs through vesicle transport, mediated by Vipp1, or by lateral movement from the PM into the TM biogenesis centres. This is followed by photosystem maturation when antenna complexes and regulatory subunits are incorporated into the complex (Nickelsen *et al*., 2011; Zak *et al*., 2001). Although early intermediates of both photosystems may be localised to the PM, the PDM is dedicated to the assembly of PSII and not PSI (Rast *et al*., 2015). Metal ion co-factors, lipids and Chl are incorporated into the growing PSII complex as it is being assembled, highlighted by the accumulation of PSII assembly factors, Chl biosynthesis enzymes and Chl precursors within the PDM (Rengstl *et al*., 2011; Schottkowski *et al*., 2009).

PSII repair, on the other hand, seems to be spatially separated from *de novo* PSII assembly as exemplified by the even distribution of D1 mRNA (*psbA*) across the TM, along with other factors involved in PSII repair, instead of accumulating in TM biogenesis centres (Komenda *et al*., 2006; Uniacke and Zerges, 2007). This may be to prevent the crossover of repair and assembly processes, reducing the risk of interference between these pathways (Nickelsen and Rengstl, 2013). Little is known about the localisation of the later stages of PSI assembly during TM biogenesis, as is the case for cyt b_6f and ATP-synthase complexes (Rast *et al*., 2015).

TM biogenesis in plants remains poorly understood. There are no structures in plants that resemble the TM biogenesis centres of cyanobacteria. This lack of research may be due to the inherent difficulty in examining the development of cellular processes in multicellular organisms compared with unicellular organisms. As single celled organisms, cyanobacteria divide frequently which inevitably means that new TM is constantly being synthesised within cellular populations. Additionally, the TM of cyanobacteria accounts for a relatively large percentage of their total biomass. In contrast, plants contain multiple differentiated cells and, once synthesised, have inherently stable TM. TM biogenesis in plants occurs during meristem differentiation (Charuvi *et al*., 2012), providing only a small window of opportunity in which to examine the process. None the less, it is generally agreed that synthesis of photosynthetic proteins most likely occurs in the stromal TM as these regions are accessible to ribosomes (Yamamoto *et al*., 1981). In support of this, PSII monomers and earlier PSII intermediates were found largely in the stromal lamellae whilst PSII dimers were confined to the grana stacks, indicating that PSII biogenesis occurs in unconstrained regions of the TM (Danielsson *et al*., 2006).

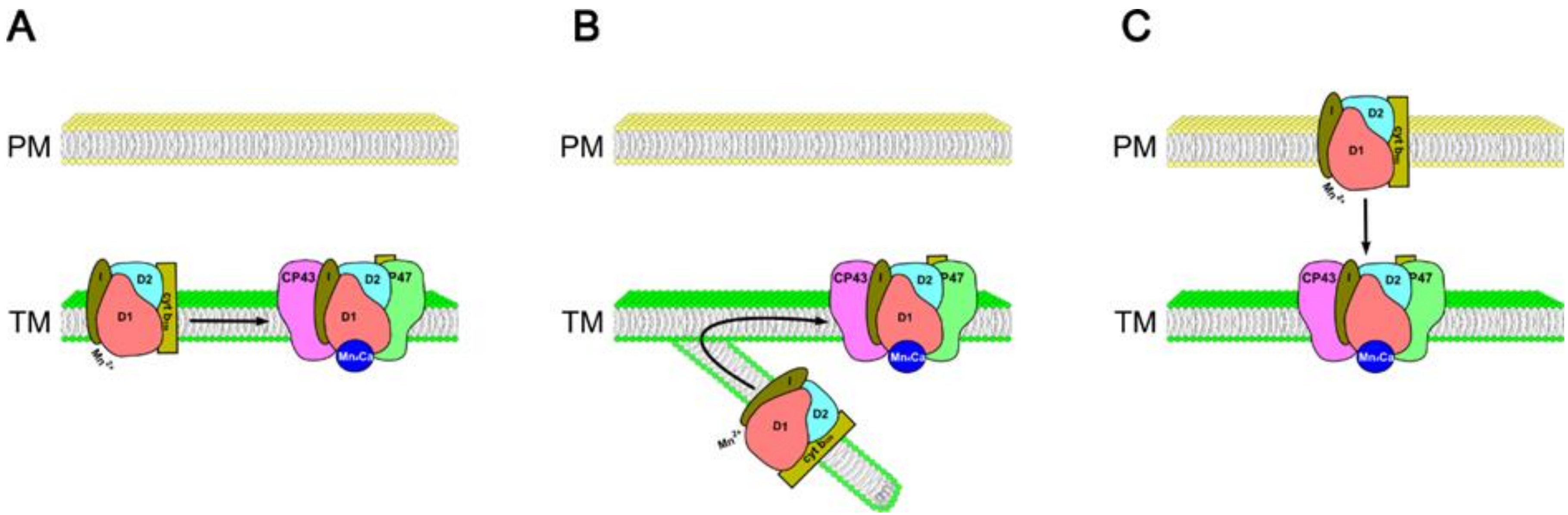


Figure 1.8: Thylakoid biogenesis. The early stages of PSII during TM biogenesis in cyanobacteria may take place entirely within the TM (A) or in specialised regions of the membranes termed TM biogenesis centres (B). Alternatively, assembly of the PSII RC complex may occur in the PM before migrating to the TM to mature (C).

1.6 Photosystem Assembly and Repair

1.6.1 Photosystem II Assembly

PSII polypeptides are synthesised and co-translationally inserted into the thylakoid membrane (Zhang *et al*., 1999) where they are subsequently assembled into multimeric complexes along with their various co-factors. This process is highly coordinated and involves numerous assembly proteins that interact transiently with specific PSII subunits or co-factors to facilitate the production of distinct intermediate complexes as PSII matures (Heinz *et al*., 2016). Although our understanding of this process is incomplete, much progress has been made in elucidating the various stages of PSII assembly in cyanobacteria via a combination of deletion mutants and protein electrophoresis techniques. Figure 1.9 highlights the main stages of PSII assembly. PSII assembly begins with the formation of the D2–cyt *b*559 and pD1-PsbI sub-complexes which merge to form the RC core complex. The transient RC47 complex is then generated by the binding of the CP47 subcomplex which consists of CP47 along with several low molecular mass subunits. Next, monomeric PSII is formed by the incorporation of the CP43 subcomplex into RC47 and the subsequent binding of the extrinsic subunits that protect the OEC, enabling photoactivation of the OEC. Finally, PSII monomers dimerise to form the completed PSII complex (Nickelsen and Rengstl, 2013). Each of the stages of PSII assembly in cyanobacteria will be discussed in more detail.

1.6.1.1 D2–cyt *b*559

PSII assembly is initiated by the accumulation of cyt *b*559 which acts as a nucleation factor, allowing the folding and insertion of D2 followed by the formation of the D2-cyt *b*559 complex (Komenda *et al*., 2004, 2008). There have been no specific assembly factors attributed to formation of the D2-Cyt *b*559 subcomplex, although Slr0286 and Slr2013 have both been suggested to play a role in the folding of D2 despite their inactivation appearing to have no detectable consequences for the cell (Kufryk and Vermaas, 2003).

1.6.1.2 pD1-psbI

The D1 precursor protein (pD1) is inserted into the membrane via the SecYEG apparatus in collaboration with the YidC insertase (Chidgey et al., 2014; Spence et al., 2004) which supports the folding and integration of the protein into the lipid bilayer (Ossenbühl *et al.*, 2004, 2006). It is believed that Chl is inserted into pD1 co-translationally, as pD1 is being integrated into the membrane. Chl may be delivered to pD1 by the nearby chlorophyll synthase complex (see Section 1.10 and 1.11), comprising chlorophyll synthase (ChlG) which is the final enzyme in the Chl biosynthesis pathway, the Chl binding protein HliD, YidC and the PSII assembly factor Ycf39 (Chidgey et al., 2014). It is hypothesized that Ycf39 localises ChlG to pD1 before chlorophyll is delivered to the protein via HliD as pD1 is being co-translationally inserted into the plasma membrane by YidC (Chidgey et al., 2014; Knoppová et al., 2014). Despite this, the handover of Chl from the chlorophyll synthase complex to the PSII assembly aparatus remains poorly understood. The structural and functional characterisation of the ChlG complex in order to better understand this process is the focus of the work presented in this thesis. Once pD1 is inserted into the membrane it can bind to the PsbI subunit forming the pD1-PsbI subcomplex (Dobáková *et al*., 2007).

1.6.1.3 Reaction-centre complex (RC)

D2–cyt b559 can act as a platform to which pD1-psbI binds and forms the RC complex (Dobáková *et al*., 2007). The binding of pD1-psbI to D2-cyt b559 requires the C-terminal processing of pD1 and coordination by several auxiliary proteins.

As discussed previously, the thylakoid biogenesis factor PratA can interact directly with the α-helical C-terminus of pD1. This interaction was mapped to residues 314–328 (Schottkowski *et al*., 2009) that are involved in stabilising the Mn_4CaO_5 cluster of the

OEC (Umena *et al*., 2011). A study showed that PratA can bind Mn^{2+} with high affinity and that delivery of Mn^{2+} to PSII was perturbed in mutants lacking PratA (Stengel *et al*., 2012). The available evidence suggests that PratA loads pD1 with the Mn^{2+} required to form the Mn_4CaO_5 cluster (Klinkert *et al*., 2004; Schottkowski *et al*., 2009; Stengel *et al*., 2012) and this process most likely occurs in the aforementioned PratA defined membrane regions (PDM) where the PM and TMs converge at TM biogenesis centres (Stengel *et al*., 2012). This hypothesis requires further clarification however (Schottkowski *et al*., 2009). Following the loading of pD1 with Mn^{2+} by PratA, the protein Ycf48 has also been shown to be required for the stabilisation of pD1 and its efficient incorporation into D2- cyt *b*559 to form the RC complex (Komenda *et al*., 2008; Plücken *et al*., 2002).

The extended C-terminal tail of pD1 must be cleaved to produce an intermediate form of D1 called iD1 (Inagaki *et al*., 2001). This is performed by the C-terminal processing protease which removes 8 of the 16 amino acids of the extended region (Komenda *et al*., 2007a) and enables the binding of D2-Cyt *b*559 (Anbudurai *et al*., 1994). The removal of this extended region has been shown to be essential to the assembly of the OEC as CtpA deletion mutants fail to form both the Mn_4CaO_5 cluster (Roose and Pakrasi, 2004; Satoh and Yamamoto, 2007) and the shielding cap consisting of the extrinsic PSII subunits (Roose and Pakrasi, 2004, 2008). In plants, the extended pD1 C-terminal domain is completely removed in one step; hence the formation of iD1 is unique to cyanobacteria. It has been hypothesised that the remaining C-terminal region of iD1 is important for targeting of the RC from the PM to the TM where the subsequent steps of PSII assembly take place (Komenda *et al*., 2007a) whereas in plants the RC is likely already present in the stromal TM where PSII biogenesis takes place (Callahan *et al*., 1989).

1.6.1.4 RC47

The RC47 complex is formed by the incorporation of a preassembled CP47 subcomplex into the RC. This step is believed to take place entirely within the TM, as CP47 has not been detected within the PM or PDM regions (Rengstl *et al*., 2011). The CP47 subcomplex is formed independently within the membrane and consists of CP47, PsbH, PsbL, PsbM, PsbT, PsbX and PsbY (Boehm *et al*., 2011, 2012). iD1 is again cleaved at residue 344 to remove the last 8 amino acids of the extended D1 C-terminus to produce mature D1 (Komenda *et al*., 2004, 2007a).

The formation of the CP47 subcomplex and its binding to the RC is mediated by the assembly factors Psb28, PAM68 and Sll0933 in cooperation with Ycf48 (Boehm *et al*., 2012; Bučinská *et al*., 2018; Rengstl *et al*., 2011). Sll0933 and PAM68 stabilises the membrane segments of CP47 as it is being co-translationally inserted into the TM by the SecY translocase in conjunction with Chl binding which is facilitated by PAM68 (Bučinská *et al*., 2018; Rengstl *et al*., 2013).

1.6.1.5 Monomeric PSII

Formation of monomeric PSII (PSII [1]) involves the binding of the CP43 subcomplex and the subsequent assembly of the OEC. The CP43 subcomplex comprises CP43, PsbK, PsbZ, and Psb30 (Boehm *et al*., 2011). The binding of this module to RC47 gives rise to a PSII [1] complex that lacks only the extrinsic subunits PsbO, PsbP, PsbQ, PsbU and PsbV that stabilise the OEC (Hopp *et al*., 1988; Nowaczyk *et al*., 2006). Once these have bound, the PSII monomer is complete. This process involves several auxiliary assembly factors such as Sll0606, which is also known to be required for CP43 attachment to RC47 (Zhang *et al*., 2010), as well as Psb28 and Psb29, which bind unassembled CP43 (Dobáková *et al*., 2009; Shi *et al*., 2012). However, the most comprehensively studied of these is Psb27 (Kashino *et al*., 2002; Shi *et al*., 2012).

Psb27 is a lipoprotein that operates in the thylakoid lumen and has been co-purified with several PSII intermediates including PSII [1], dimeric PSII (PSII [2]) via binding to CP43, as well as in a complex with unassembled CP43 (Komenda *et al*., 2012; Liu *et al*., 2011b, 2011a). Psb27 has been hypothesised to prevent binding of the extrinsic subunits to PSII to allow time for the correct assembly of the Mn_4CaO_5 cluster and C-terminal processing of D1 (Roose and Pakrasi, 2008). The available evidence also points to role of Psb27 as a more general stabilising factor as it has been shown to interact with PSI (Komenda *et al*., 2012). Additionally, its homologue in plants, LPA19, interacts with the C-terminus of pD1 and is believed to be involved in the processing of the extended tail region (Wei *et al*., 2010).

Following binding of CP43 to RC47, the correct assembly of the Mn_4CaO_5 cluster and maturation of D1, Psb27 is exchanged for the extrinsic subunits PsbO, PsbP, PsbQ, PsbU and PsbV which bind to the lumenal side of PSII and form the OEC photoprotective cap (Komenda *et al*., 2007a). When the other low molecular mass subunits are integrated into the complex is unknown. PSII [1] now provides all of the ligands required for the light-driven activation of the OEC, a process known as photoactivation (Dasgupta *et al*., 2008). This involves a rearrangement of the Ca^{2+} and Mn^{2+} ions so that, following the first instance of charge splitting by the PSII RC, oxidation of the first Mn^{2+} ion, coordinated by Asp170 of D1 (Nixon and Diner, 1994) triggers a cascade resulting in the activation of the Mn_4O_5Ca cluster. The Mn_4O_5Ca cluster can now undergo further oxidations and split H_2O. (Becker *et al*., 2011).

The assembly of PSII is completed by the dimerisation of PSII (PSII [2]), mediated by the PsbI and PsbM subunits (Kawakami *et al*., 2011), and attachment of the peripheral light harvesting antenna complexes (phychobilisomes in cyanobacteria).

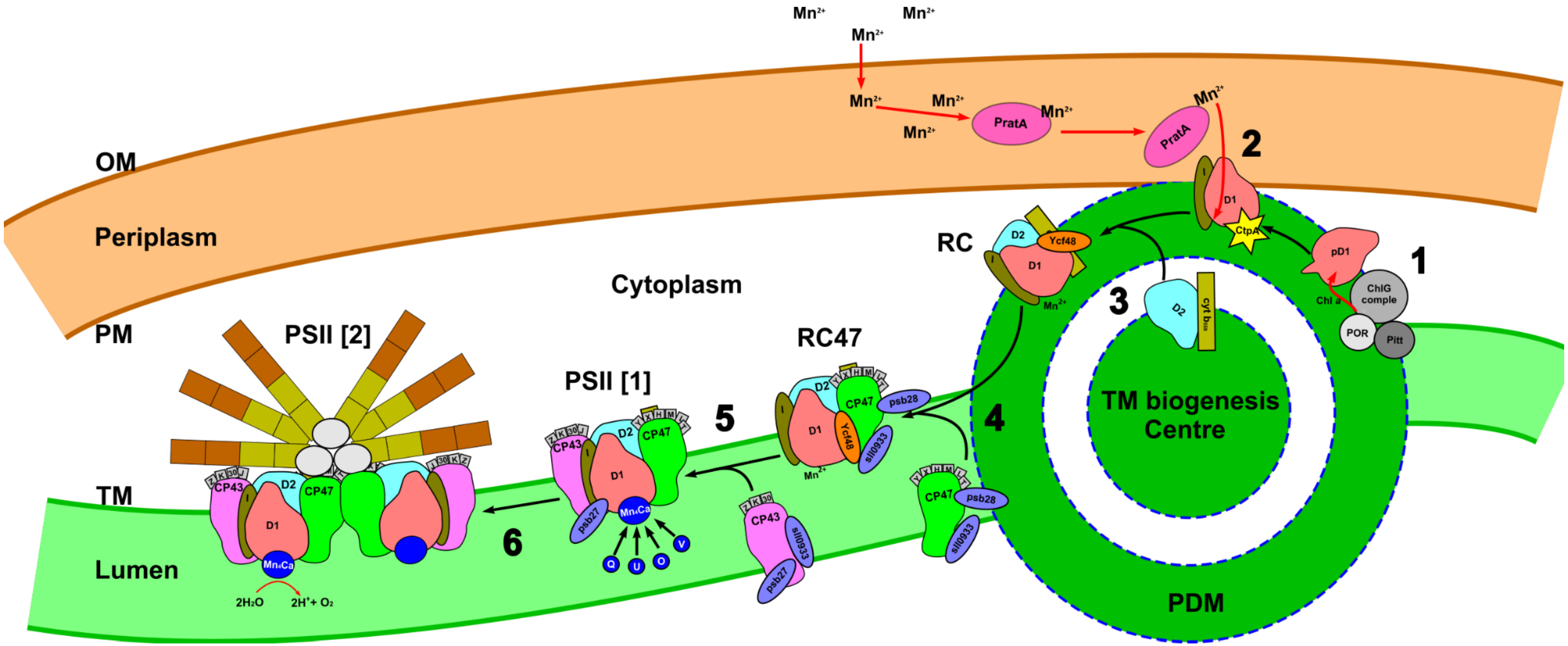


Figure 1.9: Photosystem II assembly pathway. A model of PSII assembly. Early PSII assembly intermediates arise within TM biogenesis centres and migrate into the TM to mature. pD1 is loaded with Chl co-factors by ChlG, POR and Pitt (1) before being C-terminally processed by CtpA and pre-loaded with manganese by PratA, enabling the binding of PsbI (2). The pre-formed D2-Cyt-*b*559 complex is incorporated into D1-psbI to from the RC complex (3). CP47, together with PsbH, L, M, T, X and Y bind to the RC to form RC47 (4) followed by incorporation of the CP43 module (5). Photoactivation can occur upon formation of the Mn_4Ca containing OEC, yielding PSII monomers (PSII [1]). Subsequent dimerisation of PSII and attachment of the phycobilisome yields mature PSII [2] (6).

1.6.2 PSII repair

PSII is vulnerable to damage when exposed to light of any intensity (Barber and Andersson, 1992) and will undergo photoinhibition where PSII activity decreases (Aro *et al*., 1993; Powles, 1984), followed by repair of the damaged complex. Photoinhibition is caused in one of two ways; formation of the triplet state of P680 which can generate a damaging reactive oxygen species (ROS) (acceptor side inhibition) or by an imbalance in water oxidation versus P680 oxidation so that oxidised $P680^+$ remains long lived (donor side inhibition) (Ohad *et al*., 1984). The severity of PSII photoinhibition is dictated by the balance between the rate of PSII damage, which increases in proportion to light intensity, and the rate of PSII repair (Takahashi and Murata, 2008). The rate of repair itself is deleteriously affected by environmental conditions such as low temperatures and oxidative stress (Allakhverdiev and Murata, 2004).

The D1 subunit of the PSII RC is the main component that is damaged by excess light absorption. Damaged D1 is rapidly degraded and replaced with a newly synthesised protein to alleviate the effects of photoinhibition (Ohad *et al*., 1984). PSII must be partially disassembled, the damaged D1 removed and a new copy inserted in its place before the functional complex is reassembled (Nixon *et al*., 2004). D1 turnover is a highly energy intensive process and has been calculated to require 1304 molecules of ATP per cycle (Murata and Nishiyama, 2018). PSII repair likely occurs at the site of damage as migration of damaged PSII back to the PDM region (or stromal TM in plants) would be time consuming and inefficient (Nickelsen and Rengstl, 2013).

How the cell can distinguish between a damaged and undamaged PSII subunit remains largely unclear. It has been shown, however, that when CP43 is removed from a functioning PSII complex in the dark, the D1 subunit is exposed and is selected for degradation by proteases (Krynická *et al*., 2015). In addition, the PsbO subunit of one monomer in a PSII dimer interacts with CP43 of the other monomer (Guskov *et al*., 2009) and a *Synechocystis* mutant that lacks PsbO fails to form dimeric PSII (Komenda *et al*., 2010). In light of this, is has been suggested that damage to D1 and the Mn_4O_5Ca cluster by photoinhibition causes release of PsbO which in turn causes a dissociation of CP43 from PSII to produce RC47 (Nixon *et al*., 2010). However, more research is required to further understand how damaged PSII is recognised and targeted for repair.

As D1 is a core RC protein, situated at the heart of PSII, it is inaccessible to the PSII repair machinery. PSII is therefore disassembled to isolate D1. In cyanobacteria PSII disassembly involves detachment of the CP43 and OEC subunits (Nixon *et al*., 2004). This forms a complex that resembles the RC47 intermediate complex formed during *de novo* PSII assembly. The damaged D1 is now available for degradation. All other subunits that are stripped away from the complex during PSII disassembly are recycled (Nixon *et al*., 2010).

D1 is degraded by members of the FtsH protease family in cyanobacteria which form heterodimers *in vivo* and bind to the N-terminus of D1 before digesting the subunit (Barker *et al*., 2008; Komenda *et al*., 2007b; Silva *et al*., 2003). FtsH2 in particular has been demonstrated to be important for degradation of D1 in response to D1 damage by heat stress (Kamata *et al*., 2005), UV-B (Cheregi *et al*., 2007), chemical exposure (Drath *et al*., 2008) or in cells harbouring extrinsic subunit mutants (Komenda *et al*., 2010). This, combined with the knowledge that FtsH proteins have poor unfoldase activity and target multiple unassembled PSII subunits (Komenda *et al*., 2006), would suggest that these proteases degrade PSII subunits based on the general instability of the proteins caused by unspecific damaging events (Nixon *et al*., 2010). The PSII assembly and repair factor Psb29 has been reported to interact with FtsH complexes and mediate accumulation of the FtsH heterodimers required for D1 degradation (Bečková *et al*., 2017).

In plants D1 is phosphorylated, which increases its stability (Koivuniemi *et al*., 1995) and necessitates the dephosphorylation of the subunit to enable its efficient degradation (Kato and Sakamoto, 2014; Rintamäki *et al*., 1996). D1 degradation is carried out by multiple proteases in plants (Puthiyaveetil *et al*., 2014). Among these is a family of membrane associated serine proteases called Deg proteases (Ortega *et al*., 2009) that operate within the thylakoid lumen and cleave the loops that connect the D1 transmembrane helices (Kapri-Pardes *et al*., 2007; Sun *et al*., 2007a, 2010). This activity is postulated to enhance the degradation of D1 by generating more C and N-terminal ends for FtsH to bind and digest (Kapri-Pardes *et al*., 2007; Kato and Sakamoto, 2009; Sun *et al*., 2010). There is some debate regarding the importance of Deg proteases in the breakdown of D1. D1 turnover is unaffected in *Arabidopsis (A.) thaliana* single Deg mutants grown under moderate conditions, indicating a degree of redundancy between members of the Deg family (Sun *et al*., 2007b). However, under stressful conditions when the D1 population sustains accelerated rates of damage, the D1 turnover rate is reduced (Sun *et al*., 2007b). As such, the data indicates that Deg proteases function to significantly improve the rate of FtsH mediated D1 degradation at times of high stress when the FtsH system would otherwise be overwhelmed (Nixon *et al*., 2010). On the other hand, a *Synechocystis* mutant lacking all 3 members of the Deg protease family was not impaired in D1 degradation (Barker *et al*., 2006), suggesting that FtsH proteases alone are sufficient to perform this process in cyanobacteria.

Following the removal and degradation of the damaged D1, a *de novo* D1 subunit is inserted into the RC47 complex. The mechanism behind this process in cyanobacteria is poorly understood although there is evidence to suggest that the integration of new D1 is coordinated with digestion of damaged D1. Reduced rates of D1 synthesis in *Synechocystis*, caused by mutation (Komenda *et al*., 2000) or exposure to an inhibitor (Komenda and Barber, 1995), correlate with correspondingly low rates of D1 degradation. Slr0151 has been suggested to be a PSII repair factor as it interacts with D1 and CP43, may mediate PSII assembly and disassembly, as well as increase the rates of *de novo* D1 synthesis under stressful light intensities (Yang *et al*., 2014). In addition, deletion of Ycf48 significantly reduces the rate of D1 turnover during the PSII repair cycle, implicating this protein as a PSII repair factor (Komenda *et al*., 2008).

More is known about this process in plants. cpFTSY recruits ribosomes translating *de novo* D1 subunits to the stromal TM by recognising the D1 signal recognition particle 54 (cpSRP54) (Walter *et al*., 2015a). D1 is inserted into RC47 by the SecY (Nilsson and van Wijk, 2002; Nilsson *et al*., 1999; Walter *et al*., 2015b; Zhang *et al*., 1999, 2001) translocase in collaboration with the Alb3 insertase (Klostermann *et al*., 2002;

Ossenbühl *et al*., 2004; Pasch *et al*., 2005). Several PSII repair factors facilitate this process including LPA1, TERC, PsbN, CYP38 and TEF30. Briefly, LPA1 is postulated to mediate the incorporation of D1 into PSII by acting as a chaperone (Dewez *et al*., 2009; Peng *et al*., 2006), TERC interacts with Alb3 during D1 insertion into the TM (Schneider *et al*., 2014), PsbN aids the correct folding of D1 (Torabi *et al*., 2014), CYP38 and TEF30 facilitates D1 and CP43 integration into RC47 (Fu *et al*., 2007; Muranaka *et al*., 2016; Sirpiö *et al*., 2008).

Once a new D1 subunit is in place, the CP43 subcomplex can be reassembled, followed by the reformation of the OEC complex. PSII [1] can dimerise once more, thereby producing functional PSII [2] (Theis and Schroda, 2016). MET1 has been shown to be important for CP43 and CP47 assembly as null mutants exhibited severe impairment of functional PSII assembly and accumulated isolated CP43, indicating a role of this protein in PSII assembly and/or repair (Bhuiyan *et al*., 2015). The OEC can reassemble only after CP43 has bound PSII and the C-terminal extended region of D1 is cleaved off by CtpA, as is the case in *de novo* PSII assembly (Anbudurai *et al*., 1994; Diner *et al*., 1988). In plants the completed PSII holoenzyme migrates back to the grana TM (Theis and Schroda, 2016). The main stages of the PSII repair pathway are presented in Figure 1.10.

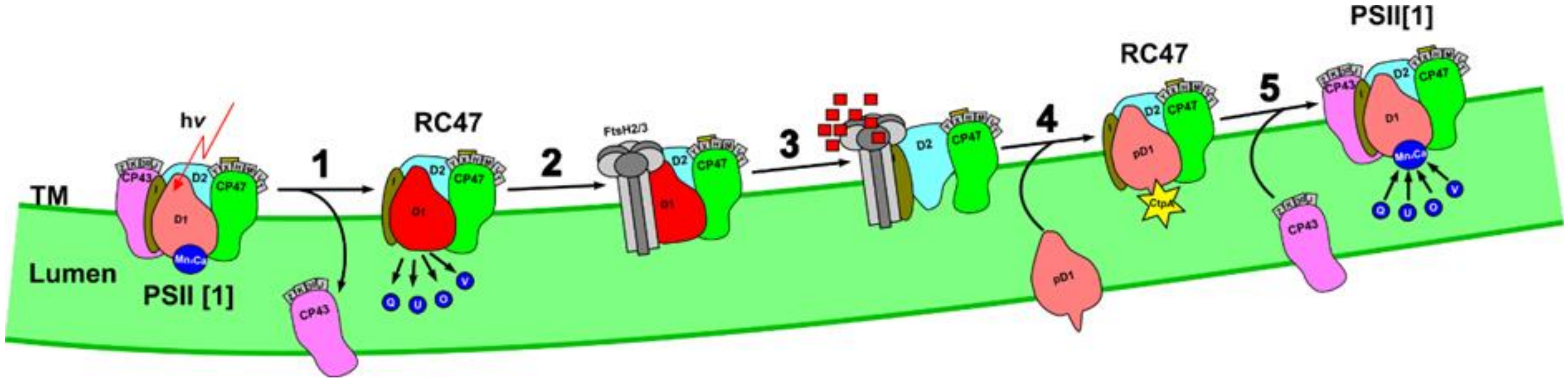


Figure 1.10: Photosystem II repair pathway. The PSII repair pathway is initiated upon photodamage of D1 and begins with detachment of the CP43 and the OEC to form RC47 (1). D1 is degraded by the FtsH2/3 complex (2-3) and a new copy of pD1 is integrated into the complex (4). The extended C-terminus of pD1 is cleaved by CtpA, enabling the rebinding of CP43, which is recycled back to the PSII complex, and reformation of the OEC (5).

1.6.3 PSI assembly and repair

Unlike PSII, the process of PSI repair has not been studied extensively due to the fact that PSI is inherently far more stable than PSII. PSI assembly research is similarly limited. Like PSII, the process is presumably subject to tight regulation and is thought to begin with the co-translational insertion of the two core subunits, PsaA and PsaB, into the membrane where they are assembled into the RC (Cai *et al*., 2010; Göhre *et al*., 2006; Schöttler *et al*., 2011). The remaining PSI subunits are integrated into the complex which subsequently forms a trimer. The exact sequence of events between RC formation and completion of the functional PSI supercomplex has not yet been accurately determined. The process seems to occur rapidly and, due to the small size of the intrinsic subunits in comparison to the large RC heterodimer, the separation of intermediate complexes based on mass is challenging (Schöttler *et al*., 2011).

PSI assembly factors have been identified in a variety of photosynthetic organisms. In *Synechocystis*, Ycf37 appears to be required to stabilise PSI assembly intermediates during periods of high-light stress (Dühring *et al*., 2006, 2007). In plants, Nellaepalli *et al.* (2018) recently showed that two conserved chloroplast proteins, Ycf3 and Ycf4, form modules that co-purify with PSI assembly intermediates. Ycf3 was previously found to be required for the attachment of PsaA and PsaD to PSI in plants, algae and cyanobacteria (Boudreau *et al*., 1997; Naver *et al*., 2001; Ruf *et al*., 1997; Schwabe *et al*., 2003). Ycf3 and Ycf4, together with other auxiliary proteins, facilitate formation of the PSI RC and the binding of accessory LHCs to the complex. Nellaepalli *et al*., (2018) proposed a model of PSI assembly in which Ycf3 binds to a partner protein, Ycf3 interacting protein (Y3IP1), to form a stable Ycf3-Y3IP1 assembly module in the TM. Y3IP1 was previously found to be exclusive to plants and to aid Ycf3 during the assembly process (Albus *et al*., 2010). Ycf3-Y3IP1 promotes the assembly of *de novo* PsaB and PsaA into the PSI RC. In addition, this process is assisted from the lumenal side by assembly factors PPD1 and PSA2 and from the stromal side by PYG7 and PSA3. PPD1 has been suggested to mediate the dimerisation of PsaA and PsaB (Liu *et al*., 2012; Roose *et al*., 2014). Pyg7, together with PSA3 (Shen *et al*., 2017), are critical for the accumulation of mature PSI (Stöckel *et al*., 2006) through interaction with PsaC

(Yang *et al*., 2017). The oligomeric Ycf4 binds to the newly formed RC and stabilises it, while facilitating the incorporation of other small PSI subunits, PsaC-F, PsaH-J and PsaL, into the RC to form the PSI core complex. Following this, seven LHCa antenna complexes bind to PSI with the exception of LHCa 2/9 which bind after PsaG and K have been integrated into the PSI-LHC subcomplex. Y3IP1 and Ycf4 then detach from the mature PSI-LHC supercomplex and are recycled, whereas Ycf3 is replaced with a newly synthesised copy.

1.7 (Bacterio)chlorophyll structure and function

1.7.1 Chlorophyll

The main light absorbing molecule of oxygenic photosynthesis is chlorophyll (Chl). Synthesised in plants, algae and cyanobacteria, these molecules absorb light in the blue and red portion of the electromagnetic spectrum and absorb poorly in the green region, giving Chl their distinctive green colour (Muneer *et al*., 2014). As one of the most abundant biological molecules on Earth, there is an estimated 10^9 tons of Chl metabolised every year and it is the only biochemical process visible from outer space (Porra, 1997; Rüdiger, 1997).

There are many variants of Chl, all of which have slightly different spectroscopic properties. They all share the same basic structure, consisting of a chlorin ring esterified to a long phytol chain. An Mg^{2+} ion is located at the centre of the macrocycle. Chl is distinguishable from other chlorins by the presence of a 5th ring beyond the 4 rings of standard chlorins, which are lettered clockwise from A to E according to International Union of Pure and Applied Chemistry (IUPAC) nomenclature (Figure 1.10A). The 6 Chl types, called Chl *a*, *b*, *c1*, *c2*, *d* and *f* vary, depending on the functional groups located at positions C2, 3, 7, 8 and 17 as well as the reduction state of the C17-18 bond. A broad range of wavelengths of light is absorbed by the various Chl pigments. The model cyanobacterium used in this work, *Synechocystis*, synthesises only Chl *a* (Figure 1.11A).

1.7.2 Bacteriochlorophyll

Bacteriochlorophylls (BChls) are light harvesting pigments closely related to Chl and are found in purple bacteria, green sulphur bacteria and heliobacteria. They differ from Chl on the C3 position, which is bonded to an acetyl group instead of a vinyl group, and at C7=C8 which is a single bond (Figure 1.11B). Like Chl, slight modifications to the functional groups produce various types of BChl.

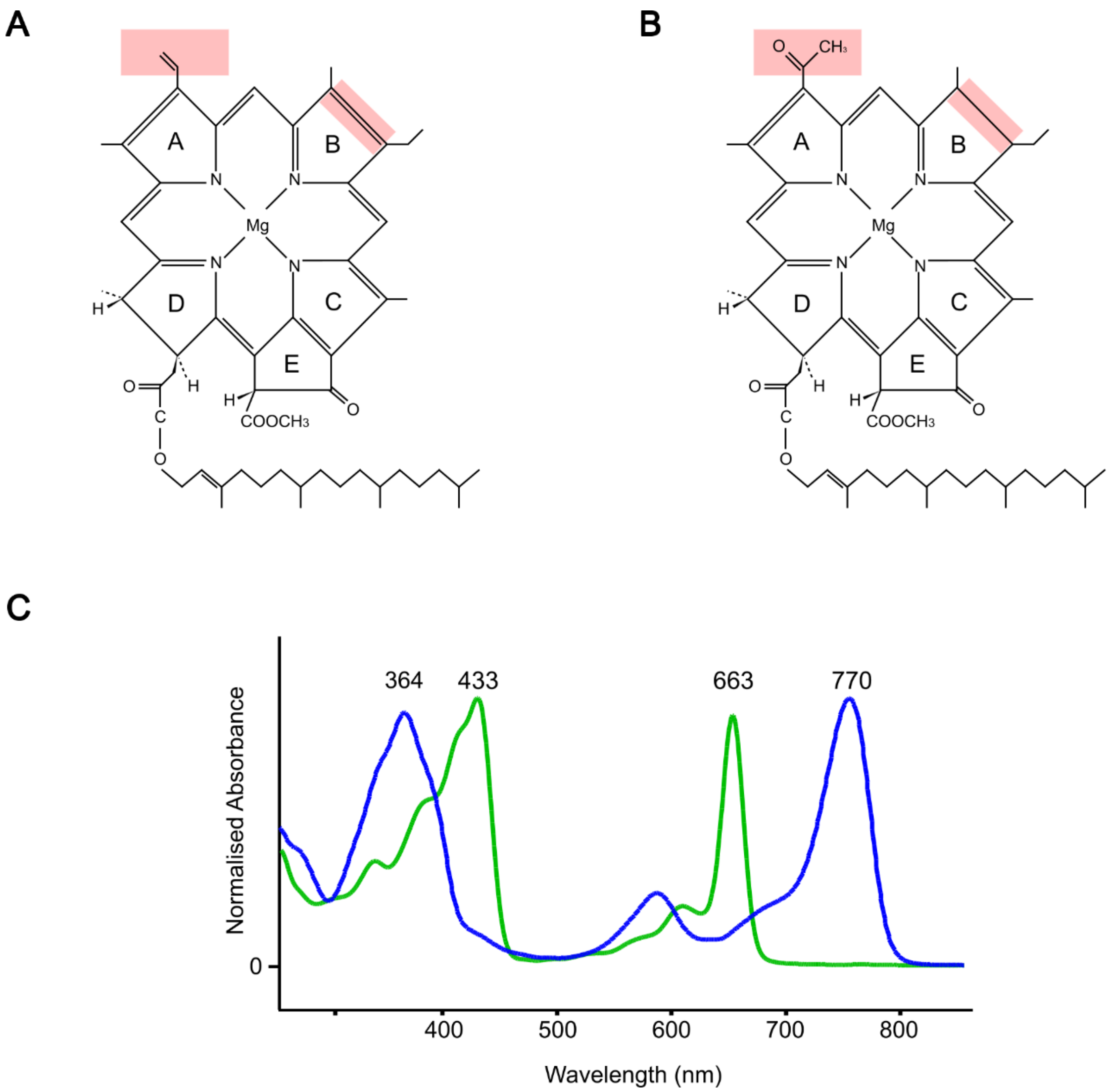


Figure 1.11: Structures of chlorophyll *a* and bacteriochlorophyll *a*. The structure of Chl *a* (A) and BChl *a* (B) highlighting the differences between the two pigments (red boxes). The absorbance spectra (C) shows the absorption profile of Chl *a* (green) relative to BChl *a* (blue).

1.8 (Bacterio)Chlorophyll biosynthesis

1.8.1 Early stages of chlorophyll biosynthesis

The process of Chl biosynthesis consists of 17 enzymatic steps catalysed by 15 enzymes. The first committed step of the pathway is the formation of 5-aminolevulinic (ALA) from L-glutamic acid (Beale and Castelfranco, 1974) (Figure 1.12) or alternatively from glycine and succinate (Beale, 2006). Condensation of two molecules of ALA, catalysed by porphobilinogen synthase, yields porphobilinogen and then four of these condense to form the open chain tetrapyrrolehydroxymethylbilane in a reaction catalysed by hydroxymethylbilane synthase. The enzyme uroporphyrinogen III synthase “flips” what will become ring D of the tetrapyrrole and closes the ring to form uroporphyrinogen III, producing the first tetrapyrrole structure.

The first branch point of the pathway is split between haem/Chl and vitamin B_{12} synthesis. C-methylation of uroporphyrinogen III shunts the molecule down the vitamin B_{12} pathway, whereas decarboxylation by uroporphyrinogen III decarboxylase (UROD) produces coproporphyrinogen III and commits the molecule to the haem/Chl synthesis pathway.

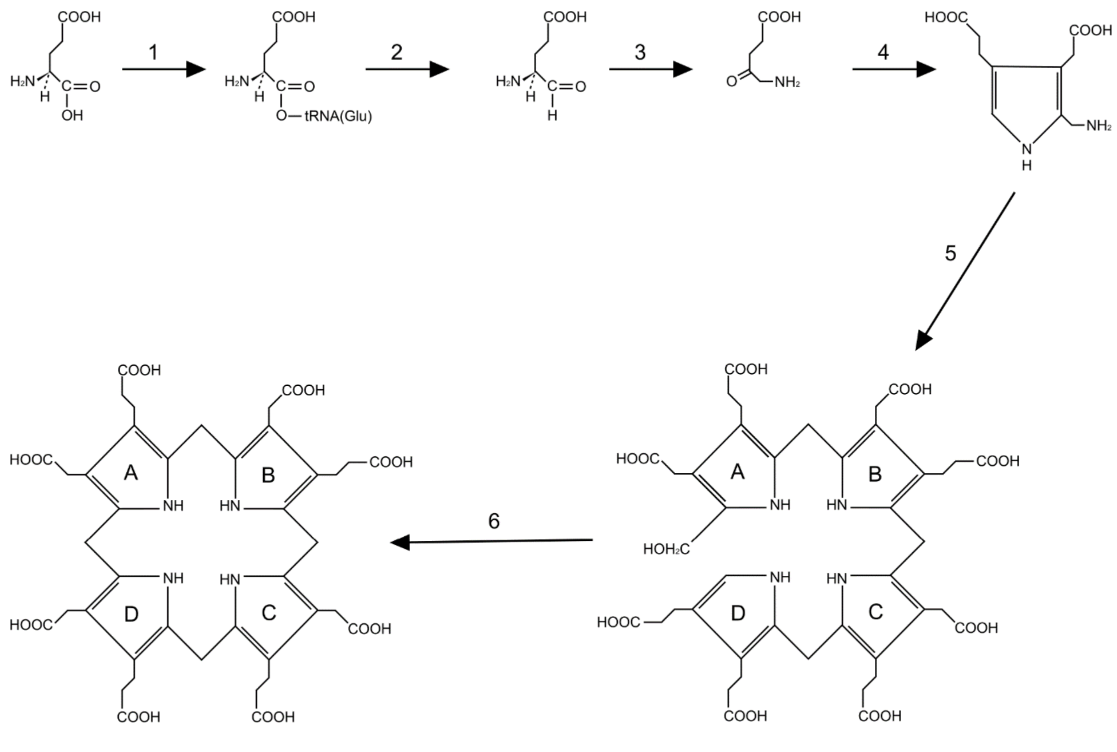


Figure 1.12: Synthesis of uroporphyrinogen III. L-glutamic acid is converted to ALA using tRNA (1-3). Two molecules of ALA are condensed to form porphobilinogen (PGB) by ALA dehydrogenase (4). Four molecules of PGB are condensed to from hydroxymethylbilane (HMB) by HMB synthase (5) and then is cyclised to uroporphyrinogen II (Urogen) by uroporphyrinogen synthase (UROS) (6).

1.8.2 Uroporphyrinogen III decarboxylase (UROD)

UROD is a cytosolic monomeric protein encoded by the gene *hemE*. The enzyme is reported to form a dimer *in vivo* which confers catalytic activity on the protein (Whitby *et al*., 1998) and allows it to remove four carboxyl groups of the carboxymethyl side chains of uroporphyrinogen III to produce coproporphyrinogen III (Elder and Roberts, 1995) (Figure 1.13). The structure of the human and tobacco orthologues have been solved by X-ray crystallography (Martins *et al*., 2001; Whitby *et al*., 1998). Following discovery of HemE in the cyanobacteria *Synechococcus sp*. PCC7 942 (Kiel *et al*., 1992), it was found that the enzyme interacts with the stress protein HptG via its N-terminus (Watanabe *et al*., 2007). HptG is a member of the Hsp90 family of proteins in prokaryotes and is important for regulating heat and oxidative stress in cyanobacteria (Hossain and Nakamoto, 2002, 2003; Tanaka and Nakamoto, 1999). It was found to decrease HemE activity both *in vivo* (Watanabe *et al*., 2007) and *in vitro* (Saito *et al*., 2008), indicating a function of HemE in regulating tetrapyrrole biosynthesis.

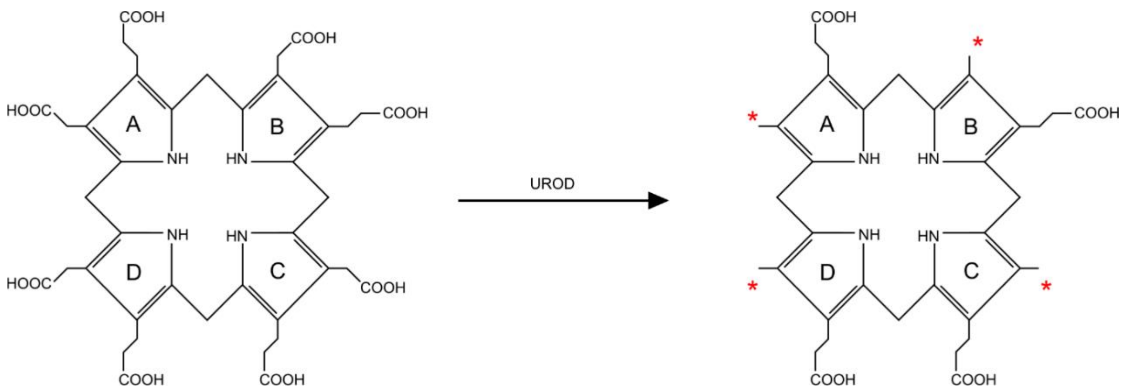


Figure 1.13: Reaction catalysed by uroporphyrinogen III decarboxylase. UROD catalyses the removal of four hydroxyl groups from Urogen to form coproporphyrinogen III (Copro). * mark sites of chemical modification.

1.8.3 Coproporphyrinogen III oxidase (CPO)

Following decarboxylation by UROD, CPO mediates oxidative decarboxylation of coproporphyrinogen-III to protoporphyrinogen-IX by catalysing removal of carboxyl from the propionate groups of rings A and B to form vinyl moieties (Sano, 1966; Sano and Granick, 1961) (Figure 1.14). Two analogous but structurally unrelated CPOs have been described in various organisms. The first, HemN, is found in prokaryotes and is oxygen independent (Troup *et al*., 1995). The enzyme is a member of the radical SAM protein family (Layer *et al*., 2002; Sofia *et al*., 2001) that contain a conserved iron-sulphur 4[Fe-S] cluster that, in combination with the co-factor S-adenosylmethionine (SAM), is essential to its catalytic function. The second, HemF, is native to both eukaryotes and prokaryotes and is dependent on oxygen for activity (Troup *et al*., 1994). This enzyme has no requirement for co-factors and thus employs a different mechanism of catalysis to HemN (Lash, 2005). The exact details of this process remain unknown despite progress in determining the structural characteristics of the enzyme (Phillips *et al*., 2004). *Synechocystis* sp. PCC 6803 contains both *hemN* and *hemF* orthologues, designated *sll1876* and *sll1185* respectively. These two proteins both demonstrated CPO activity *in vitro* and deletion of either gene caused the enzymes substrate to accumulate (Goto *et al*., 2010; O'Brian and Panek, 2002). The dual operation of *sll1876* and *sll1185* enables biosynthesis of tetrapyrroles under varying degrees of oxygen tension (Goto *et al*., 2010). A third *hemN* -like gene, annotated *sll1917*, was also found in the *Synechocystis* genome; however no phenotype was detected in a mutant strain lacking the protein and it was not found to have CPO activity *in vitro* (Goto *et al*., 2010).

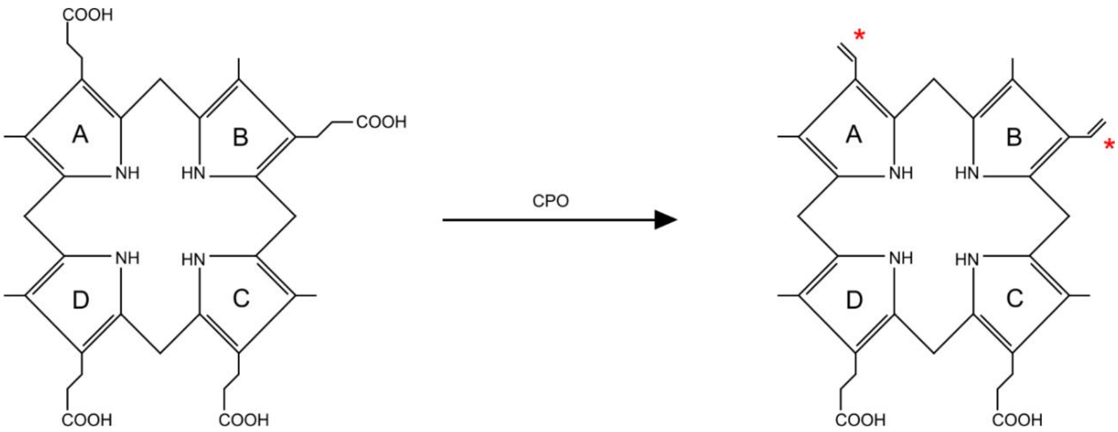


Figure 1.14: Reaction catalysed by coproporphyrinogen III oxidase. CPO catalyses the oxidative decarboxylation of Copro to protoporphyrinogen IX (Protogen) by removal of carboxyl from the propionate groups of rings A and B. * mark sites of chemical modification.

1.8.4 Protoporphyrinogen oxidase (PPOX)

The next step in the pathway is the conversion of protoporphyrinogen IX to protoporphyrin IX. This reaction can happen spontaneously and by the action of non-specific peroxidases (Jacobs and Jacobs, 1993; Lee *et al*., 1993). However, it has been documented numerous times that this step is mediated by an enzyme called PPOX. PPOX catalyses the 6 electron oxidation of protoporphyrinogen forming protoporphyrin IX (Figure 1.15).

Most organisms possess one of three known analogous PPOXs that are phylogenetically unrelated but catalyse the same reaction. The first was discovered in *Escherichia coli* and labelled *hemG* (Sasarman *et al*., 1993) followed by discovery of an orthologue in *Bacillus subtilis* (Hansson and Hederstedt, 1994) that was subsequently designated *hemY*. A third essential transmembrane orthologue was isolated from *Synechocystis* sp. PCC 6803 and annotated *hemJ* (Kato *et al*., 2010). *HemY* is native to the majority of aerobic bacteria and terrestrial plant life (Oborník and Green, 2005); Gram-negative proteobacteria possess *hemG* (Kobayashi *et al*., 2014) whilst *hemJ* is mostly found in cyanobacteria (Kato *et al*., 2010). Some organisms contain more than one isoform of the enzyme while other organisms do not contain any PPOX (that has been identified so far) (Kobayashi *et al*., 2014).

The three isoforms of PPOX differ in the co-factors they require for activity. HemY is dependent upon oxygen and flavin-adenine dinucleotide (FAD) (Corradi *et al*., 2006; Koch *et al*., 2004; Qin *et al*., 2010) whereas HemG utilises flavin mononucleotide (FMN) and can operate in aerobic and anaerobic environments (Boynton *et al*., 2009; Möbius *et al*., 2010). HemJ can function under aerobic conditions and it is not yet known if the enzyme requires additional co-factors for PPOX activity (Kato *et al*., 2010). The mechanistic details of the various PPOX activities are unclear. HemY probably transfers electrons from protoporphyrinogen IX to oxygen via its FAD co-factor (Koch *et al*., 2004) whereas HemG likely contributes the electrons to the ETC via FMN (Boynton *et al*., 2009). It has been suggested that HemJ may use haem as an electron acceptor (Kato *et al*., 2010).

Protoporphyrin IX is toxic to the cell in high concentrations and so the molecule must be tightly controlled to prevent it accumulating. In this respect, HemY is known to form a complex with the next enzyme in the Haem or Chl biosynthesis pathways, ferrochelatase (FeCH) or magnesium chelatase (MgCH) respectively (Chidgey *et al*., 2017; Ferreira *et al*., 1988; Koch *et al*., 2004; Masoumi *et al*., 2008). This enables the channelling of protoporphyrin IX from PPOX to FeCH/MgCH, preventing its release into the cell.

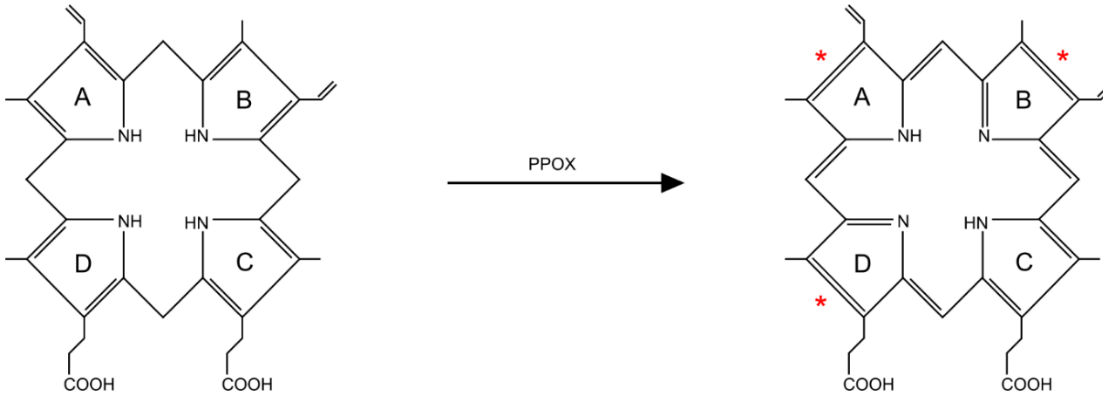


Figure 1.15: Reaction catalysed by Protoporphyrinogen oxidase. PPOX catalyses the conversion of Protogen to protoporphyrin IX (Proto) by oxidation of the porphyrin ring. * mark sites of chemical modification.

1.8.5 The haem/Chl branch point

The Chl and haem biosynthesis pathways have been shared up until and including the synthesis of protoporphyrin IX. The process is highly conserved in most organisms, with the exception of some species of archaea (Storbeck *et al*., 2010). The two pathways diverge at the point of metal ion insertion. This is the second branch point of the pathway where prospective Chl molecules are chelated with a Mg^{2+} ion and future haem molecules receive an Fe^{2+} ion. The enzyme ferrochelatase (FeCH) inserts an Fe^{2+} ion into the protoporphyrin IX macrocycle, whereas Mg^{2+} is inserted into the macrocycle by Mg chelatase (MgCH). Unlike the FeCH reaction, which has no energetic requirements, Mg^{2+} chelation requires the hydrolysis of around 14 ATP molecules to yield the energy required for the reaction to occur (Reid and Hunter, 2004). It is speculated that the energy is required to remove the hydration shell from the Mg^{2+} ion (Gibson *et al*., 1995; Jensen *et al*., 1996; Reid and Hunter, 2004). MgCH is the first enzyme exclusive to the Chl biosynthesis pathway.

In *Rba. sphaeroides*, the haem/Chl branchpoint is regulated by a protein known as PufQ (Chidgey *et al*., 2017). PufQ is under the control of the oxygen sensitive *puf* operon. Aerobic conditions suppress expression of the *puf* operon and the cells grow via aerobic respiration. FeCH is evenly distributed throughout the cell where it can produce haem for the respiratory complexes (Chidgey *et al*., 2017). When oxygen conditions decrease below a certain threshold, expression of the *puf* operon and the assembly of the photosynthetic apparatus is initiated. PufQ is synthesised and inserted into the membrane where it can interact with the membrane associated FeCH and inhibit delivery of protoporphyrin IX to the enzyme. This prompts increased delivery of the substrate to MgCH and promotes synthesis of the BChl required for photosynthesis (Chidgey *et al*., 2017).

1.8.6 Magnesium Chelatase (MgCH)

MgCH catalyses the ATP dependent conversion of protoporphyrin IX to Mg-protoporphyrin IX by inserting an Mg^{2+} ion into the chlorin macrocycle (Figure 1.16). The enzyme was discovered by mutation of the *bchI*, *bchH* and *bchD* genes of *Rba. sphaeroides* which produced a mutant incapable of synthesising Mg-protoporphyrin IX suggesting that Mg^{2+} chelation was inhibited (Coomber *et al*., 1990; Naylor *et al*., 1999). These proteins and their *Synechocystis* homologs, *chlI*, *chlH* and *chlD* respectively, were produced recombinantly in *E. coli* cells in separate experiments. *In vitro* assays confirmed their role in Mg^{2+} chelation (Gibson *et al*., 1995; Jensen *et al*., 1996). The ChlI and ChlD subunits are members of the AAA^+ superfamily (ATPases Associated with various cellular Activities) (Gibson *et al*., 1999; Jensen *et al*., 1999). ChlH is the substrate binding subunit (Gibson *et al*., 1995; Jensen *et al*., 1996; Reid and Hunter, 2004) and has been shown to contain the enzyme's active site (Sirijovski *et al*., 2006). ChlH interacts with the ChlD-ChlI complex to form the active holoenzyme. ChlI catalyses ATP hydrolysis which drives the insertion of Mg^{2+} into the macrocycle of protoporphyrin IX and promotes the subsequent disassembly of the transient MgCH holoenzyme (Jensen *et al*., 1999).

ChlD also plays a role in regulating the enzyme. The N-terminal AAA^+ domain is essential for the assembly of the MgCH complex (Adams *et al*., 2016a) whilst phosphorylation of a conserved histidine residue at the C-terminus stimulates activity (Sawicki *et al*., 2017). The ChlD subunit in *Synechocystis* facilitates a complex cooperative response of the enzyme to free Mg^{2+} concentration (Brindley *et al*., 2015) the binding of which is essential for MgCH activity (Axelsson *et al*., 2006). However, the equivalent subunit in *Thermosynechococcus elongatus* exhibits a non-cooperative response to free Mg^{2+}, highlighting the significant variation in MgCH regulation between species (Adams *et al*., 2014). A porphyrin-binding protein called Gun4 is known to regulate MgCH activity by mediating the binding of porphyrin to ChlH (Larkin *et al*., 2003; Zhou *et al*., 2012), accelerating the activity of the enzyme 10 fold (Adams *et al*., 2016b) whilst decreasing the minimum amount of Mg^{2+} needed for catalysis at low substrate concentrations (Davison *et al*., 2005).

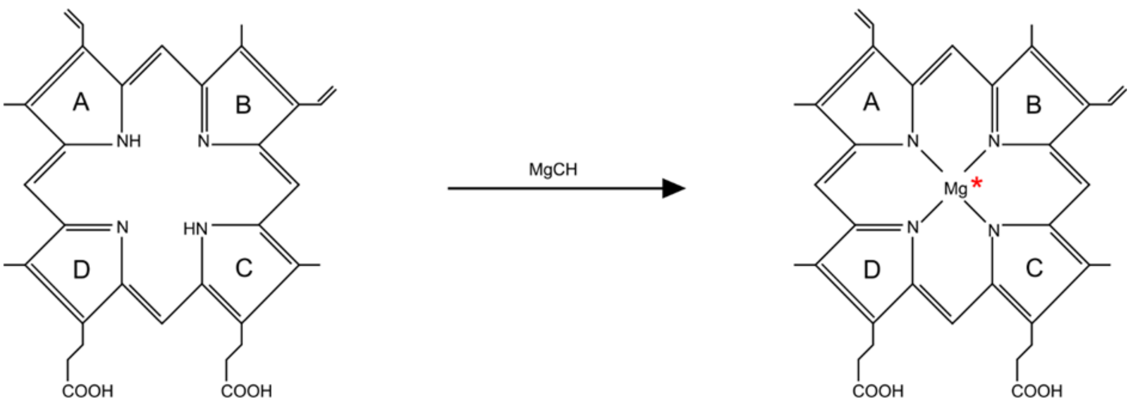


Figure 1.16: Reaction catalysed by magnesium chelatase. MgCH inserts Mg into the porphyrin ring of Proto yielding Mg- protoporphyrin IX (Mg-Proto). * mark sites of chemical modification.

1.8.7 Magnesium protoporphyrin methyltransferase (MT)

MT methylates the C13 carboxyl group of Mg-protoporphyrin IX by transferring a methyl group from S-adenosyl-L-methionine (SAM) to produce Mg-protoporphyrin monomethyl ester (MgPME) and S-adenosyl-L-homocysteine (SAH) (Gibson *et al*., 1963) (Figure 1.17). The substrates can bind in either order (Shepherd and Hunter, 2004; Shepherd *et al*., 2003). This is the first of three steps that modify the 13-propionate side chain of the chlorin macrocycle, leading to formation of the characteristic isocyclic ring E of Chl.

ChlM/*bchM* encodes MT in *Synechocystis* and *Rba. sphaeroides* respectively. *BchM* was identified by assays using recombinant *bchM* from *Rba. capsulatus* and *Rba. sphaeroides* expressed in *E.coli* (Bollivar *et al*., 1994a; Gibson and Hunter, 1994). The *Synechocystis* orthologue was identified by complementing a *Rba. capsulatus bchM* knockout mutant with a cosmid library prepared from the cyanobacteria (Smith *et al*., 1996). Sharing 29% identity with its purple bacterial homologue, the gene was subsequently labelled *chlM* in the *Synechocystis* genome.

Little is known about the mechanism of ChlM catalysis. The recent elucidation of the structure of *Synechocystis* ChlM, bound to SAM and SAH, by Cryo EM at 1.6 and 1.7 Å respectively led to the identification of a essential residues, Tyr-28 and His-139, involved in the direct transfer of the methyl group to Mg-protoporphyrin (Chen *et al*., 2014). Two flexible regions, the N-terminus and an extended alpha helix protruding from the core of the enzyme, were proposed to facilitate binding and release of the substrates (Chen *et al*., 2014). Further work is needed to fully determine the mechanistic characteristics of ChlM.

More is known about the regulatory processes employed to control ChlM activity. ChlM is stimulated by ChlH of MgCH in *Synechocystis*, purple bacteria and plants (Alawady *et al*., 2005; Hinchigeri *et al*., 1997; Johnson and Schmidt-Dannert, 2008; Sawicki and Willows, 2010; Shepherd *et al*., 2005). Low MgPMT activity is correlated with reduced ALA synthesis and MgCH activity but higher activity of FeCH. Conversely, greater MgPMT activity is correlated with increased synthesis of ALA, high MgCH and reduced FeCH activity (Alawady *et al*., 2005). This ensures efficient channelling of substrate between MgPME and MgCH (Tanaka and Tanaka, 2007) whilst controlling the rates of porphyrin synthesis and partitioning of protoporphyrin into the Chl branch of the pathway (Alawady and Grimm, 2004). In *A. thaliana*, ChlM is a target of NADPH-dependent thioredoxin reductase C (NTRC) which regulates the enzyme by stimulating MT activity in response to increased cellular photosynthetic activity (Richter *et al*., 2013). When the rate of photosynthesis decreases, the reducing power of NTRC decreases accordingly and it is unable to reduce conserved cysteine residues of ChlM, which results in destabilisation of the enzyme (Richter *et al*., 2013, 2016). The regulation of ChlM by NTRC therefore indirectly correlates the synthesis of Chl with cellular photosynthetic activity.

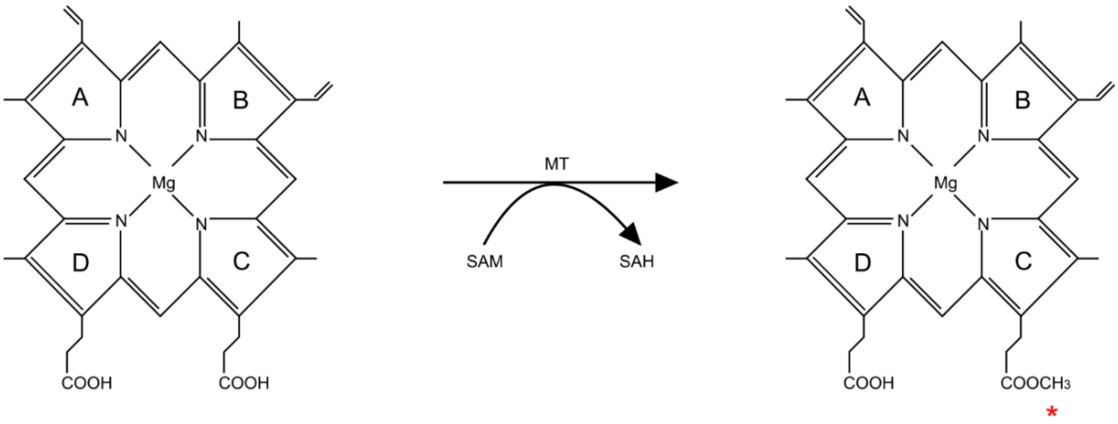


Figure 1.17: Reaction catalysed by magnesium protoporphyrin methyltransferase. MT catalyses the conversion of Mg-Proto to Mg-protoporphyrin monomethyl (MgPME) ester by transferring a methyl group from SAM to the C13 carboxyl group of Mg-protoporphyrin IX. * mark sites of chemical modification.

1.8.8 Mg-protoporphyrin monomethylester cyclase (cyclase)

The isocyclic ring E is formed by Mg-protoporphyrin monomethylester cyclase, which acts on MgPME and converts it to divinyl-protochlorophyllide (DV-Pchlide) (Figure 1.18). The closing of ring E induces a change in colour which shifts from the red of MgPME to the green of PChlide and Chl. The cyclase catalyses this reaction in three stages; firstly the hydroxylation of the C13 group to produce Mg-Protoporphyrin 6-methyl-β-hydroxypropionate, secondly the oxidation of the side chain to form Mg-Protoporphyrin 6-methyl-β-ketopropionate, and finally the formation of ring E by ligating the methylene group to the γ-meso carbon of the porphyrin ring. The enzyme uses NADPH, H^+ and O_2 as substrates at each of the three stages of catalysis (Wong *et al.*, 1985).

There are different classes of cyclase enzyme found in the various photosynthetic organisms. They differ in their requirement for an additional subunit and their dependency on oxygen. The oxygen independent cyclase consists of a single protein, encoded by the gene *bchE* in anaerobic bacteria, and utilises H_2O as a source of oxygen for activity (Hunter and Coomber, 1988; Naylor *et al*., 1999; Porra *et al*., 1996, 1998; Yang and Bauer, 1990). On the other hand, plants, cyanobacteria and some phototrophic bacteria possess an oxygen dependent cyclase that consists of two subunits called AcsF (Pinta *et al*., 2002) and Ycf54 (Chen *et al*., 2017; Hollingshead *et al*., 2012). This enzyme utilises oxygen and NADPH for catalysis (Ouchane *et al*., 2004; Pinta *et al*., 2002). The AcsF subunit is the catalytic unit of the protein although the mechanism of activity remains largely unknown. The function of Ycf54 within the cyclase is still a subject of debate but has been shown to be essential for cyclase activity in *Synechocystis* (Chen *et al*., 2017; Hollingshead *et al*., 2017) and barley (Bollivar *et al*., 2014) but not in other phototrophic bacteria, in which AcsF can operate alone (Chen *et al*., 2017). Ycf54 has been proposed to act in a structural capacity within the enzyme, stabilising the complex and promoting activity (Albus *et al*., 2012; Bollivar *et al*., 2014; Herbst *et al*., 2018). Additionally, in *A. thaliana*, Ycf54 was discovered to interact with FNR which could serve as an electron donor to AcsF in place of NADPH, providing it with the electrons required for catalysis whilst correlating Chl biosynthesis with photosynthetic activity (Herbst *et al*., 2018). Finally, another cyclase subunit was recently identified in Alphaproteobacteria and was termed BciE (Chen *et al*., 2017). Some bacteria contain both oxygen dependent and oxygen independent forms of the enzyme, enabling synthesis of BChl regardless of oxygen availability (Chen *et al*., 2016; Ouchane *et al*., 2004).

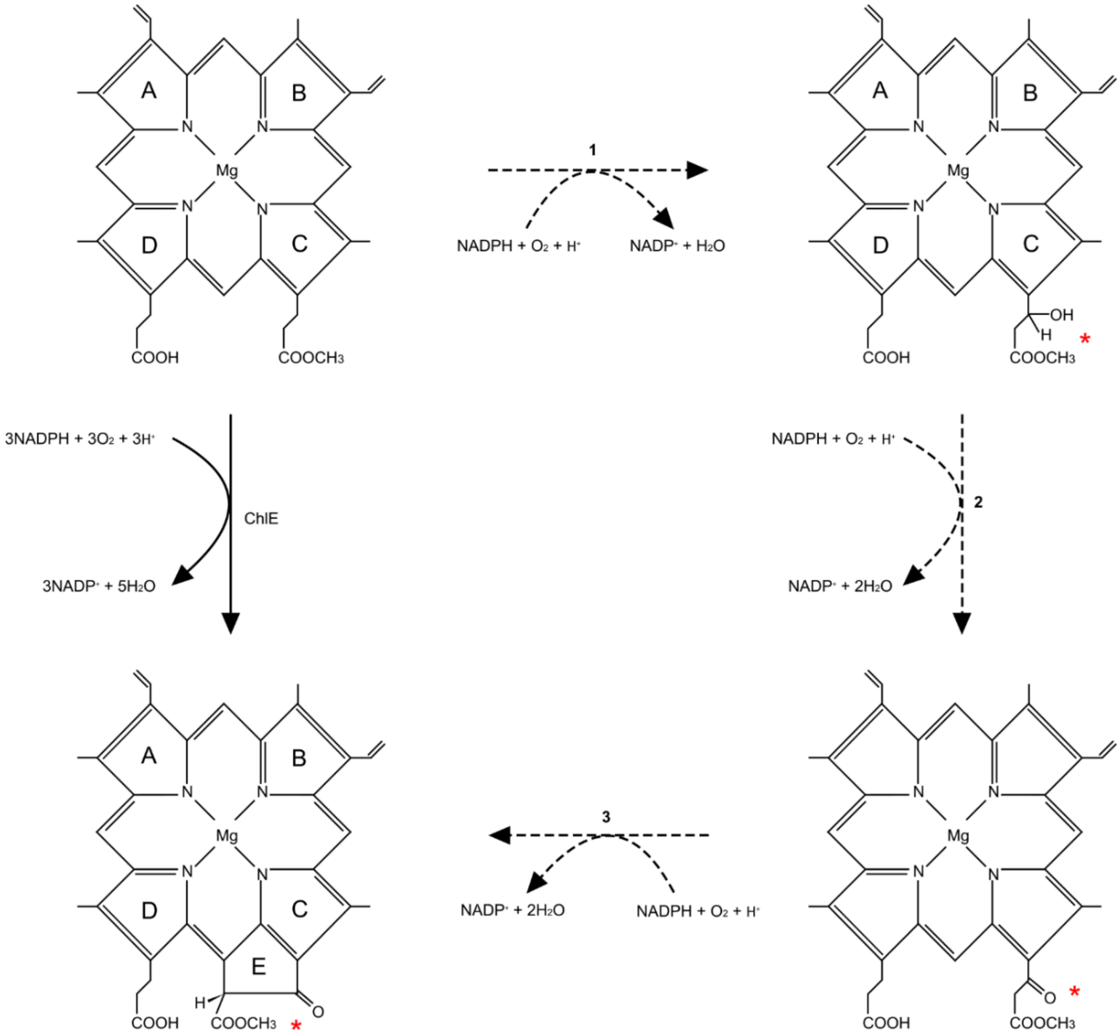


Figure 1.18: Reaction catalysed by Mg-protoporphyrin monomethylester cyclase. Formation of ring E of Chl is catalysed by ChlE, producing divinyl-protochlorophyllide (DV-PChlide) from Mg-PME. Dotted arrows indicate the reaction pathway proposed by Hollingshead *et al* (2012). * mark sites of chemical modification.

1.8.9 Protochlorophyllide (oxido)reductase

DV-Pchlide is reduced to DV-Chlorophyllide (DV-Chlide) by addition of hydrogen across the C17=C18 double bond in a chemically challenging and stereospecific reaction catalysed by one of two different enzymes; NADPH-protochlorophyllide oxidoreductase (LPOR) or protochlorophyllide reductase (DPOR) (Figure 1.19). LPOR has an absolute requirement for light to become active and is found in all Chl producing organisms, whereas DPOR can operate in the dark and is not found in angiosperms. BChl synthesising organisms appear to have only DPOR (Suzuki and Bauer, 1995). It is unclear why some organisms produce both LPOR and DPOR but this is probably due to the varying environmental conditions these organisms are exposed to. For example, one study demonstrated that PChlide reduction by LPOR increases with higher levels of light whereas DPOR contributes more under increasing oxygen tensions in the cyanobacterium *Leptolyngbya boryana* (Yamazaki *et al*., 2006).

1.8.9.1 LPOR

LPORs are oxygen independent and are theorised to have arisen in cyanobacteria around 2 billion years ago in response to the increasing oxygen concentration of earth's atmosphere (Yamazaki *et al*., 2006). The enzymes were subsequently inherited by plants although it was also obtained by some anaerobic bacteria probably by horizontal gene transfer (Kaschner *et al*., 2014).

LPOR catalyses the light driven hydride transfer from an NADPH co-enzyme to the C17 position of DV-Pchlide followed by proton transfer from a conserved tyrosine residue to C18 (Heyes and Hunter, 2002; Heyes *et al*., 2006, 2011, Menon *et al*., 2009, 2010). The enzyme undergoes a series of conformational changes associated with each of these steps (Heyes *et al*., 2007, 2008).

Only a single POR isoform has been identified within cyanobacteria whereas most phototrophs contain two isoforms, PORA and PORB, whilst *A. thaliana* has been found to contain a third isoform, PORC (Masuda and Takamiya, 2004). These appear to have no phylogenetic relationship and most likely arose independently from gene duplication events within individual species (Masuda and Takamiya, 2004). The expression of these isoforms seems to be differentially regulated (both spatially and temporally in multicellular organisms) depending on the light intensity; however the details of this regulation differs amongst photosynthetic organisms. For example, in rice the PORA isoform is expressed only early in leaf development when conditions are dark. However PORB is expressed constitutively throughout development (Kwon *et al*., 2017). It was suggested that PORA and PORB therefore became functionally distinct to aid the plant in adapting to the environment throughout the course of evolution (Kwon *et al*., 2017). Similarly the three POR isoforms in *A. thaliana* are regulated in response to varying light intensity (Su *et al*., 2001) and it is therefore believed that POR expression is regulated by circadian and diurnal rhythms to enable the adjustment of Chl production to satisfy the requirements of the cell during the light and dark periods of the day (Masuda and Takamiya, 2004). Finally, all three POR isoforms in *A. thaliana* were shown to oligomerise to increase activity (Gabruk *et al*., 2015). It is unknown whether or not this behaviour is also regulated.

1.8.9.2 DPOR

DPOR consists of three individual subunits denoted ChlL/BChl, ChlN/BchN and ChlB/BchB in Chl/BChl producing organisms respectively. The protein components of DPOR were discovered independently in 3 different *Rba. capsulatus* strains due to mutations at the respective loci of *bchL* (Yang and Bauer, 1990), *bchN* (Coomber *et al*., 1990) and *bchB* (Burke *et al*., 1993a). These genes have a high degree of similarity to the *nifH*, *nifD* and *nifK* genes that correspond to the three protein subunits of nitrogenase. NifH forms a homodimer (NifH[2]) using two cysteine residues from each partner to coordinate binding of a [4Fe-4S] cluster whilst NifD and NifK form a heterodimer (NidD-NifK) capable of forming a molybdenum-iron (MoFe) cluster. NifH[2], via its [4Fe-4S] acts as an ATP-dependent electron donor to the MoFe of NifD-NifK. This electron is then used to reduce N_2 to NH_3.

It has been documented that DPOR functions similarly to nitrogenase. ChlL can form a homodimer, [ChlL], whilst ChlN and ChlB complex to form a heterotetramer[ChlN/ChlB]$_2$. [ChlL]$_2$ binds 2 molecules of ATP and has a shared 4Fe-4S cluster that enables it to function as an ATP-dependent electron donor (Bröcker *et al*., 2008a, 2008b, 2010a). Meanwhile the [ChlN/ChlB]$_2$ group can bind the substrate (Bröcker *et al*., 2008b). These two complexes come together and, following reception of a single electron from Fd (Bröcker *et al*., 2008a), [ChlL]$_2$ transfers an electron to [ChlN/ChlB]$_2$ and hydrolyses the two ATP molecules before dissociating from

[ChlN/ChlB]$_2$. A second [ChlL] sub-complex binds [ChlN/ChlB]$_2$ and the reaction repeats to complete reduction of DV-Pchlide to produce DV-Chlide. This reaction therefore consumes four molecules of ATP (Bröcker *et al*., 2008a; Nomata *et al*., 2016). It has been reported that [ChlN/ChlB]$_2$ can associate with simultaneously with two [ChlL]$_2$ complexes to form a transient hetero-octameric holoenzyme (Bröcker *et al*., 2010b).

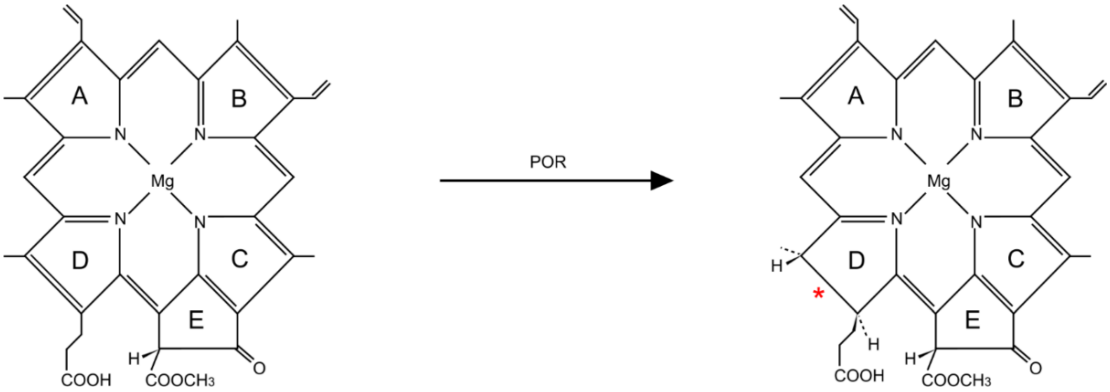


Figure 1.19: Reaction catalysed by protochlorophyllide (oxido)reductase. POR catalyses the reduction of the C17=C18 double bond of DV-Pchlide, producing divinyl-chlorophyllide (DV-Chlide). * mark sites of chemical modification.

1.8.10 C8-vinyl reductase (8VR)

8VR acts on the C8 vinyl group (8V) of the DV-Chlide macrocycle and reduces it to an ethyl group (8E) using NADPH as a reductant and restricting the π-electron system to the chlorin ring to produce MV-chlorophyllide (MV-Chlide) (Parham and Rebeiz, 1995) (Figure 1.20). The enzyme was discovered when a mutant *A. thaliana* strain, AT5G18660, accumulated 8V-Chl, indicating that the gene was an 8VR (Nagata *et al*., 2005; Nakanishi *et al*., 2005). Subsequently, the gene was expressed recombinantly in *E.coli* and *in vitro* assays demonstrated that the protein was able to produce 8E from 8V-chlide, contributing to this notion. Bioinformatics techniques led to the identification of homologs in rice (Wang *et al*., 2010), *Chlorobaculum tepidum* (Chew and Bryant, 2007) and *Rba. sphaeroides* (Canniffe *et al*., 2013). This gene was annotated as *bciA*. Although cyanobacteria do not have a *bciA* orthologue, they contain a second class of 8VR designated *cvrA* (Islam *et al*., 2008; Ito *et al*., 2008). There appears to be no relationship between 8VR and *cvrA* in terms of its distribution across phototrophic organisms and so these genes most likely evolved independently (Ito *et al*., 2008).

Plants contain both *cvrA* and *bciA* homologs in their genomes, although deletion of *bciA* in *A. thaliana* perturbed 8VR activity completely (Islam *et al*., 2008). This suggests that *cvrA* either has negligible activity, is expressed differentially depending on the tissue and/or the developmental stage of the plant, or *cvrA* no longer functions as an 8VR (Islam *et al*., 2008). In addition to this, a third class of 8VR was discovered when *bciA* was perturbed in *Rba. sphaeroides*. This protein was encoded by three genes; *bchX*, *bchY* and *bchZ,* which also function in the modification of Chlide to bacteriochlorophyllide (BChlide) (Tsukatani *et al*., 2013).

There has been much debate over whether or not 8VR acts before or after POR in the biosynthesis of Chl. It is generally agreed that 8VR acts on DV-protochlorophyllide (before POR) as its preferred substrate due to the fact that MV-PChlide, the substrate of POR, accumulates when angiosperms are grown in the dark. Despite this, almost every other Chl intermediate has been reported as a substrate of 8VR in the literature (Rebeiz *et al*., 1999). *In vitro* enzyme assays and careful monitoring of accumulation of Chl intermediates in *A. thaliana* seedlings revealed that the preferred route of 8VR, at least in *A. thaliana*, is that POR acts first, reducing DV-Pchlide to DV-Chlide, followed by the swift reduction of DV-Chlide to MV-Chlide by 8VR (Nagata *et al*., 2007). The route of Chl synthesis at this stage may differ between organisms, or differ based on other factors yet to be discerned.

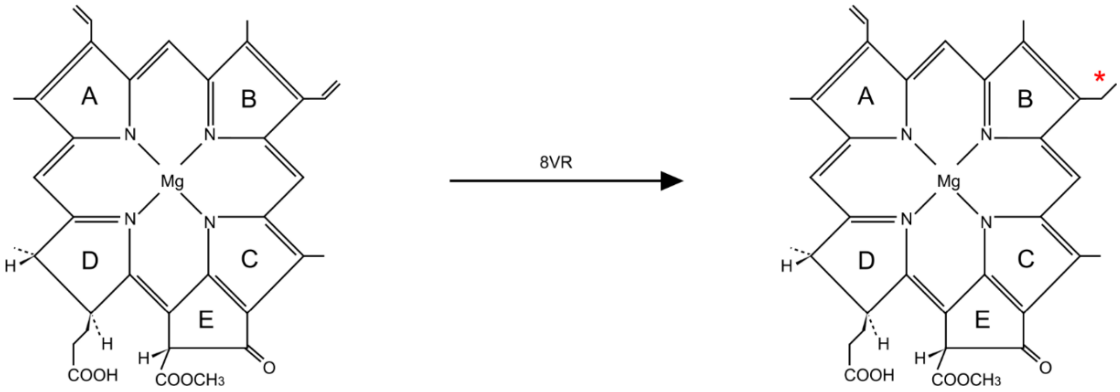


Figure 1.20: Reaction catalysed by C8-vinyl reductase. 8VR reduces the C8-vinyl group of DV-Chlide to an ethyl group, yielding chlorophyllide (Chlide). * mark sites of chemical modification.

1.8.11 Chlorophyll Synthase (ChlG)

The integral membrane protein ChlG, also known as chlorophyll synthetase in eukaryotic organisms, is a class II transferase enzyme (2007) that catalyses the esterification of Chlide with geranylgeranyl pyrophosphate (GGPP) or phytyl pyrophosphate (PPP) to the C17 position within ring D of the macrocycle to produce Chl_{GG} or Chl_{phytol} (Chl) respectively (Figure 1.21). The phytylation of Chlide with the tail moiety increases the hydrophobicity of the pigment, enabling it to be anchored into Chl binding proteins (Rudiger, 1993).

The esterification of Chlide was first reported in etiolated bean leaves (Ogawa *et al.*, 1975) and was attributed to catalysis by the enzyme chlorophyllase, which can esterify small amounts of Chlide with phytol, requiring acetone for activity *in vitro*, in addition to its primary activity as a dephytylase (removal of the tail from Chl) (Böger, 1965; Ogawa *et al.*, 1975; Shimizu and Tamaki, 1962, 1963; Willstatter, Richard ; Stoll, 1928). However it was subsequently found that another enzyme, prepared from maize shoots, was capable of esterification of Chlide in the absence of acetone (Rüdiger *et al.*, 1977). Activity of this enzyme was demonstrated in oat seedlings, when it was determined that esterification of Chlide with geranylgeraniol (GG) required exogenous ATP, whereas esterification with geranylgeranyl pyrophosphate (GGPP) did not. This enzyme was named chlorophyll synthase to distinguish it from chlorophyllase (Rüdiger *et al*., 1980). As was the case in maize shoots, ChlG did not require acetone for activity and converted over 90% of the total Chlide into Chl. In contrast, chlorophyllase converted just 1-15% of the Chlide pool (Ellsworth, 1971). The gene encoding ChlG was identified when a mutation in ORF304 of *Rba. capsulatus* caused the cell to accumulate BChlide (Bollivar *et al*., 1994b). This gene was annotated *bchG* and its activity *in vitro*, along with the *Synechocystis* homologue *chlG* (Lopez *et al*., 1996), was demonstrated by heterologous expression in *E. coli* followed by enzyme assays (Oster *et al*., 1997).

ChlG activity has been detected in mature chloroplasts isolated from greening oat seedlings (Rudiger, 1993), the stromal fraction of daffodil chromoplasts (Kreuz and Kleinig, 1981), paprika chloroplast and chromoplast (Camara, 1984), the prothylakoid and pro-lamellar fractions of etioplasts (Lütz *et al*., 1981; Rudiger, 1993) and the TM of spinach (Block *et al*., 1980; Gaubier *et al*., 1995; Soll *et al*., 1983). Activity was detected specifically in the inner membranes of plastids (Lindsten *et al*., 1990; Soll *et al*., 1983). The activity of ChlG in dark grown wheat transfers upon illumination from the prolamellar body (PLB) fraction to the prothylakoid fraction during plant development, indicating a migration of ChlG from the PLB to the developing TM during greening (Lindsten *et al*., 1993).

GGPP can be directly esterified to Chlide to from Chl_{GG} and then reduced by geranylgeranyl reductase (GGR) to form Chl (Addlesee *et al*., 1996; Chew and Bryant, 2007). Alternatively, free GGPP can be reduced by GGR to phytyl pyrophosphate (PPP) and then esterified to Chlide by ChlG (Chew *et al*., 2007). Whether GGR reduces GGPP before or after esterification to Chlide depends on the location of the substrate within the cell. Free GGPP reduction to PPP occurs in the envelope membrane of chloroplasts whereas Chl_{GG} conversion to Chl occurs in the TM where Chlide is produced (Soll *et al*., 1983). GGR is discussed in greater detail in Section 1.8.12.

The substrate specificity of ChlG for GGPP or PPP appears to depend on the growth phase of the organism and the species in question. In etiolated oat seedlings, ChlG has a preference for GGPP over PPP in a 2:1 ratio (Rüdiger *et al*., 1980; Soll and Schultz, 1981). However when this same protein was expressed recombinantly in *E.coli* it accepted both substrates equally (Schmid *et al*., 2002). The *A. thaliana* homolog which was recombinantly expressed in *E.coli* also preferred GGPP over PPP and accepted Chlide *b* as well as Chlide *a* (Oster and Rüdiger, 1997). However, the spinach homolog showed a preference for PPP over GGPP (Soll *et al*., 1983). Prokaryotic organisms, including *Rba. capsulatus* and *Synechocystis*, had a preference for PPP *in vitro* (Oster *et al*., 1997). ChlG cannot utilise BChlide as a substrate and BchG cannot utilise Chlide (Kim and Lee, 2010; Oster *et al*., 1997; Schoch *et al*., 1999). These enzymes exhibit competitive inhibition when given the "wrong" substrate *in vitro* and so the active sites of the enzymes are presumably similar (Kim and Lee, 2010). In this respect, it was shown that the I44 residue of ChlG was instrumental in determining the substrate specificity of the enzyme. When ChlG I44 was changed to F, the enzyme was capable of restoring photoautotrophic growth to a *Rba. sphaeroides* strain lacking *bchG*, indicating that the mutant ChlG enzyme was able to utilise BChlide as a substrate and produce BChl (Kim *et al*., 2016). Neither PChlide nor pheophorbide are substrates for ChlG (Rudiger, 1993). From these observations it is possible to determine that, for activity, ChlG requires the central magnesium ion of Chlide (Schmid *et al*., 2001), that the propionic group of ring D must be raised above the tetrapyrrole plane (which is not the case in PChlide) and that ring B must not be hydrogenated (ring B of BChlide is hydrogenated) (Rüdiger, 1992; Rüdiger *et al*., 1980).

ChlG catalysis proceeds via a ping-pong mechanism in which GGPP (or PPP) binds first to the enzyme and causes a conformational change in ChlG, allowing it to bind Chlide, the second substrate (Schmid *et al*., 2002). Residues 88 to 377 are catalytically active with two R residues (R91 and R161) and one C residue (C109) having been demonstrated to be critical for the enzyme's activity (Schmid *et al*., 2001). Esterification of Chlide is also temperature dependent with Chlide a conversion being 8 fold higher in plants exposed to 28°C compared with 0°C (Rudiger, 1993). This may be due to handicapping the diffusion of GGPP towards ChlG in the lipid bilayer, which is less fluid at low temperatures.

Two phases of Chlide esterification have been described in etiolated barley leaves. Upon brief illumination of the leaves, an initial rapid phase lasting 15-30 seconds produced a consistent amount of Chlide esterification, independently of the extent to which POR had reduced the protochlorophyllide pool to Chlide. This was followed by a second slow phase lasting 40 minutes. (Domanskii and Rüdiger, 2001). This was further investigated in a follow up study where it was calculated that the rapid phase converts just 15% of the Chlide pool, followed by a lag phase of 2 minutes before 85% of the remaining Chlide is converted over the course of an hour. Following incubation in the dark for 10 minutes, the ability to initiate a rapid phase was restored. The prolamellar bodies (PLB), intact during the fast phase, disaggregated during the slow phase (Domanskii *et al*., 2002). In summary it was proposed that esterification of Chlide is restricted by the limited capacity of ChlG, the time it takes for POR to release *de novo* Chlide, the diffusion of GGPP/PPP towards ChlG and by the disruption caused by the disaggregation of the PLB.

ChlG knock down and overexpression in tobacco (*Nicotiana tabacum*) decreased and increased the transcript levels of MgCH respectively. ALA synthesis was also altered in the same manner. This demonstrated that ChlG expression influences the expression of upstream Chl biosynthesis genes, indicating a role of ChlG in co-regulating the Chl biosynthesis pathway (Shalygo *et al*., 2009). In rice (*Oryza Sativa*) seedlings a ChlG mutant (*ygl1*) with reduced activity showed altered expression of some nuclear genes encoding Chl biosynthesis enzymes, similarly implicating a role of ChlG in feedback regulation of Chl biosynthesis (Wu *et al*., 2007).

At the level of translation, Chl binding proteins P700, CP43, CP47 and D1 only accumulated if *de novo* Chl was synthesised by ChlG and not if exogenous Chl was added to the translation mixture (Eichacker *et al*., 1990; Kim *et al*., 1994a). This implies that ChlG is required to channel the *de novo* Chl molecules to nascent polypeptides, thereby stabilising them (Eichacker *et al*., 1996). Subsequently a ChlG complex was discovered in *Synechocystis* and was proposed to deliver Chl co-translationally to Chl binding proteins as they are being inserted into the TM (Chidgey et al., 2014). The ChlG complex is discussed in detail in Section 1.11.

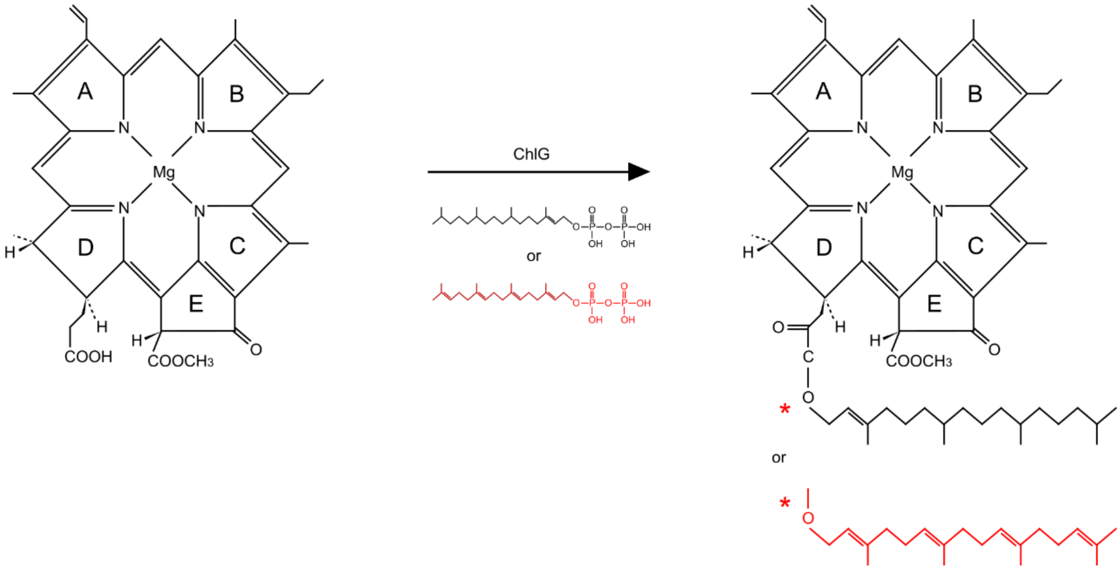


Figure 1.21: Reaction catalysed by chlorophyll synthase. ChlG catalyses the esterification of phytyl or GGPP (red) with the carboxyl group of ring D of the Chlide macrocycle, producing Chl *a*$_{GG}$ (red) or Chl *a* respectively. * mark sites of chemical modification.

1.8.12 Geranylgeranyl reductase (GGR)

GGR catalyses the NADPH and ATP dependent reduction of 3 of the 4 C=C double bonds of geranylgeranyl pyrophosphate to produce phytyl pyrophosphate (Bollivar *et al.*, 1994c) (Figure 1.22). This can occur before or after the molecule has been esterified with Chlide by ChlG (Chew *et al.*, 2007). Esterification of Chlide with GGPP results in Chl$_{GG}$, whereas esterification of Chlide with PPP yields mature Chl (Keller *et al.*, 1998). Genes encoding GGR have been identified in oxygenic phototrophs including cyanobacteria (Addlesee *et al.*, 1996) and plants (*chlP*) (Giannino *et al.*, 2004; Keller *et al.*, 1998; Tanaka *et al.*, 1999; Wang *et al.*, 2014), as well as in purple bacteria (*bchP*) (Addlesee and Hunter, 1999). Deletion of *chlP* from *Synechocystis* results in the accumulation of Chl molecule with partially reduced tail moieties which are incorporated into Chl-binding proteins and can still function in light harvesting (Shibata *et al.*, 2004; Shpilyov *et al.*, 2005; Tanaka *et al.*, 1999). PSI and PSII are functional in these mutants, however photoautotrophic growth is abolished due to the rapid degradation of the photosystems without supplementing the cell with glucose

(Shpilyov *et al*., 2005, 2013). The same growth phenotype is observed in plants lacking *chlP* (Shibata *et al*., 2004; Tanaka *et al*., 1999) in addition to an increase in sensitivity to high-light stress (Grasses *et al*., 2001). The reduced stability of the photosystems induced by integration of partially reduced Chl co-factors into the complex has been attributed to the increased rigidity of the Chl_{GG} species, due to the extra 3 C=C double bonds which may disrupt the assembly of the complexes (Shpilyov *et al*., 2005, 2013). Although purple bacteria harbouring $BChl_{GG}$ had reduced stability of the RCs (Bollivar *et al*., 1994c), akin to the situation in cyanobacteria, they are still able to grow photoautotrophically albeit at a slower rate than WT strains (Addlesee and Hunter, 1999; Harada *et al*., 2008). The full reduction of the Chl tail moiety by GGR therefore appears to be more important in oxygenic phototrophs. Full reduction of GGPP has also been implicated to be important for mediating the interactions between neighbouring Chl molecules, enabling the efficient transfer of absorbed light energy to the photosystem RC (Shpilyov *et al*., 2013).

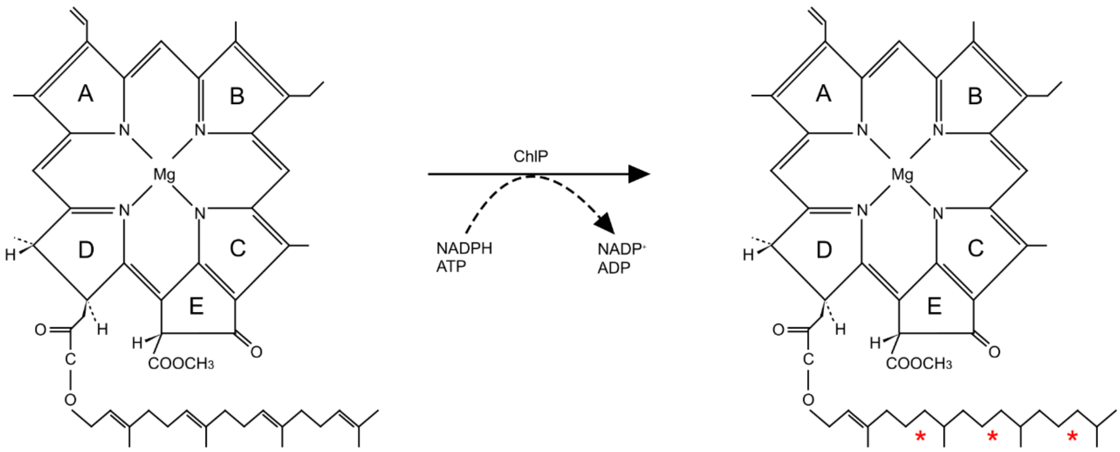


Figure 1. 22: Reaction catalysed by geranylgeranyl reductase. ChlP reduces the phytyl tail of Chl a_{GG} to produce Chl *a*. ChlP can also catalyse the reduction of GGPP to phytyl before the molecule is attached to Chlide by ChlG (not shown). * mark sites of chemical modification.

1.8.13 Bacteriochlorophyll-specific modifications

The pathway of BChl biosynthesis is analogous to the Chl a route up until, and including, the synthesis of Chlide. Following this, 3 additional modifications are made to Chlide in organisms that synthesise BChl. Firstly, the C7=C8 double bond of ring B is reduced by Chlide a reductase (COR) to produce 3-vinyl-bacteriochlorophyllide (3V-BChlide). 3V-BChlide is then converted to BChlide when the C3-vinyl group is modified to an acetyl group. Esterification of the tail moiety to BChlide by BchG completes the BChl molecule (Figure 1.23).

COR is encoded by three genes designated *bchX*, *bchY* and *bchZ* (Burke *et al*., 1993b; McGlynn and Hunter, 1993).These genes were over-expressed and purified from *Rba. capsulatus* and their reductase activity demonstrated in vitro as determined by the formation of 3V-BChlide from Chlide in the presence of co-factors ATP and the reducing agent dithionite (Nomata *et al*., 2006).

The protein products of *bchF* and *bchC* catalyse the reduction of C3-vinyl to acetyl in two subsequential steps starting with hydroxylation of the vinyl group by *bchF* followed by oxidation of the hydroxyl group to acetyl by *bchC* (Taylor *et al*., 1983; Zsebo and Hearst, 1984).

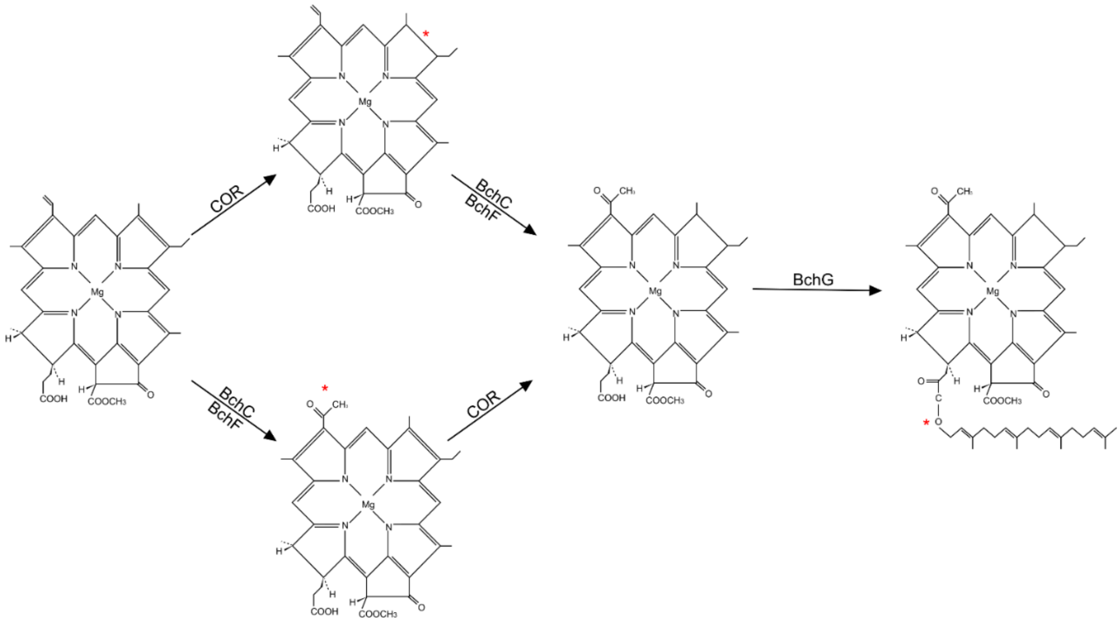


Figure 1.23: Bacteriochlorophyll-specific modifications. Chl is modified to Bacteriochlorophyll (BChl) by reduction of the C7=C8 double bond of Chlide by COR, producing 3-vinyl-bacteriochlorophyllide (3V-BChlide). Alternatively BchF and BchC can act first to convert the C3-vinyl group to an acetyl group producing 3-hydroxyethyl-chlorophyllide. Bacteriochlorophyllide (BChlide) is produced as a result of both modifications, followed by esterification of the phytyl tail to BChlide by BchG, yielding BChl. * mark sites of chemical modification.

1.9 Carotenoids

Carotenoids are the most widespread group of pigment molecules (Cazzonelli, 2011) and are produced in all photosynthetic organisms. They consist of an extended polyene chain containing between 9 and 11 double bonds, feature a delocalised π electron system and, in oxygenic phototrophs, often contain rings at each end of the molecule (Figure 1.24). They are divided into two groups; xanthophylls, which contain oxygen, and carotenes, which do not (Tóth *et al.*, 2015). Carotenoids are capable of absorbing light between the wavelengths of 450 and 570 nm and, as hydrophobic pigments, are most often found embedded in the TM lipid bilayer where they contribute to the structural integrity of the TM (Mohamed *et al.*, 2005). They are also important for the structure, synthesis and assembly of photosystems (Santabarbara *et al*., 2013; Sozer *et al*., 2011).

In cyanobacteria, carotenoids function as photoprotectors of Chl binding proteins, by quenching excess light absorbed by Chl, as well as act as light harvesting pigments in their own right (Bryant, 1994; Cazzaniga *et al*., 2012; Croce and van Amerongen, 2014; Schafer *et al*., 2006; Stamatakis *et al*., 2014). Regarding their role in photoprotection, carotenoids can quench triplet excited states of neighbouring Chl molecules by absorbing the energy and distributing it across their π electron system in order to find the lowest possible energy state (Vershinin, 1999). This prevents the generation of damaging reactive oxygen species (ROS) that arise if molecular oxygen reacts with triplet state Chl. In addition, carotenoids are able to quench ROS if they are formed. It is generally agreed that photoprotection is provided primarily by the carotenoids zeaxanthin and myxoxanthophyll in cyanobacteria (Kusama *et al*., 2015; Masamoto *et al*., 1999; Steiger *et al*., 1999).

Carotenoids are synthesised from farnesyl pyrophosphate (FPP) which is condensed with 5-carbon isoprenoid molecules to produce a 20-carbon geranylgeranyl pyrophosphate (GGPP) molecule. This is catalysed by the enzyme geranylgeranyl pyrophosphate synthase (*crtE*) (Zhu *et al*., 1997). Two GGPP molecules are condensed to form a 40-carbon phytoene compound by phytoene synthase (*crtB*). A phytoene desaturase can perform two (*crtP*) (Martínez-Férez *et al*., 1994), three or four (*crtI*) (Harada *et al*., 2001) desaturations to produce ζ- carotene, neurosporene or lycopene respectively from phytoene. Cyclisation of the ends of lycopene by either lycopene beta cyclase (*crtL*-b) or lycopene epsilon cyclase (*crtL*-e) produce the rings characteristic of β-carotene and α-carotene respectively. Further modification to the β-carotene produces zeaxanthin, catalysed by beta carotene hydroxylase (*crtR*) (Liang *et al*., 2006).

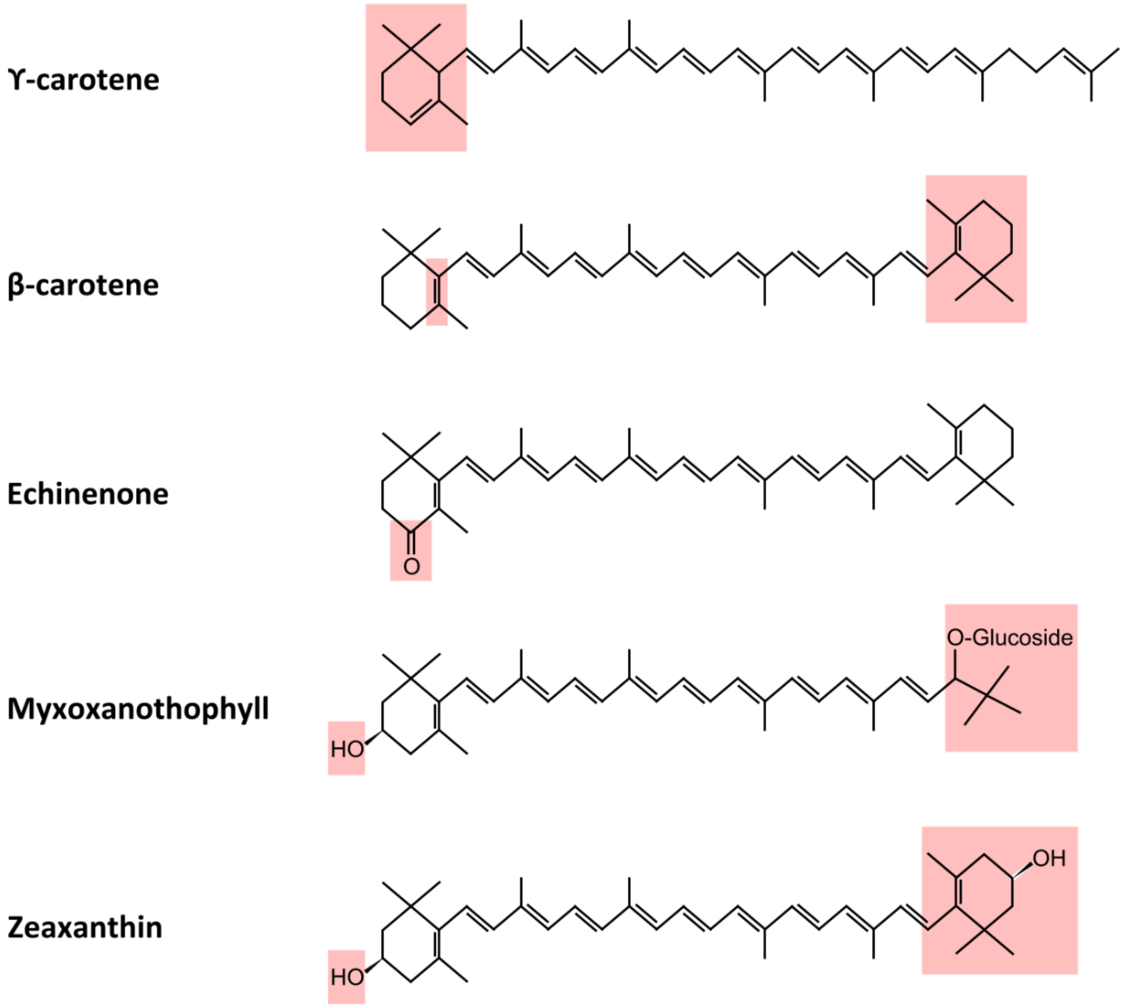


Figure 1.24: Carotenoid structures. The structures of the carotenoids produced by *Synechocystis*. Red boxes highlight the structural differences between the carotenoids.

1.10 Chlorophyll delivery/recycling during PSII metabolism

PSII assembly (Section 1.6.1) and delivery of *de novo* Chl molecules are believed to be closely coordinated to allow the efficient integration of Chl into Chl binding apoproteins co-translationally, as they are being inserted into the lipid bilayer. The insertion of Chl is essential for the correct folding of the receiving proteins (Eichacker *et al*., 1996; Müller and Eichacker, 1999) and to prevent the production of damaging reactive oxygen species generated by free Chl (Krieger-Liszkay *et al*., 2008). The availability of newly synthesised Chl also determines the rate of synthesis of Chl binding proteins (Kopečná *et al*., 2013) via regulation at the level of translation (He and Vermaas, 1998).

To reduce the time between Chl synthesis and insertion into apoproteins, it is likely that the enzymes of the Chl biosynthesis pathway are localised close to the site of PSII assembly in the TM biogenesis centres. In support of this, the protein Pitt, involved in TM membrane biogenesis, is able to interact with POR and is believed to localise it to the site of PSII assembly (Schottkowski *et al*., 2009). *Synechocystis* cells lacking the PSII assembly and repair factor Psb28 had increased rates of PSII repair and decreased PSI content. In addition, the Chl synthesis pathway was perturbed at the cyclase step as indicated by an accumulation of the cyclase substrate protoporphyrin IX methylester. This led to the proposal that delivery of Chl to Chl binding proteins, such as CP43 and PSI subunits, was impeded, implying a role for Psb28 in Chl binding to apoproteins (Dobáková *et al*., 2009).

During PSII repair, Chl is released from PSII and is recycled, firstly by removal of the tail to form Chlide, before both products are reintegrated back into the *de novo* Chl biosynthesis pathway (Vavilin and Vermaas, 2007). High-light inducible protein (Hlips) contain a highly conserved Chl *a* binding (CAB) domain and are believed to store the Chl molecules once they have been released from damaged PSII (Vavilin *et al*., 2007). A complex consisting of Ycf39, HliD and HliC was shown to be important to the formation of the PSII (Knoppová *et al*., 2014). Two members of the Hlip family, HliD and HliC, are proposed to have a role in photoprotection of PSII under stressful light conditions (Promnares *et al*., 2006) and scavenging of free Chl released during PSII repair (Yao *et al*., 2007). The Ycf39-HliP complex appears to be involved in delivery of Chl to *de novo* D1 subunits during PSII assembly/repair; not only does the complex co-purify with PSII intermediate complexes but a *Synechocystis* YCf39 null mutant was perturbed in Chl recycling (Knoppová *et al*., 2014). Moreover, Ycf39, HliD and HliC were found to co-purify with ChlG and YidC insertase, further lending support to a role of these proteins in delivery of Chl to PSII subunits as they are being co-translationally inserted into the membrane (Chidgey et al., 2014). The ChlG complex is the focus of this thesis. The following sections summarise the existing research in this area.

1.11 The chlorophyll synthase complex

In cyanobacteria, the Chl binding proteins of PSI and PSII are synthesised on TM bound ribosomes and co-translationally inserted into the photosystem complex (Frain *et al*., 2016). It is reasonable to postulate that the enzymes of the Chl biosynthesis pathway and the protein machinery responsible for photosystem assembly are co-ordinated to allow efficient Chl insertion into *de novo* photosystem polypeptides. The final enzyme in the Chl biosynthesis pathway, ChlG, forms a complex in the cyanobacteria *Synechocystis* consisting of ChlG, the high-light inducible proteins D and C (HliD/HliC), the PSII assembly factor Ycf39 and the membrane insertase YidC (Chidgey et al., 2014; Niedzwiedzki et al., 2016).

1.11.1 High-light inducible proteins

All oxygenic phototrophs, in addition to light harvesting complexes, contain light harvesting like (LIL) proteins (Engelken *et al*., 2010; Jansson, 2008; Neilson and Durnford, 2010a). This group includes the one-helix proteins (OHP) (Lil2/6) and the early light-induced proteins (ELIP) (Lil1) of plants as well as the high-light induced proteins (Hlips) found in cyanobacteria. All feature a conserved Chl a/b binding (CAB) domain reminiscent of that found in LHCs (Adamska *et al*., 1999; Funk, 2001; Staleva *et al*., 2015; Storm *et al*., 2008). The Lil proteins function in the photoprotection, assembly and regulation of the photosystems (Beck *et al*., 2017; Hutin *et al*., 2003; Lohscheider *et al*., 2015; Niyogi *et al*., 2004; Peers *et al*., 2009) as well as regulation of pigment biosynthesis (Rossini *et al*., 2006; Sobotka *et al*., 2008; Takahashi *et al*., 2014; Tanaka *et al*., 2010; Xu *et al*., 2002) and pigment carriers/stabilisers (Hernandez-Prieto *et al*., 2011; Knoppová *et al*., 2014; Yao *et al*., 2007). The HliPs were shown to significantly increase the half-life of Chl in *Synechocystis*, lending further support to a role of these proteins in recycling Chl by binding free pigment, released by PSII as the complex is being disassembled for repair, and shepherding it into the dephytylationphytylation cycle before it is incorporated back into Chl binding proteins (Vavilin *et al*., 2007).

Cyanobacteria contain five HliPs designated HliA-HliD (Dolganov *et al*., 1995) with the 5^{th} being fused to the C-terminus of ferrochelatase (He *et al*., 2001; Kufryk *et al*., 2008). They are single helix integral membrane proteins and are believed to be the ancestors of plant LHC proteins (Neilson and Durnford, 2010b; Yurina *et al*., 2013). Upregulated under high-light intensities, the HliPs function to photoprotect the photosystems (and and Vermaas, 1999; Dolganov *et al*., 1995) and are important to the survival of the cell under stressful conditions, including low temperature and nutrient starvation (He *et al*., 2001; Mikami *et al*., 2002). Cyanobacteria cells that lack HliA-HliD have increased sensitivity to increasing light intensities and were unable to quench excess light absorbed by nearby Chl pigments (Havaux *et al*., 2003). In addition, HliPs can scavenge free Chl molecules that are dangerous to the cell due to their tendency to generate damaging ROS (Xu *et al*., 2004). Finally, HliPs have a role in the regulation of Chl biosynthesis (Havaux *et al*., 2003; Xu *et al*., 2004).

In *Synechocystis*, HliA and HliB are functionally complementary (Kufryk *et al*., 2008; Xu *et al*., 2004) and, due to their high degree of sequence simililarity, are thought to have arisen from a gene duplication event (He *et al*., 2001; Kufryk *et al*., 2008). Loss of both of these proteins was found to be lethal to the cell in high-light conditions (Wang *et al*., 2008). Both were found to associate with, and stabilise, PSI trimers (Akulinkina *et al*., 2015; Wang *et al*., 2008) and monomers as well as the CP47 subunit of PSII (Yao *et al*., 2007) in response to light stress.

HliC and HliD form homodimers *in vivo*; the former binds to four Chl and two β-carotene pigments (Shukla *et al*., 2018a). whereas HliD binds 6 Chl molecules and 2 β-carotene (Niedzwiedzki *et al*., 2016; Staleva *et al*., 2015) (Figure 1.25). HliD binds these pigments in a configuration that allows transfer of absorbed light energy from Chl to β-carotene which dissipates the energy as heat (Llansola-Portoles *et al*., 2017; Staleva *et al*., 2015). Both HliPs form a complex with the putative short-chain dehydrogenase Ycf39 which associates with *de novo* PSII RC, RC47 or the isolated CP43 subcomplex before it is incorporated into RC47 (Komenda and Sobotka, 2016). It is postulated that

Ycf39-HliD/HliC bind to these PSII intermediates to convey photoprotection to the proteins, stabilising them as they are assembled during *de novo* PSII biogenesis or PSII repair.

Both HliD and HliC form part of a larger complex with ChlG, Ycf39 and the YidC insertase (Chidgey et al., 2014; Niedzwiedzki et al., 2016). Chidgey *et al.* (2014) found that ChlG and HliD form a stable core to this complex to which the other components can bind. It was shown that the carotenoids bound to the HliPs are able to effectively quench light absorbed by this complex and prevent its damage (Niedzwiedzki *et al*., 2016). In addition to the Hlip bound β-carotene, the ChlG-HliP complex also contained myxoxanthophyll and zeaxanthin, the binding of which was attributed to ChlG (Chidgey et al., 2014; Niedzwiedzki et al., 2016). The function of these carotenoids was hypothesised to be of a structural nature, acting at the interface between ChlG and the HliPs, facilitating their binding (Niedzwiedzki *et al*., 2016). HliC has recently been shown to mediate the remodelling of the ChlG complex by facilitating release of Ycf39 from the complex in response to high-light stress (Shukla *et al*., 2018b). The authors propose that the Ycf39, once released from ChlG, forms a complex with an HliC-HliD heterodimer, producing a trimeric pigment-protein complex (Ycf39-HliD-HliC) that had been previously identified (Knoppová *et al*., 2014). The ChlG-HliD "core" complex (Chidgey et al., 2014) is able to then bind to PSI trimers and facilitate recycling of Chl molecules released during PSII repair, using PSI as a store of these pigments, whilst the Ycf39-HliD-HliC complex can bind to PSII repair intermediates and photoprotect them (Shukla *et al*., 2018b).

In plants, the *lil* genes are upregulated in response to stressful high-light conditions, in negative correlation to the *LHC* genes which are reduced (Klimmek *et al*., 2006). This was shown to be the case in *A. thaliana* when *OHP1* (Jansson *et al*., 2000) and *OHP2* (Andersson *et al*., 2003) expression increased in response to high-light intensity. OHP1 associates with, and photoprotects, the RC of PSII and PSII assembly proteins during the biogenesis and repair of the complex (Myouga *et al*., 2018), whereas OHP2 was observed to accumulate in PSI (Andersson *et al*., 2003). Recently, OHP1 was found to dimerise with OHP2, akin to the HliD-HliC heterodimer of *Synechocystis*, and stabilise the protein HCF244, the *A. thaliana* homologue of Ycf39 in *Synechocystis* (Section 1.11.3) (Hey and Grimm, 2018). Chidgey *et al.* (2014) demonstrated that HliD and HliC form a complex with Ycf39 and bind ChlG and YidC. However, no ChlG was detected in the eqfuivalent OHP1/OHP2-HCF244 complex in *A. thaliana*; although an interaction of ChlG with OHP-HCF244 may exist *in vivo,* detergent solubilisation could have removed ChlG from the complex.

Knockout of OHP2 was found to perturb Chl biosynthesis and destabilise the PSII core. In addition, HCF244 production was completely abolished in OHP2 null strains, leading to the conclusion that free HCF244 is unstable and requires anchoring to the TM by OHP2 (Hey and Grimm, 2018).

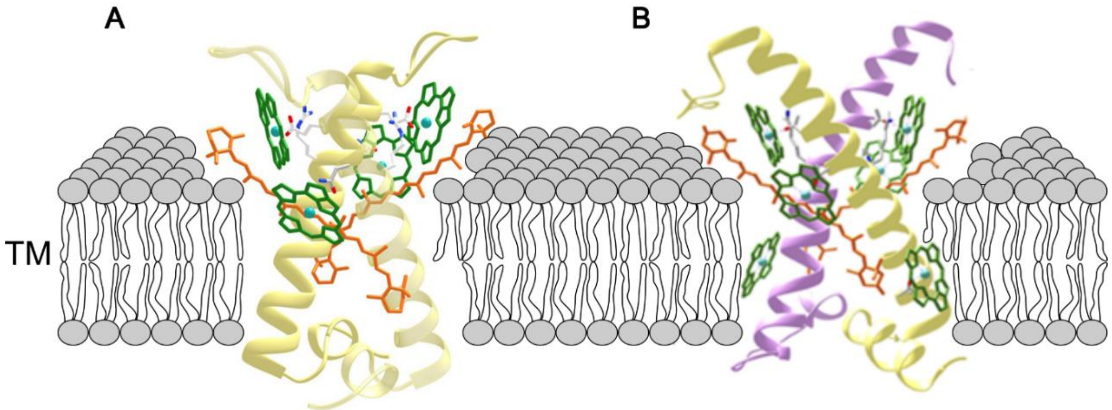


Figure 1.25: Structural model of high-light inducible proteins C and D. Model of an HliC dimer (A) bound to 4 Chl (green) and 2 β-carotene (orange) pigments (modified from Shukla, Llansola-Portoles, *et al*. 2018) and an HliD dimer (B) bound to 6 Chl and 2 β-carotene (modified from Staleva *et al.* (2015)).

1.11.2 YidC insertase

Bacteria contain two main membrane protein transport systems, the Sec translocase and YidC insertase, which promote the transport of hydrophilic membrane extrinsic domains across the membrane and facilitate the insertion of the membrane spanning domains into the lipid bilayer. YidC is a member of the evolutionarily conserved YidC/Oxa1/Alb3 integral membrane protein family. These proteins are involved in the insertion and assembly of membrane protein complexes into bacterial, mitochondrial and chloroplasts respectively, either alone or in concert with the Sec translocase (Dalbey *et al*., 2014).

The protein has been well characterised in *E. coli* and consists of six transmembrane helices (TMHs) and a single periplasmic domain, with TMHs 2-5 having been identified as essential to the proteins function (Jiang *et al*., 2003; Sääf *et al*., 1998). YidC binds to its substrates via these TMHs, inducing a conformational change in YidC, which aids the partitioning of the substrates TMHs into the lipid bilayer (Chen *et al*., 2002; Winterfeld *et al*., 2009). Little is known about the mechanism of YidC mediated membrane insertion, however it has been suggested that YidC shields hydrophobic residues within TMHs from the polar lipid head groups of the bilayer and charged residues of the membrane extrinsic domains as they are being translocated across the membrane (Kol *et al*., 2008).

Many membrane proteins feature a signal recognition particle (SRP) that functions as a targeting sequence directing nascent proteins to the membrane insertion apparatus (Akopian *et al*., 2013). When operating alone, YidC appears to not require that a substrate contain an SRP for membrane insertion; instead the hydrophobicity of the substrates TMH seems to be a key factor in determining the substrate specificity of YidC (Ernst *et al*., 2011; Zhu *et al*., 2013).

A subset of membrane proteins require both YidC and the Sec translocase for insertion. YidC forms a complex with the Sec apparatus via binding to the SecY subunit at the lateral gate of the protein where substrate TMHs escape laterally into the lipid bilayer (Sachelaru *et al*., 2013). YidC bind the TMHs as they escape the Sec translocon and mediates their assembly into helix bundles as they diffuse into the membrane (Beck *et al*., 2001). YidC has also been shown to directly bind to ribosomes and accept nascent polypeptides as they insert into the membrane (Jia *et al*., 2003; Kedrov *et al*., 2013; Seitl *et al*., 2014).

In addition to its role a membrane insertase, YidC has been shown to exhibit protein foldase activity (Nagamori *et al*., 2004). Nagamori *et al*. 2004 desmonstrated that lactose permease (LacY) did not require YidC for insertion into the membrane but did require YidC for correct folding, most likely by transient interaction during membrane insertion.

YidC/Alb3 has been shown to be essential for thylakoid membrane biogenesis in cyanobacteria, algae and plants as discussed in Section 1.5 (Göhre *et al*., 2006; Spence *et al*., 2004). In the alga, *Chlamydomonas reinhardtii*, there are two *alb3* homologs, Alb3.1 and Alb3.2. Alb3.1 knockout strains were able to grow photoautotrophically but has a severe reduction in the accumulation of LHCI, LHCII, PSII complexes while Chl levels are reduced to 30% WT levels. However, the amounts of PSI, Cyt b_6f and ATP synthase complexes were unaffected in the knockout strain and up to 10% WT levels of LHCII remained in the cell, prompting the conclusion that Alb3.1 is involved in, but not absolutely required for the assembly of light harvesting complexes (Bellafiore *et al*., 2002). The incorporation of the PSII subunit D1 into PSII complex during assembly is also perturbed in an Alb3.1 deletion mutant, although the insertion of D1 into the thylakoid membrane was not impeded, indicating a role of Alb3.1 as a PSII assembly factor (Ossenbühl *et al*., 2004). Alb3.2 was found to be capable of forming a complex with Alb3.1, PSII and PSI RC subunits as well as the thylakoid biogenesis protein VIPP1. Alb3.2 depletion resulted in reduced levels of PSI and PSII, indicating that this proteins is also involved in photosystem assembly/repair, perhaps partially overlapping with the function of Alb3.1 (Bellafiore *et al*., 2002; Göhre *et al*., 2006).

In plants, loss of Alb3 is more severe than in algae, causing severe depletion of Chl and retardation of thylakoid biogenesis, resulting in premature death during the seedling stage of plant development (Sundberg *et al*., 1997). Further research showed that, like Alb3.1 in algae, Alb3 in plants is required for integration and assembly of LHCs within the thylakoid membrane, although the insertion of other proteins, such as PSII subunits PsbX, PsbW and PsbY, did not depend on Alb3 (Moore *et al*., 2000a, 2003; Woolhead *et al*., 2001). LHC assembly by Alb3 is also not reliant on the Sec apparatus although an interaction between Alb3 and SecY has been demonstrated, possibly required for the insertion/folding of other chloroplast protein complexes (Klostermann *et al*., 2002; Moore *et al*., 2003). Interactions of Alb3 with CP43 and PsaA, core subunits of PSII and PSI respectively, have also been identified (Pasch *et al*., 2005). A second isoform of Alb3, named Alb4, was identified in the plant Arabidopsis thaliana and shown to have a role in thylakoid membrane biogenesis. A deficit in the production of Alb4 caused aberrant chloroplast and thylakoid formation but did not impact the assembly of LHCs (Gerdes *et al*., 2006). Alb4 was later shown to be required for the proper assembly of ATP synthase (Benz *et al*., 2009).

In *Synechocystis*, the *yidC* gene (*slr1471*) cannot be fully deleted, suggesting an essential role for YidC in cyanobacteria. The partially segregated strain had reduced levels of photosynthetic pigments and photosynthetic productivity in addition to retarded thylakoid biogenesis (Spence *et al*., 2004). Modification of the enzyme by addition of green fluorescent protein (GFP) to the C-terminus of YIdC resulted in a *Synechocystis* strain with an increased sensitivity to light and accumulated the D1 precursor, pD1, leading the authors to conclude that PSII assembly/repair was impaired and that YidC is involved in the integration of D1 into the complex during these processes (Ossenbühl *et al*., 2006).

YidC was found to co-purify with a FLAG-tagged ChlG protein when the latter was purified from *Synechocystis* thylakoid membranes. The discovery of an association between YidC and ChlG led to the hypothesis that, as Chl binding proteins are being translated by the ribosome, YidC/Alb3 fixes the polypeptides into a configuration that allows for the insertion of newly synthesised Chl molecules from neighbouring ChlG (Chidgey et al., 2014; Sobotka, 2014). This may occur during a distinct pause in translation that appears to coincide with Chl binding to the polypeptide (Kim *et al*., 1994b). The discovery of a ChlG-YidC interaction presented the first evidence of a direct link between Chl biosynthesis and photosystem assembly.

1.11.3 Ycf39

Ycf39 is a member of a family of short-chain alcohol dehydrogenases that feature an N-terminus NAD(P)H binding motif which lacks a tyrosine residue critical to enzyme function (Chidgey et al., 2014). Ycf39 is a peripheral membrane protein and contains no membrane spanning domain, and no catalytic function has been assigned. Ycf39 has been found to form a complex with HliD in which the HliD component binds β-carotene and Chl and is capable of quenching Chl fluorescence (Knoppová *et al*., 2014; Staleva *et al*., 2015). This Ycf39-HliD complex has been shown to be important for the early stages of PSII assembly. The Ycf39-HliD complex binds to the DE loop of the PSII precursor complex pD1 and seems to be restricted in its role to the biosynthesis of the D1 subunit. It has been postulated that the ChlG-Ycf39-HliD complex could arrive at the site of PSII assembly and bind pD1 as it is being co-translationally inserted into the membrane by YidC during which time Chl, provided by nearby ChlG, can be bound to the polypeptide. There is a pause in pD1 synthesis where Ycf39 can bind to pD1 and Chl can be inserted into the polypeptide (Knoppová *et al*., 2014). The discovery of a ChlG-Ycf39-HliD-YidC complex lends support to this hypothesis (Chidgey et al., 2014). It is possible that the role of Ycf39 in inserting Chl into pD1 is redundant as a Ycf39 knockout mutant is viable, albeit the strain is more sensitive to photoinhibition (Knoppová *et al*., 2014). Under such photodamaging conditions Chl pigments are released from the photosystems and must be recycled back to the membrane. In this respect there is some evidence to suggest that Ycf39 is required for Chl recycling. A Ycf39 knockout mutant also lacking PSI, and the PSII subunits CP43 and CP47, rapidly depletes reserves of the Chl precursor Mg-protoporphyrin under high-light conditions whilst D1 levels decreased by 30%. Chl precursor and D1 quantities remained stable in a control strain containing Ycf39, indicating that lack of Ycf39 impeded reuse of Chl (Knoppová *et al*., 2014). Release of Ycf39 from ChlG under high-light may deter the channelling of *de novo* Chl pigment to D1 and instead promotes the re-use of Chl released from damaged photosystems. Once released from the ChlG complex the Ycf39 may also become available to form a complex with HliD and act as a scavenger of free Chl molecules that would otherwise generate damaging singlet oxygen species (Komenda and Sobotka, 2016).

As previously discussed in Section 1.11.1, the homologue of Ycf39 in plants is HCF244. Deletion of this gene from *A. thaliana* lead to growth retardation and Chl bleaching, in addition to perturbation of D1 synthesis (Link *et al*., 2012). HCF244 forms a complex with OHP1/2 in *A. thaliana* that is postulated to be important for PSII assembly and repair (Hey and Grimm, 2018; Link *et al*., 2012).

1.11.4 Minor components

FLAG-tagged ChlG pulldown eluates, purified from *Synechocystis* thylakoid membranes and analysed by mass spectrometry, lead to the identification of other minor components of the ChlG complex. These included Sll1167, PSI core subunits PsaA and PsaB, the SecY subunit of the Sec translocase and the ribosome subunit Rpl1 (Chidgey et al., 2014). Sll1167 is related to the AmpH family of proteins, which in bacteria is postulated to be involved in the maturation and remodelling of the peptidoglycan layer (González-Leiza *et al*., 2011). Although a role for Sll1167 has not been elucidated in cyanobacteria, Chidgey *et al*. 2014 speculated that the protein may be involved in thylakoid biogenesis due to its interaction with the ChlG complex. The co-purification of SecY and Rpl1 with ChlG, as well as YidC, is indicative of a continuous link between chlorophyll biosynthesis and the co-translational insertion of Chl-binding proteins. The presence of PsaA/B, the major Chl binding proteins of PSI, suggests the ChlG complex may be localised in the vicinity of PSI.

1.12 Thesis aims:

Thylakoid biogenesis in oxygenic photosynthetic organisms relies on the biosynthesis of Chl pigments and their incorporation into macromolecular protein complexes called photosystems where they can perform the primary reaction of photosynthesis, the absorption of light. The Chl biosynthesis and photosystem assembly processes must be linked in order to ensure the efficient delivery of Chl pigments to the photosystem assembly apparatus where Chl can be co-translationally inserted into Chl-binding proteins and assembled into functioning photosystems. The Chl handover mechanism remains poorly understood. In the cyanobacterium *Synechocystis*, a protein-pigment complex comprising of the terminal enzyme of the Chl biosynthesis pathway, chlorophyll synthase (ChlG), high light inducible proteins C and D (HliC/HliD), a photosystem assembly factor Ycf39, and the protein insertase YidC/Alb3 was identified as having an important role in this process.

The work presented in this thesis aims to build on the current understanding of the function of the ChlG at the interface between Chl biosynthesis and photosystem assembly by further structurally and functionally characterising the complex. The third chapter in this thesis investigates whether a similar ChlG complex exists in higher photosynthetic organisms, such as plants and algae, and whether ChlG homologs from these organisms can functionally compliment the deletion of the essential *chlG* gene from *Synechocystis*. In the fourth chapter, the role of the carotenoid components of the ChlG complex were elucidated by construction of *Synechocystis* strains lacking these pigments, followed by characterisation of the ChlG complexes isolated from these backgrounds. A chemical cross-linking approach was taken in chapter five in order to determine the orientation and arrangement of the ChlG complex in the thylakoid membranes of *Synechocytis*. The N-terminal domain of ChlG was found to localise in close proximity to the other members of the complex. This domain was sequentially truncated in chapter six and the resulting ChlG proteins isolated and examined for binding partners to determine if this domain was important for formation of the ChlG complex. The enzyme activity of the truncated ChlG variants were also tested. In the final chapter, methods for the recombinant production of plant ChlG in *E. coli*, and production of the enzymes substrate chlorophyllide *a*, were developed with the aim of enabling characterisation of the protein by enzyme kinetic, mutagenesis and structural analysis. To demonstrate the viability of this approach, six point mutations were made to the enzyme, targeting residues predicted to be important for enzyme activity or substrate specificity by substituting them for alternate amino acids. These ChlG variants were produced in *E. coli* and tested for enzyme activity.

Chapter 2: Materials and Methods

2.1 Standard buffers, reagents and media.

All solutions were made using distilled ultra-pure water from a Milli-Q® machine (Millipore). LB and autoinduction growth media were made according to the manufacture instructions. All solutions were sterilised either by autoclaving for 20 minutes (15 psi, 121 °C). Once the growth media had cooled below 50 °C, heat sensitive reagents e.g. antibiotics, vitamins and sugar solutions, were added. These, and other reagents not suitable for sterilisation by autoclave, were filtered through 0.2 µm filters. Growth media used in this study are listed in Table 1.

2.2 *Escherichia coli* strains, growth and plasmids

Escherichia (*E*) *coli* cells were purchased from Promega. The *E. coli* strains used in this study are listed in Table 2. Cells were grown on Luria-Bertani (LB) agar (Formedium) at 37 °C overnight. JM109 cells grown in liquid LB medium were agitated at 250 rpm. When grown in liquid, BL21 (DE3) cells were either grown in LB (280 rpm), autoinduction (Formedium) (180 rpm) or Terrific Broth 300 rpm media at various temperatures as specified in the relevant sections. All media were supplemented with the appropriate antibiotic(s) at the following concentrations: kanamycin 30 µg mL^{-1}, ampicillin 100 µg mL^{-1} and streptomycin 20 µg mL^{-1}. *E. coli* strains were stored at -80 °C in 50% (v/v) LB-glycerol.

A list of the plasmids used in this study is presented in Table 3.

2.3 *Synechocystis* sp. PCC 6803

Synechocystis sp. PCC 6803 (hereafter *Synechocystis*) cells were kindly provided by Roman Sobotka of Centre Algatech, Institute of Microbiology, Czech Academy of Sciences. *Synechocystis* strains used in this study are listed in Table 2.

Strains were grown at 30 ^oC in a rotary shaker (150 rpm) in BG11 media supplemented with 10 mM TES-KOH pH 8.2 and moderate illumination (30-50 µmol photons $m^{-2}s^{-1}$). For growth on plates 1.5% (w/v) agar and 0.3% (w/v) sodium thiosulphate were added. Photoheterotrophic growth media contained 5 mM glucose. Zeocin (2.5-20 µg mL^{-1}) and kanamycin (5-40 µg mL^{-1}) were included where appropriate. For purification of protein complexes (Section 2.11) cultures were grown photoautotrophically with 100 µmol photons $m^{-2}s^{-1}$ illumination in 8 L vessels bubbled with sterile air and mixed by a magnetic stirrer. *Synechocystis* stocks were re-suspended in BG11 liquid media plus 10% (v/v) DMSO and flash frozen in liquid nitrogen before storage at -80 °C.

2.4 *Rhodobacter sphaeroides* strains and growth

Rhodobacter (*R.*) *sphaeroides* strains used in this study are listed in Table 2. *R. sphaeroides* were grown semi-aerobically in M22 media at 34 °C. Liquid cultures were grown as above with agitation at 150 rpm. When antibiotic selection was required, kanomycin was added to a concentration of 30 µg mL^{-1}. *R. sphaeroides* strains were stored at -80 °C in 1:1 LB:Glycerol.

2.5 Chemically competent *E. coli* cells

Chemically competent JM109 cells were purchased from Promega. Chemically competent BL21 (DE3) cells were made by preparation of a 50 mL liquid LB culture grown to an OD_{600} of 0.6 at 37 °C with agitation at 180 rpm. The cells were pelleted by centrifugation (4000 x*g*, 4 °C) and washed twice in 25 mL of 0.1 M $MgCl_2$ pre-chilled on ice. The cells were re-suspended in 1 mL 0.1 M $CaCl_2$ 20% (v/v) glycerol and decanted into 25 µL aliquots before being flash frozen in liquid nitrogen and stored at -80 °C.

2.6 Genetic transformation of cells

2.6.1 Chemical transformation of *E. coli* JM109 cells

A JM109 cell aliquot (20 µL) was thawed on ice, mixed with 2 µL of plasmid DNA and incubated in ice for 30 min. Cells were heat shocked at 42 °C for 45 s and then incubated in ice for a further 90 s. 1 mL of LB was added to the cells and incubated at 37 °C for 1 h with agitation at 250 rpm. The cells were pelleted by centrifugation (6000 x*g*, 4 °C), re-suspended in 100 µL of LB and plated on LB-agar supplemented with the appropriate antibiotic. Colonies were incubated overnight at 37 °C. Transformants were screened by PCR using TAQ polymerase (Section 2.7.2).

2.6.2 Transformation of *Synechocystis*

1-2 mL of the recipient *Synechocystis* strain was pelleted or a scraping of cells from a BG11 plate re-suspended in 100 µL of sterile H_2O and transferred to a 1.5 mL microcentrifuge tube and centrifuged at 8000 x*g* for 10 minutes. Cell pellets were re-suspended in 100 µL of sterile H_2O. 10-50 ng of plasmid or linear DNA was added to the re-suspended cells and incubated at 30 °C with 40 µmol photons $m^{-2}s^{-1}$ illumination for 1 h. Suspensions were agitated by hand every 15 minutes to ensure the cells remained suspended throughout the medium. Cells were plated onto BG11 agar and incubated overnight at 30 °C with 40 µmol photons $m^{-2}s^{-1}$ illumination. Cells were transferred onto BG11 agar with a low concentration of the appropriate antibiotic: streptomycin 7.5 µg mL^{-1}, erythromycin 7.5 µg mL^{-1}, kanamycin 5 µg mL^{-1}, zeocin 2 µg ml^{-1}, chloramphenicol 12.5 µg mL^{-1}. Established colonies were sequentially plated onto BG11 agar supplemented with double the concentration of antibiotic as the last culture up up to a concentration of 30-50 µg mL^{-1} Full segregation of the construct was confirmed by PCR using the appropriate primers (Section 2.7.2). Primers are listed in Table 4.

2.7 Nucleic acid manipulation

2.7.1 Preparation of plasmid DNA

A plasmid mini prep kit was used to purify plasmid DNA (Nippon Genetics Co., Ltd. or Quigen) according to the instructions provided by the manufacturer.

2.7.2 Polymerase chain reaction (PCR)

Amplification of DNA by PCR was achieved using Q5 High Fidelity DNA Polymerase or Taq DNA Polymerase. Q5 reactions contained 25 µL of 2x Q5 High-Fidelity Master Mix (New England BioLabs), 10 ng template DNA (plasmid or synthetic gene fragment), 0.25 ng of each primer and were made up to 50 µL final volume with QH_2O. Primers were produced by Invitrogen and diluted to 125 ng μL^{-1}. Artificial gene fragments were produced by Integrated DNA Technologies®. For amplifying genes from *Synechocystis*, a small cell scraping was re-suspended in 50 µL of QH_2O and 1 µL added to each reaction in place of template DNA. Taq reactions contained 10 µL 2x MyTAQ Hot Start Red Mix (Bioline), a small scraping of cells in place of template DNA and made up to 20 µL with QH_2O.

PCR conditions were as follows: 3 minute initial denaturation phase at 95 °C followed by 30 cycles of annealing (30 s), extension (72 °C, 10 seconds Kb^{-1}) and denaturation (95 °C, 30 seconds) and ending in a final extension phase lasting 2 minutes at 72 °C. The temperature used during the annealing phase was specific to the primers used for each reaction. PCR reactions were purified using a PCR clean up kit (QUIGEN) or analysed by agarose gel electrophoresis (Section 2.7.4).

A list of the primers used in this study are presented in Table 4 respectively.

2.7.3 Restriction Enzyme Digests

Restriction enzymes were purchased from Promega or New England Biolabs. Reaction buffers were selected according to the manufacturer's instructions. Reactions were carried out in 20 µL reactions consisting of 4 µL QH_2O, 2 µL reaction buffer, 2 µL BSA, 10 µL DNA (plasmid, synthetic gene fragment or purified PCR product), and 1 µL of each restriction enzyme. The reactions were incubated at 37 °C for 2 h.

2.7.4 Agarose gel electrophoresis

DNA fragments from PCR and restriction enzyme digests were analysed by electrophoresis on 0.8% (w/v) agarose gels made in 1x TAE (0.04 M tris-acetate, 1 mM EDTA) with 0.5 mg mL^{-1} ethidium bromide. DNA samples were mixed with DNA loading dye (0.03% bromophenol blue, 0.03% xylene cyanol, 60% glycerol, 60 mM EDTA in 10 mM Tris-HCl pH 7.6) and electrophoresis carried out at 85-90V alongside Hyperladder I (Bioline) to enable size estimation. Gels were visualised under UV light.

2.7.5 Recovery of DNA from agarose gels

DNA was recovered from agarose gel using a gel extraction kit from Quigen or Nippon Genetics Co., Ltd according to the instructions provided. The DNA was eluted in 30-50 µL of QH_2O.

2.7.6 Ligation of DNA fragments

T4 DNA ligase buffer and T4 DNA ligase enzyme from New England Biolabs were used in 10 µL ligation reactions according to manufacturer's instructions. A 3:1 molar ratio of insert to vector was used although other ratios were attempted if the ligation was unsuccessful. Reactions contained 1 µL 10x ligation buffer, 1 µL T4 DNA ligase and the appropriate quantities of insert and vector.

2.7.7 DNA sequencing

30-50 ng purified DNA construct, along with 50 µL of the appropriate primers diluted to 200 nM in QH_2O, were sent to GATC biotech for sequencing.

2.8 Mutagenesis

2.8.1 3xFLAG-tagging of chlorophyll synthase genes

To N-terminally FLAG-tag chlorophyll synthase genes, the genes were inserted into the pNPD-FLAG plasmid ((Hollingshead *et al*., 2012), which contains regions of homology to the upstream promoter and downstream regions of the *psbAII* gene encoding a redundant copy of a PSII subunit. Chlorophyll synthase genes were digested and cloned into the *Not*I and *Bgl*II sites of the plasmid such that they were in frame with an N-terminal 3xFLAG tag. The resulting vectors were transformed into *Synechocystis* as described in Section 2.6.2. The chlorophyll synthase genes were integrated into the *Synechocystis* genome at the *psbAII* locus by homologous recombination, placing the tagged-construct under the control of the *psbAII* promoter.

For C-terminal FLAG-tagging of chlorophyll synthase genes, the gene was inserted into the pCPD-FLAG plasmid (Chidgey et al., 2014) essentially as described above except that the 3xFLAG-tag is located downstream of the multiple cloning site.

2.8.2 Deletion of *Synechocystis* genes

Deletion of native genes in *Synechocystis* was achieved by replacement of part of the gene with an antibiotic resistance cassette using a linear mutagenesis construct. The antibiotic resistance cassette was placed between nucleotide sequences homologous to the upstream and downstream regions of the gene to be deleted and transformed into *Synechocytis* as described in Section 2.6.2. Following homologous recombination, the gene was replaced by the antibiotic resistance cassette. Segregation of genome copies was performed as outlined in Section 2.6.2.

2.8.3 Mutagenesis of *R. sphaeroides*

Mutagenesis of *R. sphaeroides* was achieved using the pK18mobsacB plasmid which contains both kanamycin resistance cassette and *sacB* selection markers. SacB catalyses the biosynthesis of sucrose into larger polysaccharides that are toxic to the host cell. Regions of DNA immediately proximal to the 5' and 3' ends of the target gene were cloned into the multiple cloning site of this vector and transformed into electrocompetent S17-1 *E. coli* cells (kindly provided by Judith Armitage, University of Oxford).

The S17-1 cells were conjugated with the recipient strain of *R. sphaeroides* which obtained the plasmid from S17-1 via lateral gene transfer. A 50 mL culture of the recipient *R. sphaeroides* strain was grown, centrifuged at 6000 x*g* for 10 minutes and re-suspended in 200 µL of LB medium before a loop of S17-1 colonies was added to this cell suspension. After mild agitation, the cell suspension was transferred to an LB agar plate in 50 µL droplets and grown overnight at 34 °C. The entire cell mass was transferred onto M22 agar supplemented with 30 µg mL^{-1} kanamycin and incubated at 34 °C until transconjugant colonies were visible. These colonies were cultured in liquid M22 medium supplemented with 30 µg mL^{-1} kanamycin overnight and serially diluted from 1:100 to 1:10,000 before being plated onto M22 agar supplemented with 10% (w/v) sucrose. Emerging colonies were screened for kanamycin resistance by replica plating onto two M22 agar plates, one containing just sucrose and the other containing both sucrose and kanamycin. Colonies that grew on the latter were screened by PCR using MyTaq Hot Start Red Mix DNA polymerase (Bioline) (Section 2.7.2).

2.8.4 Site-directed mutagenesis

Site directed mutagenesis of an *E. coli* pET28a plasmid containing a plant *chlG* gene was carried out using a Quickchange II site directed mutagenesis kit (Agilent) according to the manufacture instructions. PCR reactions contained 125 ng of the appropriate primers, 1 µL dNTPs, 5µL 10x buffer, 5-50 ng plasmid template, 1 µL DMSO, 1 µL PfuUltra HF DNA polymerase and made up to 50 µL using QH_2O.

PCR conditions were as follows: 16 cycles of a denaturation phase lasting 30 s at 95 °C, an annealing phase lasting 1 minute at 55 °C and an extension phase lasting 7 minutes at 68 °C.

The reactions were cooled to room temperature and 1 µL Dpn1 was added and incubated for 1 h at 37 °C. 50 µL of XL1-Blue cells was placed in a 15 mL falcon tube and 1 µL of plasmid added and left on ice for 30 minutes before being heat-shocked at 42 °C for 45 seconds followed by a recovery period on ice for 2 minutes. 500 µL of NZY medium, pre-heated to 42 °C was added to the cells and left incubating for 1 h at 37 °C with agitation at 180 rpm. 100 µL of the cells suspension was plated onto LB medium containing the appropriate antibiotic.

2.9 Protein Analysis

2.9.1 Determination of protein concentration

Protein concentration was determined by absorbance at 280 nm and converted into an approximate protein concentration using the following equation (Gill and von Hippel, 1989):

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

Where n_{Trp}, n_{Tyr} and n_{Cys} are the numbers of tryptophan, tyrosine and cystine residues and M_r is the predicted molecular weight of the protein.

2.9.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE carried out on precast 12% Bis-Tris NuPage gels (Invitrogen) according to the manufacture instructions. Samples were run alongside Bio-rad precision plus marker and proteins stained with Coomassie Brilliant Blue (Bio-Rad) or, if higher protein resolution was required, silver stained using a silver stain kit (Bio-Rad). Alternatively, proteins were transferred to PVDF membranes for immunoblotting (Section 2.9.3).

2.9.3 Immunoblotting

Proteins separated by SDS-PAGE were transferred to polyvinylidene fluoride membranes (Novex). The membrane was activated by soaking in methanol for 20 seconds and then equilibrated in transfer buffer (10 mM $NaHCO_3$, 10 mM $NaCO_3$, 10% methanol) along with 2x filter papers and 2x porous pads. These were assembled while submerged in transfer buffer into a transfer cassette, sandwiching the gel and membrane between the two porous pads layered with filter paper. The transfer cassette and buffer were placed in a transfer tank and a current applied across them at 30 mA overnight or 350 mA for 1 h at 4 °C with stirring.

The membrane was blocked in TBS (50 mM Tris/HCl, 150 mM NaCl, pH 7.6) with 5% (w/v) milk powder and 0.2% (w/v) Tween 20 for 1 h at room temperature. Primary antibodies were diluted to the appropriate concentration (generally 1:1000 to 1:10,000) in TBS + 0.05% Tween 20. The membranes were incubated with the selected primary antibody for 6 hours at room temperature or overnight at 4 °C with gentle agitation. The membranes were washed three times for 10 minutes each in TBS + 0.05% Tween 20 before being incubated with the appropriate secondary antibody conjugated with horseradish peroxidase (Sigma-Aldrich) and diluted to the recommended concentration in TBS + 0.05% Tween 20 for 1 h at room temperature. The membrane was washed three times for 10 minutes in TBS + 0.05% Tween 20.

The membrane was briefly soaked in WESTAR ETA C 2.0 chemiluminescent substrate (Cyanagen) and imaged using an Amersham Imager 600 (GE Healthcare). ChlG and Ycf39 primary antibodies, kindly provided by Roman Sobotka (Centre Algatech, Trebon, Czech Republic), were raised against synthetic peptides 89-104 and 311-322 respectively. The antibody raised against the R117- S384 recombinant fragment of YidC was provided by Jörg Nickelsen (Ludwig-Maximilians-University, Munich, Germany). HliD antibodies were purchased from Agrisera (Vännäs, Sweden) and FLAG antibodies from Sigma-Aldrich.

2.9.4 Mass spectrometry sample preparation

Protein samples were prepared for mass spectrometry by concentration to 50-100 µL using 10 kDa NMWL centrifugal filters (Millipore) according to the manufacturer's instructions. Samples were transferred to Lobind Eppendorf tubes and proteins precipitated using a 2-D clean-up kit from GE Healthcare according to the manufacturer's instructions. The tubes were centrifuged 16,000 x*g* for 5 minutes to pellet the protein which was then re-dissolved in 20 µL UT buffer (8M urea, 100 mM Tris-HCl, pH 8.5) with brief sonication.

1 µL aliquots were taken from each sample and diluted with 9 µL of HPLC grade H_2O (Fisher Chemical) for protein estimation at 280 nm using a Nanodrop spectrophotometer. 50 µg of protein was transferred to fresh Lo-bind Eppendorfs and adjusted to a final volume of 10 µL with UT buffer.

The samples were reduced by addition of 1 µL of 50 mM TCEP and incubated for 30 minutes at 37 °C. For S-alkylation 1 µL of 100 mM iodoacetamide in 100 mM Tris-HCl pH 8.5 was added to the protein sample before incubation for 30 minutes at room temperature in the dark. 2 µL of trypsin-endoproteinase Lys-C mix (Promega, 1 µg/µL in 50 mM acetic acid) was added to the protein sample and incubated at 37 °C for 2 h. The reaction was diluted with 75 µL of 50 mM Tris-HCl pH 8.5 containing 10 mM $CaCl_2$ and incubated overnight at 37 °C. 5 µL of 10% trifluoroacetic acid (TFA) was added to the samples before they were desalted using a Pierce® C-18 spin column (Thermo

Scientific) according to the manufacturer's instructions. The peptide eluates were dried by vacuum centrifugation and stored at -20 °C.

2.9.5 Mass spectrometry

Proteins were precipitated from cell lysates and FLAG eluates using a 2-D clean-up kit (GE Healthcare) according to the manufacturer's instructions and recovered by centrifugation at 16,000 x*g* for 10 min. The pellets were dissolved in 10 - 20 µL 8 M urea in 100 mM Tris-HCl pH 8.5 and 1 µL taken for protein determination by Nanodrop analysis at 280 nm. Reduction of disulphide bonds was carried out by the addition of 5 mM tris(2-carboxylethyl)phosphine-HCl and incubation at 37 °C for 30 minutes. Cysteine side-chains were then S-alkylated of by the addition of 10 mM methyl methanethiosulphonate (stock solution at 100 mM in isopropanol) at room temperature for 30 minutes. Trypsin/endoproteinase Lys-C mix (Promega) was added at an enzyme:substrate ratio of 1:25 by mass and the protein digested at 37 °C for 2 h. After dilution with 7.5 vol 50mM Tris-HCl pH 8.5, 10 mM $CaCl_2$ digestion was continued overnight. The digests were desalted using C18 spin columns (Thermo Scientific) following the manufacturer's instructions and 500 ng analysed by nanoflow liquid chromatography (Dionex UltiMate 3000 RSLCnano system) coupled online to a Q Exactive HF quadrupole-Orbitrap mass spectrometer (Thermo Scientific).The chromatograph was programmed to deliver a linear gradient of 97% solvent A (0.1% formic acid in water) to 10% solvent B (0.08% formic acid in 80% acetonitrile) over 5 min followed by 10% - 50% solvent B over 75 min at 300 nL/min. The mass spectrometer was programmed for automated data dependent acquisition with each full scan at 120 000 resolution followed by a maximum of ten dependent product ion scans at 30 000 resolution. The mass spectra were processed with Byonic v. 2.9.38 (Protein Metrics) with parameters set to default except that methylthio-Cys and Met oxidation were specified as fixed and variable (common) modifications respectively. Protein identifications were made by searching the Cyanobase *Synechocystis* sp. PCC 6803 proteomic database (http://genome.annotation.jp/cyanobase/Synechocystis/genes.faa, downloaded on

29/01/2016). The identification of crosslinked peptides was performed was enabled by specifying DSS for both DSS and BS3 (identical linkage chemistry) and manually creating a rule for NHS-LC- and sulpho-NHS-diazirine with the following syntax: 'diaz / 195.125929 @ K,C,S,T,L,I,V | xlink; Xlink:diaz / 213.136443 @ K | common2 % Dead end'. Crosslinked peptide assignments produced by Byonic were curated visually and only those complying with the feasibility criteria stated in Chapter 5 Table 2 were considered.

2.9.6 Quantitative mass spectrometry

Quantitative protein mass-spectrometry was performed by mixing FLAG-immunoprecipitation eluates (Section 2.11.2) with 30 pmol of an ^{15}N-labelled internal standard in the form of an artificial protein consisting of concatenated proteotypic tryptic peptides belonging to ChlG, HliD and Ycf39 (Shukla *et al*., 2018b) expressed in *E. coli* following the method described in (Qian *et al*., 2013). Proteins were isolated by precipitation using a GE Healthcare 2-D clean-up kit, dissolved in 8M urea, 100 mM Tris-HCl pH 8.5 and 50 µg reduced, S-alkylated, digested and the resultant tryptic peptides desalted as described in Section 2.9.4.

500 ng of tryptic peptides was analysed by nanoflow liquid chromatography (Dionex UltiMate 3000 RSLCnano system) coupled online to a Q Exactive HF mass spectrometer (Thermo Scientific).

Proteotypic peptides were identified in the mass spectra using Mascot Daemon v 2.5.1 running with Mascot Server v 2.5.1 (Matrix Science) searching a *Synechocystis* sp. PCC 6803 proteome database and picomolar amounts of FLAG-ChlG, HliD and Ycf39 were calculated from the relative intensities of the ^{14}N and ^{15}N isotopomers of their respective proteotypic peptide ions.

2.10 Chemical crosslinking

2.10.1 *In vivo* cross-linking

A 1 L cell culture of a *Synechocystis* strain containing an N-terminally FLAG-tagged chlorophyll synthase gene was grown as outlined in Section 2.3 until the optical density of the culture, measured at 750 nm, reached 0.7. The culture was harvested by centrifugation at 17,700 x*g* and washed 3x in PBS. The wet weight of the cell pellet was measured and re-suspended to a density of 4 mL/g in BPER reagent (Thermo Scientific) for treatment with a water soluble cross-linker or in PBS for a DMSO soluble cross-linker. Cell suspensions were incubated with cross-linker (see below) in the dark for 10 minutes on a rotator at room temperature.

In this work, four different cross-linking reagents supplied by Thermo Scienctific were used according to the instructions provided. Two water soluble cross-linkers bis-sulfosuccinimidyl-suberate (BS3) and sulfosuccinimidyl 6-(4,4'-azipentanamido)hexanoate (sulfo-LC-SDA) were dissolved in BPER to a final concentration of 180 mM. Two hydrophobic cross-linkers disuccinimidyl suberate (DSS) and succinimidyl 6-(4,4'-azipentanamido)hexanoate (LC-SDA) were dissolved in DMSO to 180 mM. The chosen cross-linking reagent was immediately added to the cell suspension to a final concentration of 45 mM, unless stated otherwise, and the suspension incubated in the dark for 40 minutes at room temperature with gentle rotation. Cross-linking reactions were quenched by addition of 1 M Tris-HCl pH 8.5 to a concentration of 50 mM followed by a further 5 minute period of incubation as described above. Cells were pelleted by centrifugation at 5000 x*g* for 5 minutes, washed twice in PBS and re-suspended to the pre-wash volume in PBS.

Procedures involving the use of the heterobifunctional cross-linking reagents LC-SDA and sulfo-LC-SDA required additional steps. The cell suspension was placed into a weighing boat such that the cell layer was as thin as possible. The cells were placed under a 365 nm UV light at a distance of 5 cm and irradiated for 15 minutes with gentle agitation at room temperature.

Cells were pelleted at 5000 x*g* for 5 minutes and stored at -20 °C. Alternatively, the thylakoid membrane fraction and FLAG immunoprecipitation procedures were immediately performed in order to enrich for cross-linked FLAG-ChlG complex. These were performed essentially as described in Sections 2.11.1 and 2.11.2 respectively except that a 200 µL anti-FLAG-M2 agarose (Sigma-Aldrich) column was used during the FLAG immunoprecipitation procedure.

2.10.2 *In vitro* cross-linking

An 8 L cell culture of a *Synechocystis* strain containing an N-terminally FLAG-tagged chlorophyll synthase gene was grown as outlined in Section 2.3 until the optical density of the culture measured at 750 nm reached 0.7. Isolation of the thylakoid membrane fraction and FLAG immunoprecipitation procedures were performed according to Sections 2.11.1 and 2.11.2 respectively with the exception of the omission of 10% (w/v) glycerol from the FLAG buffer (25 mM sodium phosphate pH 7.4, 10 mM $MgCl_2$ and 50 mM NaCl and Roche EDTA-free protease inhibitor cocktail) in experiments utilising the heterobifunctional cross-linking reagents LC-SDA and sulfo-LC-SDA.

To reduce the concentration of FLAG peptide to <10 µM, the 400 µL FLAG immunoprecipitation eluate was diluted tenfold in FLAG buffer supplemented with 0.04% (w/v) β-DDM and passed through a Vivaspin 10 kDa NMWL spin concentrator (Sartorius) according to the manufacturer's instructions until the volume returned to 400 µL. For experiments using LC-SDA and sulfo-LC-SDA, the β-DDM concentration was reduced from 0.04 to 0.008% (w/v) using the same technique except that the eluate was diluted with FLAG buffer without additional β-DDM. The eluate was concentrated to 70 µL.

Cross-linking reagents were prepared as described in Section 2.10.1 and added to the eluate to a final concentration of 45 mM. Reactions were performed as above up until and including quenching of the cross-linking reactions using excess Tris-HCl (Section 2.10.1). After a 5 minute incubation with 50 mM Tris-HCl, the cross-linking reactions were diluted tenfold in FLAG buffer supplemented with 0.01% β-DDM and passed through a 10,000 NMWL cellulose spin concentrators (Microcon) by centrifugation at 14,000 x*g* according to the manufacturer's instructions until the volume reached 70 µL.

Experiments using LC-SDA and sulfo-LC-SDA were incubated under UV light as outlined in Section 2.10.1.

Cross-linked eluates were transferred to Lobind Eppendorf tubes and stored at -20 °C until ready for use in mass spectrometry experiments (Section 2.9.5).

2.10.3 Cross-linking of the thylakoid fraction prepared from *Synechocystis*

The thylakoid fraction was isolated from an 8 L culture *Synechocystis* cells expressing FLAG-tagged ChlG as outlined in Section 2.11.1 up until the point of solubilisation and using FLAG buffer without the addition of glycerol. The thylakoid pellet was washed 3x in FLAG buffer and re-suspended in the same buffer to a volume of 5 mL.

The cross-linking reagent sulfo-LC-SDA were directly dissolved in 2.4 mL of thylakoid membranes to a final concentration of 45 mM. The tube was wrapped in foil and incubated for 30 mins in the dark at room temperature with gentle agitation. The reaction was quenched by addition of 50 mM Tris-HCl as described previously (Section 2.10.1) and the thylakoid membranes pelleted by centrifugation at 16,600 x*g* for 30 minutes. The pellet was washed twice in FLAG buffer and re-suspended to the original volume of 1.2 mL in the same buffer. The sample was placed in a weighing boat and irradiated with UV as described previously (Section 2.10.1).

The sample was solubilised in 1.5% (w/v) β-DDM for 1 h at 4 °C with gentle agitation before being used in a small scale FLAG immunoprecipitation experiment essentially as described in Section 2.11.2 except that a 200 µL anti-FLAG-M2 agarose (Sigma-Aldrich) column was used.

2.11 Protein purification

2.11.1 Preparation of *Synechocystis* thylakoid membrane fraction

Synechocystis cells expressing FLAG-tagged ChlG were grown to an OD_{750} of ≈ 0.7 and harvested by centrifugation at 17,700 *g* at 4°C for 20 min. The subsequent procedures were performed either in the dark or under green light. Pellets from 4 or 8 L cultures were washed and re-suspended in FLAG-buffer (25 mM sodium phosphate pH 7.4, 10 mM $MgCl_2$ and 50 mM NaCl, 10% (w/v) glycerol and EDTA-free Protease Inhibitor (Roche)), mixed with an equal volume of 0.1 mm glass beads (BioSpec) and broken in a Mini-Beadbeater-16 (BioSpec) in 8 cycles lasting 55 seconds each. Cells were cooled on ice for 5 minutes between cycles. Soluble and membrane proteins were separated by centrifugation at 48,400 x*g*, 4°C for 30 minutes and the membrane fraction was re-suspended in 10 mL FLAG-buffer with 2 % (w/v) n-dodecyl-β-maltoside (β-DDM; Anatrace). The thylakoid membrane fraction was solubilised at 4 °C for 1 h with gentle agitation. Insoluble material was pelleted by centrifugation at 48,400 x*g*, 4°C for 30 minutes and the supernatant containing the solubilised thylakoids was retained.

2.11.2 Purification of FLAG-tagged proteins from *Synechocystis*

The solubilised thylakoid fraction prepared as described in Section 2.11.1 was diluted twofold in FLAG buffer and applied to a 0.3 mL anti-FLAG-M2 agarose (Sigma-Aldrich) column equilibrated in FLAG buffer. The flow through was retained and reapplied to the column twice, for a total of three bindings. The resin was washed with 20 resin volumes of FLAG buffer supplemented with 0.04% (w/v) β-DDM to remove contaminating proteins. FLAG-tagged proteins were eluted by re-suspension of the resin in 400 µL of the same buffer containing 187.5 mg/mL of 3xFLAG peptide (Sigma-Aldrich). This resin suspension was incubated for 1 h at 4 °C with gentle agitation and separated from the protein eluate by filtration through a 0.22 µm spin column. The flow-through containing the FLAG-tagged protein was collected and stored at -80 °C.

2.11.3 Preparation of *E. coli* lysates

E. coli cultures prepared as described in Section 2.2, were centrifuged at 6000 x*g* for 10 minutes at 4 °C and the cell pellets re-suspended in HEPES binding buffer (25mM HEPES, 0.5 M NaCl, 5 mM imidazole, pH 7.5) to a density of 2 mL/g. Cells were incubated with 50 mg of DNase I, lysozyme and tablet of EDTA-free Protease Inhibitor (Roche) per L of culture for 30 minutes at room temperature with gentle agitation. The cells were cooled on ice for 10 minutes before being lysed by two passes through a French Press at 124 MPa with cooling on ice between cycles. Larger volumes of cell suspension were filtered through cheese cloth and lysed by two passes through a cell disrupter at 25 MPa at 4 °C with cooling on ice between cycles. Cell lysates were centrifuged at 8000 x*g*, 4 °C for 20 minutes to remove unbroken cells and the supernatant was retained. Cell lysates were stored at -80 °C.

2.11.4 Preparation of *E. coli* cytoplasmic membrane fraction

E .coli lysates, prepared as outlined in Section 2.11.3, were centrifuged at 130,000 x*g*, 4 °C for 1 h to pellet the cytoplasmic membrane fraction. The supernatant was discarded and the pellets re-suspended in 10 mL HEPES binding buffer using a cell homogeniser.

The membrane fraction was incubated with 1.5% (w/v) β-DDM for 1 h with gentle agitation at 4°C. The solubilised membranes were diluted three fold and centrifuged at 130,000 x*g*, 4 °C for 1 h to remove insoluble material. The supernatant containing the solubilised membrane fraction was retained.

2.11.5 Purification of His-tagged proteins by Ni^{2+} IMAC

Recombinant His-tagged chlorophyll synthase was purified from the *E. coli* cytoplasmic membrane fraction, prepared as described in Section 2.11.4, by gravity flow immobilised metal affinity chromatography using a Ni^{2+} NTA Agarose (Qiagen). The membrane fraction was applied three times to a column pre-equilibrated in HEPES binding buffer. Approximately 200 µL of resin was used per L of cell culture from which the membrane fraction was isolated. The column was then washed firstly with 20 column volumes of HEPES binding buffer supplemented with 0.04% (w/v) β-DDM followed by washing with HEPES buffer containing progressively higher concentrations of imidazole. The imidazole concentrations used were optimised for each experiment, however in general three wash buffers containing 20 mM, 50 mM and 100 mM imidazole, each supplemented with 0.04% (w/v) β-DDM, were used. Proteins were eluted by incubation of the resin with 400 µL of HEPES elution buffer (25mM HEPES, 0.2M NaCl, 400 mM Imidazole, pH 7.5, 0.04% (w/v) β-DDM) for 10 minutes with gentle agitation before the solution was passed through a 0.22 µM 0.22 µm spin column which separated the resin from the protein elution. Eluates were stored at -80 °C.

2.11.6 HPLC gel filtration of protein complexes

Proteins were separated on a BioSep-SEC-S 3000 (Phenomenex) gel filtration column on an Agilent-1200 HPLC system. β-DDM was added to the samples to a final concentration of 1% (w/v). The running buffer was 25 mM HEPES, pH 7.5 supplemented with 0.2% (w/v) β-DDM. The samples were run at 0.2 mL min^{-1} and the absorbance monitored at 280 nm and 440 nm. Sample fractions were collected and stored at -20°C. The column was calibrated using gel filtration markers (29 kDa – 700 kDa) separated using the method described above.

2.12 Chlorophyll synthase activity assays

Chlorophyll synthase assays were performed using FLAG column eluates, obtained from *Synechocystis* cells containing FLAG tagged ChlG, or clarified *E. coli* lysates containing recombinant *Arabidopsis thaliana* ChlG. The substrate chlorophyllide *a* was purified from the growth medium of a *R. sphaeroides* mutant (Section 7.3.4) according to the method outlined in Section 2.14.3 or by chlorophyllase assay (Section 2.13) and dissolved in methanol.

ChlG samples were incubated with 20 µM geranylgeranyl-pyrophosphate (Sigma-Aldrich) and, where appropriate, 30 µM chlorophyllide *a* for 1 h at 30 °C with agitation at 180 rpm in the dark. Assays were stopped by addition of 2 volumes of methanol and centrifuged at 16,000 x*g* for 5 minutes and the supernatant containing the pigments retained. The supernatant was analysed on an Agilent-1200 HPLC system as described in Section 2.14.4.

2.13 Chlorophyllase assays

Clarified *E. coli* lysates containing recombinant chlorophyllase (CLH-1) (Section 7.3.6) were prepared according to section 2.11.3. with the exception that cells were re-suspended in 50 mM MOPS buffer. The enzymes substrate, chlorophyll *a*, was prepared as described in Section 2.14.1. The CLH-1 assays were performed essentially as described by (Tsuchiya *et al*., 1997). The E. coli lysates were used in 125 µL reactions containing 50 µM LDAO and 50 µM Chl (in acetone). Reactions were started by addition of Chl and incubated at 30 °C for 45 minutes. Reactions were quenched by addition of 200 µL acetone. The reactions were centrifuged at 16,000 x*g* for 5 minutes and the supernatant containing the pigments retained. The supernatant was analysed on an Agilent-1200 HPLC system as described in Section 2.14.4.

2.14 Pigment analysis

2.14.1 Extraction of (bacterio)chlorophyll *a* from whole cells

Synechocystis or *R. sphaeroides* cells were pelleted by centrifugation at 16,600 x*g* for 5 minutes. The cell pellets were washed in PBS three times and the supernatant discarded before the wet weight of the cells was measured. The following steps were performed in the dark. The cell pellets were re-suspended in 3 pellet volumes of a 7:2

mixture of acetone and methanol and agitated in a Mini-Beadbeater-16 (BioSpec) for 55 seconds before a period of cooling on ice for 5 minutes. The cells were pelleted at 16,600 x*g* for 5 minutes and the supernatant retained. 30 µL of 5 M NaCl was added to the solution and 1 volume of hexane. The suspension was agitated by manual shaking and centrifuged for 2 minutes at 16,600 x*g*. The hexane layer containing the (bacterio)chlorophyll *a* was decanted into 1 mL microcentrifuge tubes and dried by vacuum centrifugation (Eppendorf) for 10 minutes.

2.14.2 Extraction of pigments from *Synechocystis* cells or protein preparations

Synechocystis cells were pelleted by centrifugation at 14,500 x*g* for 5 minutes at 4 °. The cells were washed three times in QH_2O and re-suspended in 75% methanol before incubation at room temperature in the dark for 5 minutes. The cells were pelleted and the supernatant retained. This process was repeated using 80% methanol and the supernatant added to the existing extract.

Pigments were extracted from protein preparations by vortexing with 9 volumes of methanol followed by incubation at room temperature for 20 minutes in the dark. Insoluble material was pelleted by centrifuging at 16,600 x*g* for 5 minutes and the supernatant retained.

The pigment extracts was concentrated where necessary by evaporation under a stream of nitrogen gas.

2.14.3 Extraction of pigments from growth media

Cell cultures of the relevant *R. sphaeroides* strain were grown as described in Section 2.4 except that the M22 growth medium was supplemented with 0.2% (v/v) Tween 80 to enable the diffusion of pigments out of the cells and into the growth medium. The following steps were performed in the dark. The cells were pelleted by centrifugation at 6000 x*g* and the growth medium mixed vigorously with diethyl ether and acetonitrile in a 3:2:1 ratio (media : diethyl ether : acetonitrile) before being allowed to partition into two phases in a separation funnel. The top phase containing the pigments was decanted into a boiling flask and concentrated using a rotary evaporator. Once the volume was sufficiently reduced, the solvent containing the pigments was decanted into 1 mL microcentrifuge tubes and dried by vacuum centrifugation.

2.14.4 Separation of chlorophyll *a* and chlorophyllide *a*

Separation of pigments extracted from ChlG assays (Section 2.12) was performed by reverse-phase HPLC using an Agilent-1200 series HPLC system and a Nova-Pak C18, 4-µm particle size, 3.9 × 150 mm column (Waters). Solvent A consisted of 350 mM ammonium acetate and 30% methanol and solvent B consisted of 100% methanol. Pigments were eluted with a linear gradient of solvent B (0 to 100%, 15 min) followed by 100% solvent B at a flow rate of 1 mL min^{-1} at 40 °C for 5 minutes monitoring the absorbance of the column eluate at 665 nm and fluorescence (440 nm excitation; 670 nm emission). Relative quantification of pigments was performed by integration of the corresponding chromatograph peaks.

2.14.5 HPLC separation of chlorophyll precursors

Chlorophyll precursors extracted from *Synechocystis* cells (Section 2.14.2) or growth medium (Section 2.14.3) were analysed by reverse-phase HPLC on a Nova-Pak C18, 4 µM particle size, 3.9 x 150mm column (Waters) using an Agilent-1200 series HPLC system. Solvent A consisted of 350 mM ammonium acetate and 30% methanol and solvent B consisted of 100% methanol. Pigments were eluted in a linear gradient of solvent B from 65% to 75% over 35 minutes at a flow rate of 1 mL $minute^{-1}$ at 40°C followed by washing in 100% solvent B. The pigment content of the eluate was monitored by absorbance at 433 nm, 440 nm, 632 nm, 663 nm and fluorescence (440 nm excitation, 670 nm emission).

Preparative scale reverse-phase HPLC was performed as above using a UniverSil C18, 5 µM particle size, 10 x 150 mm (Fortis). The flow rate was increased to 3.5 mL minute^{-1} and 3.5 mL fractions collected.

2.14.6 Separation of pigments by HPLC

Pigments extracted from *Synechocystis* cells or protein preparations, as described in Section 2.14.2, were separated by reverse-phase HPLC on a Nova-Pak C18, 4 µM particle size, 3.9 x 150mm column (Waters) using an Agilent-1200 series HPLC system. 100% ethyl acetate was used as solvent A and acetonitrile/H_2O/trimethylamine (9 : 1 : 0.01, v/v/v) was used as solvent B. After 2 minutes of washing in solvent B a linear gradient of 0–100% ethyl acetate was applied over 15 min followed by washing in this solvent for 5 min at a flow rate of 1 mL min^{-1} at 40 °C. Absorbance was monitored at 450, 492 and 665 nm and chlorophyll and carotenoid species were identified by their absorption spectra and retention time which have been previously published (chlorophyll *a*, zeaxanthin, myxoxanthophyll, echinenone and β-carotene - Lagarde and Vermaas, 1999; deoxy-myxoxanthophyll - Graham, 1998; Lagarde and Vermaas, 1999b).

2.15 UV-visible absorbance spectroscopy

UV-visible absorbance spectra of cells, membranes and protein complexes were measured in a Cary 60 UV-Vis spectrophotometer (Agilent) at room temperature with appropriate media/buffer baseline correction.

Table 1: Growth media

Growth Media	Reagents
LB/LB agar	Ready mixed from FORMEDIUM, prepared according to manufacturer's instructions
Autoinduction LB	Ready mixed from FORMEDIUM, prepared according to manufacturer's instructions
Terrific Broth	**Stock solutions** Phosphate buffer (1 L) 23.1 g pottasium dihydrogen phosphate 125.4 g dipottasium phosphate **Terrific Broth (1 L)** 12 g Tryptone 24 g Yeast extract 4 mL Glycerol 100 mL phosphate buffer added after cooling
BG11	**Stock solutions** Trace Minerals (1 L): 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 Bg11 (1 L): 149.6 g sodium nitrate 7.49 g magnesium sulphate 3.60 g calcium chloride 0.6 g citric acid 0.10 g EDTA (0.56 ml of 0.5 M pH 8.0 stock) 100 ml trace minerals stock 1000x Iron (1 L): 6 g ferric ammonium citrate 1000x Phosphate (1 L): 30.5 g dipotassium hydrophosphate 1000x Carbonate(1 L): 20 g sodium carbonate 1 M TES/KOH pH 8.2 1 M Glucose **1x BG11 liquid medium (1 L)** 10 ml 100x Bg11 1 ml each 1000x stock. After autoclaving TES and glucose (if required) added to a final concentration of 10 mM and 5 mM respectively **BG11 agar (1 L)** 1x Bg11 15 g bacto-agar 3 g sodium thiosulfate TES and glucose (if required) added after melting

Growth Media	Reagents
M22	**Stock solutions** <u>Solution C</u>: 40 g nitriloacetic acid 96 g magnesium chloride 13.36 g calcium chloride 0.5 g EDTA 1.044 g zinc chloride 1.0 g ferrous chloride 0.36 g manganese chloride 0.037 g ammonium heptamolybdate 0.031 g cupric chloride 0.0496 g cobaltous nitrate 0.0228 g boric acid <u>10x M22 (4L):</u> 122.4 g potassium dihydrogen orthophosphate 120.0 g dipotassium hydrogen orthophosphate 100.0 g lactic acid 20 g ammonium sulphate 20 g sodium chloride 173.7 g sodium succinate 10.8 g sodium glutamate 1.6 g aspartic acid 800 ml solution C <u>10,000x Vitamins (100 ml):</u> 1 g Nicotinic acid 0.5 g Thiamine 0.1 g 4- aminobenzoic acid 10 mg biotin <u>Casamino acids (1 L):</u> 50 g Casein Hydrosylate **1x M22 (1 L)** 100 ml 10x M22 20 ml casamino acids Vitamins added after autoclaving **1x M22 agar (1 L)** 100 ml 10x M22 15 g bacto-agar Vitamins added after cooling

Table 2: Strains

***E. coli* strains**

Strain	Abbreviation	Properties	Source
S17-1	-	RP4-2 (Tc::Mu, Nm::TN7) integrated into the chromosome: *thi prohsdR hsdM*$^+$ *recA* Tp^R (Sm^R)	Simon *et al.* (1983)
JM109	-	*end*A1 *gln*V44 *thi*-1 relA1 *gyr*A96 recA1 *mcr*B$^+$ Δ(*lac-pro*AB) e14-[F'traD36 *pro*AB$^+$ *lac*I^q lacZΔM15] *hsd*R17 ($r_K^- m_K^-$)	Promega
XL1-Blue supercompetent cells	-	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^r)].	Agilent
BL21 (DE3)	BL21	(F$^-$ *ompT* $r_B^- m_B^-$) + bacteriophage DE3	Studier and Moffat (1986)
BL21 (DE3) pET28a::At_*chlG*	AtChlG	BL21 (DE3) strain transformed with a pET28a vector containing an *Arabidopsis chlG* gene codon optmised for expression in *E. coli* and in frame with an N-terminal 6x His-tag	This study
BL21 (DE3) pET28a::At_*chlG* Q46E	Q46E	BL21 (DE3) transformed with pET28a::At_*chlG* Q46E vector containing a point mutation in codon 46 of At_*chlG* resulting in a residue change from Q to E, kan^R	This study
BL21 (DE3) pET28a::At_*chlG* P54F	P54F	BL21 (DE3) transformed with pET28a::At_*chlG* P54F vector containing a point mutation in codon 54 of At_*chlG* resulting in a residue change from P to F, kan^R	This study
BL21 (DE3) pET28a::At_*chlG* L56P	L56P	BL21 (DE3) transformed with pET28a::At_*chlG* L56P vector containing a point mutation in codon 56 of At_*chlG* resulting in a residue change from L to P, kan^R	This study
BL21 (DE3) pET28a::At_*chlG* V60Y	V60Y	BL21 (DE3) transformed with pET28a::At_*chlG* V60Y vector containing a point mutation in codon 60 of At_*chlG* resulting in a residue change from V to Y, kan^R	This study
BL21 (DE3) pET28a::At_*chlG* N99A	N99A	BL21 (DE3) transformed with pET28a::At_*chlG* N99A vector containing a point mutation in codon 99 of At_*chlG* resulting in a residue change from N to A, kan^R	This study
BL21 (DE3) pET28a::At_*chlG* A225M	A225M	BL21 (DE3) transformed with pET28a::At_*chlG* A225M vector containing a point mutation in codon 225 of At_*chlG* resulting in a residue change from A to M, kan^R	This study
BL21 (DE3) pET21a::*CLH-1*	-	BL21 (DE3) transformed with pET21a::*CLH-1* vector containing an *Arabidopsis CLH-1* gene codon optimised for expression in *E. coli*, kan^R, Amp^R	This study
BL21 (DE3) pET21a::His-*CLH-1*	-	BL21 (DE3) transformed with pET21a::His-*CLH-1* vector containing *CLH-1* in frame with an N-terminal 6x His-tag, kan^R, Amp^R	This study

***Rba. sphaeroides* strains**

Strain	Abbreviation	Properties	Source
2.4.1	-	Wild-type	S. Kaplan, university of Texas
Δ*bchC*/*bchX*	Δ*bchCX*	In frame deletion of the *bchC* and *bchX* genes	Chidgey (2014)
Δ*bchC*/*bchX*/*bchF*	Δ*bchCXF*	In frame deletion of the *bchC*, *bchX* and *bchF* genes	Chidgey (2014)
Δ*bchC*/*bchX*/*bchF* *	Δ*bchCXF**	Δ*bchCX* strain transformed with pK18mobsacB-Δ*bchF** to introduce partial frameshift deletion of *bchF* gene	This study

***Synechocystis* strains**

Strain	Abbreviation	Properties	Source
Synechocystis sp. PCC 6803	WT	Glucose tolerant wild-type strain of *Synechocystis* sp. PCC 6803	Dr R. Sobotka, Institute of microbiology, Třeboň
psbAII::FLAG-6803_*chlG*	-	N-terminally FLAG tagged copy of *Synechocystis chlG* in place of *psbAII* gene, kanamycin resistant (kan^R).	This study
psbAII::FLAG-6803_*chlG* /Δ*chlG*	FLAG-6803	*psbAII*::FLAG-6803_*chlG* strain in which native *chlG* gene is replaced with a zeocin resistance (zeo^R) cassette, kan^R.	This study
psbAII::FLAG-7002_*chlG*	-	N-terminally FLAG tagged copy of *Synechococcus* sp. PCC 7002 *chlG* in place of *psbAII* gene, kan^R	This study
psbAII::FLAG-7002_*chlG* /Δ*chlG*	FLAG-7002	*psbAII*::FLAG-7002_*chlG* strain in which the native *chlG* gene has been deleted, kan^R, zeo^R.	This study
psbAII::FLAG-Cr_*chlG*	-	N-terminally FLAG tagged copy of *Chlamydomonas reinhardtii chlG* in place of *psbAII* gene, kan^R.	This study
psbAII::FLAG-Cr_*chlG*/Δ*chlG*	FLAG-Cr	*psbAII*::FLAG-Cr_*chlG* strain in which the native *chlG* gene has been deleted, kan^R, zeo^R.	This study
psbAII::FLAG-At-*chlG*	-	N-terminally FLAG tagged copy of *Arabidopsis thaliana chlG* in place of *psbAII* gene, kan^R	This study
psbAII::FLAG-At-*chlG*/Δ*chlG*	FLAG-At	*psbAII*::FLAG-At-*chlG* strain in which the native *chlG* gene has been deleted, kan^R, zeo^R.	This study
psbAII::FLAG-*bchG*	-	N-terminally FLAG tagged copy of *Rhodobacter sphaeroides bchG* in place of *psbAII* gene, kan^R.	This study
psbAII::FLAG-*bchG*/Δ*chlG*^{NS}	-	*psbAII*::FLAG-*bchG* strain with non-segregated deletion of the native *chlG* gene, kan^R, zeo^R.	This study
Δ*psbB*	-	*psbB* (*slr0906*) deletion strain, zeo^R. Cannot grow under photoautotrophic conditions.	Cereda *et al*., 2014
psbAII::FLAG-6803_*chlG*_Δ1-11	Δ1-11	N-terminally FLAG tagged *Synechocystis* ChlG lacking first 11 codons in place of *psbAII* gene, kanamycin resistant (kan^R)	This study
psbAII::FLAG-6803_*chlG*_Δ1-23	Δ1-23	N-terminally FLAG tagged *Synechocystis chlG* lacking first 23 codons in place of *psbAII* gene, kanamycin resistant (kan^R)	This study
psbAII::FLAG-6803_*chlG*_Δ1-32	Δ1-32	N-terminally FLAG tagged *Synechocystis chlG* lacking first 32 codons in place of *psbAII* gene, kanamycin resistant (kan^R)	This study
psbAII::FLAG-6803_*chlG*_Δ1-39	Δ1-39	N-terminally FLAG tagged *Synechocystis chlG* lacking first 39 codons in place of *psbAII* gene, kanamycin resistant (kan^R)	This study
psbAII::FLAG-6803_*chlG*_Δ1-45	Δ1-45	N-terminally FLAG tagged *Synechocystis chlG* lacking first 45 codons in place of *psbAII* gene, kanamycin resistant (kan^R)	This study
psbAII::FLAG-6803_*chlG*_Δ1-51	Δ1-51	N-terminally FLAG tagged *Synechocystis chlG* lacking first 52 codons in place of *psbAII* gene, kanamycin resistant (kan^R)	This study
psbAII::FLAG-6803 *chlG*_Δ1-11/Δ*chlG*	Δ1-11/Δ*chlG*	*psbAII*::FLAG-6803_*chlG*_Δ1-11 strain in which native *chlG* gene is replaced with a zeocin resistance (zeo^R) cassette, kan^R.	This study
psbAII::FLAG-6803 *chlG*_Δ1-23/Δ*chlG*	Δ1-23/Δ*chlG*	*psbAII*::FLAG-6803_*chlG*_Δ1-23 strain in which native *chlG* gene is replaced with a zeocin resistance (zeo^R) cassette, kan^R.	This study
psbAII::FLAG-6803 *chlG*_Δ1-32/Δ*chlG*^{NS}	-	*psbAII*::FLAG-6803_*chlG*_Δ1-32 strain with non-segregated deletion of the native *chlG* gene, kan^R, zeo^R.	This study
psbAII::FLAG-6803 *chlG*_Δ1-39/Δ*chlG*^{NS}	-	*psbAII*::FLAG-6803_*chlG*_Δ1-39 strain with non-segregated deletion of the native *chlG* gene, kan^R, zeo^R.	This study
psbAII::6803_*chlG*-FLAG	-	C-terminally 3xFLAG-tagged *Synechocystis* sp. PCC 6803 *chlG* inserted in place of *psbAII* ; Kan^R	This study
psbAII::6803 *chlG*-FLAG/Δ*chlG*^{NS}	-	Partially-segregated deletion of *chlG* (merodiploid) in *chlG*.f background; Kan^R and Zeo^R	This study
Δ*crtR*	-	Deletion of *crtR* in *psbAII*::FLAG-6803_*chlG* /Δ*chlG* background; Kan^R, Zeo^R and Ery^R	This study
Δ*cruF*	-	Deletion of *cruF* in *psbAII*::FLAG-6803_*chlG* /Δ*chlG* background; Kan^R, Zeo^R and Sm^R	This study
Δ*crtR*/Δ*cruF*	-	Deletion of *cruF* in Δ*crtR* background; Kan^R, Zeo^R, Ery^R and Sm^R	This study
Δ*chlP*	-	Deletion of *chlP* in WT background; Kan^R, Zeo^R	Hitchcock *et al*., 2016

Table 3: Plasmids

Empty vectors

Plasmid	Properties	Source
pPD-FLAG	pBluescript II KS+ based vector containing a 3xFLAG encoding sequence. Km^R	Dr P. Davison and Dr D. Canniffe, University of Sheffield
pPD-CFLAG	A modified version of the pPD-FLAG vector enabelling C-terminal tagging of inserts. Km^R	Dr R. Sobotka, Institute of microbiology, Třeboň
pK18mobsacB	An allelic exchange suicide vector mobilised by *E. coli* S17-1. Allows blue white screening for inserts. Km^R	Novagen
pET28a	pBR322-based expression vector with the T7 promoter and terimator, His_6-tag and thrombin linker. Km^R	Novagen
pET21a	pBR322-based expression vector with the T7 promoter and terimator, His_6-tag and thrombin linker. Km^R, Amp^R	Novagen

Modified vectors

Plasmid	Properties (brackets indicate primers used to amplify DNA insert)	Source
pPD-NFLAG-6803_*chlG*	pPD-FLAG vector containing a N-terminally 3xFLAG-tagged *Synechocystis chlG* gene (1/2) flanked by *psbAII* up- and downstream regions	Chidgey *et al*., 2014
pPD-NFLAG-7002_*chlG*	pPD-FLAG vector containing a N-terminally 3xFLAG-tagged *Synechococcus chlG* gene (3/4) flanked by *psbAII* up- and downstream regions	This study
pPD-NFLAG-Cr_*chlG*	pPD-FLAG vector containing a N-terminally 3xFLAG-tagged *Chlamydomonas chlG* gene (5/6) flanked by *psbAII* up- and downstream regions	This study
pPD-NFLAG-At-*chlG*	pPD-FLAG vector containing a N-terminally 3xFLAG-tagged *Arabidopsis chlG* gene (7/8) flanked by *psbAII* up- and downstream regions	This study
pPD-NFLAG-*bchG*	pPD-FLAG vector containing a N-terminally 3xFLAG-tagged *bchG* gene (19/20) flanked by *psbAII* up- and downstream regions	This study
pPD-NFLAG-6803 *chlG*_Δ1-11	pPD-NFLAG-6803_*chlG* vector containing a modified *chlG* gene lacking the first 11 5' codons (39/43)	This study
pPD-NFLAG-6803 *chlG*_Δ1-23	pPD-NFLAG-6803_*chlG* vector containing a modified *chlG* gene lacking the first 23 5' codons (40/43)	This study
pPD-NFLAG-6803 *chlG*_Δ1-32	pPD-NFLAG-6803_*chlG* vector containing a modified *chlG* gene lacking the first 32 5' codons (41/43)	This study
pPD-NFLAG-6803 *chlG*_Δ1-39	pPD-NFLAG-6803_*chlG* vector containing a modified *chlG* gene lacking the first 39 5' codons (42/43)	This study
pPD-NFLAG-6803 *chlG*_Δ1-45	pPD-NFLAG-6803_*chlG* vector containing a modified *chlG* gene lacking the first 45 5' codons (44/43)	This study
pPD-NFLAG-6803 *chlG*_Δ1-51	pPD-NFLAG-6803_*chlG* vector containing a modified *chlG* gene lacking the first 51 5' codons (45/43)	This study
pPD-CFLAG-6803_*chlG*	pPD-FLAG vector containing a C-terminally 3xFLAG-tagged *Synechocystis chlG* gene (37/38) flanked by *psbAII* up- and downstream regions	This study
pK18mobsacB-Δ*bchF**	*bchF* knock out construct containing regions of homology to sequence upstream (48/49) and within (50/51) the gene.	This study
pET28a::At_*chlG*	pET28a plasmid contaning an N-terminally 6xHis-tagged *Arabidopsis chlG* gene (46/47) codon optimised for expression in *E. coli*.	This study
pET28a::At_*chlG* Q46E	pET28a::At_*chlG* plasmid contaning a modified *Arabidopsis chlG* gene in which codon 46 has been changed from CAA to GAA (55/56)	This study
pET28a::At_*chlG* P54F	pET28a::At_*chlG* plasmid contaning a modified *Arabidopsis chlG* gene in which codon 54 has been changed from CCA to TTT (57/58)	This study
pET28a::At_*chlG* L56P	pET28a::At_*chlG* plasmid contaning a modified *Arabidopsis chlG* gene in which codon 56 has been changed from TTA to CCA (59/60)	This study
pET28a::At_*chlG* V60Y	pET28a::At_*chlG* plasmid contaning a modified *Arabidopsis chlG* gene in which codon 60 has been changed from GTG to TAC (61/62)	This study
pET28a::At_*chlG* N99A	pET28a::At_*chlG* plasmid contaning a modified *Arabidopsis chlG* gene in which codon 99 has been changed from AAT to GCT (63/64)	This study
pET28a::At_*chlG* A225M	pET28a::At_*chlG* plasmid contaning a modified *Arabidopsis chlG* gene in which codon 225 has been changed from GCA to ATG (65/66)	This study
pET21a::*CLH-1*	pET21a plasmid contaning an *Arabidopsis CLH-1* gene (52/54) codon optimised for expression in *E. coli*.	This study
pET21a::His-*CLH-1*	pET21a plasmid contaning an N-terminally 6xHis-tagged *Arabidopsis CLH-1* gene (53/54) codon optimised for expression in *E. coli*.	This study

Table 4: Primers

Primers in order of appearance (brackets indicate primer number)

Primer	Sequence (5' - 3')	Restriction site	Source
Synechocystis chlG Forward (1)	GCGCTCTAGA**CATATG**GTGAGCAAGGGCGAGGAGCTGTTCAC	NdeI	Chidgey *et al.*, 2014
Synechocystis chlG Reverse (2)	GCGCTCTAGA**GCGGCCGC**CTTGTACAGCTCGTCCATGCCGAGAGTGAT	NotI	Chidgey *et al.*, 2014
Synechococcus chlG Forward (3)	GGAATTC**GCGGCCGC***A*CCCAATGACGAGTGGTTTTC	NotI	This study
Synechococcus chlG Reverse (4)	GGAATTC**AGATCT**TTAGGAAATCCCCGCATGGC	BglII	This study
Chlamydomonas chlG Forward (5)	GGAATTC**GCGGCCGC**AGCGATGAACCAACAGGCC	NotI	This study
Chlamydomonas chlG Reverse (6)	GGAATTC**AGATCT**CTATAATGCACCAGCAGC	BglII	This study
Arabidopsis chlG Forward (7)	GGAATTC**GCGGCCGC**AGCTGCAGAAACAGACACC	NotI	This study
Arabidopsis chlG Reverse (8)	GGAATTC**AGATCT**CTAGTGTTGGCTGGCCAA	BglII	This study
psbAII-upstream (9)	AAACGCCCTCTGTTTACCCA	-	This study
psbAII-downstream (10)	TCAACCCGGTACAGAGCTTC	-	This study
chlG-upstream Forward (11)	TCGTTGAGCGGGAGAGTTTG	-	Chidgey *et al.*, 2014
chlG-upstream Reverse (12)	ACATTAATTGCGTTGCGCTCACTGCGGCTTCATCAACTGAAGACG	-	Chidgey *et al.*, 2014
chlG-downstream Forward (13)	CAACTTAATCGCCTTGCAGCACATCAAGTCCTTGCCGGTGATGT	-	Chidgey *et al.*, 2014
chlG-downstream Reverse (14)	AATTGACCAACCATTTCTCC	-	Chidgey *et al.*, 2014
Zeo internal forward (15)	CGGGTCGCGCAGGGCGAAC	-	Chidgey *et al.*, 2014
Zeo internal reverse (16)	CAACTTAATCGCCTTGCAGCACATCAAGTCCTTGCCGGTGATGT	-	Chidgey *et al.*, 2014
chlG-locus Forward (17)	TGCAAGCAACCCGTTACCA	-	This study
chlG-locus Downstream (18)	CCCTTTAGATTTTAGGACGGCGA	-	This study
bchG Forward (19)	AATCAG**GCGGCCGCA**TCTGTGAACCTGTCTCTCCATCCC	NotI	This study
bchG Reverse (20)	AATCAG**AGATCT**TCAGGGTAAGACCTCGAGCCC	BglII	This study
Ery^R Forward (21)	TATTCAATAATCGCATCCGATTGCAG	-	This study
Ery^R Reverse (22)	CACACGAAAAACAAGTTAAGGGATGC	-	This study
crtR Upstream Forward (23)	TGTCAAGTCTGTTGACCAAAAGAG	-	This study
crtR Upstream Reverse (24)	CTGCAATCGGATGCGATTATTGAATAGCACCGAACAACTAAAACAAAGC	-	This study
crtR Downstream Forward (25)	GCATCCCTTAACTTGTTTTTCGTGTGCTTGGTATCAGTACCAAAACACC	-	This study
crtR Downstream Reverse (26)	TCCGTCAATACACCATCTGGC	-	This study
crtR-locus Forward (27)	TGTCAAGTCTGTTGACCAAAAGAGC	-	This study
crtR-locus Reverse (28)	TCCGTCAATACACCATCTGGC	-	This study

aadA Forward (29)	ACCGAGTGAGCTAGCTATTTG	-	This study
aadA Reverse (30)	TTATTTGCCGACTACCTTGGTG	-	This study
cruF Upstream Forward (31)	AACAAACTCCCACAACACCTC	-	This study
cruF Upstream Reverse (32)	CTGCAATCGGATGCGATTATTGAATAACCACCAGCCATAGACC	-	This study
cruF Downstream Forward (33)	GCATCCCTTAACTTGTTTTTCGTGTGTTGCCATGGTGATGAGC	-	This study
cruF Downstream Reverse (34)	AGTTGAGTCTTTTACACTCGATCG	-	This study
cruF-locus Forward (35)	AACAAACTCCCACAACACCTC	-	This study
cruF-locus Reverse (36)	AGTTGAGTCTTTTACACTCGATCG	-	This study
chlG-FLAG Forward (37)	ATATGGG**CATATG**TCTGACACACAAAATACCGGCCAAAAC	NdeI	This study
chlG-FLAG Reverse (38)	TAAGG**ACTAGT**AATCCCCGCATGGCCTAG	SpeI	This study
*chlG*_Δ1-11 Forward (39)	AGAATTC**GCGGCCGC**AGCCAAGGCTCGGCAGTTACTGG	NotI	This study
*chlG*_Δ1-23 Forward (40)	AGAATTC**GCGGCCGC**AGCCCCGGGGGAAAGTTCC	NotI	This study
*chlG*_Δ1-32 Forward (41)	AGAATTC**GCGGCCGC**AATTCGTCTTCAGTTGATGAAGC	NotI	This study
*chlG*_Δ1-39 Forward (42)	AGAATTC**GCGGCCGC**ACCCATCACTTGGATTCCCCTG	NotI	This study
chlG Reverse (43)	AGTTC**AGATCT**TCAAATCCCCGCATGGCCTAGG	BglII	This study
*chlG*_Δ1-45 Forward (44)	AGAATTC**GCGGCCGC**ACTGATCTGGGGGGTGGTCTG	NotI	This study
*chlG*_Δ1-51 Forward (45)	AGAATTC**GCGGCCGC**ATGTGGGGCCGCTTCTTCC	NotI	This study
At_*chlG* Forward (46)	TAAGC**CATATG**GCTGCAGAAACAGACACCG	NdeI	This study
At_*chlG* Reverse (47)	ATTCG**GAATTC**CTAGTGTTGGCTGGCCAATG	EcoRI	This study
bchF upstream forward (48)	CCG**GAATTC**CGGCGCGACACAAGCCCG	EcoRI	This study
bchF upstream reverse (49)	CGCGTCGCATCTGCATCTTGAGGTTCGCCTTCC	-	This study
bchF downstream forward (50)	CTCAAGATGCAGATGCGACGCGCTGGACCCTG	-	This study
bchF downstream reverse (51)	CCCC**AAGCTT**GCTGGCCCGTGAAGAGCG	HindIII	This study
At_*CLH-1* Forward (52)	ATATGGG**CATATG**GCGGCCATCGAAGAC	NdeI	This study
At_His-*CLH-1* Forward (53)	TAAGG**CTCGAG**GACGAAGATGCCAGAGGCTTC	XhoI	This study
At_*CLH-1* Reverse (54)	TAAGG**CTCGAG**CTAGACGAAGATGCCAGAGGCTTC	XhoI	This study
At_*chlG* Q46E Forward (55)	AGTGGAAGATTCGACTCGAACTAACCAAGCCCGTG	-	This study
At_*chlG* Q46E Reverse (56)	TCACGGGCTTGGTTAGTTCGAGTCGAATCTTCCAC	-	This study

At_*chlG* P54F Forward (57)	ACCAAGCCCGTGACTTGGTTTCCTTTAGTGTGGGGCGTG	-	This study
At_*chlG* P54F Reverse (58)	CACGCCCCACACTAAAGGAAACCAAGTCACGGGCTTGG	-	This study
At_*chlG* L56P Forward (59)	GTGACTTGGCCACCTCCAGTGTGGGGCGTGGTG	-	This study
At_*chlG* L56P Reverse (60)	CACCACGCCCCACACTGGAGGTGGCCAAGTCAC	-	This study
At_*chlG* V60Y Forward (61)	ACCTTTAGTGTGGGGCTACGTGTGTGGTGCCGCGGCGTC	-	This study
At_*chlG* V60Y Reverse (62)	ACGCCGCGGCACCACACACGTAGCCCCACACTAAAGGTG	-	This study
At_*chlG* N99A Forward (63)	ACCGGGTATACTCAGACTATTGCTGACTGGTACGATCGCGATATTG	-	This study
At_*chlG* N99A Reverse (64)	CAATATCGCGATCGTACCAGTCAGCAATAGTCTGAGTATACC	-	This study
At_*chlG* A225M Forward (65)	TCCATCGCAGGGTTGGGCATTATGATCGTTAATGATTTCAAATCTGTAG	-	This study
At_*chlG* A225M Reverse (66)	CTACAGATTTGAAATCATTAACGATCATAATGCCCAACCCTGCGATGG	-	This study

Chapter 3: Production of functional plant and algal chlorophyll synthases in cyanobacteria indicates a conserved interaction with the YidC/Alb3 membrane insertase

3.1 Summary

The *Synechocystis* ChlG-HliD-Ycf39-YidC complex acts at the interface between the chlorophyll biosynthesis and photosystem assembly pathways to coordinate delivery of *de novo* chlorophyll pigments to the chlorophyll binding proteins of the thylakoid membrane (Chidgey et al., 2014). To gain an insight into the ubiquity of such an assembly complex in higher photosynthetic organisms, *chlG* genes from the cyanobacterium *Synechococcus* sp. PCC 7002, the alga *Chlamydomonas reinhardtii* and plant *Arabidopsis thaliana* were N-terminally FLAG tagged and produced in a cyanobacterial host, *Synechocystis* sp. PCC 6803. The enzymes were retrieved by FLAG immunoprecipitation and analysed for binding partners by SDS-PAGE and immunoblots.

Production of the foreign chlorophyll synthases allowed subsequent deletion of the native cyanobacterial *chlG* gene, with all mutant strains capable of autotrophic growth under normal conditions. Analysis of the purified protein complexes revealed that the *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* ChlG proteins do not associate with cyanobacterial HliD or Ycf39, whereas the cyanobacterial homolog from *Synechococcus* sp. PCC 7002 was capable of forming the same complex as the native enzyme. The interaction with YidC was maintained for both eukaryotic enzymes, indicating that a ChlG-YidC/Alb3 complex may be evolutionarily conserved in algae and higher plants.

3.2 Introduction

Photosynthetic organisms, such as plants, algae and cyanobacteria, utilise chlorophyll (Chl) pigments situated in membrane intrinsic photosystems to absorb solar energy, which is used for charge separation to drive ATP and NADPH production. The structures of PSI and PSII show that the Chl molecules are specifically arranged within the photosystems to allow highly efficient light capture and energy transfer (Caspy and Nelson, 2018; Jordan *et al*., 2001; Umena *et al*., 2011). The complexity and hydrophobicity of these protein-pigment complexes makes their assembly dependent on numerous auxiliary factors and that the insertion of Chl and other cofactors must be coordinated with the assembly process (Nickelsen and Rengstl, 2013; Yang *et al*., 2015).

Chlorophyll synthase (ChlG) is the terminal enzyme of the Chl biosynthesis pathway and catalyses the conversion of chlorophyllide (Chlide) to Chl by addition of the phytyl tail moiety to the propionate residue at the C17 position on ring D of the Chlide macrocycle (Figure 3.1). ChlG is discussed at length in Section 1.8.11. In a previous study, Chidgey *et al* (2014) investigated handover of Chl from ChlG to nascent light-harvesting polypeptides. ChlG in *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*) was N-terminally FLAG-tagged and used as bait in FLAG-immunoprecipitation experiments to retrieve an enzymatically active protein-pigment complex containing the high-light inducible protein HliD, the photosystem II assembly factor Ycf39 and the membrane insertase YidC, as well as associated carotenoids and the Chl precursor Chlide. Size exclusion chromatography revealed that ChlG was found to bind tightly to HliD in a ChlG-HliD ‘core’ complex, with Ycf39 and YidC more loosely associated. This study provided the first evidence for a link between Chl biosynthesis and YidC-dependent co-translational insertion of nascent light harvesting polypeptides into membranes. Pigment binding to this complex was found to be mediated by the HliD component. HliD belongs to a conserved family of high-light induced proteins (Hlips) (Komenda and Sobotka, 2016), the members of which share significant sequence similarity with plant Chl *a*/*b* binding proteins and possess a conserved Chl *a*-binding (CAB) motif (Bhaya *et al*., 2002). Ycf39 is an atypical short-chain dehydrogenase reported to have a role in photosystem II (PSII) assembly (Knoppová *et al*., 2014) and is often found closely associated with HliD (Knoppová *et al*., 2014; Staleva *et al*., 2015). YidC belongs to the evolutionarily conserved YidC/OxaI/Alb3 family of membrane insertase proteins with homologs found in bacteria, mitochondria and chloroplasts. YidC aids the folding and partitioning of nascent transmembrane polypeptides into the phospholipid bilayer of cells, often in association with the SecYEG apparatus (Beck *et al*., 2001; Nagamori *et al*., 2004). Thylakoid membrane biogenesis in cyanobacteria and higher photosynthetic organisms is known to be dependent on YidC (Göhre *et al*., 2006; Spence *et al*., 2004). HliD, YidC and Ycf39 are discussed in greater detail in Sections 1.11.1, 1.11.2 and 1.11.3 respectively.

Because of the high degree of similarity between cyanobacterial photosystems and those of higher photosynthetic organisms, it is likely that equivalent Chl handover systems operate in algae and plants. In this chapter, foreign ChlG enzymes from a cyanobacteria, algae and plant were N-terminally FLAG-tagged and produced in *Synechocystis* with the aim of investigating whether these homologs can complement the deletion of the native *chlG* gene and form a ChlG complex similar to that reported by (Chidgey et al., 2014).

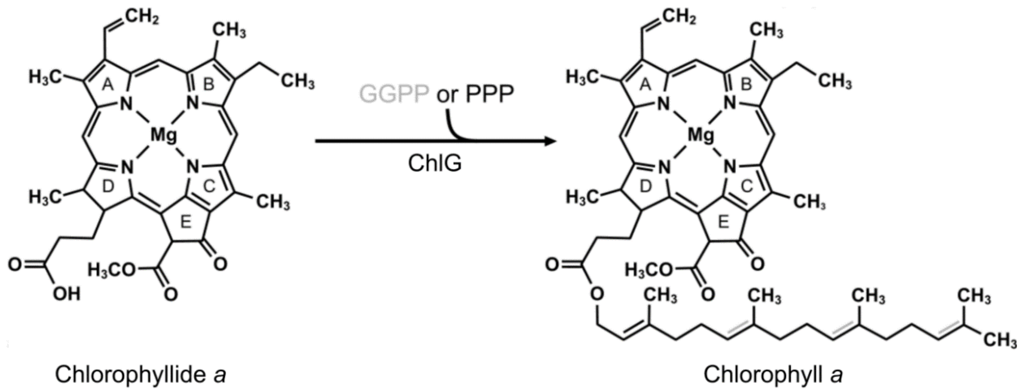


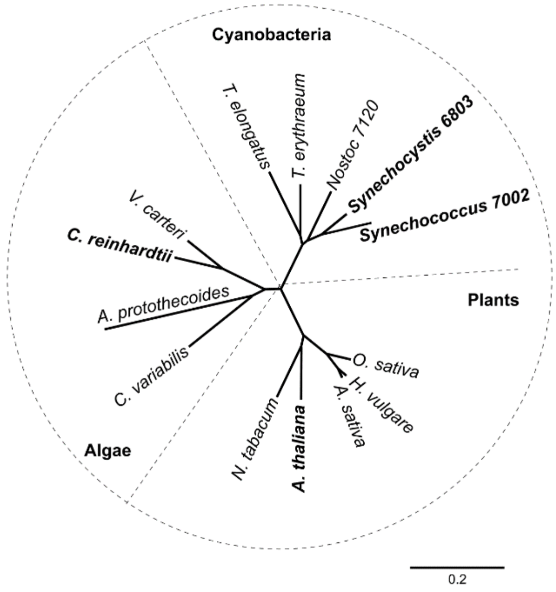
Figure 3.1: The reaction catalysed by chlorophyll synthase. Chlorophyll synthase (ChlG) catalyses the esterification of chlorophyllide with either geranylgeranyl pyrophosphate (GGPP) or phytyl-pyrophosphate (PPP) resulting in geranylgeranylated chlorophyll *a* (Chl a_{GG}) or phytylated chlorophyll *a* (Chl *a*). Three carbon-carbon double bonds (shown with grey lines) in the geranylgeranyl tail of Chl a_{GG} are sequentially reduced to phytol by the geranylgeranyl diphosphate reductase ChlP.

3.3 Results

3.3.1 Generation of *Synechocystis* strains producing foreign FLAG-tagged chlorophyll synthases

To investigate whether ChlG proteins from algae and higher plants are functional in a cyanobacterial system and to observe interactions between these eukaryotic ChlGs and the cyanobacterial HliD, Ycf39 and YidC proteins, a collection of mutant strains was generated in which foreign *chlG* genes were added *in trans* to *Synechocystis*. Protein phylogeny shows that the cyanobacterial, algal and plant *chlG* genes form separate branches (Figure 3.2), thus enzymes from another cyanobacterium, *Synechococcus* sp. 7002 (hereafter Syn 7002); the model green alga, *Chlamydomonas reinhardtii*; and the model plant species, *Arabidopsis thaliana* were chosen. The *A. thaliana* and *C. reinhardtii chlG* genes lacking the sequence coding for the N-terminal chloroplast transit peptides (57 or 43 amino acids respectively, according to the ChloroP 1.1 Server (Emanuelsson *et al*., 1999)), were codon optimised for expression in *Synechocystis* and commercially synthesised (Integrated DNA Technologies). The *chlG* gene from Syn 7002 was amplified from genomic DNA. The genes were cloned into the pPD-*N*FLAG plasmid (Hollingshead *et al*., 2012) such that they were in frame with sequence encoding an N-terminal 3xFLAG tag. The pPD-*N*FLAG-*chlG* plasmid described in Chidgey *et al.* (2014) was used to replace the *psbAII* gene with the FLAG-tagged *Synechocystis chlG* gene. The existing *Synechocystis* FLAG-*chlG* (hereafter FLAG-6803) strain described in Chidgey *et al.* (2014) was produced in a cell line shown to be unsuitable for further work (Tichý *et al*., 2016). These vectors were therefore introduced into another glucose tolerant cell line (Nixon background) (Williams, 1988) (kindly provided by Roman Sobotka) by natural transformation and, following homologous recombination into the *Synechocystis* genome at the *psbAII* locus, the tagged-construct replaced the *psbAII* gene so that expression was under the control of the *psbAII* promoter (Hollingshead *et al*., 2012) (Figure 3.3A). Segregation of genome copies was confirmed by PCR using primers flanking the *psbAII* locus (Figure 3.3B).

A



B

```
6803      ------------------------------MSDTQNTGQNQAKARQLLGMKGAAPGESSI
7002      MPNDEWFSFTLFFCTKYSNPDAMTETPNPDTKPATAPEEQGSKARQLLGMKGAAGGETSI
Cr        --------------------------MAMNQQATEEKSDTNSAARQMLGMKGAA-LETDI
At        -------------------MAAETDTDKVKSQTPDKAPAGGSSINQLLGIKGAS-QETNK
                                         .         :  .*:**:***:  *:.

6803      WKIRLQLMKPITWIPLIWGVVCGAASSGGYIWS-VEDFLKALTCMLLSGPLMTGYTQTLN
7002      WKIRLQLMKPITWIPLIWGVVCGAASSGGYVWG-VEDFLKAMTCMLLSGPLLTGYTQTLN
Cr        WKIRVQLTKPVTWIPLIWGVACGAAASGHYQWNNPTQIAQLLTCMMMSGPFLTGYTQTIN
At        WKIRLQLTKPVTWPPLVWGVVCGAAASGNFHWT-PEDVAKSILCMMMSGPCLTGYTQTIN
          ****:** **:** **:***.****:** : *     :. : : **::*** :******:*

6803      DFYDRDIDAINEPYRPIPSGAISVPQVVTQILILLVAGIGVAYGLDVWAQHDFPIMMVLT
7002      DFYDREIDAINEPYRPIPSGAISVPQVVTQILVLLGSGIGLSYLLDLWAGHDFPVMLVLT
Cr        DWYDREIDAINEPYRPIPSGRISERDVIVQIWVLLLGGIGLAYTLDQWAGHTTPVMLQLT
At        DWYDRDIDAINEPYRPIPSGAISEPEVITQVWVLLLGGLGIAGILDVWAGHTTPTVFYLA
          *:***:************** **  :*:.*: :** .*:*::  ** ** *  * :: *:

6803      LGGAFVAYIYSAPPLKLKQNGWLGNYALGASYIALPWWAGHALFGTLNPTIMVLTLIYSL
7002      VVGCFIAYIYSAPPLKLKQNGWLGNYALGASYIALPWWAGHALFGTLTPTVMVVTLIYSF
Cr        IFGSFISYIYSAPPLKLKQSGWAGNYALGSSYIALPWWAGQALFGTLTLDVMALTIAYSL
At        LGGSLLSYIYSAPPLKLKQNGWVGNFALGASYISLPWWAGQALFGTLTPDVVVLTLLYSI
          : *.:::************.** **:***:***:******:******.  ::.:*: **:

6803      AGLGIAVVNDFKSVEGDRQLGLKSLPVMFGIGTAAWICVIMIDVFQAGIAGYLIY-VHQQ
7002      AGLGIAVVNDFKSVEGDRQLGLKSLPVMFGVGTAAWICVLMIDIFQVGIAGYLVS-IHEQ
Cr        AGLGIAIVNDFKSIEGDRQMGLQSLPVAFGVDTAKWICVSTIDVTQLGVAAYLAWGLHEE
At        AGLGIAIVNDFKSVEGDRALGLQSLPVAFGTETAKWICVGAIDITQLSVAGYLLA-SGKP
          ******:******:**** :**:**** **  ** ****  **: * .:*.**     :

6803      LYATIVLLLLIPQITFQDMYFLRNPLENDVKYQASAQPFLVFGMLATGLALGHAGI----
7002      LYATILLLLVIPQITFQDMYFLRDPLKNDVKYQASAQPFLVLGMLVAGLAMGHAGIS---
Cr        LYGAVLLALILPQIYFQYKYFLPDPIANDVKYQASAQPFLVFGLLTAGLACGHHVNAVAA
At        YYALALVALIIPQIVFQFKYFLKDPVKYDVKYQASAQPFLVLGIFVTALASQH-------
           *.  :: *::*** **  *** :*:  *************:*::.:.**  *

6803      --------      324     % sequence identity relative to 6803
7002      --------      355     7002: 86%
Cr        AASAAGAL      341     Cr:   66%
At        --------      331     At:   66%
```

Figure 3.2: Protein phylogeny of chlorophyll synthases from representative cyanobacteria, algae and plants. (A) The chlorophyll synthases used in this study are shown in bold. The scale bar indicates the number of amino acid substitutions per site. (B) Sequence alignment of ChlG homologs from *Synechocystis* (6803), *Synechococcus* (7002), *Chlamydomonas reinhardtii* (Cr) and *Arabidopsis thaliana*, aligned using Clustal Omega. The sequence identities of 7002, Cr and At in comparison to 6803 were calculated using BLAST software from NCBI.

A

B

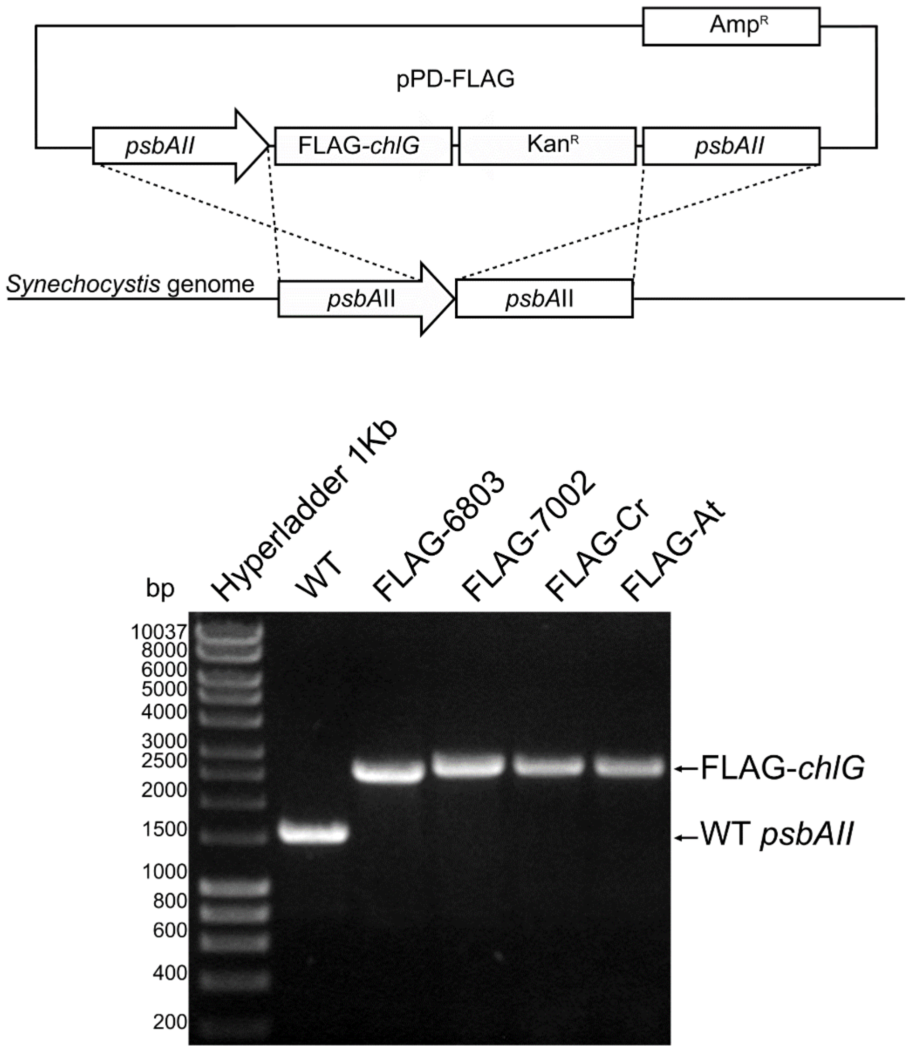


Figure 3.3: Generation of *Synechocystis* strains expressing foreign chlorophyll synthase genes. (A) Constructs encoding 3xFLAG tagged chlorophyll synthases were inserted in place of the *psbAII* gene in the *Synechocystis* genome. (B) PCR amplification of the *psbAII* locus of WT and transformant FLAG-*chlG Synechocystis* strains.

3.3.2 Algal and plant chlorophyll synthases can functionally replace the native enzyme

To see if the foreign FLAG-tagged ChlG enzymes were functional in the recombinant *Synechocystis* strains, deletion of the native *chlG* (slr0056) was attempted. As *chlG* is an essential gene, successful deletion would indicate that the foreign ChlG enzymes are able to replace the function of the WT protein. Deletion of the native *chlG* gene was achieved using the linear mutagenesis construct described by Chidgey *et al* (2014). This construct replaced most of the *chlG* gene with a zeocin resistance cassette (*zeo*^{R}) (Figure 3.4A). Segregation was carried out by selection on zeocin and successful recombinants checked by PCR screening (Figure 3.4B). The resultant strains: *psbAII::*3xFLAG-*Syn* 7002_*chlG*/Δ*chlG* (hereafter FLAG-7002), *psbAII::* 3xFLAG-*C. reinhardtii*_*chlG*/Δ*chlG* (hereafter FLAG-Cr), *psbAII::* 3xFLAG-*A. thaliana*_*chlG*/Δ*chlG* (hereafter FLAG-At) were capable of photoheterotrophic and photoautotrophic growth (Figure 3.5A-B) and have absorption profiles similar to that of WT cells (Figure 3.5C), indicating that the foreign ChlG proteins can replace the function of the endogenous enzyme, at least under standard growth conditions.

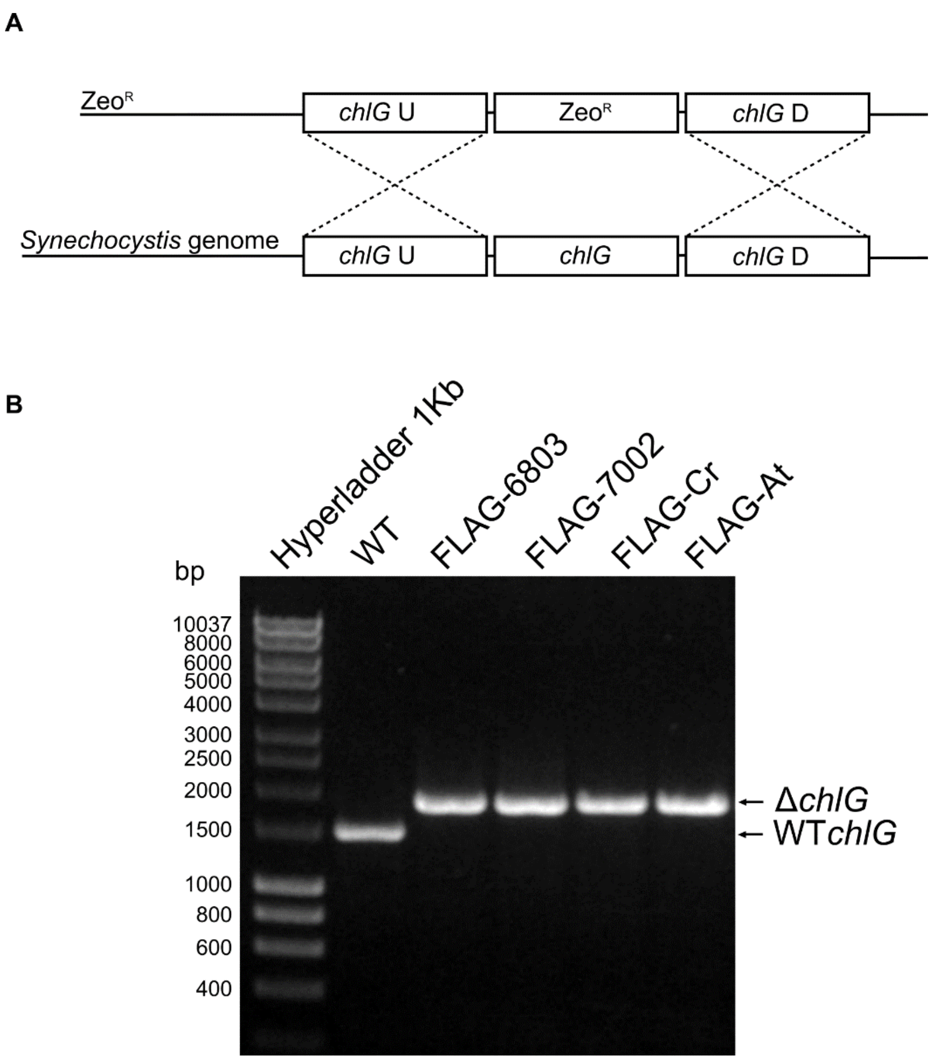


Figure 3.4: Deletion of *chlG* from the FLAG-*chlG* mutant strains. (A) The native *chlG* gene was deleted from the strains producing foreign FLAG-tagged ChlGs by replacement with a zeocin resistance cassette (Zeo^R). (B) Complete segregation at *chlG* locus was confirmed by PCR with primers flanking the integration sites.

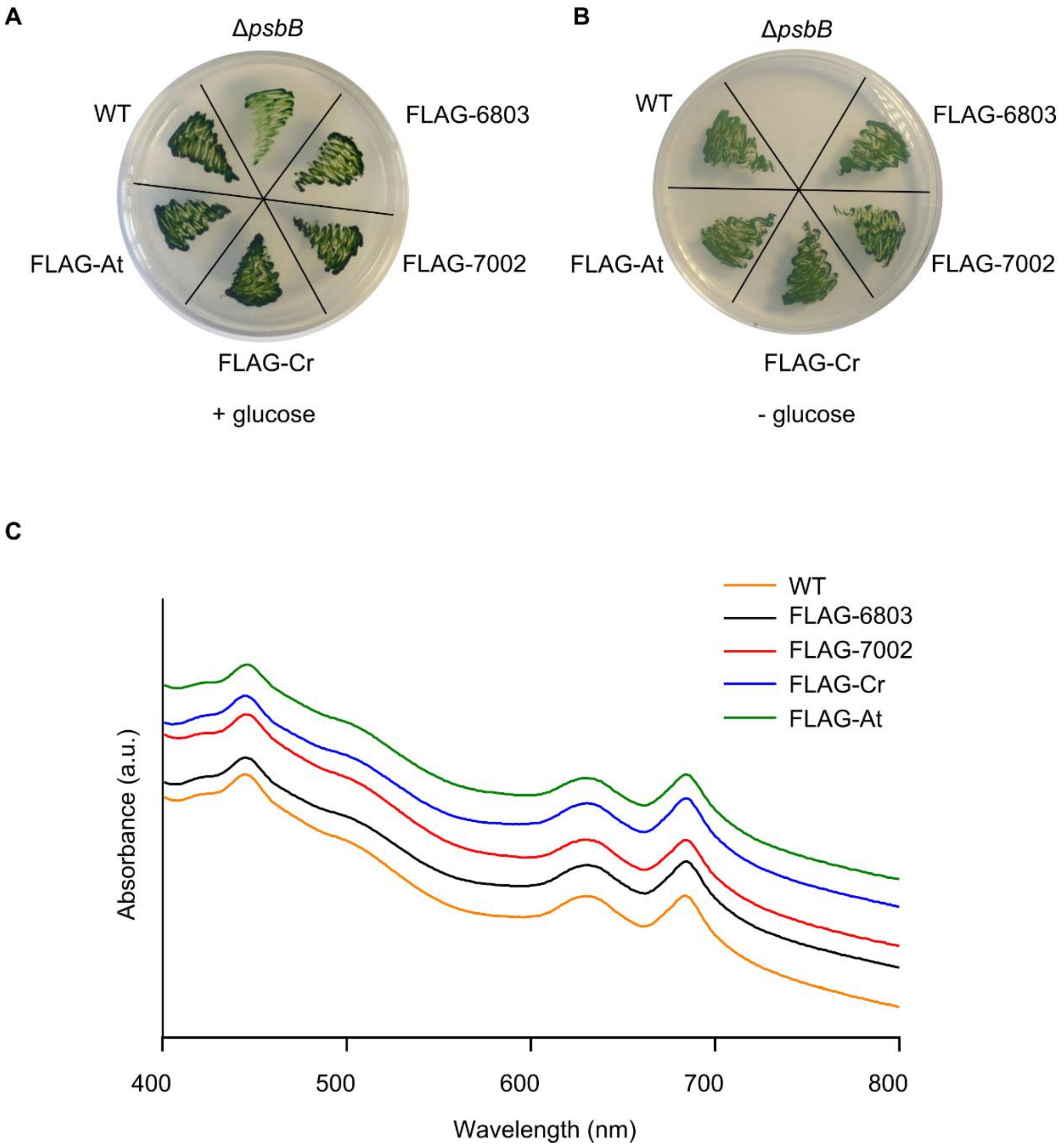


Figure 3.5: Growth and whole cell spectra of *Synechocystis* strains. (A-B) Mixotrophic (A) and photoautotrophic (B) growth of *Synechocystis* strains generated in this chapter. For comparison, a *Synechocystis* Δ*psbB* mutant that cannot grow under photoautotrophic conditions is used as a control (Cereda *et al*., 2014). (C) Whole cell absorbance spectra of *Synechocystis* strains. Spectra are normalised at 575 nm and offset to allow individual traces to be distinguished.

3.3.3 Chlorophyll biosynthesis is unaffected in the mutant strains

Although production of the plant and algal chlorophyll synthases allow inactivation of the native enzyme, and the recombinant *Synechocystis* strains are capable of photoautotrophic growth and produced WT levels of photosynthetic complexes, the effects of the replacement of the *Synechocystis* enzyme with homologs from *A. thaliana* and *C. reinhardtii* may have been more subtle. As the terminal enzyme in the Chl biosynthesis pathway, any adverse affects of replacing native *chlG* with foreign homologs could affect preceding steps in the pathway resulting in the accumulation of Chl biosynthetic intermediates. Pigments were extracted with excess methanol and the relative concentrations of Chl and its precursors were determined by reverse-phase HPLC. The levels of Chl precursors were broadly comparable between all strains (Figure 3.6). The ratio of monovinyl-chlorophyllide (MV-chlide) to divinyl-chlorophyllide (DV-chlide) varied slightly between the mutants, with the ratio being greater in FLAG-6803 (Figure 3.6A), smaller in the FLAG-7002 and FLAG-At strains (Figure 3.6B and 3.6D) and approximately equal in FLAG-Cr strain (Figure 3.4C). However, these pigments vary in concentration within WT cells and so the slight discrepancy between strains is unlikely to be due to the substitution of *chlG* with the foreign homologs (Roman Sobotka, personal communication). There was also no significant difference between the Chl content of the strains (Table 1). Taken together, it can be concluded that the replacement of native ChlG with plant and algal enzymes maintains a functioning Chl biosynthesis pathway under normal growth conditions.

Table 1: Chlorophyll content of strains used in this study.

Strain	Chlorophyll (µg ml^{-1} OD_{750}^{-1})
WT	5.0±0.20
FLAG-6803	5.1±0.18
FLAG-7002	5.2±0.22
FLAG-Cr	4.6±0.13
FLAG-At	4.9±0.26

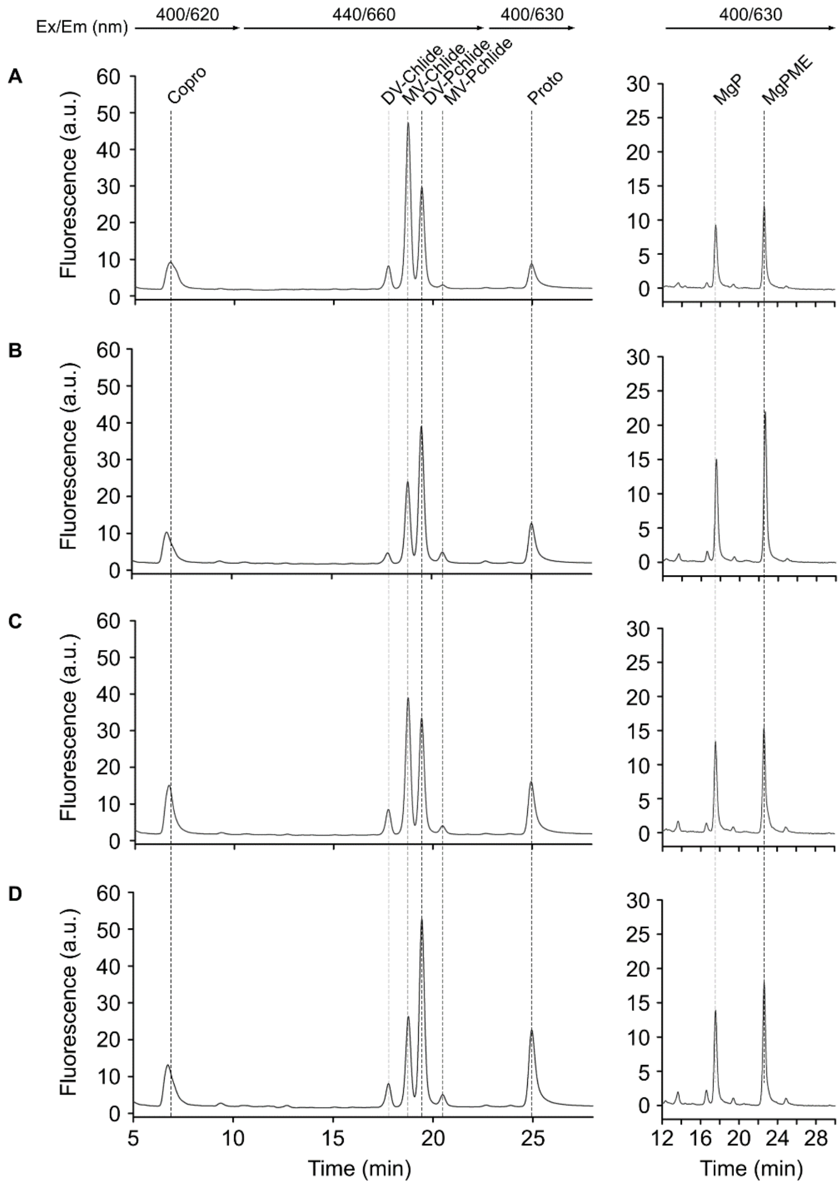


Figure 3.6: Analysis of chlorophyll precursors from the FLAG-6803 (A), FLAG-7002 (B), FLAG-Cr (C) and FLAG-At (D) strains of *Synechocystis*. Pigments were extracted from cell pellets and separated by reverse phase HPLC (Section 2.14.5). The left and right hand chromatograms were recorded simultaneously using two fluorescence detectors. The excitation (Ex) and emission (Em) wavelengths used are shown above the chromatograms. Copro = coproporphyrin IX; DV-Chlide = divinyl chlorophyllide *a*; MV-Chlide = monovinyl chlorophyllide *a*; DV-PChlide = divinyl protochlorophyllide *a*; MV-PChlide = monovinyl protochlorophyllide *a*; Proto = protoporphyrin IX; MgP = Mg-protoporphyrin IX; MgPME = Mg-protoporphyrin IX monomethyl ester.

3.3.4 Growth of the FLAG-At strain is impeded in cold conditions

All of the mutant strains were capable of normal growth under moderate light and temperatures. As the *A. thaliana chlG* homolog is phylogenetically the most distant from the *Synechocystis* homolog, the FLAG-At strain was subjected to a variety of stress conditions to try and find a growth phenotype for this mutant in comparison to the FLAG-6803 strain. FLAG-At and FLAG-6803 cells were cultured on solid and in liquid media and exposed to high-light (800 µmol photon m^{-2} s^{-1}), fluctuating light (30 second pulses of 100 µmol photon m^{-2} s^{-1} light followed by 2 minutes of no light) and cold (20 °C) stress conditions. The only condition under which FLAG-At cells grew more slowly than the FLAG-6803 strain was when cultured on BG11 agar at 20 °C (Figure 3.7). This phenotype was tested in two separate experiments, when cells were grown photoautotrophically (Figure 3.7A) and mixotrophically (Figure 3.7B). In both instances there was a significant reduction in the growth rate of the FLAG-At strain in comparison to FLAG-6803 WT cells.

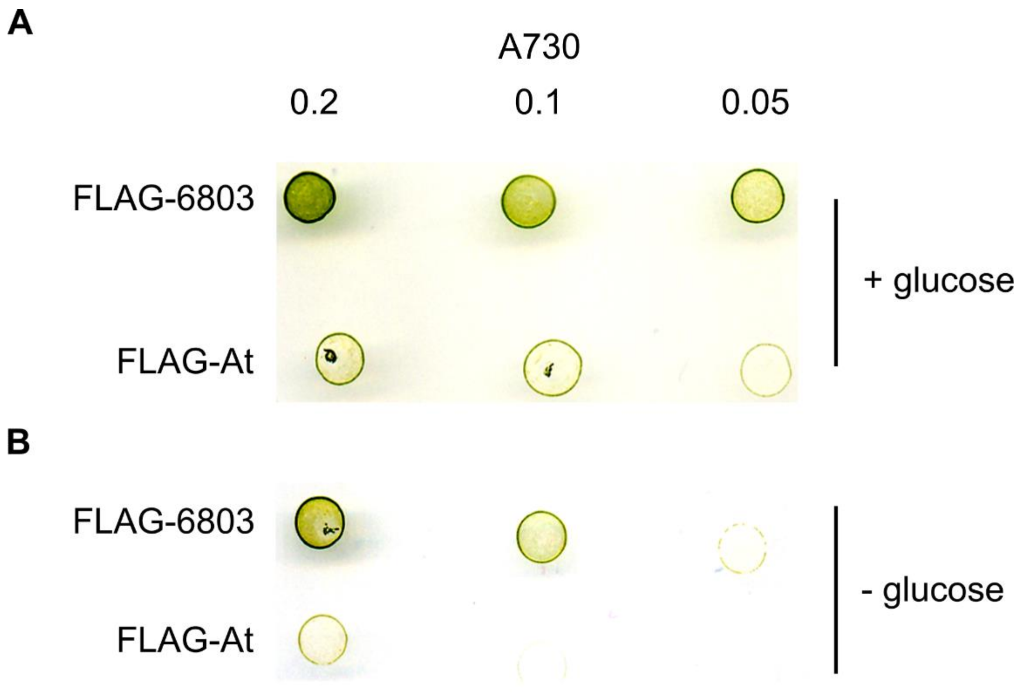


Figure 3.7: Drop-growth assay of FLAG-6803 and FLAG-At at 20 °C with 100 µmol photon m-2 s-1 of illumination. FLAG-6803 and FLAG-At cells grown under normal conditions in liquid medium were serially diluted to various cell densities, plated onto BG11 agar and grown photoheterotrophically (A) and photoautotrophically (B) at 20 °C.

3.3.5 The algal and plant chlorophyll synthases co-purify with YidC but not HliD or Ycf39

Membrane fractions were prepared from the FLAG-6803, FLAG-7002, FLAG-Cr and FLAG-At strains and solubilised in β-DDM for FLAG-immunoprecipitation experiments. A WT control was included to check for any non-specific protein interactions with the anti-FLAG resin (Figure 3.10). The presence of the bait protein in the resultant eluates was confirmed by SDS-PAGE (Figure 3.8A). In all cases, a band approximately 30 kDa in size was visible and confirmed to be FLAG-ChlG by immunoblot using antibodies raised against the FLAG-tag (Figure 3.8B). A band was also visible around the 10 kDa marker in the FLAG-6803 and FLAG-7002 but not in the plant or algal ChlG eluates. This was confirmed as HliD by immunoblot (Figure 3.8B). Also visible, particularly in the FLAG-At and FLAG-Cr eluates, were bands around 60 kDa in size (marked with asterisk in Figure 3.8A). These gave a positive signal when probed with both FLAG and ChlG antibodies indicating that they are ChlG dimers (FLAG-ChlG[2]) (Figure 3.8B).

Spectrophotometric analysis of the immunoprecipitation eluates shows that the FLAG-6803 and FLAG-7002 eluates, both of which were visibly yellow-green, contain Chl (436 nm, 674 nm) and carotenoids (487 nm, 515 nm) (Figure 3.8C). The FLAG-Cr and FLAG-At eluates appeared colourless to the naked eye and contained negligible levels of pigment when analysed spectroscopically (Figure 3.8C).

In order to further analyse the components of the FLAG-immunoprecipitation eluates, immunoblot experiments were carried out in which samples were interrogated with antibodies specific to HliD, Ycf39 and YidC (Figure 3.8B). The FLAG-6803 and FLAG-7002 eluates both yielded positive signals for all antibodies (see Section 3.3.8 for discussion of Ycf39 in FLAG-7002), whilst FLAG-Cr and FLAG-At eluates contained only YidC. YidC, Ycf39 and HliD were all detected in the membranes of cells producing the eukaryotic enzymes (Figure 3.9), ruling out that the absence of Ycf39 or HliD in the eukaryotic immunoprecipitations is due to any drastic cellular reduction in the amount of the HliD. These results indicate that HliD/Ycf39 are responsible for the pigmentation of the native complex, but are dispensable for ChlG activity, and that the ChlG-YidC/Alb3 interaction is likely to be conserved in Chl producing organisms.

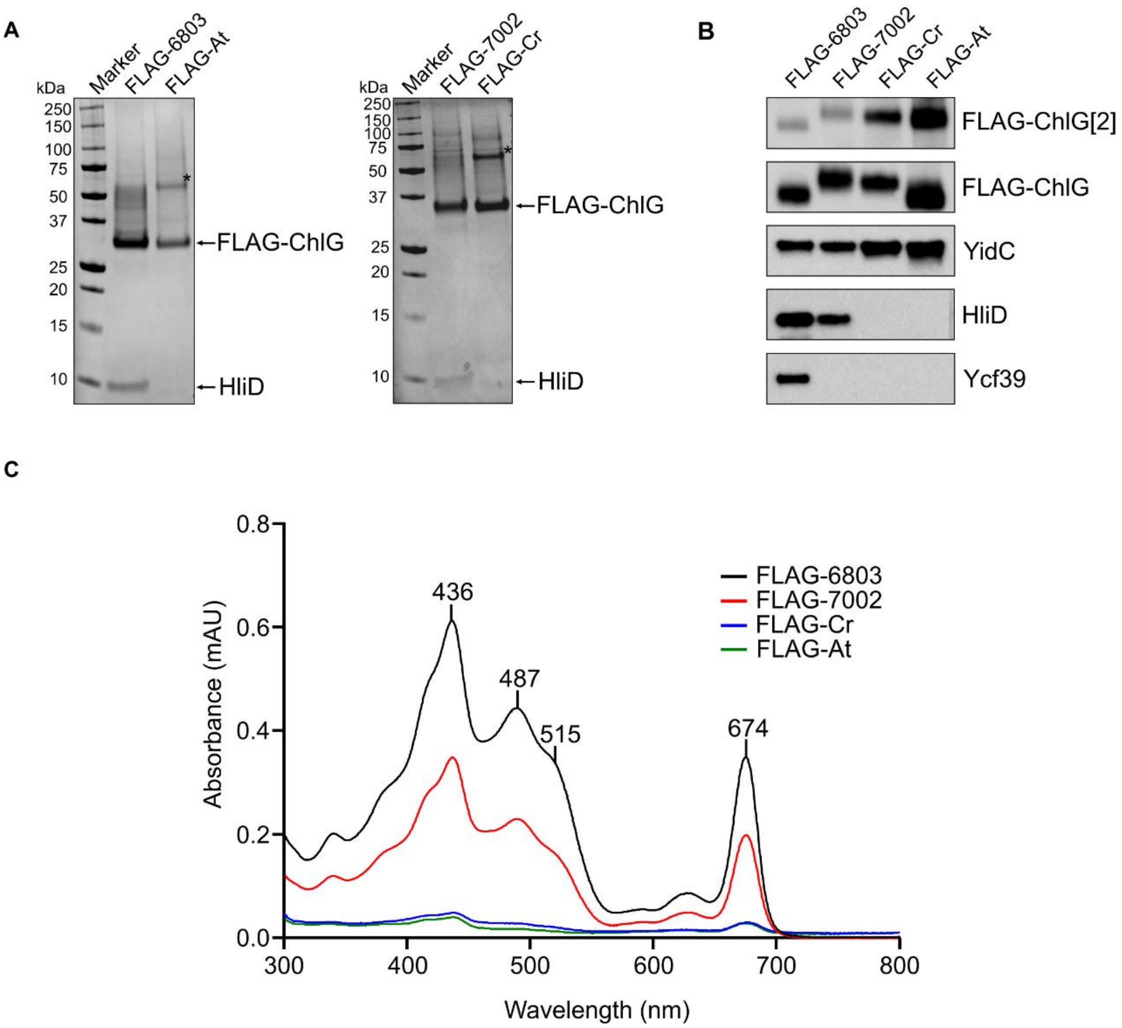


Figure 3.8: Purification of FLAG-ChlG from *Synechocystis* strains and identification of interacting proteins. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG and previously reported interaction partners YidC, Ycf39, and HliD. Data from a single experiment are presented but are representative of three biological replicates, with the exception of the 7002-ChlG interaction with Ycf39 (see Section 3.3.8. further explanation). The asterisk (*) in panel (A) indicates a prominent protein band in the eukaryotic enzyme eluates that cross reacts with both anti-FLAG (B-top panel) and anti-ChlG antibodies, indicating a ChlG dimer. (C) Absorption spectra of FLAG-immunoprecipitation eluates.

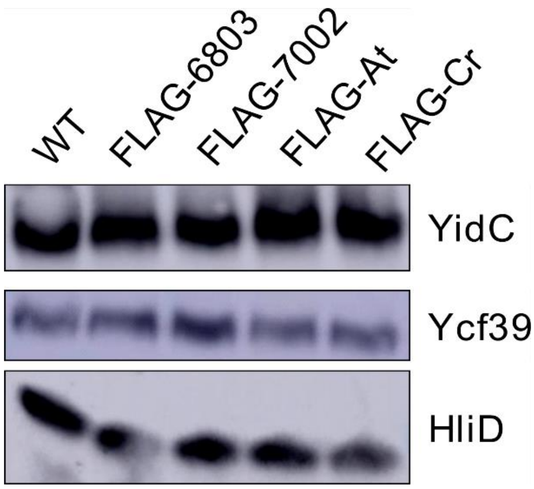


Figure 3.9: Immunodetection of YidC, Ycf39 and HliD in solubilised membranes. Membranes prepared from cells grown under standard illumination were probed with antibodies raised against YidC, Ycf39 and HliD.

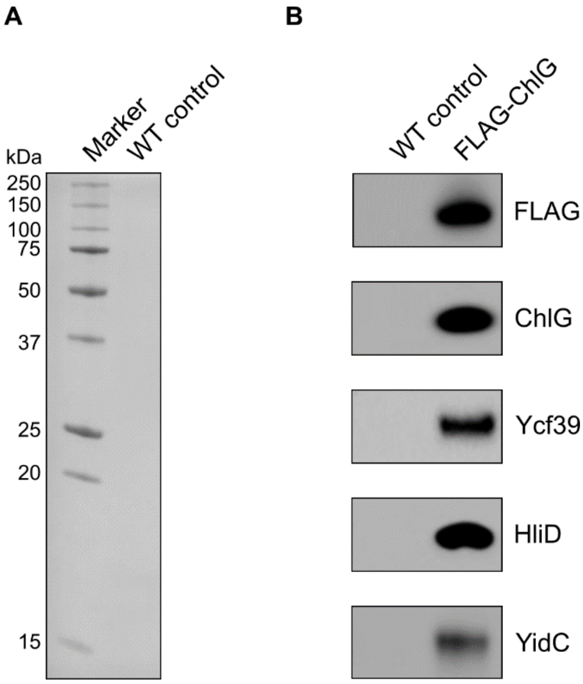


Figure 3.10: SDS-PAGE and immunoblot analysis of control FLAG-immunoprecipitations from solubilised WT membranes. Parallel immunoprecipitations were performed using solubilised membranes prepared from FLAG-6803 or WT *Synechocystis* cells and the eluates were analysed by SDS-PAGE (A) and immunoblotting (B). None of the proteins identified in the FLAG-ChlG immunoprecipitation complex was detected in the WT negative control, confirming that HliD, Ycf39 and YidC do not non-specifically interact with the anti-FLAG resin.

3.3.6 Further analysis of the FLAG-6803 and FLAG-At complexes by gel filtration

The FLAG-ChlG assemblies from the FLAG-6803 and FLAG-At eluates were further analysed by gel filtration chromatography and monitored for protein and carotenoids by absorption at 280 nm and 440 nm respectively. The data shows that the FLAG-6803 complex separated into three major protein-pigment sub-complexes, in agreement with Chidgey *et al* (2014) (Figure 3.11A). Elution fractions that correlated with the absorbance peaks were collected and analysed for FLAG-ChlG, HliD, Ycf39 and YidC by immunoblotting. A small population of FLAG-ChlG appears to be associated with just YidC whilst the majority of the FLAG-ChlG eluted together with YidC, Ycf39 and HliD. A third sub-complex consist of mostly FLAG-ChlG, Ycf39 and HliD. The results confirm that the pigment is predominantly associated with HliD containing sub-complexes in the FLAG-6803 elution as the major pigment peaks correlate with fractions containing HliD. A small pigment peak that eluted early from the column did not appear to coincide with fractions containing HliD, however this is most likely due to poor HliD antibody cross-reactivity or association of these pigments with PSI (Figure 3.11A). Further immunoblot analysis of these fractions using antibodies raised against a PSI subunit should be performed to confirm this.

In contrast, the FLAG-At complex eluted as a single peak and contained no pigments (Figure 3.11B), in agreement with the results obtained from the spectral analysis of the FLAG-AtChlG immunoprecipitation eluate. Immunoblots confirmed the absence of HliD and Ycf39 but the presence of YidC in the FLAG-At complex (Figure 3.11B). The immunoblot signals for FLAG-At and YidC correlated well, indicating the formation of a single major AtChlG-YidC complex, in comparison to the situation with FLAG-6803 ChlG.

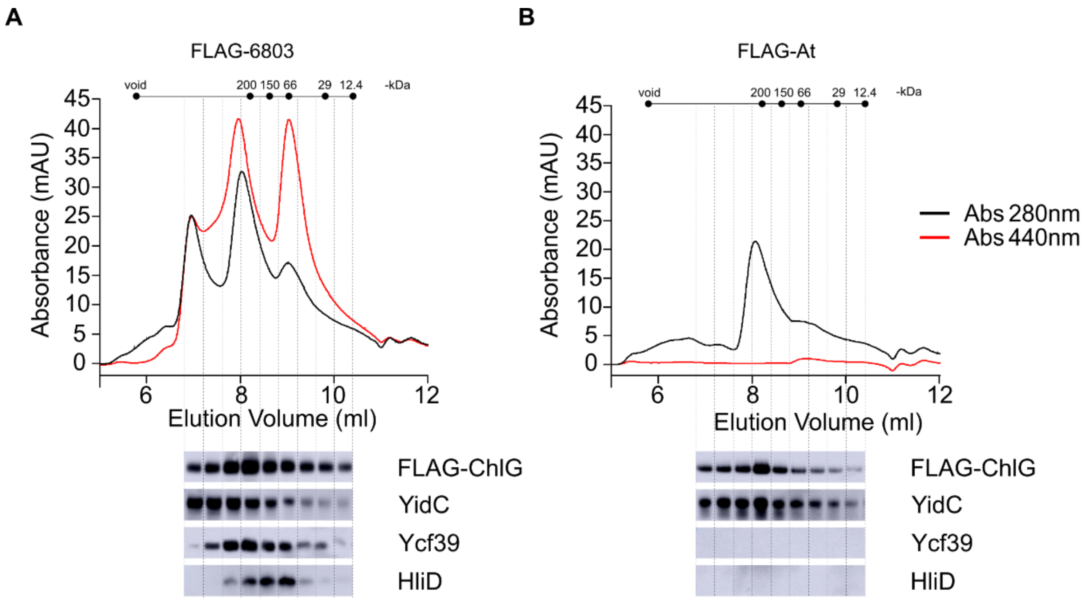


Figure 3.11: Gel filtration of the FLAG-6803 and FLAG-At immunoprecipitation complexes. HPLC gel filtration chromatography separation of purified cyanobacterial and plant FLAG-ChlG eluates. Elution of pigment and protein were monitored at 440 (red line) and 280 (black line) nm respectively. Immunoblot analysis of the HPLC elution fractions are shown below the traces.

3.3.7 Pigment analysis of the FLAG-7002 complex

The pigments associated with the FLAG-6803 and FLAG-7002 complexes were extracted in excess methanol and analysed by reverse-phase HPLC. The complexes were found to contain the carotenoids myxoxanthophyll, zeaxanthin and β-carotene in addition to Chl *a* as reported previously (Niedzwiedzki *et al*., 2016) (Figure 3.12). The profiles were similar between the FLAG-6803 (Figure 3.12A) and FLAG-7002 (Figure 3.12B) ChlG complexes indicating that carotenoid binding to the FLAG-7002 complex is not impeded by the replacement of the *Synechcocystis* ChlG with the Syn 7002 homologue. This is despite the fact that Syn 7002 and *Synechocystis* contain different variants of myxoxanthophyll with the former producing myxol-2' fucoside and the latter myxol-2' dimethylfucoside (Graham and Bryant, 2009).

3.3.8 The interaction between ChlG and Ycf39 is abolished by high-light

In order to verify the results presented in the previous sections, all immunoprecipitation experiments were repeated a minimum of three times with independently grown cultures. In all cases the results were highly consistent (data not shown) with the exception of the presence of Ycf39 in the FLAG-7002 eluates. All three eluates contained pigment and were spectrally very similar (Figure 3.13A), but only two of the replicates contained Ycf39 (Figure 3.13 insert panel). Chidgey *et al*. (2014) showed that Ycf39 dissociates from the FLAG-6803 complex when the cells are exposed to high-light stress (800 µmol photon m^{-2} s^{-1}), although the interactions with YidC and HliD are maintained. The observation that Ycf39 is also capable of dissociating from the FLAG-7002 complex lends further support to the existing evidence showing that the interaction between ChlG and Ycf39 is abolished at high-light intensity. Furthermore, the complexes retrieved from the individual immunoprecipitation experiments were spectrally consistent, independent of Ycf39 binding. This is in agreement with earlier results demonstrating that Ycf39 has no pigment binding properties (Knoppová *et al*., 2014; Llansola-Portoles *et al*., 2017; Shukla *et al*., 2018b; Staleva *et al*., 2015).

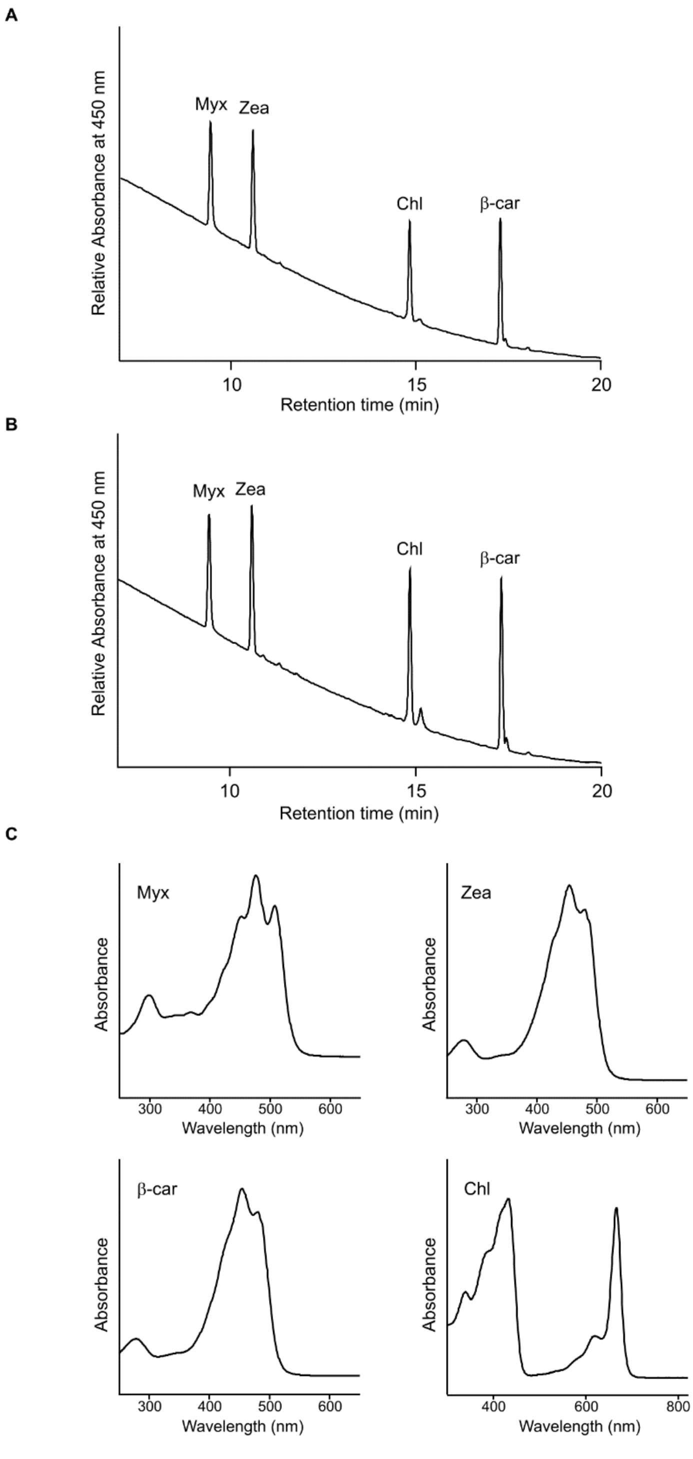


Figure 3.12: Analysis of the pigment content of immunoprecipitation eluates for the FLAG-6803 (A) and FLAG-7002 (B) strains. Pigments were extracted in methanol and separated by reverse phase HPLC. The profiles were similar for both the cyanobacterial complexes, with myxoxanthophyll (Myx), zeaxanthin (Zea), β-carotene (β-car) and chlorophyll (Chl) all present. Pigments were identified by their retention time and absorbance spectra (C).

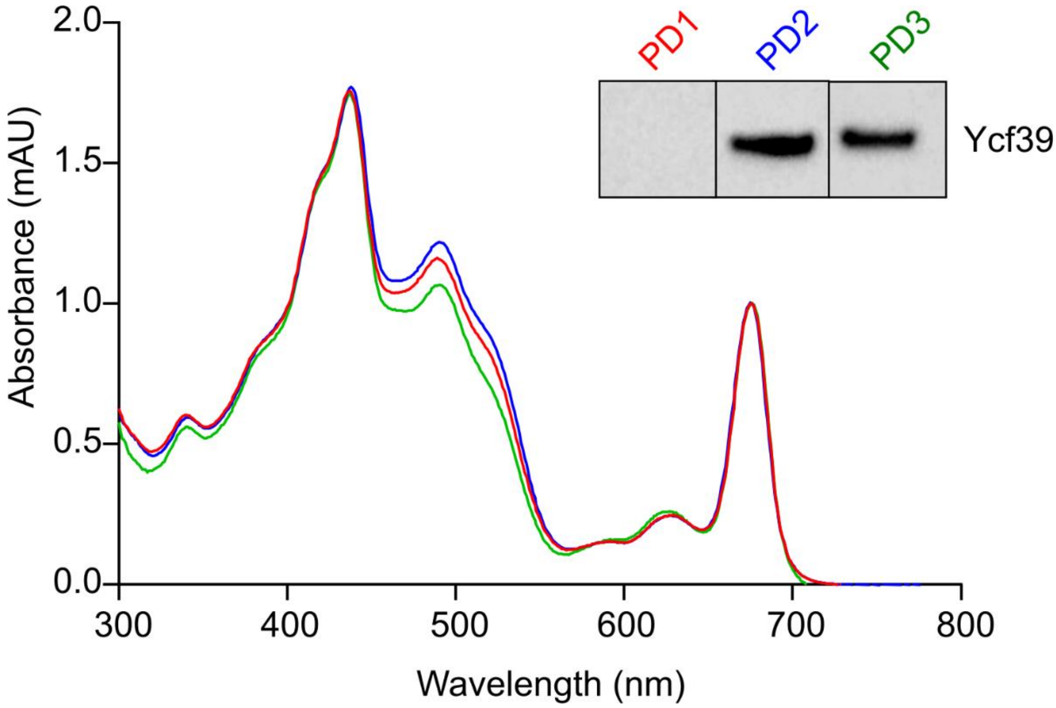


Figure 3.13: Absorbance spectra of three independent FLAG-7002 ChlG immunoprecipitation eluates. The inset panel shows immunodetection of Ycf39 in the same eluates.

3.3.9 Bacteriochlorophyll synthase interacts with *Synechocystis* YidC but not HliD or Ycf39

Bacteriochlorophyll synthase (BchG) is the purple bacterial version of the ChlG found in oxygenic organisms, and it catalyses the addition of a phytyl tail to bacteriochlorophyllide (BChlide) to produce bacteriochlorophyll (BChl). Although *Synechocystis* and *A. thaliana* ChlG enzymes have a sequence identity of 66%, the latter was able to functionally complement *Synechocystis* ChlG despite being unable to form the native complex with HliD and Ycf39 when heterologously produced in this organism. The *Synechocystis* and *A. thaliana* ChlG enzymes only have 35% sequence identity with BchG. To test whether BchG is able to functionally complement *Synechocystis* ChlG, the *bchG* gene from *Rhodobacter sphaeroides,* engineered to encode an N-terminally FLAG-tagged protein, was cloned into the *Synechocystis* genome by homologous recombination at the *psbAII* locus, confirmed by PCR (Figure

3.14A) and sequencing. The native *Synechocystis chlG* gene could not be deleted from the genome of the resulting BchG$^+$ strain, indicating that BchG was unable to functionally complement ChlG (Figure 3.14B), as expected from the results with purified enzymes (Oster *et al*., 1997). Immunoprecipitation of the FLAG-BchG protein followed by analysis of the eluate by SDS-PAGE (Figure 3.14C) and immunoblots using anti-FLAG antibodies (Figure 3.14D) confirmed production of the recombinant protein in the cyanobacterial host. Immunoblots using antibodies raised against ChlG, Ycf39, HliD and YidC showed that FLAG-BchG interacted only with YidC, as was the case with the plant and algal ChlG homologs (Figure 3.14E).

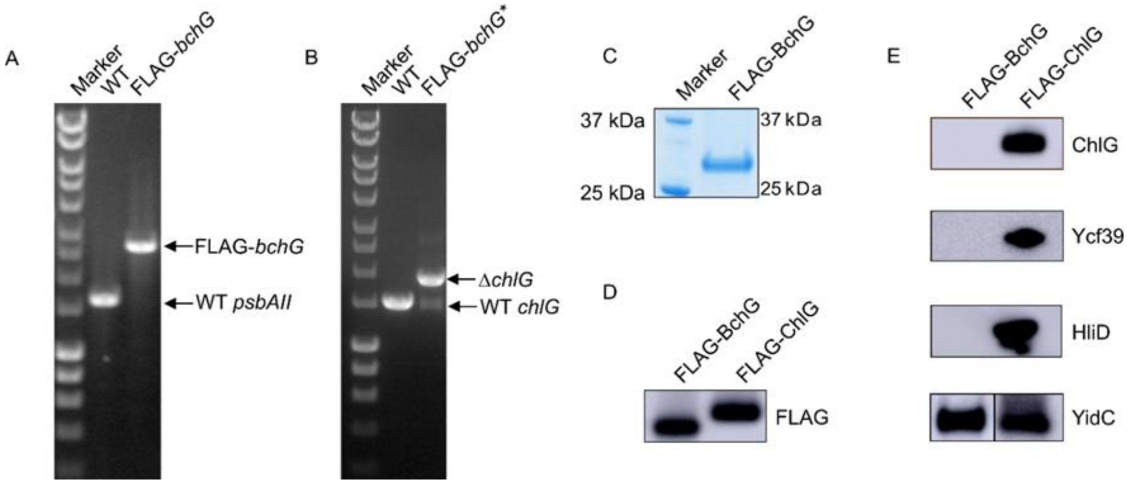


Figure 3.14: Production of the *Rhodobacter sphaeroides* 2.4.1 bacteriochlorophyll synthase (BchG) in *Synechocystis* does not allow full deletion of the native *chlG* gene. (A) The *Rhodobacter sphaeroides* 2.4.1 *bchG* gene alsoencoding an N-terminal 3xFLAG tag was inserted at the *psbAII* locus. (B) Subsequent attempts to delete the native *chlG* gene resulted in a non-segregated merodiploid strain containing chromosomes with both WT and mutated copies of *chlG*. (C-D) FLAG-tagged BchG was produced, as confirmed by SDS-PAGE (C) and anti-FLAG immunblots (D) of FLAG-immunoprecipitation eluates. (E) Of the three major interaction partners that co-elute with *Synechocystis* FLAG-ChlG, only YidC was detectable in the FLAG-BchG eluates.

3.4 Discussion

3.4.1 ChlG from algae and plants are able to complement the function of the native protein when heterologously produced in *Synechocystis*

The processes of Chl biosynthesis and photosystem assembly in phototrophic organisms appear to be closely coordinated to ensure efficient channelling of newly produced Chl pigments to *de novo* photosystem polypeptides, where they can be co-translationally combined, integrated into the thylakoid membrane, and assembled into functioning photosystems (Chidgey et al., 2014). The final enzyme in the Chl biosynthesis pathway, ChlG, catalyses the esterification of a hydrophobic alcohol moiety to Chlide, producing mature Chl. The results presented here demonstrate that ChlG enzymes from the alga *C. reinhardtii*, the plant *A. thaliana* and the cyanobacterium *Synechococcus* 7002 can complement the function of the native ChlG enzyme when produced in the model cyanobacterium *Synechocystis*.

The whole cell spectra and Chl and Chl precursor content of the engineered strains were comparable to the FLAG-6803 and WT strains, indicating that the esterification of Chlide is unaffected by substitution of the native ChlG protein for algal and plant homologs, or the change in genomic location of the *chlG* gene to the *psbAII* locus. Expression of the native gene from the *psbAII* promoter results in levels of protein similar to those in the WT organism (Chidgey et al., 2014). In tobacco plants, it has been reported that ChlG imposes a degree of feedback control of the Chl biosynthesis pathway by influencing levels of ALA synthesis and expression of the branchpoint enzyme magnesium chelatase (Shalygo *et al*., 2009). Additionally, the synthesis of photosystem polypeptides P700, CP43, CP47 and D1 is only possible when *de novo* Chl synthesis (requiring ChlG) is unimpeded (Eichacker *et al*., 1990, 1996; Kim *et al*., 1994a). As no growth phenotype was apparent in the mutant ChlG stains under normal conditions, it can be assumed that the Chl biosynthesis pathway and assembly of the photosystems are unaffected in these mutants. This indicates that the foreign ChlG proteins are able to fulfil the essential biochemical functions of the WT protein, at least under standard laboratory conditions of non-stressed growth.

3.4.2 The interaction between plant and algae ChlG with YidC/Alb3 is maintained

The Chl-binding proteins that constitute photosystems I and II are synthesised by thylakoid associated ribosomes and co-translationally inserted into the membrane (Kim *et al*., 1994b; Zhang *et al*., 1999). During this process photosystem assembly proteins must insert the Chl produced by ChlG into to the photosystem apoproteins. Chidgey *et al*. (2014) previously identified a ChlG complex in *Synechocystis* consisting of ChlG, the high-light inducible protein (HliD), the photosystem II assembly factor Ycf39 and the membrane insertase YidC. To test whether this complex is still formed in the strains producing foreign synthases, the heterologous enzymes were N-terminally FLAG tagged and retrieved from solubilized membranes by FLAG immunoprecipitation. The elution fractions were analysed by SDS-PAGE and immunoblotting.

YidC was detected in the immunoprecipitation eluates for all enzymes, including the non-complementing BchG. In the phototrophic bacterium *Rba. sphaeroides* the light harvesting 1 photosystem assembly factor LhaA was found to co-migrate in CN-PAGE with the integral membrane protease FtsH, BchG and YidC (Mothersole *et al*., 2016), so (bacterio)chlorophyll synthase-YidC associations might be widespread in phototrophs. YidC/Alb3 is a member of the evolutionally conserved protein family of membrane insertases (Samuelson *et al.*, 2000; Beck *et al.*, 2001) and is essential for thylakoid membrane biogenesis in cyanobacteria, algae and plants (Spence *et al.*, 2004; Göhre *et al.*, 2006). In *Chlamydomonas*, Alb3 is critical to the assembly of photosystems, and in plants it promotes the assembly of light-harvesting complexes (Moore *et al.*, 2000; Göhre *et al.*, 2006). In cyanobacteria, YidC is believed to assist the partitioning of the transmembrane segments of polypeptides into the lipid bilayer in concert with the SecYEG translocon (Beck *et al.*, 2001). The discovery of an association between YidC and ChlG led to the hypothesis that YidC fixes Chl-binding proteins into a configuration that allows for the insertion of newly synthesised Chl molecules from the neighbouring ChlG (Chidgey et al., 2014; Sobotka, 2014). Unlike HliD, which is visible on stained gels, YidC is not observable by Coomassie Blue staining, indicating that it is present in the complex at significantly less than a 1:1 ratio with ChlG/HliD. Nonetheless, the observed interaction between algal/plant ChlG proteins and cyanobacterial YidC provides evidence that these proteins may form similar interactions with Alb3 in their native organisms, implying that coordinated delivery of Chl to nascent light harvesting polypeptides via ChlG-YidC/Alb3 interactions is conserved among photosynthetic organisms. *A. thaliana* contains a paralog of Alb3 called Alb4, which is required for chloroplast biogenesis (Gerdes *et al*., 2006) and thylakoid protein targeting (Bédard *et al*., 2017); it is possible the plant ChlG may also interact with this protein although this needs to be tested. Co-immunoprecipitations with solubilised plant thylakoids using antibodies raised to the *Arabidopsis* or Spinach ChlG will allow the *in vivo* partner proteins of the plant enzymes to be confirmed.

3.4.3 Plant and algal ChlG enzymes do not bind HliD or Ycf39

HliD associates with the FLAG-6803 and FLAG-7002 ChlG enzymes but was absent from the plant and algal immunoprecipitation complexes (Figure 3.3). HliD is a member of a family of small one-helix pigment binding proteins called high-light inducible proteins (Hlips) that also includes HliA, HliB and HliC (Funk and Vermass, 1999). Hlips are rapidly upregulated in response to high-light and associate with Chl-binding proteins such as members of the Chl biosynthesis pathway and PSII polypeptides (He *et al*., 2001; Komenda and Sobotka, 2016). A *Synechocystis* mutant lacking all four Hlips was found to be highly sensitive to increased irradiance, leading to the hypothesis that Hlips play a role in photoprotection of Chl-binding proteins (Sinha *et al*., 2012; Xu *et al*., 2004). HliD is known to bind both β-carotene and Chl *a* allowing the dissipation of absorbed light energy as heat by Chl to β-carotene energy transfer, lending further support to this theory (Staleva *et al*., 2015).

It is possible that in algae and plants similar interactions occur with other small Hlip-like proteins (Beck *et al*., 2017). For example, in plants light-harvesting-like (LIL)3 binds pigments and is involved in the latter stages of Chl biosynthesis, interacting with ChlP and protochlorophyllide oxidoreductase (Hey *et al*., 2017; Mork-Jansson *et al*., 2015;

Tanaka *et al*., 2010); Hey *et al*. did not find a LIL3-ChlG interaction in *Arabidopsis* but suggested that another protein may adopt this function. Similarly, two one-helix proteins (OHP1 and OHP2), were found to form dimers in *A. thaliana*, akin to the HliD-HliC heterodimer of *Synechocystis*. This complex bound to and stabilised HCF244, the *Arabidopsis* homologue of Ycf39 in *Synechocystis* (Hey and Grimm, 2018). Immunoprecipitation experiments showed that the OHP1-OHP2-HCF244 complex did not co-elute with ChlG, however, the authors did not dismiss the possibility of an interaction with ChlG *in vivo*.

The *Synechocystis* strains generated in this study which contain *Arabidopsis* and *Chlamydomonas* ChlG display no obvious phenotype under high-light conditions, despite the lack of a ChlG-HliD interaction. This is consistent with earlier results demonstrating that deletion of HliD from *Synechocystis* produces no phenotype when cells are grown under high-light (He *et al*., 2001). This indicates that the role of HliD within the ChlG complex is not essential to cyanobacteria under high-light conditions *in vivo*, despite the fact that HliD has been shown to quench light energy absorbed by the ChlG-HliD complex *in vitro*, preventing photodamage(Niedzwiedzki *et al*., 2016). The role of HliD within the ChlG complex *in vivo* therefore requires further investigation.

Like HliD, the peripheral membrane protein Ycf39 was not present in the plant and algal ChlG immunoprecipitations. Ycf39 is a member of a family of atypical short-chain alcohol dehydrogenases that feature an N-terminal NAD(P)H binding motif which lacks a tyrosine residue critical to enzyme function (Kallberg *et al*., 2002). Members of this family have diverse functions in different phototrophs. The *A. thaliana* homologue, HCF244, has been shown to be important for translational initiation of *psbA* mRNA encoding the PSII core subunit D1, whereas in *Synechocystis* it has been suggested that Ycf39 acts as a quinone chaperone for PSII (Ermakova-Gerdes and Vermaas, 1999; Link *et al*., 2012). In addition, Ycf39 forms a complex with HliD in which the HliD component binds β-carotene and Chl and is capable of quenching Chll fluorescence(Knoppová *et al*., 2014; Staleva *et al*., 2015). This Ycf39-HliD complex is important for the early stages of PSII assembly, binding to the DE loop of the PSII precursor complex pD1 (Knoppová *et al*., 2014). The discovery of the ChlG-Ycf39-HliD-YidC complex led to the hypothesis that ChlG, associated with HliD and Ycf39, binds pD1 as it is being co-translationally inserted into the membrane by YidC, during which time Chl provided by ChlG can be bound to the polypeptide (Chidgey et al., 2014; Knoppová et al., 2014). The data presented here indicates that a Ycf39-ChlG interaction is not necessary for photosystem assembly in *Synechocystis* under normal growth conditions as the produciton of foreign algal and plant ChlG abolishes the association between ChlG and Ycf39. It is possible that the role of Ycf39 in Chl insertion into pD1 is redundant as a Ycf39 knockout mutant is viable, albeit the strain is more sensitive to photoinhibition (Knoppová *et al*., 2014).

3.4.4 Syn 7002 ChlG binds HliD and interacts transiently with Ycf39

Immunoprecipitations of the FLAG-tagged ChlG proteins showed that only the most closely related ChlG from Syn 7002 eluted with YidC, HliD and Ycf39. Syn 7002 has close homologs of the *Synechocystis* Ycf39 (SYNPCC7002_A0216) and HliD (SYNPCC7002_A0858) and it is likely that the same ChlG-Hlip-Ycf39-YidC complex is conserved in this and other related cyanobacteria.

Ycf39 was lost from the FLAG-7002 ChlG complex in one of the three immunoprecipitations performed using this strain (Figure 3.12). The loss of Ycf39 from the FLAG-6803 complex has been previously reported (Chidgey et al., 2014), who showed that, under high-light stress, Ycf39 dissociated from the FLAG-6803 complex. The HliD and YidC components of the complex remained bound and were therefore not dependent on the presence of Ycf39. Analysis of the light-shocked FLAG-6803 complex by size exclusion chromatography resulted in a different elution profile, in comparison with complexes purified from cells grown under normal light. The data indicated a rearrangement to form larger MW complexes in response to light stress, while a small amount of the “original” complex was maintained. Although this was not examined using the FLAG-7002 complex in this work, the loss of Ycf39 from the FLAG-7002 complex is akin to the results reported by Chidgey *et al*. (2014).

HliC is involved in the remodelling of the *Synechocystis* ChlG complex in response to high-light stress by facilitating release of Ycf39 and HliD from ChlG (Shukla *et al*., 2018b). The release of the Ycf39-HliD sub-complex from ChlG may enable its binding to PSII repair intermediates for photoprotection, whilst promoting *de novo* synthesis of D1. Under photo-damaging conditions, Chl pigments are released from the photosystems and must be recycled back to the membrane. There is some evidence to suggest that Ycf39 is required for Chl recycling; a Ycf39 knockout mutant lacking PSI, and the PSII subunits CP43 and CP47 rapidly depleted reserves of the Chl precursor Mg-protoporphyrin under high-light conditions and D1 levels decreased (Knoppová *et al*., 2014). It can be speculated that release of Ycf39-HliD from ChlG under high-light deters the channelling of *de novo* Chl pigment to D1 and instead promotes the re-use of Chl released from damaged photosystems. The Ycf39-HliD sub-complex could therefore act as a scavenger of free Chl molecules that would otherwise generate damaging singlet oxygen species (Komenda and Sobotka, 2016).

The pigment content of the FLAG-7002 complex remained consistent in each of the immunoprecipitation experiments, regardless of Ycf39 association. This is in agreement with the literature, as Ycf39 has no pigment binding properties of its own; this role is instead attributable to HliD and HliC, which have been irrefutably shown to bind β-carotene and Chl *a* independent of association with Ycf39 (Knoppová *et al*., 2014; Llansola-Portoles *et al*., 2017; Shukla *et al*., 2018b; Staleva *et al*., 2015). These pigments, along with myxoxanthophyll and zeaxanthin, were present within the FLAG-7002 complex in the same ratios as the FLAG-6803 complex. Myxoxanthophyll and zeaxanthin have been postulated to bind at the interface between HliD and ChlG, mediating their association within the FLAG-6803 complex (Niedzwiedzki *et al*., 2016). The presence of both the carotenoid and HliD components in the FLAG-7002 complex indicates that the carotenoid mediated interactions within the complex are likely conserved in Syn 7002, even though Syn 7002 produces a different type of myxoxanthophyll (discussed in chapter 4).

3.4.5 A *Synechocystis* strain harbouring recombinant ChlG from *A. thaliana* is cold sensitive

The FLAG-At strain exhibited a growth deficiency phenotype in comparison to WT when grown photoautotphically or photoheterotrophically on BG11 agar at a lower temperature of 20 °C, temperatures more likely to be experienced by cells in their natural environment. The biochemical mechanisms behind this observation are yet to be studied. One possibility is that the mutant strain is unable to adapt to a drop in temperature due to the lack of an interaction between FLAG-At and HliD and/or Ycf39. Ycf39 has been shown to be significantly upregulated in response to cold stress (22 °C) in *Synechocystis* suggesting that it perhaps has a role in adaption of the cell to cold stress (Kopf *et al*., 2014; Suzuki *et al*., 2001). Recently, a *Synechocystis* Δ*hliC* mutant was found to be sensitive to cold and high-light stress (Shukla *et al*., 2018b). The study showed that HliC is upregulated under cold/high-light conditions and binds to the PSI associated ChlG complex by forming a dimer with the HliD component of the complex. This causes the ChlG complex to dissociate and release Ycf39. The authors speculate that Ycf39 forms a complex with HliD/HliC heterodimers and together bind to PSII assembly intermediates, protecting them from photodamage. Meanwhile, the ChlG-HliD/HliC complex participates in the re-utilisation of the Chl released during PSII repair, perhaps channelling them to PSI monomers for storage. Applying this scenario to the FLAG-At strain, the lack of an interaction between FLAG-At ChlG and HliD/Ycf39 could prevent the switch from *de novo* Chl synthesis to Chl recycling during times of cell stress, resulting in the growth phenotype observed. However, further work is required to better understand the role of Ycf39 and HliPs within the ChlG complex, in particular during times of cell stress.

3.5 Future work

The results from this study have indicated that the ChlG-YidC interaction, native to cyanobacteria, is likely conserved in plants and algae. It is possible that, although HliD and Ycf39 did not co-purify with the plant and algal ChlG in the heterologous host, other proteins may bind to these ChlG homologs within their native organisms. However, this has yet to be conclusively demonstrated by retrieval of an equivalent complex from a higher photosynthetic organism. ChlG from *A. thaliana* will be purified from the thylakoid fraction by immunoprecipitation using ChlG antibodies and analysed for any interaction partners by SDS-PAGE and mass spectrometry. This could also be achieved by encoding a FLAG-tagged ChlG within the *A. thaliana* genome followed by immunoprecipitation experiments, analgous to the method described in this study with *Synechocystis*.

Although FLAG-7002 produced in *Synechocystis* could form the native *Synechocystis* ChlG complex, suggesting the complex is conserved in cyanobacteria, an equivalent complex in Syn 7002 has never been reported. A similar experiment to FLAG immunoprecipitations has been attempted, whereby a gene encoding His-tagged Syn 7002 was cloned into the pAQ1EX-Pcpc vector (Xu *et al*., 2011) was expressed in Syn 7002. However, no recombinant protein was isolated from these cells by Ni^{2+} IMAC and subsequent analysis showed that the plasmid copy of the gene rapidly acquired frame-shift mutations and deletions, indicating that high-level production of ChlG from the strong *cpcA* promoter was deleterious to the cell. The purification of Syn 7002 ChlP was achieved using the same vector approach, indicating that the problem is specific to ChlG. As such, FLAG-tagging the native *chlG* gene within the Syn 7002 genome, followed by FLAG pulldown experiments, should be attempted next.

The FLAG-At strain exhibited a defective growth phenotype when cultured on plates at 20 °C. The same phenotype has not yet been tested in liquid cultures of this strain. The method described by Shukla, Jackson, *et al.* (2018), in which a *Synechocystis* Δ*hliC* mutant was exposed to high-light (600 µmols photons s^{-1} m^{-2}) and temperatures of 18 °C in liquid culture, could be applied to the FLAG-At strain. The cold shocked cells could then be analysed for Chl precursor concentrations and PSI/PSII levels to try and elucidate the reasons for the arrested growth phenotype.

Results presented in this chapter have been published:

Proctor MS, Chidgey JW, Shukla MK, Jackson PJ, Sobotka R, Hunter CN & Hitchcock A. Plant and algal chlorophyll synthases function in *Synechocystis* and interact with the YidC/Alb3 membrane insertase. *FEBS Lett.*

Results presented in this chapter have been submitted for publication:

Shukla MK, Jackson PJ, Moravcová L, Zdvihalová B, **Proctor MS,** Brindley AA, Dickman MJ, Hunter CN & Sobotka R. Cyanobacterial LHC-like proteins control formation of the chlorophyll-synthase-Ycf39 complex. *Mol. Plant.* Under review

Chapter 4: Xanthophylls mediate the interaction between chlorophyll synthase and high-light inducible protein D in the cyanobacterium *Synechocystis* sp. PCC 6803

4.1 Summary

In *Synechocystis*, the terminal enzyme of the chlorophyll biosynthesis pathway, chlorophyll synthase (ChlG), forms a protein-pigment complex with a high-light inducible protein D (HliD), chlorophyll *a* (Chl) and the carotenoids zeaxanthin, myxoxanthophyll and β-carotene. HliD binds β-carotene and Chl in a 3:1 ratio, whereas, the zeaxanthin and myxoxanthophyll are only present in the ChlG-HliD complex. These two pigments are speculated to be located at the interface between ChlG and HliD and to mediate their interaction. To test this theory, the carotenoid biosynthesis pathway was perturbed at the point of zeaxanthin and myxoxanthophyll production by deletion of the gene *crtR* from a strain of *Synechocystis* containing an N-terminal FLAG-tagged *chlG* gene. FLAG-ChlG was immunoprecipitated from the Δ*crtR* mutant and analysed for HliD binding by immunoblot. The eluate exhibited a significant reduction in the concentration of HliD and contained a new carotenoid, the myxoxanthophyll precursor deoxy-myxoxanthophyll, which accumulated in this strain. Subsequent removal of this pigment by deletion of the gene *cruF* from the Δ*crtR* background abolished the ChlG-HliD interaction completely in the resulting Δ*crtR*/Δ*cruF* strain, confirming the role of zeaxanthin and/or myxoxanthophyll in mediating the formation of the ChlG-HliD complex. Abolishing the synthesis of only myxoxanthophyll, by generation of a Δ*cruF* single mutant, restored the association of HliD and ChlG to levels comparable to WT, indicating that zeaxanthin alone was capable of facilitating this interaction. Analysis by gel filtration showed that these FLAG-ChlG complexes rearranged into larger assemblies as the pigment and HliD content of the eluates decreased.

4.2 Introduction

In the model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*), chlorophyll synthase (ChlG) associates with high-light inducible proteins (Hlips) C and D (HliC and HliD), the PSII assembly factor Ycf39 and the YidC/Alb3 insertase, forming the ChlG-YidC-HliD-HliC-Ycf39 complex (hereafter ChlG complex) (Chidgey et al., 2014). The ChlG complex contains Chl and carotenoids in the approximate molar ratio Chl (6): zeaxanthin (3/2.1): β-carotene (1/1.1): myxoxanthophyll (1/0.6) (Chidgey et al., 2014; Niedzwiedzki et al., 2016). HliD and HliC bind Chl *a* and β-carotene in a 3:1 (Niedzwiedzki *et al*., 2016; Staleva *et al*., 2015) and 2:1 (Shukla *et al*., 2018a) ratio respectively. It has been demonstrated that it is the β-carotene bound to the Hlips that acts as the quencher of excited Chl *a* in the ChlG-HliC/D complex, with myxoxanthophyll and zeaxanthin proposed to bind at the interface between ChlG and HliD/C (Niedzwiedzki *et al*., 2016); specific removal of Ycf39, which itself does not bind pigments, does not alter the pigment composition of the complex (Knoppová *et al*., 2014; Proctor *et al*., 2018; Shukla *et al*., 2018b). The heterologous production and purification of plant, *Arabidopsis* (*A.*) *thaliana*, and algal, *Chlamydomonas* (*C.*) *reinhardti*, ChlG homologs in *Synechocystis* demonstrated that these enzymes are functional, allowing deletion of the otherwise essential native *chlG* gene (Proctor *et al*., 2018) (Chapter 3). However, these eukaryotic enzymes did not co-immunoprecipitate with HliD or carotenoids, strengthening the hypothesis that the interaction between cyanobacterial ChlG and HliD is mediated by zeaxanthin/myxoxanthophyll. Both *A. thaliana* and *C. reinhardtii* synthesise zeaxanthin (Lohr *et al*., 2005; Ruiz-Sola and Rodríguez-Concepción, 2012) but myxoxanthophylls are uniquely found in cyanobacteria (Graham and Bryant, 2009).

In phototrophs, carotenoids are important for light harvesting, photoprotection of Chl pigments, and structural stabilisation of proteins and membranes (Frank and Cogdell, 1996). For example, in cyanobacteria all-*trans* β-carotene is an important component of both photosystem I (Jordan *et al*., 2001) and photosystem II (Umena *et al*., 2011), and the *Synechocystis* PSI trimeric supercomplex has recently been shown to contain 72 carotenoids, with 7 echinenones, 4 zeaxanthins, 2 canthaxanthins, and one 3’-

hydroxyechinenone in addition to 58 β-carotenes (Malavath *et al*., 2018). β-carotene is also present in the cytochrome b_6f complex of oxygenic phototrophs (Kurisu *et al*., 2003). Keto-carotenoids (e.g. echinenone, 3′-hydroxy-echinenone and cantaxanthin) are essential for nonphotochemical dissipation of excess energy from phycobilisomes (PBS) by the water-soluble orange carotenoid protein (OCP) (Kerfeld *et al*., 2017).

Cyanobacterial carotenoids are C_{40} molecules synthesised from 8 isoprene molecules. *Synechocystis* accumulates four major carotenoid species, β-carotene, and the xanthophylls, zeaxanthin, myxoxanthophyll and echinenone (Lagarde and Vermaas, 1999), and can also produce 3′-hydroxychinenone, cantaxanthin and synechoxanthin (χ,χ-caroten-18,18′-dioic acid) (Graham and Bryant, 2008). An overview of the cyanobacterial carotenoid biosynthesis pathway is presented in Figure 4.1. The products of the MEP (2C-methyl-D-erythritol 4-phosphate)/non-mevalonate pathway, the isoprene isomers isopentyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), are first condensed by geranylgeranyl pyrophosphate synthase (GGDS; CrtE) to form geranyl diphosphate (GPP).

Condensation of GGP with two additional IPP units (catalysed by GGDS) forms the major carotenoid precursor geranylgeranyl diphosphate (GGPP), two molecules of which are condensed by phytoene synthase (CrtB) (Martínez-Férez *et al*., 1994) to produce 15-*cis*-phytoene. The sequential activities of phytoene desaturase (CrtP) (Martínez-Férez and Vioque, 1992), ζ-carotene isomerase (Z-ISO) (Chen *et al*., 2010) and ζ-carotene desaturase (CrtQ) (Breitenbach *et al*., 1998) catalyse the introduction of four double bonds and an isomerisation, dehydrogenating 15-*cis*-phytoene to 7,9,7',9'-tetra-*cis*-lycopene, which is subsequently converted to all-*trans* lycopene by prolycopene isomerase (CrtH) (Masamoto *et al*., 2001). Lycopene is cyclised by the lycopene cyclases, CruA and CruP (Maresca *et al*., 2007; Xiong *et al*., 2017), forming γ-carotene and β-carotene; β-carotene can be further converted to zeaxanthin or echinenone by β-carotene hydroxylase (CrtR) (Masamoto *et al*., 1998) or β-carotene ketolase (CrtO) (Fernández-González *et al*., 1997), while β-carotene desaturase/methyltransferase (CruE) catalyses the first dedicated step in synechoxanthin biosynthesis (Graham and Bryant, 2008).

Cyanobacteria also produce monocyclic myxoxanthophylls glycosylated with different sugar moieties at the C-2′-hydroxyl group of the ψ end of myxol (3′,4′-didehydro-1′,2′-dihydro-β,ψ-carotene-3,1′,2′-triol) (Graham and Bryant, 2009). The C-1′-hydroxylase CruF first catalyses the hydroxylation of the tertiary carbon of the ψ end of γ-carotene, forming 1′-hydroxy-γ-carotene that is converted into 3-dehydroxymyxol (plectaniaxanthin) by introduction of 3′,4′ double bond and a 2′-OH group by an unidentified biochemical pathway. CrtR then acts on 3-dehydroxymyxol to produce myxol, which is glycosylated at the 2′ position hydroxyl group by 2'-O-glycosyltransferase (CruG) to form myxoxanthophyll (Graham and Bryant, 2009). The myxoxanthophyll species produced in *Synechocystis* is myxol-2′ dimethylfucoside (Takaichi *et al*., 2001), thus at least one methyltransferase is also required in this organism.

In order to investigate the role of the non-quenching xanthophylls in the cyanobacterial ChlG-HliD-Ycf39-YidC complex, a series of mutant strains with altered carotenoid contents were generated; a β-carotene hydroxylase (Δ*crtR*; sll1468) null mutant that cannot synthesise myxoxanthophyll or zeaxanthin but can still produce β-carotene and echinenone along with a new carotenoid species, 3-dehydroxy-myxoxanthophyll (deoxy-myxoxanthophyll) (Lagarde and Vermaas, 1999). Deletion of sll0814, which encodes a homologue of the *Synechoccocus* sp. PCC 7002 (Syn 7002) C-1′-hydroxylase (CruF) that catalyses the first committed step in myxoxanthophyll biosynthesis (Graham and Bryant, 2009), resulted in a strain that specifically lacks myxoxanthophyll but still makes zeaxanthin, β-carotene and echinenone. Finally, a double Δ*crtR*/Δ*cruF* mutant that only makes β-carotene and echinenone was analysed.

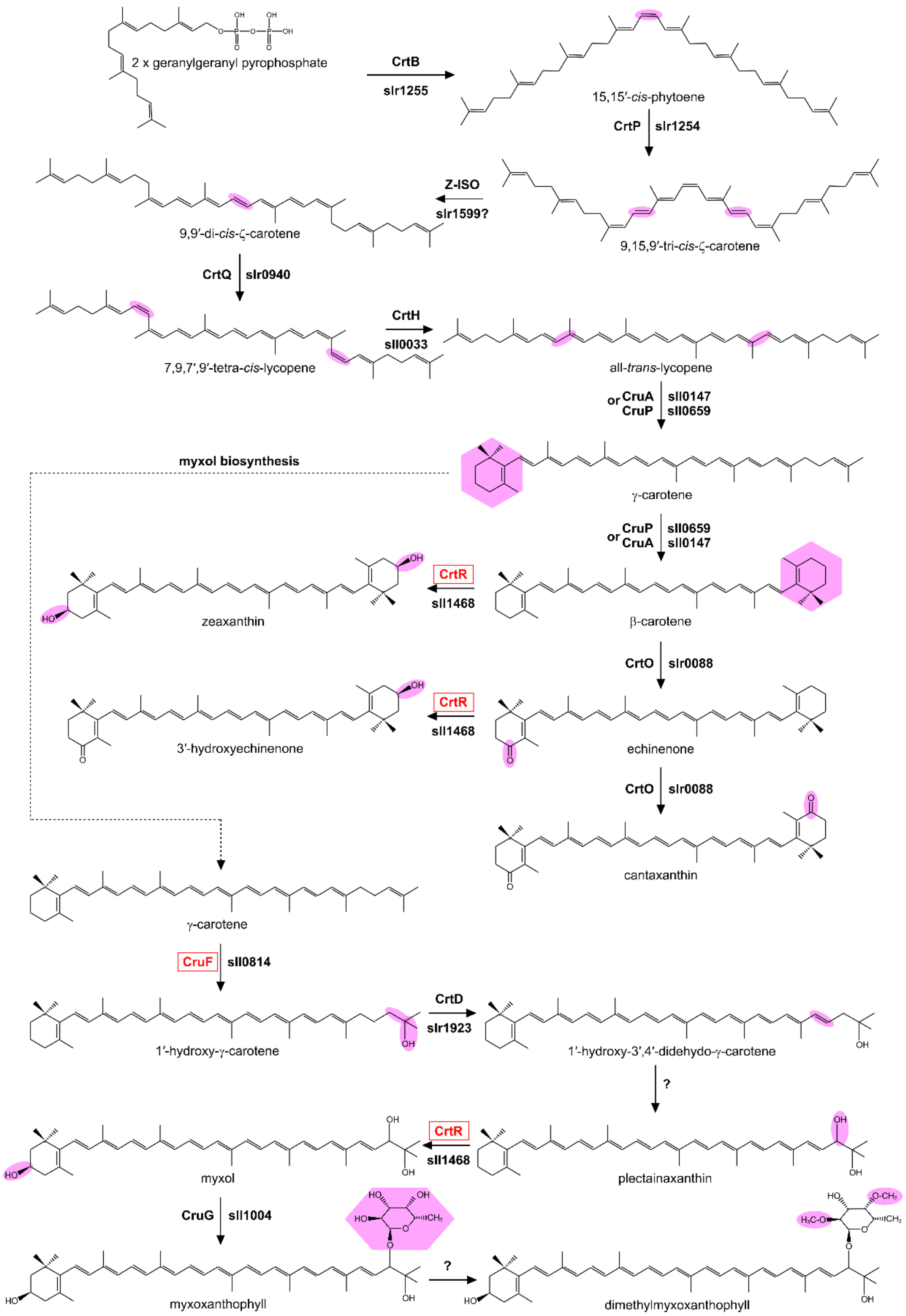


Figure 4.1: Carotenoid biosynthesis in *Synechocystis*. The genes targeted for deletion in this study, *crtR* and *cruF*, are boxed in red. *crtR* catalyses three reactions, producing zeaxanthin from β-carotene, 3'-hydroxyechinenone from echinenone and myxol from plectainaxanthin. *cruF* catalyses the first myxol biosynthesis specific reaction, producing 1'-hydroxy-ϒ-carotene from ϒ-carotene. Purple boxes indicate sites of chemical modification of the carotenoid.

4.3 Results

4.3.1 Deletion of the *crtR* gene from *Synechocystis* abolishes synthesis of myxoxanthophyll and zeaxanthin

In order to investigate the role of xanthophylls as constituents of the ChlG complex, the synthesis of zeaxanthin and myxoxanthophyll was abolished in the FLAG-*chlG* Δ*chlG* (hereafter FLAG-*chlG*) strain (details of this strain are presented in chapter 3). This was achieved by deletion of the *crtR* gene encoding the enzyme β-carotene hydroxylase, which converts β-carotene to zeaxanthin and myxol from 3-dehydroxymyxol during the synthesis of myxoxanthophyll. The central 555 bp of the 939 bp *crtR* gene (sll1468) was replaced with an erythromycin-resistance cassette (*ery*^{R}) in the FLAG-*chlG* background using a linear mutagenesis construct generated by OLE-PCR (Figure 4.2A). Segregation of genome copies was achieved by sequential plating with increasing antibiotic concentration and confirmed by PCR (Figure 4.2B).

Pigments from FLAG-*chlG* Δ*crtR* whole cells (hereafter Δ*crtR*) were extracted and analysed by reverse-phase HPLC (Figure 4.3A) in comparison to the FLAG-*chlG* (Figure 4.3B). Pigments were identified by their absorbance spectra (Figure 4.3C). Comparison of the Δ*crtR* carotenoid profile to that of the FLAG-*chlG* parent strain showed two pigments eluting at 9.5 and 10.5 minutes in the FLAG-*chlG* sample that were not present in Δ*crtR*. These corresponded to myxoxanthophyll and zeaxanthin, respectively, confirming that the synthesis of these two carotenoids was prevented in the Δ*crtR* strain. However, the Δ*crtR* profile showed the accumulation of a new carotenoid species, previously identified as deoxy-myxoxanthophyll (myxoxanthophyll lacking the hydroxyl group on the β-ring) (Graham, 1998; Lagarde and Vermaas, 1999). The production of deoxy-myxoxanthophyll in the Δ*crtR* strain indicates that the 2'-O-glycosyltransferase (CruG) can glycosylate the hydroxyl group at the 2′ position of 3-deoxy-myxol, and CrtR can subsequently act on 3-deoxy-myxoxanthophyll. Furthermore, as the Syn 7002 Δ*cruG* mutant accumulates myxol (Graham and Bryant, 2009), CrtR can also add a hydroxyl group to the non-glycosylated 3-deoxy-myxoxanthophyll analogue 3-deoxy-myxol.

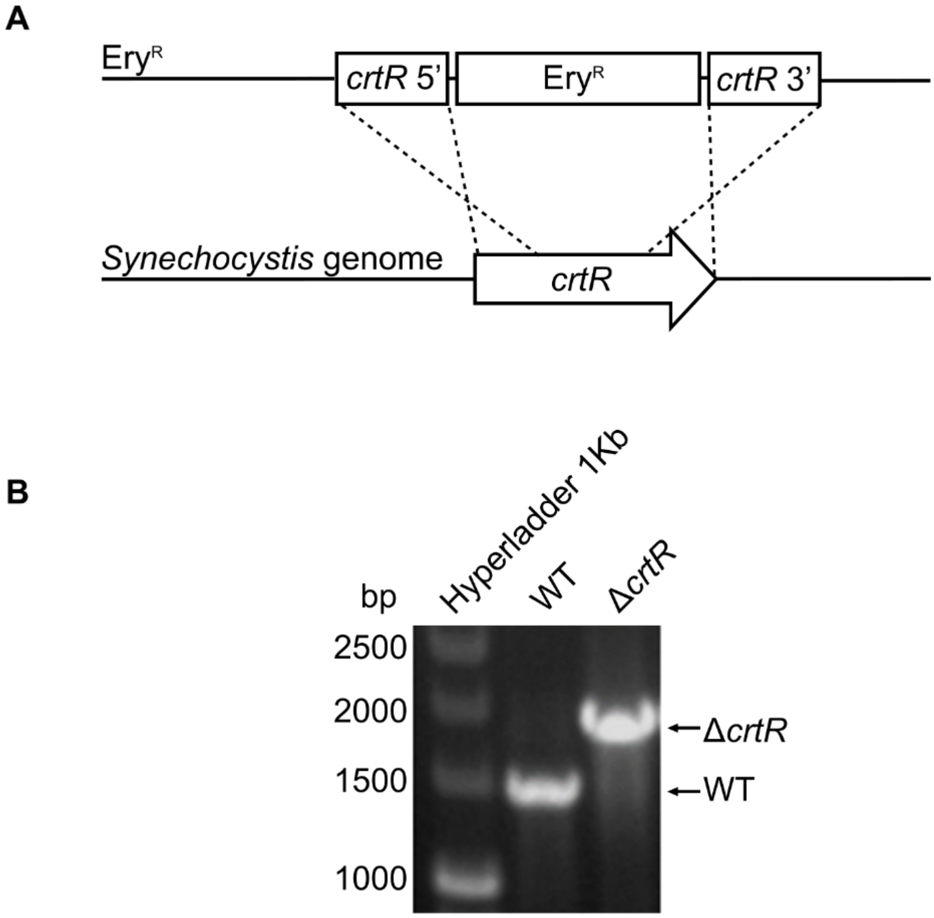


Figure 4.2: Deletion of *crtR* from the FLAG-*chlG* Δ*chlG* strain. (A) The *crtR* gene was partially replaced by an erythromycin resistance cassette using a linear mutagenesis construct containing regions homologous to the *crtR* upstream and downstream regions within the *Synechocystis* genome. (B) Complete segregation at the *crtR* locus was confirmed by PCR with primers flanking the integration site.

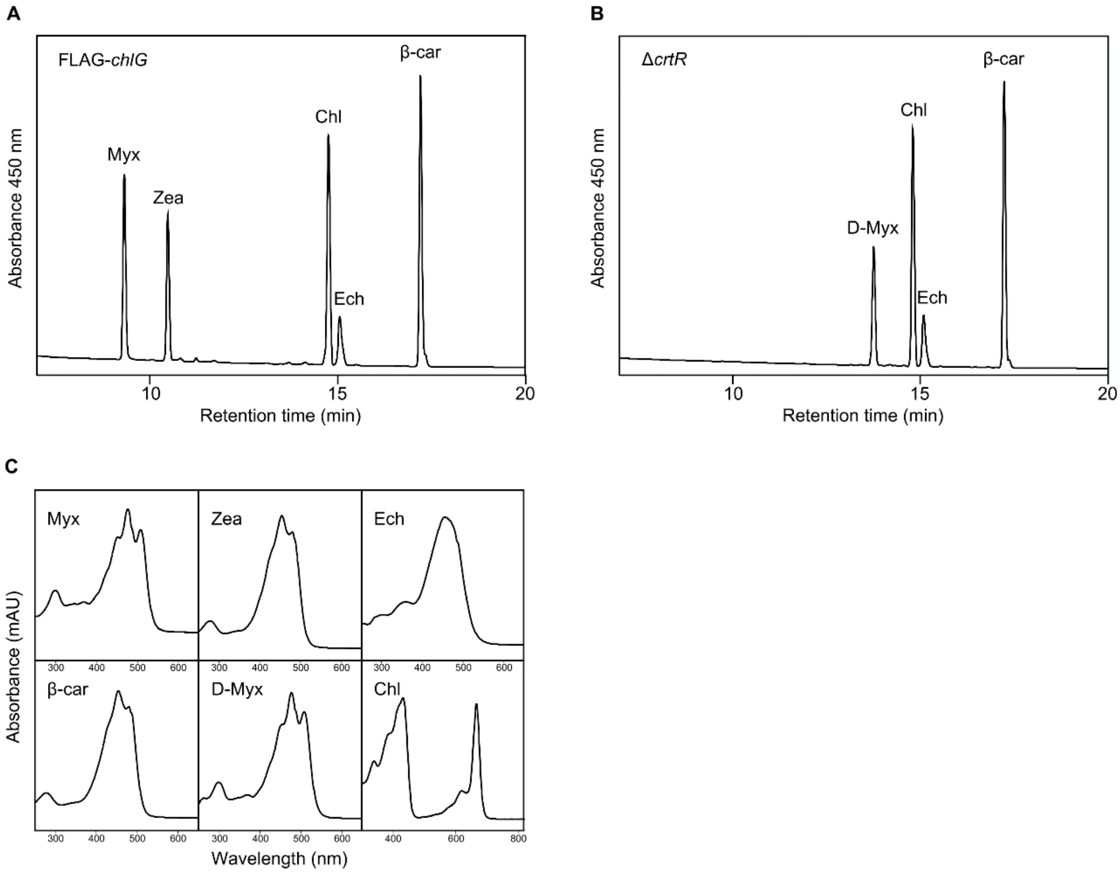


Figure 4.3: Analysis of the pigment content of FLAG-*chlG* Δ*chlG* (A) and FLAG-*chlG* Δ*chlG* Δ*crtR* (B) strains. Pigments were extracted in methanol and separated by reverse-phase HPLC. (A) Myxoxanthophyll (Myx), zeazanthin (Zea), β-carotene (β-car) and chlorophyll (Chl) were all present within the FLAG-*chlG* Δ*chlG* strain. (B) Myx and Zea were absent from the Δ*crtR* strain and deoxy-myxoxanthophyll (D-Myx) accumulated. (C) Pigments were identified by their retention time and absorbance spectra.

4.3.2 Deletion of *crtR* significantly impedes the formation of the ChlG-HliD complex

To analyse the effects of abolishing the synthesis of myxoxanthophyll and zeaxanthin on the ChlG complex, FLAG-ChlG was purified from the Δ*crtR* background by FLAG immunoprecipitation alongside a FLAG-*chlG* control. The eluate from the Δ*crtR* strain was not coloured to the naked eye and the absorbance spectra revealed a clear reduction in pigmentation, especially in the carotenoid region, compared to the visibly orange coloured eluate from the FLAG-*chlG* strain (Figure 4.4D).

The eluates were separated by SDS-PAGE, and similar amounts of FLAG-ChlG were clearly visible at approximately 30 kDa (Figure 4.4A). A band of slightly higher molecular weight, visible in the FLAG-*chlG* eluate but not in the Δ*crtR* sample, is tentatively assigned as Ycf39. Similarly, a band approximately 10 kDa in size that was observed in the FLAG-*chlG* eluate, attributed to HliD, was absent from Δ*crtR*. The identities of these bands were confirmed by immunoblotting using the appropriate antibodies (Figure 4.4B). Ycf39 was absent from the Δ*crtR* mutant and the levels of HliD greatly diminished in comparison to the FLAG-*chlG*. On the other hand, the concentration of YidC within the Δ*crtR* eluate appeared to be greater than in the FLAG-*chlG* (Figure 4.4B) control. The absence of HliD and Ycf39 in the Δ*crtR* eluate could have been due to the loss of these two proteins during the isolation process. However, co-elution of HliD and Ycf39 with the FLAG-*chlG* control indicates that this is not the case and that it is the loss of carotenoids that prevents the co-immunopreciptiation of Ycf39 and HliD with FLAG-ChlG when purified from the Δ*crtR* strain. The concentration of HliD in solubilised thylakoid membranes were comparable in both strains demonstrating that accumulation of HliD is unaffected by the deletion of *crtR* (Figure 4.4C).

Quantitative protein mass-spectrometry (Section 2.9.6.) was used to determine the picomolar amounts of FLAG-ChlG, HliD and Ycf39 in the eluates, allowing the calculation and comparison of the ChlG:HliD:Ycf39 stoichiometries between each strain. The ChlG:HliD:Ycf39 ratio changed from 1:1.83:0.24 in the presence of *crtR* to 1:0.46:0.03 when *crtR* was absent (Table 1), thus there is an approximately a 4-fold decrease in HliD and an 8-fold reduction in Ycf39 per ChlG.

The pigments in the Δ*crtR* and FLAG-*chlG* eluates were extracted in an excess of methanol and analysed by reverse-phase HPLC (Figure 4.5). The FLAG-*chlG* sample showed the expected profile with clear peaks for myxoxanthophyll, zeaxanthin, Chl and β-carotene and a small amount of echinenone, analogous to the pigment profile presented by Chidgey *et al*. (2014) (Figure 4.5A). Conversely, there was less Chl and β-carotene in the Δ*crtR* eluate, along with a trace amount of deoxy-myxoxanthophyll

(Figure 4.5B) and a small amount of echinenone as before. These altered pigment levels accompany the significant reduction in HliD protein within the Δ*crtR* eluate.

Taken together these data indicate that myxoxanthophyll and/or zeaxanthin are important for facilitating the association of ChlG and HliD. The myxoxanthophyll precursor deoxy-myxoxanthophyll may be responsible for residual binding of HliD, although the level of this pigment was very low compared to myxoxanthophyll and zeaxanthin in FLAG-*chlG*. Alternatively, some HliD may bind to ChlG via β-carotene/echinenone or independently of carotenoids.

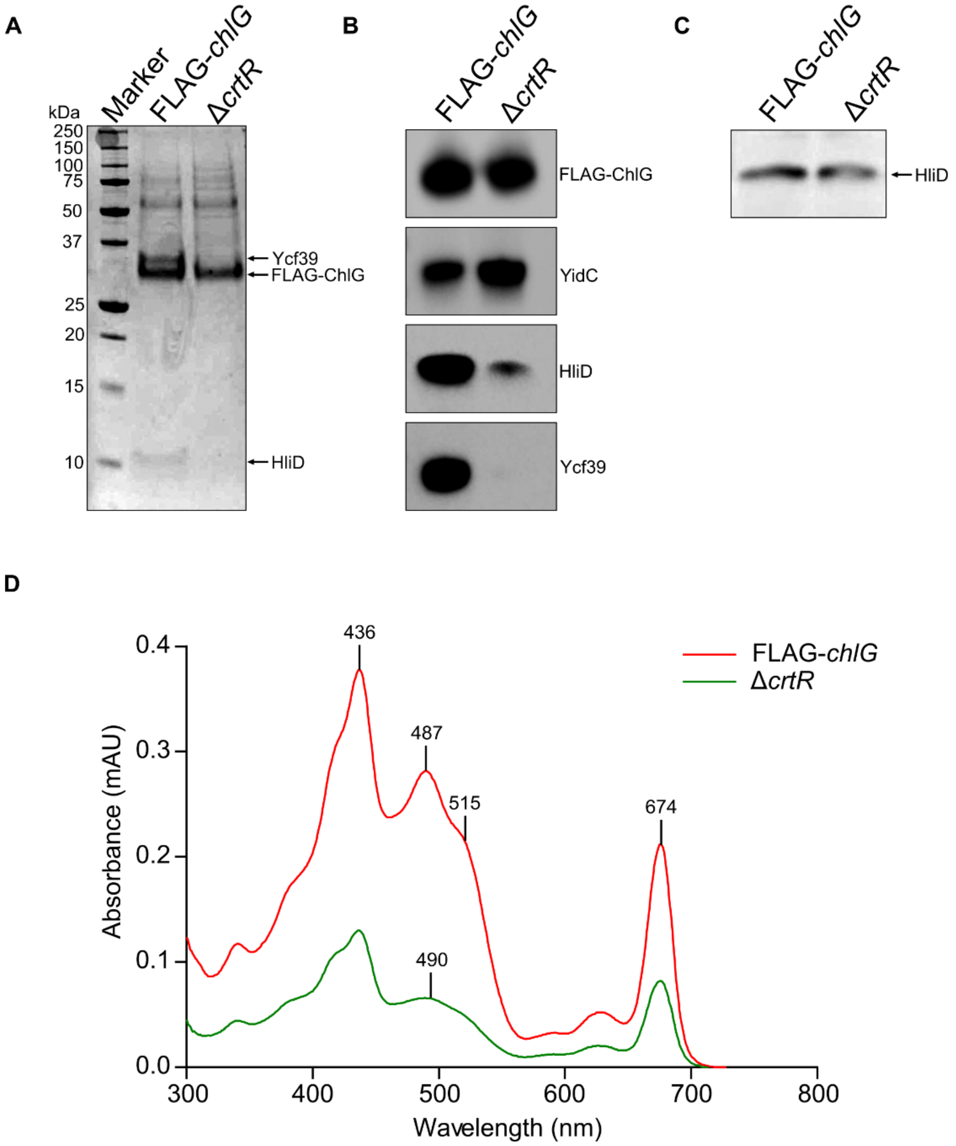


Figure 4.4: Purification of FLAG-ChlG from *Synechocystis* FLAG-*chlG* and Δ*crtR* strains. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG and ChlG interaction partners YidC, Ycf39, and HliD. (C) Immunodetection of HliD in solubilised thylakoid membranes prepared from cells grown under standard illumination. (D) Absorption spectra of FLAG-immunoprecipitation eluates.

Table 1: Quantification of FLAG-ChlG, HliD and Ycf39 proteins in FLAG-ChlG immunoprecipitation eluates. The solubilised membrane fractions from *Synechocystis* cells expressing FLAG-*chlG* in FLAG-*chlG* and Δ*crtR* backgrounds were used in immunoprecipitation experiments and the proteins within the eluates quantified mass spectrometry. Picomolar amounts of FLAG-ChlG, HliD and Ycf39 were calculated by averaging the relative intensities of the ^{14}N and ^{15}N isotopomers of their respective proteotypic peptide ions.

Strain	Protein	Quantity (pmols ± SD)	Stoichiometry (ratio per 1 ChlG ± propagated SD)
FLAG-*chlG*	ChlG	155.18 ± 1.01	1.00
	HliD	283.70 ± 57.53	1.82 ± 0.37
	Ycf39	36.73 ± 12.47	0.24 ± 0.08
Δ*crtR*	ChlG	161.51 ± 1.48	1.00
	HliD	73.29 ± 19.66	0.45 ± 0.12
	Ycf39	4.86 ± 2.10	0.03 ± 0.01

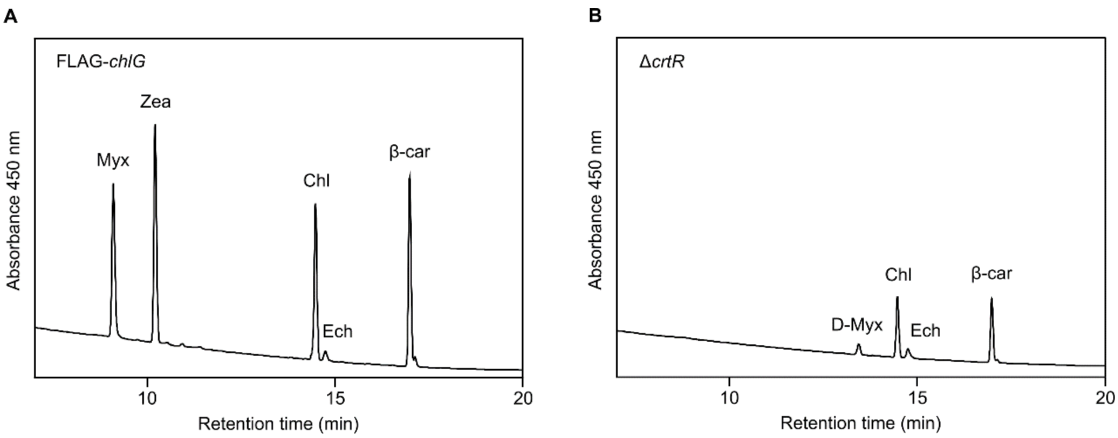


Figure 4.5: Analysis of the pigment content of immunoprecipitation eluates purified from FLAG-*chlG* (A) and Δ*crtR* (B) strains. Pigments were extracted in methanol and separated by reverse-phase HPLC. (A) Myxoxanthophyll (Myx), zexzanthin (Zea), β-carotene (β-car), echinenone (Ech) and chlorophyll (Chl) were all present within the FLAG-*chlG* strain. (B) Myx and Zea were both absent from Δ*crtR* strain and levels of chlorophyll and β-carotene were significantly reduced in comparison to FLAG-*chlG* whereas the concentration of echinenone was consistent. A small amount of deoxy-myxoxanthophyll (D-myx) was present.

4.3.3 Deletion of the *cruF* gene from *Synechocystis* abolishes synthesis of myxoxanthophyll

In order to determine whether myxoxanthophyll, zeaxanthin or both carotenoids are required for mediating the association of ChlG with HliD, it was necessary to selectively abolish the synthesis of just one of these pigments. It is not possible to generate a strain that makes myxoxanthophyll but not zeaxanthin as all the enzymes necessary for zeaxanthin biosynthesis are also required to synthesise myxoxanthophyll (Figure 4.1). However, the reverse situation was achieved by deletion of *cruF*, allowing the effect of the specific loss of myxoxanthophyll on the ChlG complex to be investigated. The 912 bp *cruF* gene was deleted from the FLAG-*chlG* background by replacing the central 568 bp with the *aadA* gene from pCDFDuet-1 (Novagen) (Figure 4.6A). Following segregation of genome copies by sequential plating on media with increasing streptomycin concentration, successful deletion of *cruF* was confirmed by PCR (Figure 4.6B).

The carotenoid content of the FLAG-*chlG* Δ*chlG* Δ*cruF* mutant (hereafter Δ*cruF*) was extracted from whole cells in excess solvent and analysed by reverse-phase HPLC in comparison to the Δ*crtR* and FLAG-*chlG* profiles (Figure 4.7). The carotenoid profiles of the FLAG-*chlG* and Δ*crtR* strains were as previously described. The Δ*cruF* mutant contained carotenoids in comparable ratios to FLAG-*chlG* with the exception of myxoxanthophyll, which was absent from the strain. This confirmed the successful specific elimination of myxoxanthophyll synthesis from the Δ*cruF* mutant.

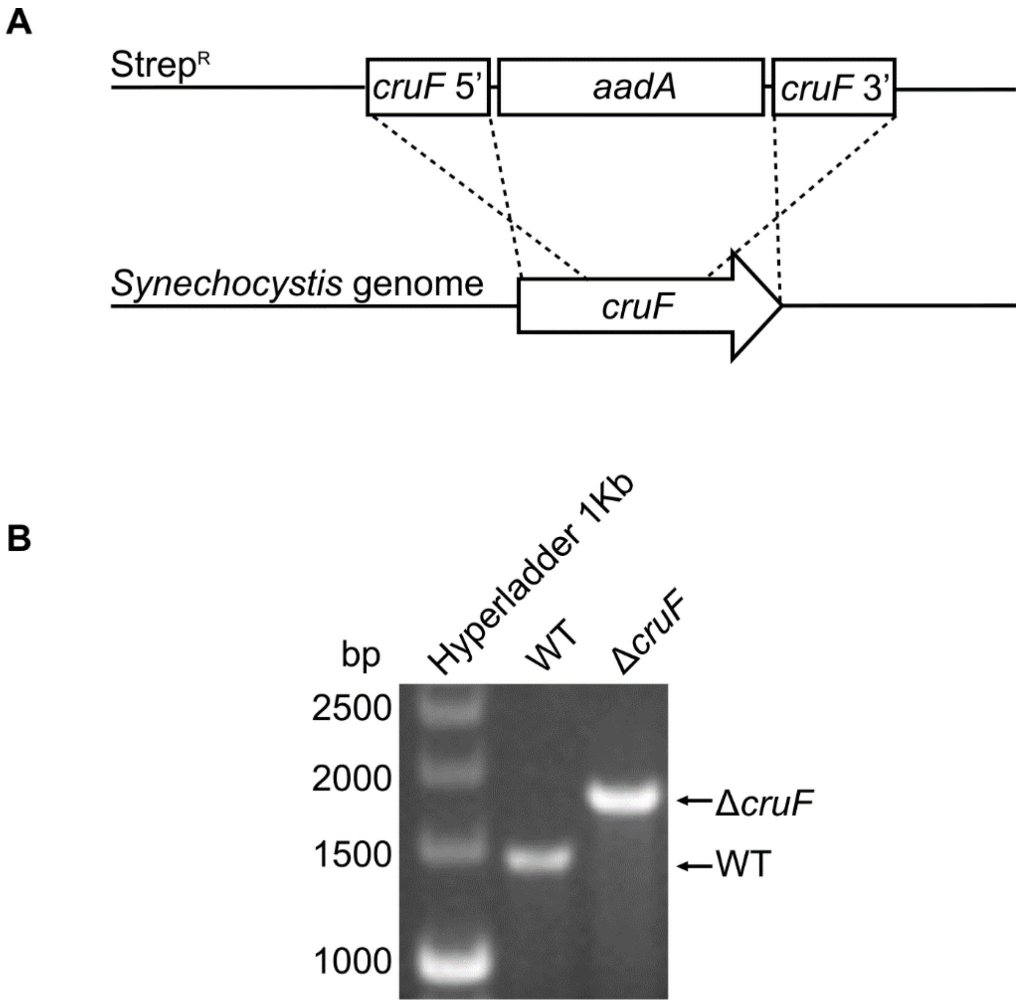


Figure 4.6: Deletion of the *cruF* gene from the FLAG-*chlG* Δ*chlG* strain. (A) The *cruF* gene was deleted from the FLAG-*chlG* Δ*chlG* background by partial replacement with a streptomycin resistance cassette. (B) Complete segregation at the *cruF* locus was confirmed by PCR with primers flanking the integration site.

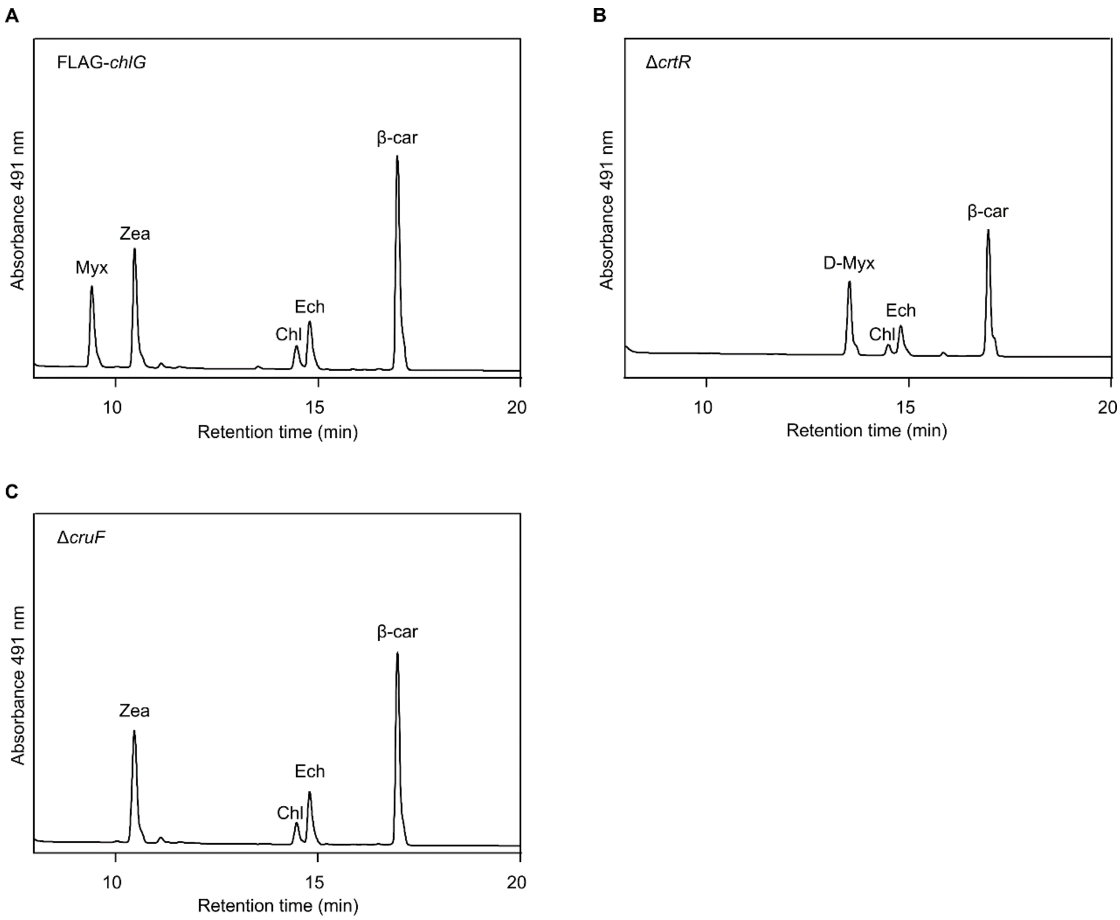


Figure 4.7: Analysis of the pigment content of FLAG-*chlG* Δ*chlG* (A) FLAG-*chlG* Δ*chlG* Δ*crtR* (B) and FLAG-*chlG* Δ*chlG* Δ*cruF* (C) strains. Pigments were extracted in methanol and separated by reverse-phase HPLC. (A-B) The WT and Δ*crtR* pigment profiles were comparable to the profiles in Figure 4.8. (C) The Δ*cruF* elution profile did not contain any Myx but the ratio of the other pigments was comparable to that of the FLAG-*chlG*.

4.3.4 The interaction of ChlG and HliD is unaffected by deletion of *cruF*

To analyse the effects of selectively removing myxoxanthophyll from the ChlG complex, FLAG-ChlG was purified from the Δ*cruF* and Δ*crtR* backgrounds. The immunoprecipitation eluate from the Δ*crtR* strain was not coloured, as observed previously, but the Δ*cruF* eluate was orange to the naked eye. The absorbance spectra of the Δ*cruF* eluate was comparable to that of the FLAG-*chlG* strain (Figure 4.8C).

The eluates from both strains were separated by SDS-PAGE and analysed for HliD, Ycf39 and YidC by immunoblotting, as described above. A clear band corresponding to FLAG-ChlG was visible by SDS-PAGE in both samples, confirmed by immunoblotting using anti-FLAG antibodies, however, the concentration of FLAG-ChlG in the Δ*cruF* was approximately double that of Δ*crtR* (Figure 4.8A and 4.8B). YidC was present in both strains, indicating that the association of ChlG and YidC is not dependent on carotenoids as previously surmised. The HliD content of the Δ*cruF* eluate was significantly higher than in Δ*crtR*, with the HliD levels being comparable to those of the FLAG-*chlG* Δ*chlG*, whereas Ycf39 was detectable within the Δ*cruF* eluate although the concentration appeared lower than in the FLAG-*chlG* sample. However, the interaction of Ycf39 with ChlG has been shown to be sensitive to the conditions in which the cells are grown, e.g. light intensity, and so can vary between cultures (Proctor *et al*., 2018).

The pigment contents of the Δ*cruF* and Δ*crtR* eluates were analysed by reverse-phase HPLC (Figure 4.9). The profiles of the Δ*cruF* and FLAG-*chlG* samples were similar, shown in Figure 4.5A, with the exception of a missing myxoxanthophyll peak (Figure 4.9B). A zeaxanthin peak was prominent, demonstrating the continued association of this pigment with the ChlG complex independent of the presence of myxoxanthophyll. A large β-carotene peak was also observed, indicative of the presence of HliD in agreement with the immunoblot analysis. It can be concluded that zeaxanthin alone is capable of mediating the formation of the ChlG-HliD complex and is capable of remaining associated to the ChlG-HliD complex in the absence of myxoxanthophyll.

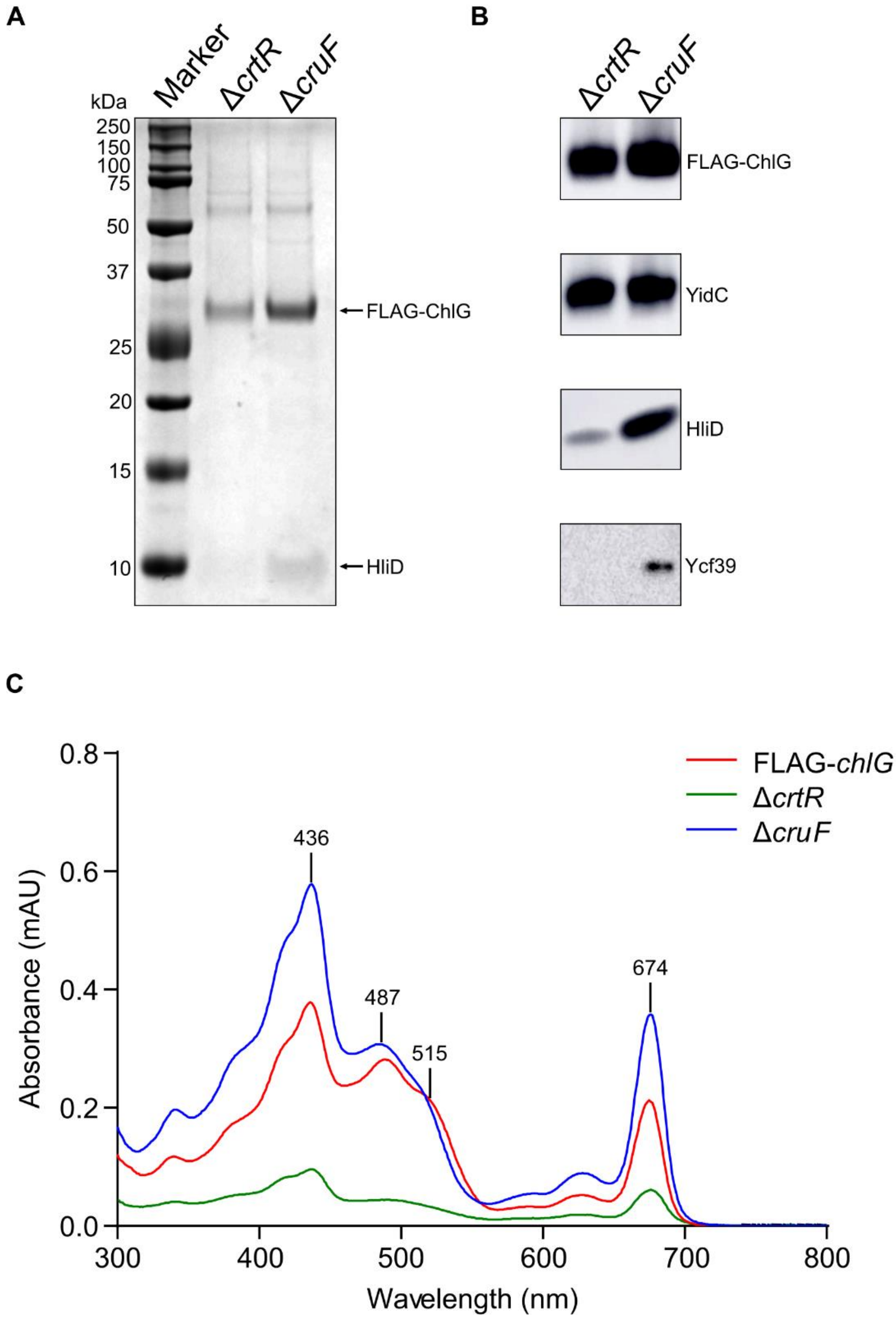


Figure 4.8: Purification of FLAG-ChlG from *Synechocystis* Δ*crtR* and Δ*cruF* strains. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG and ChlG interaction partners YidC, Ycf39, and HliD. (C) Absorption spectra of FLAG-immunoprecipitation eluates.

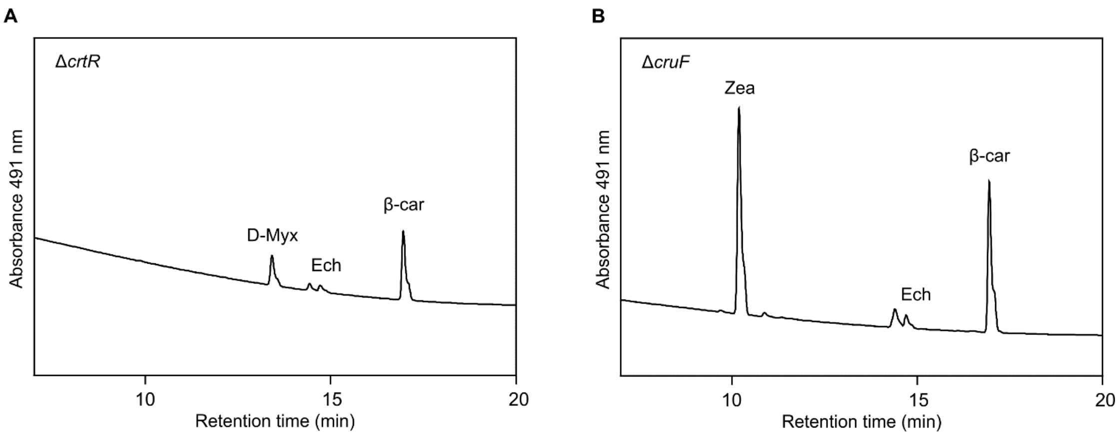


Figure 4.9: Analysis of the pigment content of immunoprecipitation eluates purified from the Δ*crtR* (A) and Δ*cruF* (B) strains. Pigments were extracted in methanol and separated by reverse-phase HPLC. (A) Myx and Zea were absent and D-Myx was present in the Δ*crtR* eluate as previously observed. (B) Myx was absent but Zea present in the Δ*cruF* strain and levels of β-car were comparable that of the FLAG-*chlG* strain.

4.3.5 A *Synechocystis* Δ*crtR*/Δ*cruF* mutant cannot synthesise zeaxanthin, myxoxanthophyll or deoxy-myxoxanothophyll

The results described so far demonstrate that the absence of zeaxanthin and myxoxanthophyll in a Δ*crtR* mutant perturbs the association of ChlG with HliD. However, the immunoprecipitate obtained from the Δ*crtR* strain still contained a small amount of HliD, indicating that the ChlG-HliD complex could still form to a lesser extent in this strain. This was considered to be most likely due to the accumulation of the myxoxanthophyll precursor deoxy-myxoxanthophyll. To confirm that this pigment was responsible for the residual HliD binding, it was necessary to prevent the accumulation of deoxy-myxoxanthophyll in the Δ*crtR* strain, so *cruF* was deleted from this background, generating the strain FLAG-*chlG* Δ*crtR* Δ*cruF* (hereafter Δ*crtR*/Δ*cruF*) (Figure 4.10). This strain, in addition to being unable to synthesise zeaxanthin or myxoxanthophyll, no longer accumulated deoxy-myxoxanthophyll (Figure 4.11).

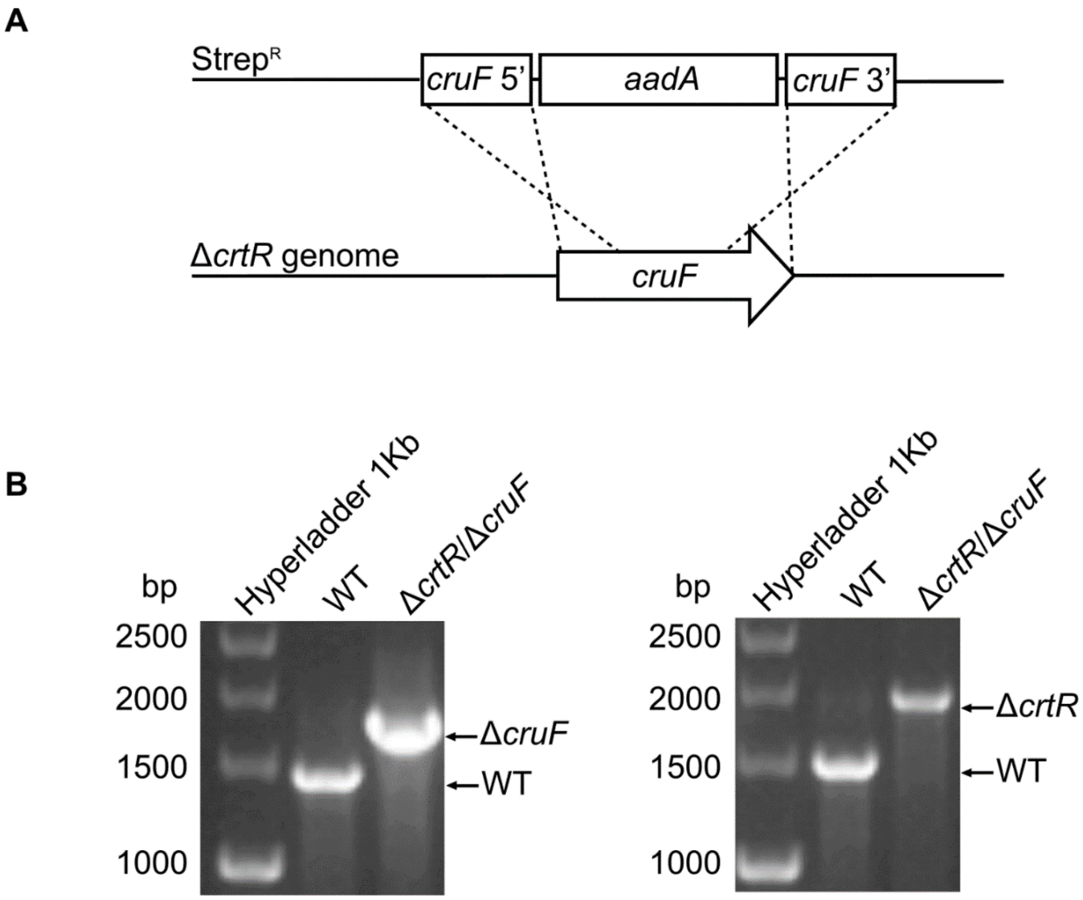


Figure 4.10: Deletion of the *cruF* gene from the FLAG-*chlG* Δ*chlG* Δ*crtR* strain. (A) The *cruF* gene was deleted from the FLAG-*chlG* Δ*chlG* Δ*crtR* background by partial replacement with a streptomycin resistance cassette usng a linear mutagenesis construct. (B) Complete segregation at the *cruF* (left) and *crtR* (right) locus was confirmed by PCR with primers flanking the integration site.

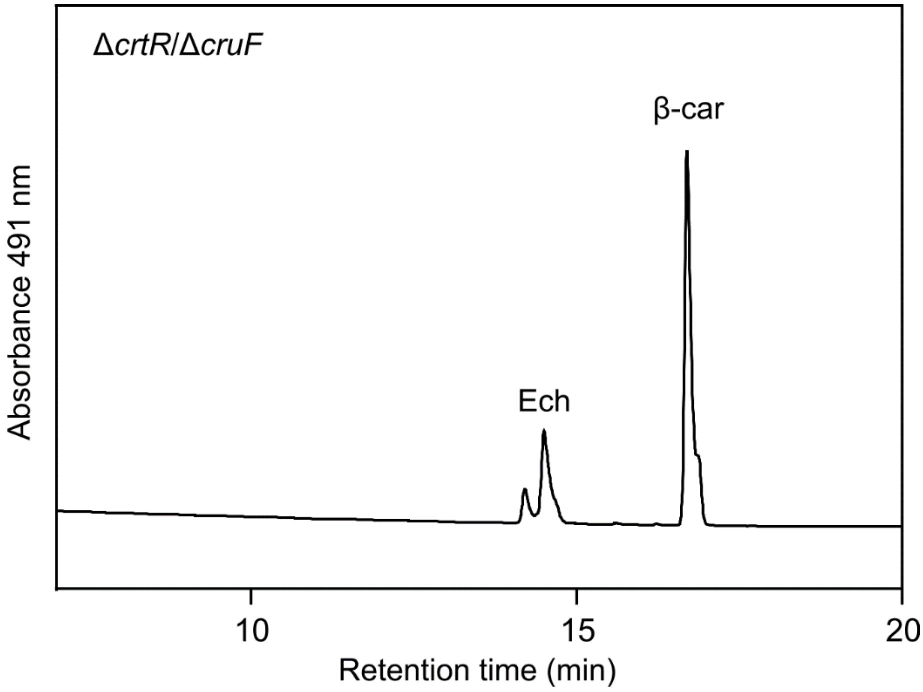


Figure 4.11: Analysis of the pigment content of Δ*crtR*/Δ*cruF* strain. Pigments were extracted in methanol and separated by reverse-phase HPLC. Myx, Zea and D-Myx were absent from the strain.

4.3.6 The interaction between ChlG and HliD is completely abolished in the absence of myxoxanothophyll, deoxy-myxoxanthophyll and zeaxanthin

To analyse the effects of preventing the synthesis of zeaxanthin, myxoxanthophyll and deoxy-myxoxanthophyll, the FLAG-*chlG* complex was purified from the Δ*crtR*/Δ*cruF* background. The immunoprecipitation eluate was not coloured to the naked eye and the absorbance spectra revealed that there was a clear reduction in the pigment content in comparison to the FLAG-*chlG* and Δ*cruF* eluates described previously, with levels even lower than that of the Δ*crtR* strain (Figure 4.12D). The pigments were extracted from the Δ*crtR*/Δ*cruF* eluate and analysed by reverse-phase HPLC (Figure 4.13). The profile revealed that a small amount of β-carotene remained in the eluate although deoxy-myxoxanthophyll was absent as expected.

The Δ*crtR*/Δ*cruF* eluate was separated by SDS-PAGE and analysed by immunoblot, probing for FLAG-ChlG and HliD (Figure 4.12A and 4.12B). FLAG-ChlG was present in the Δ*crtR*/Δ*cruF* eluate although there appeared to be a slight reduction in size between the FLAG-ChlG proteins from the Δ*crtR* mutant, and in particular the double Δ*crtR*/Δ*cruF* mutant, both of which have significant xanthophyll deficiencies. Immunoblot analysis failed to detect HliD from the Δ*crtR*/Δ*cruF* eluate in contrast to the WT and Δ*cruF* samples, which were similar, and the Δ*crtR* eluate that contained a significantly reduced but detectable quantity of HliD. To rule out the possibility that HliD does not accumulate in the Δ*crtR*/Δ*cruF* mutant, thylakoid membranes from this strain were prepared as described in Section 2.11.1 and analysed for HliD by imunnoblot (Figure 4.12C). A positive immunoblot singal was attained from the Δ*crtR*/Δ*cruF* sample in addition to positive signals for FLAG-*chlG*, Δ*crtR* and Δ*cruF* strains, indicating that HliD production is unaffected in these backgrounds.

Considering all of the data presented thus far, zeaxanthin is capable of independently mediating the association of ChlG with HliD. However, it is likely that myxoxanthophyll may also be capable of mediating this association to some extent, given the fact that a small population of HliD remains bound to ChlG when the myxoxanthophyll precursor deoxy-myxoxanthophyll accumulates in the cell and that in the absence of deoxy-myxoxanthophyll no HliD is detected in the FLAG-ChlG eluate by immunoblotting.

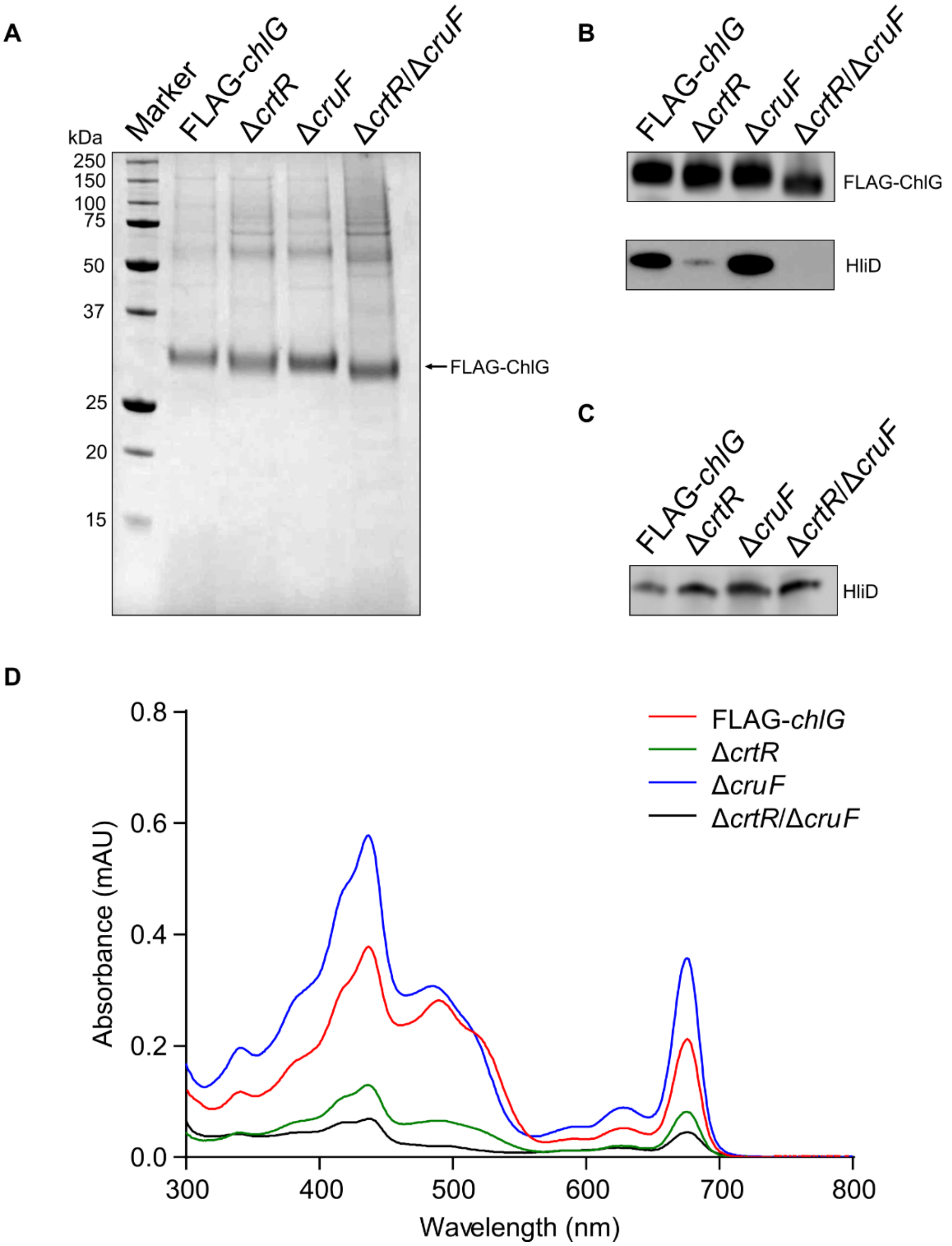


Figure 4.12: Purification of FLAG-ChlG from the Δ*crtR*/Δ*cruF* strain. (A) FLAG-immunoprecipitation eluates from FLAG-*chlG* Δ*chlG*, FLAG-*chlG* Δ*chlG* Δ*crtR* (Δ*crtR*), FLAG-*chlG* Δ*chlG* Δ*cruF* (Δ*cruF*) and FLAG-*chlG* Δ*chlG* Δ*crtR* Δ*cruF* (Δ*crtR*/Δ*cruF*) were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG and HliD. (C) Thylakoid membranes were solubilised in β-DDM and analysed for HliD content by immunoblot. (D) Absorption spectra of FLAG-immunoprecipitation eluates.

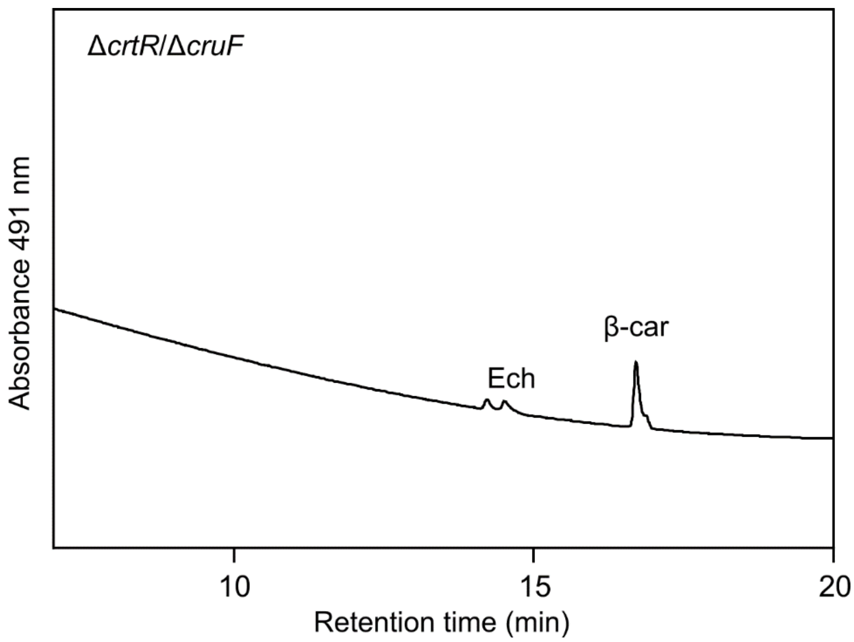


Figure 4.13: Analysis of the pigment content of the Δ*crtR*/Δ*cruF* immunoprecipitation eluate. Pigments were extracted in methanol and separated by reverse-phase HPLC. (A) Myx, Zea and D-Myx were absent.

4.3.7 The absence of carotenoids and HliD results in high molecular weight FLAG-ChlG sub-complexes

The FLAG immunoprecipitation eluates obtained from the FLAG-*chlG* Δ*chlG*, Δ*cruF*, Δ*crtR* and Δ*cruF*/Δ*crtR* strains were separated by HPLC gel filtration chromatography, monitoring absorbance at 280 nm (for protein) and 440 nm (for carotenoids) (Figure 4.14). The HPLC profiles of the samples changed significantly, with the protein component of the eluates eluting earlier as the carotenoid content decreased. The FLAG-*chlG* and Δ*cruF* samples (Figure 4.14A and 4.14B) contained the highest concentration of carotenoids, compared to Δ*crtR* and Δ*crtR*/Δ*cruF* (Figure 4.14C and 4.14D), in which most of the protein eluted in the void volume. The FLAG-*chlG* complex eluted in three distinct peaks, whereas an additional fourth peak appeared earlier in the HPLC profile in the Δ*cruF* sample, at ~6.4 ml. This protein peak did not co-elute with carotenoids and increased in intensity within the Δ*crtR* profile, becoming the major peak in the Δ*crtR*/Δ*cruF* sample. The shift in the protein absorbance profile is representative of a reorganisation of the ChlG complex into higher molecular weight species in response to decreasing carotenoid and/or HliD binding. The gel filtration fractions corresponding to protein peaks were collected and analysed by immunoblotting using antibodies raised against the FLAG-tag. The immunoblots revealed a shift of the FLAG-ChlG protein containing fractions in agreement with the 280 nm elution profile.

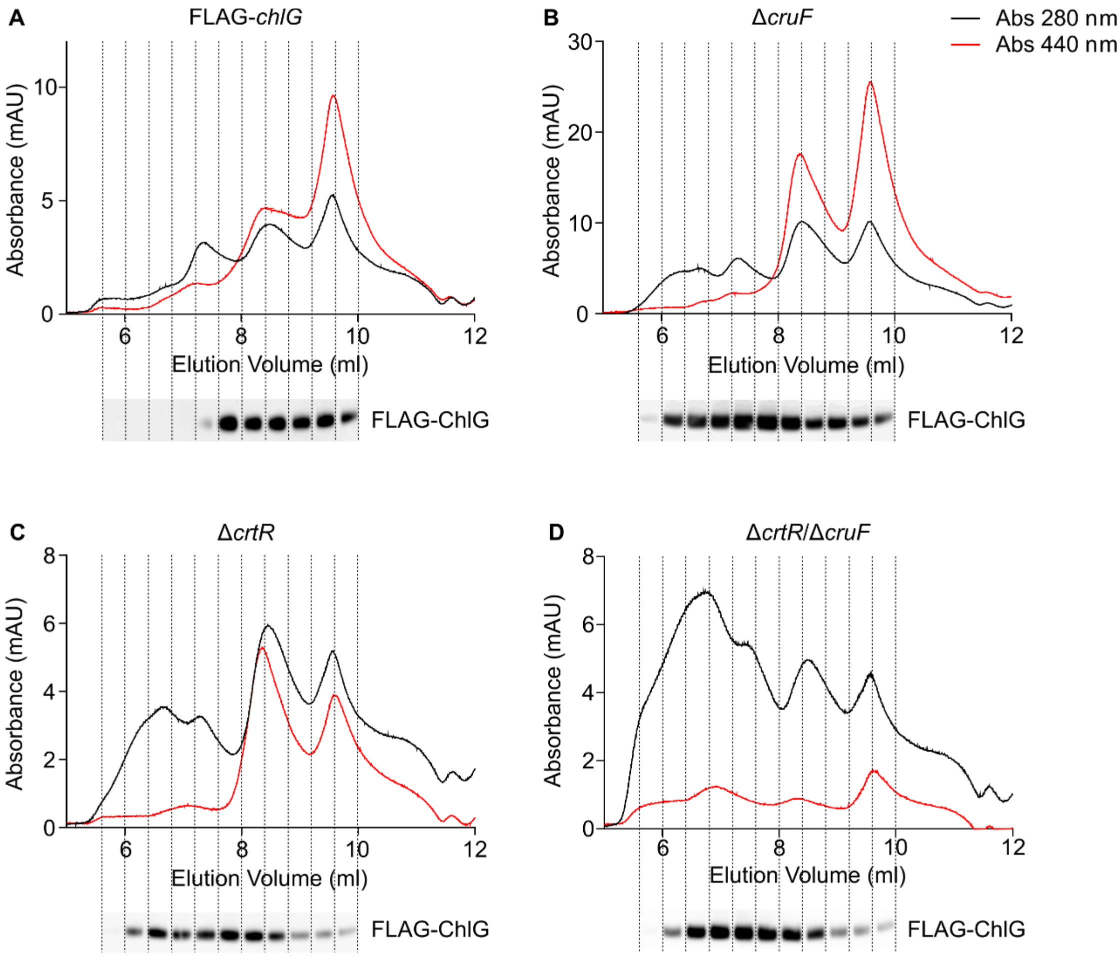


Figure 4.14: Gel filtration of the FLAG-ChlG eluates from the different carotenoid mutants. HPLC gel filtration chromatography separation of purified FLAG-ChlG complexes from *Synechocystis* FLAG-*chlG* Δ*chlG* (FLAG-*chlG*), FLAG-*chlG* Δ*chlG* Δ*cruF* (Δ*cruF*), FLAG-*chlG* Δ*chlG* Δ*crtR* (Δ*crtR*) and FLAG-*chlG* Δ*chlG* Δ*cruF* Δ*crtR* (Δ*cruF*/Δ*crtR*) strains. Elution of pigment and protein were monitored at 440 (red line) and 280 (black line) nm respectively. Immunoblot analyses of the FLAG-ChlG content of the HPLC elution fractions, using an anti-FLAG antibody, are shown below the traces.

4.4 Discussion

4.4.1 Zeaxanthin and myxoxanthophyll facilitate the interaction between ChlG and HliD

The carotenoids myxoxanthophyll and zeaxanthin bind to ChlG and were hypothesised to mediate the association of HliD with ChlG, forming the ChlG-HliD 'core' complex in *Synechocystis* (Chidgey et al., 2014; Staleva et al., 2015). To test this hypothesis, the synthesis of these carotenoids was prevented in *Synechocystis* strains containing an N-terminally FLAG-tagged ChlG protein that was used as bait in immunoprecipitation experiments, and the resulting eluate was analysed for HliD content by immunodetection.

Removal of both zeaxanthin and myxoxanthophyll by deletion of *crtR* resulted in a significant reduction in HliD binding to ChlG, possibly caused by the accumulation of the glycosylated CrtR substrate deoxy-myxoxanthophyll, which lacks the hydroxyl group on the β-ring. A small amount of this precursor was present within the FLAG-ChlG immunoprecipitate purified from the Δ*crtR* strain. The similarity of this carotenoid to myxoxanthophyll suggests that it enabled the binding of the small amount of HliD to the FLAG-ChlG protein, detected by immunoblot. To test whether this pigment was responsible for the residual ChlG-HliD complexion, the *cruF* gene was deleted from the Δ*crtR* background, producing a Δ*crtR*/Δ*cruF* double mutant that was not capable of synthesising zeaxanthin, myxoxanothophyll or deoxy-myxoxanthophyll. HliD was not detectable by immunoblot in the FLAG-ChlG complex purified from this background, indicating that deoxy-myxoxanthophyll is responsible for the residual ChlG-HliD interaction.

The residual β-carotene and Chl within the FLAG-ChlG immunoprecipitate purified from the Δ*crtR*/Δ*cruF* strain would normally signify the presence of HliD. However, no HliD was detected in the immunoprecipitation eluate by immunoblot. ChlG itself likely contains residual Chl, or its substrate Chlide, within the active site of the enzyme but has not been shown to bind any alternate pigments (Chidgey et al., 2014; Shukla et al., 2018a). HliC co-purifies with FLAG-ChlG and has been shown to bind both Chl and β-carotene in a 2:1 ratio (Chidgey et al., 2014; Shukla et al., 2018a). However, the interaction of HliC with ChlG appears to be dependent upon the presence of HliD, and so is unlikely to be responsible for the residual β-carotene observed in the Δ*crtR*/Δ*cruF* eluate (Shukla *et al*., 2018b). ChlG has also been documented to co-purify with small amounts of PSI momomer (Chidgey et al., 2014; Shukla et al., 2018b). Cyanobacterial PSI contains 22 β-carotene and 96 Chl *a* molecules (Wang *et al*., 2004). Although no experiments were performed in order to detect the presence of PSI within the FLAG-ChlG complex purified from the Δ*crtR*/Δ*cruF* background, it is likely that this accounts for the residual β-carotene remaining in the eluate and contributes to the signal produced by Chl, a hypothesis which should be tested in future experiments.

Taken together, it can be concluded that zeaxanthin and/or myxoxanthophyll mediates the formation of the ChlG-HliD 'core' complex in *Synechocystis*. Whether or not carotenoids facilitate formation of the ChlG-HliD complex in other cyanobacteria remains to be seen. All cyanobacteria contain homologs of HliD, however, the various species of cyanobacteria make different carotenoids. For example, *Synechococcus* sp. PCCC 7002 synthesises a different species of myxoxanthophyll called myxol-2' fucoside, whereas, *Synechocystis* makes myxol-2' dimethylfucoside (Graham and Bryant, 2009). It is likely that myxol-2' fucoside could mediate an equivalent ChlG-HliD interaction in *Synechococcus* sp. PCC 7002, consistent with the results of the previous chapter where heterologously produced FLAG-7002 ChlG co-immunoprecipitates with the same protein-pigment complex as the native *Synechocystis* enzyme (Proctor *et al*., 2018). In plants, an interaction of ChlG with the plant equivalents of Hlips, one helix proteins (OHPs), has not been demonstrated despite recent investigation (Hey and Grimm, 2018). Furthermore, the *Arabidopsis thaliana* and algal *Chlamydomonas reinhardtii* FLAG-ChlG proteins, when heterologously expressed in *Synechocystis*, were found not to bind either HliD or carotenoids (Proctor *et al*., 2018) (Chapter 3). It is possible therefore that a carotenoid mediated interaction between ChlG and HliD is limited to cyanobacteria.

4.4.2 Zeaxanthin is able to mediate the ChlG-HliD interaction in the absence of myxoxanthophyll

The ChlG-HliD interaction could be mediated by either zeaxanthin or myxoxanthophyll alone, or both may be required together for the complex to form. The selective removal of each of these pigments in turn would have been a means to determine this. Unfortunately, it is not possible to generate a *Synechocystis* strain that does not produce myxoxanthophyll but continues to synthesise zeaxanthin (Figure 4.1). However, a mutant *Synechocystis* strain that produces zeaxanthin but not myxoxanthophyll was generated by deletion of the gene *cruF*, enabling the effects of the specific loss of myxoxanthophyll on the ChlG complex to be examined. This strain was found to retain the association of ChlG and HliD in levels comparable to that of WT. As such, it can be concluded that zeaxanthin alone is enough to facilitate the formation of the ChlG-HliD complex. However, the possibility that myxoxanthophyll can also independently mediate the ChlG-HliD interaction cannot be dismissed. Considering the fact that small quantities of the myxoxanthophyll precursor, deoxy-myxoxanthophyll, enabled residual amounts of HliD to remain associated to ChlG in the Δ*crtR* strain, it is possible that this is the case. Although this is yet to be conclusively demonstrated, it can be speculated that there is a degree of redundancy in the function of myxoxanthophyll and zeaxanthin within the ChlG complex, with either carotenoid being able to facilitate the binding of HliD to ChlG independent of the other. The precise binding sites, binding constants and stoichiometries of these carotenoids to the ChlG complex have also yet to be established. If the roles of zeaxanthin and myxoxanthophyll overlap, there may not be independent binding sites for each pigment; individual ChlG enzymes could instead arbitrarily bind these carotenoids within the cell.

4.4.3 Ycf39 cannot bind ChlG in the absence of zeaxanthin/myxoxanthophyll and HliD

The putative short chain alcohol dehydrogenase protein, Ycf39, associates with ChlG in *Synechocystis* (Chidgey et al., 2014). The function of Ycf39 within the ChlG complex has yet to be determined; however, it has been shown to form a separate complex with HliD that is involved in the assembly of PSII (Knoppová *et al*., 2014). Ycf39 has been shown to be lost from the ChlG complex when the *Synechocystis* cells are subjected to high-light stress (Proctor *et al*., 2018). The mechanism behind this loss has recently been uncovered. Under high-light stress, the high-light inducible protein C (HliC) is upregulated and binds to the ChlG complex via HliD, causing Ycf39 to dissociate from the complex and leaving it free to form a new complex with a HliC/HliD heterodimer. The Ycf39-HliC-HliD complex can bind to PSII assembly intermediates and photoprotect them during the PSII assembly/repair process (Shukla *et al*., 2018b).

In the present study, the ChlG-HliD interaction, abolished by the inhibition of myxoxanthophyll and zeaxanthin biosynthesis, was accompanied by the simultaneous loss of Ycf39 from the ChlG complex. HliC cannot directly bind to ChlG but instead interacts by forming a heterodimer with the HliD component, triggering the dissociation of Ycf39 (Shukla *et al*., 2018b). Without HliD, the HliC protein is presumably absent from the FLAG-ChlG complex purified from the Δ*crtR*/Δ*cruF* strain, although this was not confirmed and should be investigated in future work. It is therefore likely that the interaction of Ycf39 with ChlG is dependent on the presence of HliD, as postulated previously (Proctor *et al*., 2018). It follows that HliC/Ycf39 dependent remodelling of the ChlG complex in response to high-light stress are also absent in the Δ*crtR*/Δ*cruF* strain.

4.4.4 Abolishing the binding of HliD to ChlG remodels the ChlG complex

Analysis of the FLAG-ChlG eluates purified from the WT, Δ*cruF*, Δ*crtR* and Δ*cruF*/Δ*crtR* strains showed a rearrangement of the ChlG complexes as their carotenoid and HliD content decreased, progressively forming larger sub-complexes that eluted earlier during size-exclusion chromatography separation. The progressive loss of HliD from the complexes purified in this study may cause oligomerisation of ChlG into the higher molecular weight complexes, or ChlG may associate more with PSI in the absence of xanthophylls/Hlips. Alternatively, the YidC component of the complex may associate with higher frequency to the ‘naked’ ChlG proteins, increasing the molecular weight. Further analysis of the gel filtration fractions by immunoblots targeting HliD, YidC and Ycf39 would further shed light on the change in composition of the ChlG complexes as xanthophyll carotenoids are removed.

4.5 Future work

To supplement the immunoblot analysis of these complexes from the carotenoid deficient mutants, precise quantification of the protein constituents of ChlG complexes purified from the Δ*cruF* and Δ*crtR*/Δ*cruF* strains should be performed by mass spectrometry to enable comparison with the existing data for WT and Δ*crtR* eluates presented in this chapter (Table 1).

The carotenoids bound to HliD have been shown to be able to quench light absorbed by the ChlG complex and have been hypothesised to confer photoprotection (Niedzwiedzki *et al*., 2016). The effect of light on the enzyme activity of the FLAG-ChlG complexes purified with and without bound carotenoids and HliD should be tested. This could be achieved by purifying the FLAG-ChlG complexes from WT and Δ*crtR*/Δ*cruF* *Synechocystis* strains followed by measuring the enzyme kinetics of the eluate in both light and dark conditions.

Although it was not possible to selectively abolish the synthesis of zeaxanthin from *Synechocystis*, the reverse scenario was achieved resulting in a strain specifically lacking myxoxanthophyll. The absence of myxoxanthophyll appeared to have no impact on the ChlG complex which remained associated with HliD, demonstrating that zeaxanthin alone is able to mediate the formation of the ChlG-HliD complex. Whether or not myxoxanthophyll can also mediate this interaction independently of zeaxanthin remains to be seen. As the generation of a ChlG complex containing myxoxanthophyll but not zeaxanthin cannot be achieved *in vivo*, the *in vitro* reconstitution of the ChlG-HliD interaction by provision of exogenous myxoxanthophyll to solubilised membranes from the Δ*crtR*/Δ*cruF* strain will be attempted. The FLAG-ChlG bait protein can then be immunoprecipitated to see if HliD co-purifies. Controls using zeaxanthin can also be performed.

Previous data suggest that a ChlG-HliD complex is also present in *Synechococcus* sp. PCC 7002 and is mediated by carotenoids (Proctor *et al*., 2018), but this was based on heterologous production of the *Synechococcus* sp. PCC 7002 ChlG in *Synechcystis* and should be confirmed *in vivo*.

A Δ*hliC Synechocystis* mutant has been shown to exhibit an increased sensitivity to high-light and cold stress since the HliC dependent remodelling of the ChlG complex is absent in these cells (Shukla *et al*., 2018b). In the present study, the prevention of zeaxanthin and myxoxanthophyll biosynthesis in a Δ*crtR*/Δ*cruF Synechocystis* mutant resulted in the loss of HliD, Ycf39 and presumably HliC from the ChlG complex. The remodelling of the ChlG complex by HliC is therefore also likely to be abolished. Exposure of this strain to high-light growth conditions, in a fashion analogous to the experiments described in Shukla, Jackson, *et al.* (2018), should be performed in order to analyse any potential phenotype exhibited by this mutant. It is worth noting, however, that the effects of abolishing zeaxanthin and myxoxanthophyll biosynthesis in this mutant are unlikely to be exclusive to the ChlG complex, thus any observed phenotype exhibited by this strain cannot be attributed solely to the disruption of ChlG-HliD 'core' complex formation but could be due to wider pleiotropic effects.

Finally, a small amount of echinenone was consistently present in all of the FLAG-ChlG complex variants purified from the mutant strains generated in this study. This may be a contaminant as it was not reported in the study of Chidgey *et al*. (2014). The effect of the removal of this pigment from *Synechocystis* on the ChlG complex should be examined. In Syn 7002, deletion of the gene encoding β-carotene 4-ketolase, *crtW*, resulted in a loss of echinenone and 3'-hydroxyechinenone (Zhu *et al*., 2010). *Synechocystis* produces an alternative FAD dependent β-ionone ring ketolase, CrtO

(Fernández-González *et al*., 1997). A *Synechocystis* mutant lacking the *crtO* gene has been generated in order to investigate this issue.

Chapter 5: Characterisation of the ChlG-YidC-HliC-HliD-Ycf39 complex via *in vivo* and *in vitro* chemical cross-linking

5.1 Summary

The terminal enzyme of the chlorophyll biosynthesis pathway in *Synechocystis*, chlorophyll synthase (ChlG), forms a complex with the YidC insertase, high-light inducible proteins C and D (HliC/HliD) and the atypical short chain dehydrogenase Ycf39. The ChlG complex is speculated to operate at the interface between chlorophyll biosynthesis and photosystem assembly, delivering *de novo* chlorophyll pigments to chlorophyll-binding proteins as they are being co-translationally inserted into the thylakoid membrane and assembled into functioning photosystems. Structural characterisation of the arrangement and stoichiometry of this complex would give a greater insight into the mechanisms behind this process.

In the present study, a chemical cross-linking approach was taken to form covalent bonds between the protein members of the ChlG complex. The cross-linking process was optimised *in vivo* and used to cross-link FLAG-tagged ChlG enzymes in intact *Synechocystis* cells before FLAG-ChlG was isolated from the cells for analysis by mass spectrometry. *In vitro* cross-linking experiments using isolated FLAG-ChlG complexes were performed to complement the *in vivo* study. The results from these two approaches were largely consistent. Subsequent analysis of the cross-linked complex by mass spectrometry revealed regions of interaction between proteins. Cross-links were identified between the N-terminal domain of ChlG and the central and C-terminal regions of Ycf39. This domain of ChlG was also observed to form cross-links with the cytoplasmic domains of both YidC and HliC. The C-terminal cytoplasmic domain of YidC formed cross-links with Ycf39, HliC and HliD. HliC and HliD both formed cross-links with Ycf39. Taken together, the results were used to generate a tentative model of the ChlG complex.

5.2 Introduction

Analysis of the tertiary and quaternary structures of proteins, as well as interaction partners and co-factors, is essential for determining their biological functions. Historically this has been achieved by methods such as X-ray crystallography, NMR and cryo-EM, although each comes with its own set of challenges. Crystallisation of proteins still relies largely on a trial and error approach, often with little success, especially with large protein complexes and integral membrane proteins. NMR is limited to smaller protein assemblies and cryo-EM to larger. These techniques often require large amounts of protein, presenting yet another challenge to structural analysis by these methods. However, in recent years, techniques have been developed that use mass spectrometry to generate low resolution structural information on proteins that can be used alone or in combination with these more traditional approaches. These mass spectrometry methods theoretically require smaller sample quantities given the enhanced sensitivity of instruments developed within the last 10 years and are broadly applicable to all biological systems (Leitner *et al*., 2012).

Among these mass spectrometry based approaches, chemical cross-linking of proteins, in combination with enzymatic digestion, has been increasingly employed to elucidate the three dimensional structure of a protein and the arrangement of protein subunits within a complex. This approach consists of covalently linking the functional groups of two peptides from within the same protein, or a protein interaction partner, using a cross-linking reagent (for examples, see below). This fixes structural features by introducing covalent bonds and, in the case of bifunctional cross-linkers with a spacer of defined length, additionally imposes constraints on the maximum distances possible between the cross-linked side chains. The peptides are then enzymatically digested and analysed by mass spectrometry. Three different types of bifunctional cross-link modification of peptides can be identified. A monolink is when an amino acid has been modified with a cross-linker which is subsequently hydrolysed at the free end without reacting with a nearby side chain. An intralink occurs when an amino acid has been cross-linked to another amino acid within the same peptide. When two amino acids belonging to separate peptides are cross-linked, this is known as an interlink. The distance restraints imposed by the cross-linker provides information on the proximity of the amino acid residues and, when enough cross-linked peptides have been identified, can aid the elucidation of the three dimensional structure of the protein or protein complex.

There are various forms of cross-linking reagent available for studying protein structure and interactions. Cross-linking reagents that contain a photoreactive group, such as diazirine, form covalent bonds with interacting peptides upon exposure to UV radiation. However, the most commonly used cross-linking reagents are homobifunctional N-hydroxysuccinimide (NHS) esters such as disuccinimidyl suberate (DSS) or bis(sulfosuccinimidyl) suberate (BS3), which react with amines i.e the free N-terminus and the functional group of lysine by nucleophilic attack, releasing the NHS or sulfo-NHS group in the process (Figure 5.1A and 5.1B). In addition to lysine, NHS esters have also other side reactions under various conditions (Kalkhof and Sinz, 2008). It is therefore important to know all possible amino acids that can react with the cross-linker to enhance interpretation.

Enrichment of cross-linked peptides is often an essential step in the design of cross linking experiments involving complex sample compositions, such as entire proteomes. This is because the yield of cross-linked target peptide is often low, due to the presence of other proteins and molecules that compete for the cross-linking reagent within the sample. Using a large excess of cross-linker is both expensive and will impede the analysis of the sample by mass spectrometry due to incomplete digestion of excessively cross-linked proteins. Enrichment strategies for cross-linked peptides may target the peptide of interest directly, such as purification by immunoprecipitation or affinity tags, or may enrich for cross-linked peptides in general such as by cation exchange chromatography. New cross-linking reagents are currently in development that aids this step of the process, including the synthesis of affinity tagged cross-linkers. These reagents have been described, for example Chu *et al.*, 2006, but have not yet been commercialised.

Once the sample has been cross-linked, it is necessary to subject the peptides to enzymatic cleavage before analysis by mass spectrometry. Trypsin remains the most popular proteolytic enzyme, having been used in the majority of studies to date as it produces peptide fragments with ideal properties for mass spectrometry, such as length, charge state and dissociation behaviour for product ion scanning. NanoLC-MS/MS is most commonly employed to identify cross-linked peptides as this method is capable of detecting cross-linked peptides at low abundance in a background of unmodified peptides.

In the field of photosynthetic research, cross-linking techniques have been applied in order to elucidate the binding sites of the electron carrier ferredoxin (Fd) to PSI and NADPH-cytochrome *c* reductase in spinach thylakoid membranes (TM) (Merati and Zanetti, 1987; Zilber and Malkin, 1988), as well as the binding sites of flavodoxin to PSI in cyanobacteria (Muhlenhoff *et al*., 1996). *In vivo* cross-linking of whole *Synechocystis* cells revealed the association of the phycobilisome, PSI and PSII into a megacomplex that facilitates the simultaneous transfer of captured light energy to the reaction centres of both photosystems (Liu *et al*., 2013). A novel interaction between PSI subunits PsaF and PsaE was discovered by *in vitro* cross linking of *Synechocystis* PSI particles and TM (Armbrust *et al*., 1996). In *Rhodobacter sphaeroides*, the interaction between cytochrome c_2 and the reaction centre complex was mapped to the M subunit of the latter and enabled the kinetics of electron transfer between the two to be elucidated (Friedel Drepper *et al*., 1997). Time resolved cross-linking experiments in *Rhodopseudomonas capsulata* (now *Rhodobacter capsulatus*) were performed to examine the near neighbour relationships between photosynthetic peptides. The results revealed that the synthesis of bacteriochlorophyll and bacteriochlorophyll binding proteins are coordinated at the level of translation (Drews *et al*., 1983). A separate study using the same system was performed with the aim to reveal the topographical organisation of the photosynthetic complexes for efficient energy transfer (Takemoto *et al*., 1982).

In *Synechocystis*, a chlorophyll synthase (ChlG) protein-pigment complex acts at the interface between the chlorophyll biosynthesis and photosystem assembly pathways (see Section 1.11 for discussion). The arrangement and stoichiometry of the ChlG complex has not yet been elucidated. Determination of the binding sites between its protein members may shed light on the individual functions of these components and better our understanding of how this complex functions within the photosynthetic unit. In this chapter, cross-linking techniques were used in combination with mass spectrometry to gain an insight into the spatial organisation and contact sites between the proteins of the ChlG complex within the thylakoid membrane of *Synechocystis. In vitro* cross-linking experiments were performed on FLAG-ChlG complex purified from *Synechocystis* via FLAG-immunoprecipitation and analysed by nanoLC-MS/MS. Additionally, *in vivo* cross-linking experiments were optimised to elucidate how this protein assembly is arranged in the thylakoid membrane.

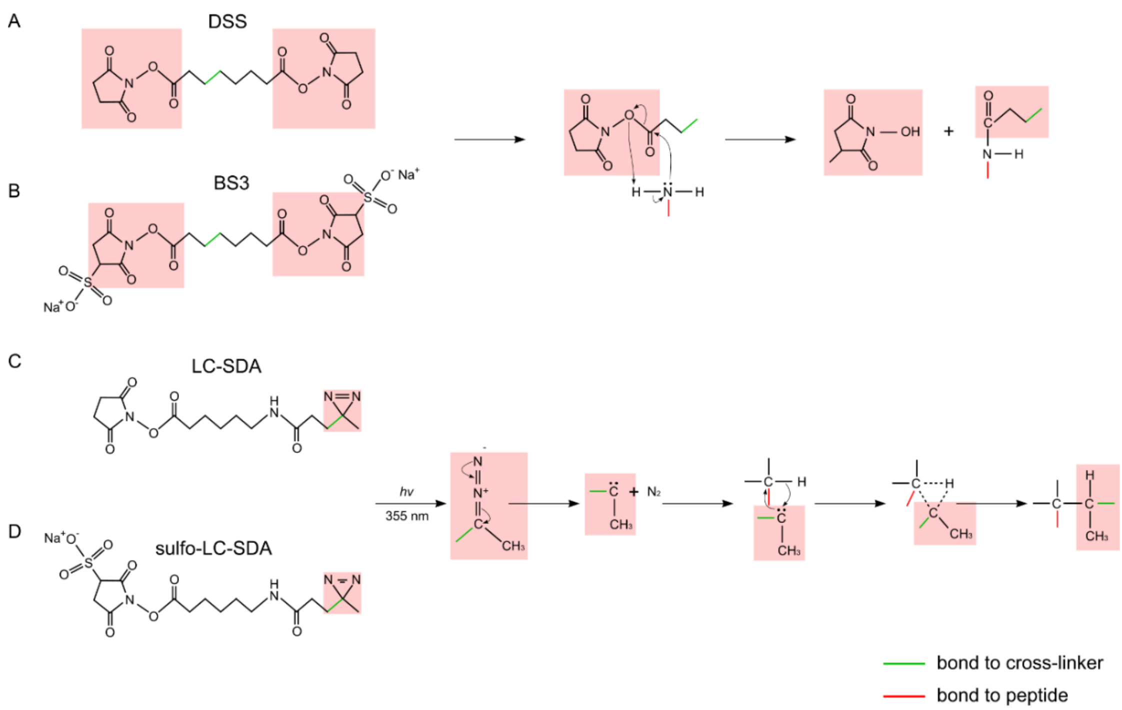


Figure 5.1: Structures of DSS, BS3, LC-SDA and sulfo-LC-SDA. Homobifunctional disuccinimidyl suberate (DSS) and its water soluble counterpart bis(sulfosuccinimidyl)suberate (BS3) react with the primary amine groups of proteins and form amide bonds, releasing N-hydroxysuccinimide. (C-D) Heterobifunctional cross-linking reagents LC-SDA and water soluble variant LC-SDA have a amine-reactive N-hydroxysuccinimide (NHS) ester at one end and a diazirine group at the opposite end that, once activated with UV light (355 nm), can theoretically react with any amino acid side chain.

5.3 Results

5.3.1 *In vitro* cross-linking of FLAG-ChlG using BS3, DSS and LC-SDA

To determine the sites of interaction between the components of the ChlG-YidC-Ycf39-HliD complex, the proteins were cross-linked using the homobifunctional NHS-ester cross-linking reagents BS3 and DSS and the heterobifunctional NHS ester/diazirine functionalised sulfo-LC-SDA and LC-SDA. Since BS3 and DSS react primarily with amino groups, the diazirine-based reagent was used to potentially extend the sequence coverage of cross-linked peptides. The NHS ester ends of these cross-linkers react with primary amines whilst the diazirine end, once photoactivated by UV light, forms a reactive carbene species that can insert into sterically unhindered C-H, O-H and N-H bonds in any amino acid side chain within the spacer arm distance of 12.5Å (Figure 5.1 C-D). These heterobifunctional cross-linkers are therefore not reliant on the close proximity of two lysine side chains in order to from a heterolink, increasing the potential for wider sequence coverage.

The FLAG-ChlG complex from *Synechocystis* was isolated by immunoprecipitation and subjected to *in vitro* cross-linking experiments as described in Materials and Methods Sections 2.11.2 and 2.10.2. Following cross-linking, the proteins were digested with a combination of endoproteinase Lys-C and trypsin (described in Section 2.9.4) and the peptide fragments analysed by nano-LC-MS/MS (Section 2.9.5). The mass spectra were analysed for cross-links between ChlG (slr0056), YidC (slr1471), HliD (ssr1789), HliC (ssl1633) and Ycf39 (slr0399) using the Byonic search engine from Protein Metrics. Putative cross-linked peptides identified by this software were checked for legitimacy by visual curation according to the criteria outlined in Table 2. The cross-linking reagents that produced valid results were BS3 and LC-SDA which are water and DMSO soluble respectively. No cross-linked peptides were detected after treatment with the DMSO soluble DSS or the water soluble sulfo-LC-SDA. Therefore, in this experiment there is no clear relationship between the solubility properties of the cross-linking reagent and reaction with detergent-solubilised proteins. Examples of mass spectra pertaining to cross-links between ChlG and its partner proteins are shown in Figure 5.2.

ChlG, YidC, HliD and HliC are all predicted to contain transmembrane helices, however structures for these proteins are unavailable to date. These proteins were modelled using five different software applications, TMHMM, PRED-TMR, TMpred, HMMTOP and Uniprot, designed to predict protein secondary structure and transmembrane regions. The numbers of transmembrane regions predicted for each protein are presented in Table 1. There was disagreement between the results depending on the software used; ChlG was predicted to have between 6 and 9 TMH regions, but 8 TMH regions have been adopted in this study given that the structural models of both the *Synechocystis* and *Arabidopsis thaliana* ChlG homologs, presented in chapters 5 and 7 respectively, contain 8 TMHs. For YidC, 3 TMHs were assumed on the basis of the number given by 3 of the 5 prediction algorithms. HliD and HliC were classified as both having 1 TMH, despite predictions of zero TMHs by a number of algorithms, because previous studies have already determined this to be the case due to their high sequence homology to the first or third transmembrane helix of the plant light-harvesting complexes, believed to be the evolutionary descendants of Hlips in higher phototrophic organisms (Dolganov *et al*., 1995; Komenda and Sobotka, 2016). The orientations of ChlG, YidC, HliD and HliC within the thylakoid membrane were not apparent from the results produced using these software applications. In all cases, there was an almost equal probability that the N- or C-terminal membrane-extrinsic domains were either cytoplasmic or thylakoid luminal. The transmembrane regions were mapped to the protein sequences and used in conjunction with the cross-linking results to inform a tentative model of the ChlG complex (Figure 5.8).

The following cross-links were detected: ChlG-ChlG (Figure 5.2A), ChlG-Ycf39 (Figure 5.2B), ChlG-YidC (Figure 5.2C), ChlG-HliC (Figure 5.2D) and HliD-ChlG (Figure 5.2E) shown in Figure 5.8 as red and green dashed lines. In cases where product ion spectra are populated with ions that represent sufficient peptide sequence coverage, it was possible to identify cross-linked amino acid side chains unambiguously. However, because cross-linked peptides frequently dissociate inefficiently and/or idiosyncratically in the mass spectrometer during product ion scanning, it is only possible to deduce the sequence region involved in the cross-linkage. These scenarios are indicated in Figure 5.2.

Figure 5.8 shows a ChlG-ChlG interaction between amino acid side chains located in the N-terminal FLAG extension and two interactions within the predicted N-terminal cytoplasmic domain (red dashed lines). Because the cross-linked peptides identified in these cases were different, it is impossible to determine whether these interactions are within a single ChlG or between two neighbouring ChlG proteins. There is also evidence for an interaction between this cytoplasmic domain and the predicted TMH7-TMH8 extra-membranous domain (red dashed line). Since this domain would be, according to the current model, located in the thylakoid lumen, such an interaction would be improbable. This discrepancy highlights the need to acquire further data to enable refinement of the model. Alternatively, this interaction may be artefactual, resulting from inversion of a sub-population of ChlG complexes, with respect to neighbouring proteins, within the detergent micelles before reaction with the cross-linker. Detergent-related anomalies may be eliminated by performing cross-linking reactions *in vivo* (Section 5.3.3).

Consistent with the current models, a cross-link was detected between the TMH7-TMH8 extra-membrane loop domain of ChlG and the predicted N-terminal thylakoid luminal domain of HliD (green dashed line). Similarly, the N-terminal cytoplasmic domain of ChlG also cross-linked to YidC in the predicted TMH1-TMH2 cytoplasmic loop domain (red dashed line). The N-terminus of the FLAG extension on ChlG was cross-linked to the predicted N-terminal cytoplasmic domain of HliC (red dashed line).

A putative cross-link was also identified between the predicted cytoplasmic domain of HliD and the predicted TMH2-TMH3 sequence of YidC located in the thylakoid lumen (green dashed line). This contradiction can be explained by the use of detergents as above or it is possible that this cross-link was formed during YidC assisted insertion of HliD into the membrane when these two domains may be in close proximity. As a YidC-HliD interaction has not been reported in the literature, this scenario requires further investigation**.**

Table 1: TMH predictions. The amino acid sequences of ChlG (slr0056), YidC (slr1471), HliD (ssr1789), HliC (ssl1633) and Ycf39 (slr0399) were analysed by 5 different software applications to predict the number and location of transmembrane helices within the query sequence.

	Number of TMHs predicted				
Protein	TMHMM	PRED-TMR	TMpred	HMMTOP	Uniprot
ChlG	7	7	8	9	6
YidC	3	3	5	5	3
Ycf39	0	0	0	0	0
HliD	1	0	1	1	1
HliC	0	0	1	1	Not specified

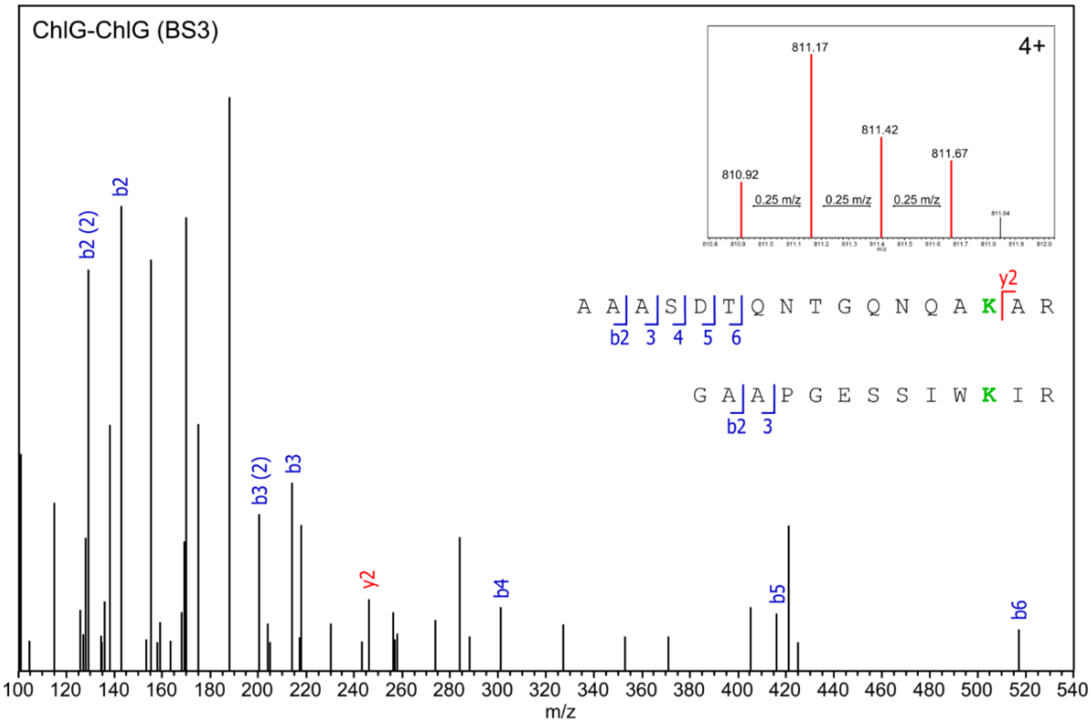
Table 2: Criteria for the selection of cross-linked peptides. Criteria to guide the visual curation of cross-linked peptides identified using the search engine Byonic (Protein Metrics).

Criteria
2-D Posterior Error Probability score below 0.1
C-terminus of putative cross-linked peptides is K or R
The peptide sequences are consistent with the cross-linking chemistry
Observed precursor ion m/z is equal to predicted m/z to two decimal places
Tryptic cleavage sites are consistent with sequence and cross-linking chemistry
Charge state of the precursor peptide ion is correct
The precursor peptide ion is consistent with the ion selected for MS/MS
The precursor peptide has the expected isotopomer peak distribution

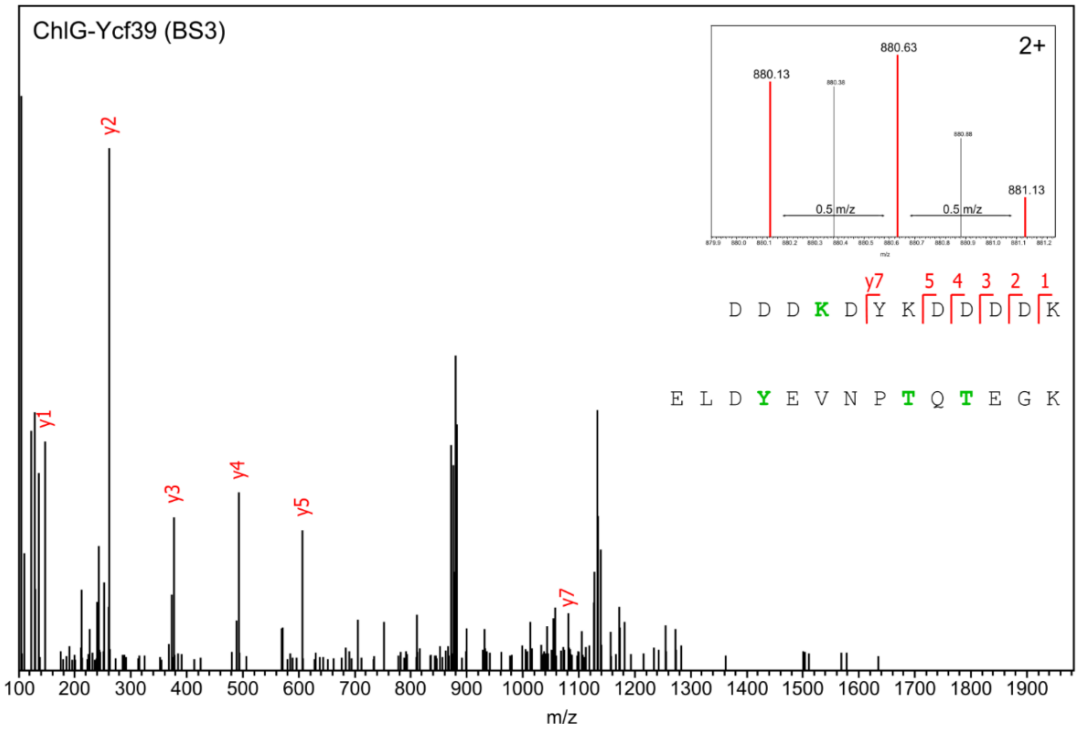
Table 3: *In vitro* BS3 and LC-SDA cross-linked peptides. Analysis by mass spectrometry of a ChlG complex isolated from 4 L cells after in vitro chemical cross-linking with DSS, BS3, LC-SDA and sulfo-LC-SDA. Proteins 1-2 pertains to the left and right proteins in the pair respectively. Putative cross-linked residues are highlighted in bold. PSM (peptide spectrum match) refers to the number of times a peptide was identified within the mass spectra. K indicates monolink modification of a K side chain.

Cross-linked Proteins 1-2	Cross-linker	Peptide in protein 1	Peptide in protein 2	Calculated m/z	Observed m/z	PSM
ChlG-ChlG	BS3	MDYKDDDD**K**DYK	DDDD**K**DYKDDDDK	1102.4368	1102.4375	1
ChlG-ChlG	BS3	QLLGM**K**GAAPGESSIWK	AAASDTQNTGQNQA**K**AR	915.2127	915.2131	6
ChlG-ChlG	BS3	AAASDTQNTGQNQA**K**AR	AAASDTQNTGQNQA**K**AR	900.9391	900.9394	1
ChlG-ChlG	BS3	NPLENDV**K**YQASAQPFLVFGMLATGLALGHAGI	AAASDTQNTGQNQA**K**AR	1063.1419	1063.1427	1
ChlG-ChlG	BS3	AAASDTQNTGQNQA**K**AR	GAAPGESSIW**K**IR	810.9143	810.9149	1
ChlG-Ycf39	BS3	DDDD**K**DYKDDDDK	ELD**Y**EVNP**T**Q**T**EGK	880.1310	880.1307	16
ChlG-Ycf39	BS3	MDY**K**DDDDK	AVQ**K**AGIK	1117.5485	1117.5475	1
ChlG-YidC	BS3	QLLGM**K**GAAPGESSIWKIR	DDPA**K**QQEEMA**K**VMK	1021.5295	1021.5307	2
ChlG-HliC	BS3	MDY**K**DDDDK	MNNENS**K**FGFTAFAENWNGR	1205.8524	1205.8466	3
YidC-YidC	BS3	ITQPLM**K**ER	**K**MR	426.4890	426.4888	1
HliD-YidC	BS3	SEELQPNQTPVQEDP**K**FGFNNYAEK	E**K**TS	876.4191	878.4201	1
YidC-HliD	LC-SDA	GSPFS**DINYTVDL**QILPQEQVER	FGFNNYAE**K**LNGR	1093.8026	1093.8080	1
HliD-ChlG	LC-SDA	M**S**EELQPNQTPVQEDPK	**NPLENDV**K	1031.5022	1031.5079	1

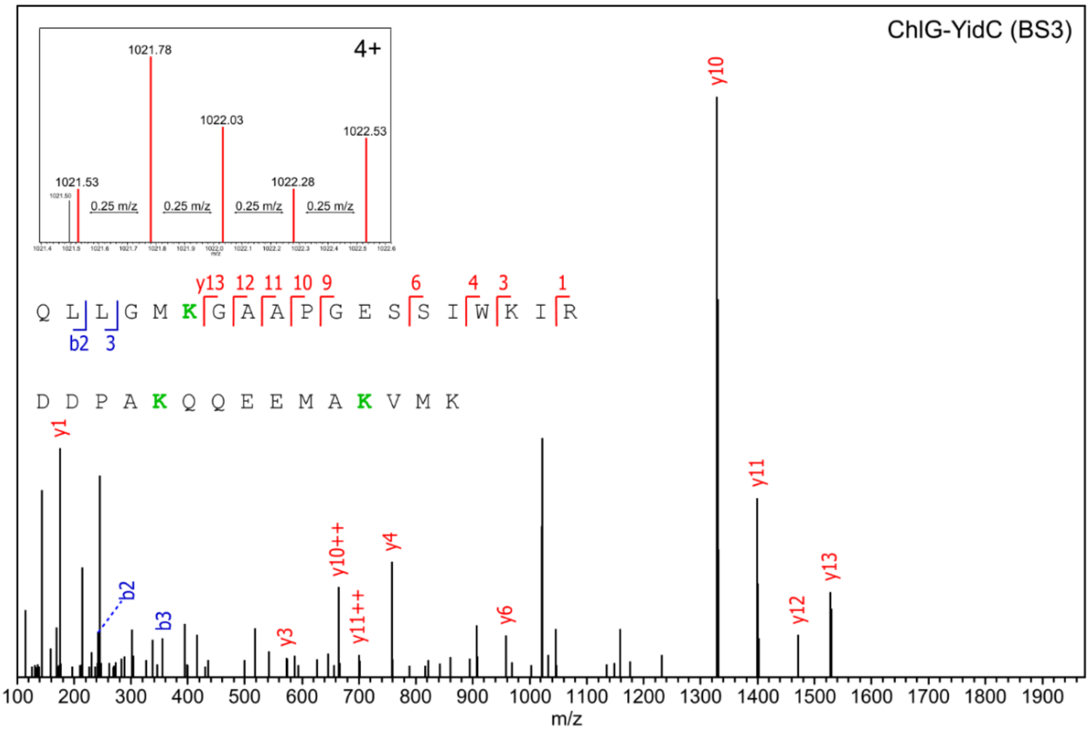
A


ChlG-ChlG (BS3)
4+
811.17
811.42
811.67
810.92
0.25 m/z
A A A S D T Q N T G Q N Q A K A R
G A A P G E S S I W K I R
b2 (2)
b2
b3 (2)
b3
y2
b4
b5
b6
m/z

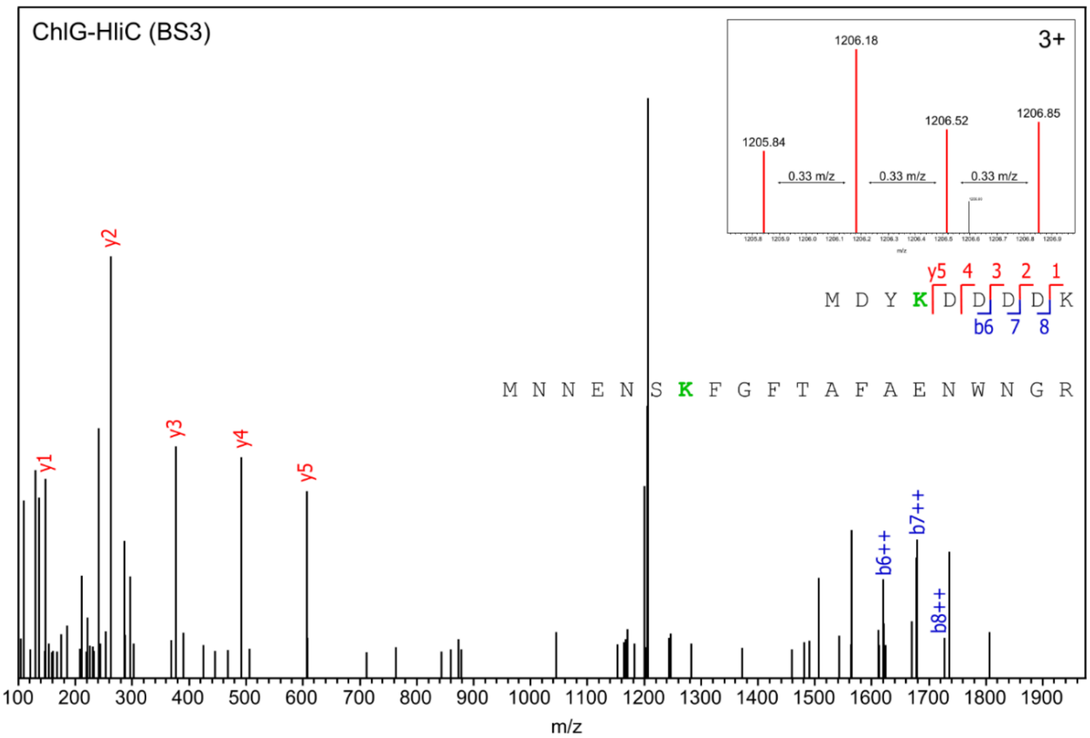
B


ChlG-Ycf39 (BS3)
2+
880.63
880.13
881.13
0.5 m/z
D D D K D Y K D D D D K
E L D Y E V N P T Q T E G K
y1
y2
y3
y4
y5
y7
m/z

C


ChlG-YidC (BS3)
4+
1021.78
1022.03
1022.53
1021.53
1022.28
0.25 m/z
Q L L G M K G A A P G E S S I W K I R
D D P A K Q Q E E M A K V M K
y1
b2
b3
y3
y10++
y11++
y4
y6
y10
y11
y12
y13
m/z

D


ChlG-HliC (BS3)
3+
1206.18
1206.52
1206.85
1205.84
0.33 m/z
M D Y K D D D D K
M N N E N S K F G F T A F A E N W N G R
y1
y2
y3
y4
y5
b6++
b7++
b8++
m/z

E

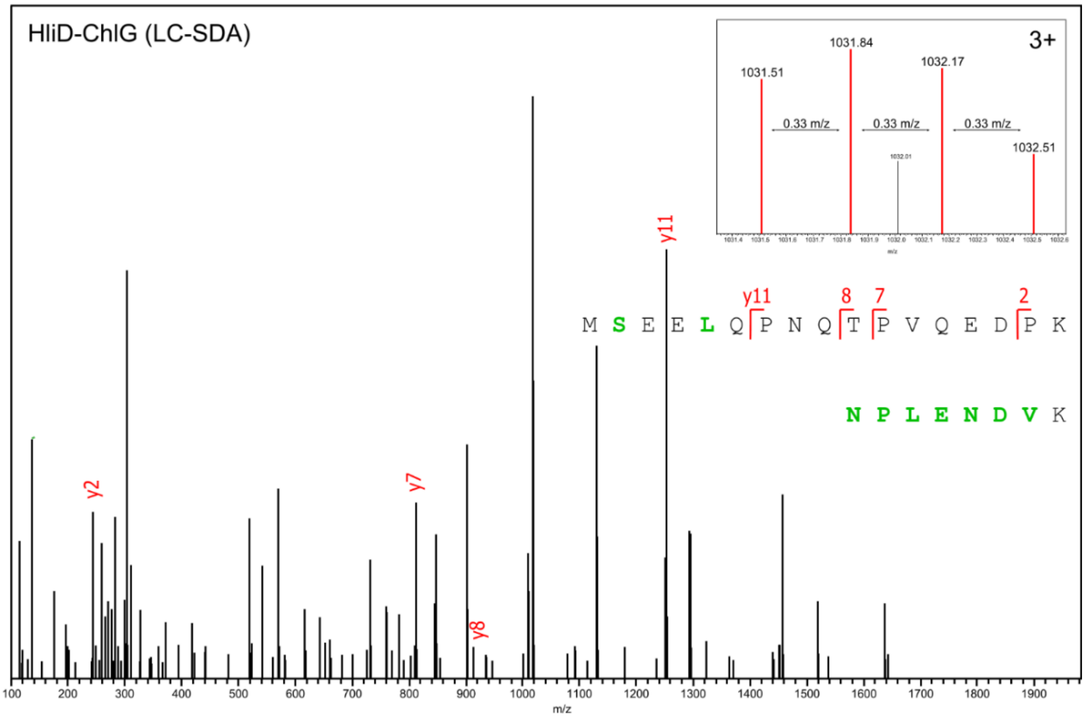


Figure 5.2: Analysis of FLAG-ChlG eluate, cross-linked with BS3, DSS and LC-SDA *in vitro*, by mass spectrometry. FLAG-ChlG was purified from 4 L of *Synechocystis* cells and treated with BS3, DSS and LC-SDA cross-linking reagents before being analysed by nanoLC-MS/MS. Cross-linked peptides were identified using the search engine Byonic (Protein Metrics). MS/MS spectra representing (A) ChlG-ChlG, (B) ChlG-YidC, (C) ChlG-Ycf39, (D) ChlG-HliC and (E) HliD-HliD among others (Table 3) were analysed to determine the sites of interaction between these proteins. Residues shown in bold green are potential cross-link sites. MS spectra (panels) were used to confirm the charge state and isotopomer peak distribution of the precursor ion selected for MS/MS.

5.3.2 Optimisation of *in vivo* cross-linking in WT *Synechocystis* intact cells

Immunoprecipitation of the FLAG-ChlG complex led to the identification of HliD, HliC and Ycf39 as interaction partners (Chidgey et al., 2014). Cross-linking of this complex *in vitro* with homobifunctional NHS ester and heterobifunctional NHS ester/diazirine reagents provided information to complement the current model for the TMH arrangement of these proteins, albeit in detergent micelles rather than thylakoid membranes. Since this matrix is artificial an *in vivo* cross-linking strategy, utilising intact *Synechocystis* cells, was investigated. *In vivo* cross-linking of intact *Synechocystis* cells has been previously reported (Liu *et al.*, 2013). In this earlier study of photosystem-phycobilisome interactions, the DMSO soluble cross-linking reagent dithiobis[succinimidylpropionate] (DSP) was applied directly to an intact cell suspension. Entry of the reagent into the cells relied on the ability of DMSO to permeabilise the cell wall and membrane. The aim of this section is to extend the range of cross-linking reagents employed, optimising concentrations and investigating the feasibility of using water soluble reagents in addition to their DMSO soluble counterparts.

Cross-linking reagents were added to WT *Synechocystis* cells in varying concentrations with the aim of determining the optimal concentration of cross-linker to use in order to achieve maximal yields of cross-linked peptides without oversaturating the sample (highly cross-linked proteins are more resistant to tryptic digestion). Freshly prepared *Synechocystis* cells were re-suspended to a concentration OD_{750}=0.7 in B-PER reagent. The commercially available B-PER reagent (Thermo Fisher Scientific) has been shown to permeabilise the thick outer cell membrane of cyanobacterial species *Synechococcus* sp. PCC 7002 and *Synechocystis* in order to increase the efficiency of *in vivo* enzyme assays. The pores formed by B-PER in the membrane of these cells aid the diffusion of the enzyme substrates into the cell, negating the need to lyse the cells before enzyme assays could be performed (Rasmussen *et al*., 2016). It was theorised that pre-treatment of the *Synechocystis* cells used in this study would facilitate the entry into the cells of, in particular, the water soluble cross-linker in order to react with the proteins in the thylakoid membrane and cytoplasm. As such, once re-suspended in B-PER, the cells were left incubating in the dark for 10 minutes according to the method of Rasmussen *et al* (2016), after which the cross-linking reagent was added to the cell suspension, following the method outlined in Section 2.10.1. Once the cross-linking reaction had been performed, excess cross-linker was quenched and the cells washed. The supernatants of the samples containing DMSO (DSS and LC-SDA) were highly pigmented, suggesting that the DMSO/B-PER combination had solubilised the intracellular proteins which had subsequently leaked out of the cell into the medium.

The cells were lysed and the extracted proteins digested with endoproteinase-Lys-C and trypsin and the peptide fragments analysed by nanoLC-MS/MS (Sections 2.9.4 and 2.9.5).

The data obtained by nanoLC-MS/MS was scanned for putative cross-links using Byonic software. The software was configured to identify cross-link modifications to the four most abundant phycobiliproteins: phycocyanin alpha (sll1578) and beta (sll1577), allophycocyanin alpha (slr2067) and beta (slr1986), as identified by proteomic analysis (Figure 5.3); and four photosystem apoproteins: P700 Ia (slr1834) and Ib (slr1835), and PSII D1 (slr1311) and core light harvesting protein (slr0906). Putative cross-links identified by Byonic were screened manually as described previously (Table 2) to select definitive cross-linked peptides. The number of cross-linked peptides identified was used as an estimate of the efficiency of the cross-linking reactions to determine the best conditions to use in scale up experiments with the FLAG-ChlG mutant *Synechocystis* strain (Figure 5.4).

Figure 5.4 shows that sulfo-LC-SDA was by far the most effective cross-linking reagent, producing 199 putative cross-links involving the selected high abundance proteins in the *Synechocystis* thylakoid membranes. Increasing the sulfo-LC-SDA concentration from 20 to 45 mM had little impact on the number of cross-links produced. Of the eight proteins screened for cross-link modification, slr0906 appeared to be the most susceptible to reaction with sulfo-LC-SDA (105 cross-links). DSS was the second most efficient reagent, producing 30 cross-links when used in 20 and 45 mM concentrations, however at 5 mM concentration 25 cross-links were identified, indicating that increasing the concentration of cross-linker above 5 mM had little impact upon cross-linking efficiency. Modification by DSS was almost entirely limited to phycobiliproteins with the majority of cross-links being detected on slr2067 and a few on sll1578 and slr1986. A single cross-link was detected on slr0906, the only photosystem polypeptide to be modified using DSS. BS3 was the least effective of the three cross-linking reagents used. Of the eight proteins, the majority of the cross-links were detected on slr2067 and slr1986. Like DSS, increasing the concentration of BS3 did not increase the number of cross-links; 5 mM produced 9 cross-links and 45 mM produced 10 cross-links.

In the experiment with the hydrophobic NHS ester/diazirine reagent, LC-SDA no cross-links were detectable presumably because of the significant depletion of intra-cellular proteins, as deduced from the strong colouration of the supernatants after pelleting the cells. It can therefore be concluded that the majority of the protein was solubilised by the DMSO used in these reactions which subsequently leaked out of the pores created in the cellular outer-membrane by B-PER. As the same effect was observed in the DSS experiment; these results indicate the incompatibility of B-PER and DMSO when used in combination in *in vivo* cross-linking reactions. B-PER was therefore omitted from all subsequent cross-linking reactions utilising DMSO soluble reagents.

In addition to enabling the determination of the ideal cross-linker and reaction conditions for subsequent experiments, the cross-linking data obtained above was also scanned for putative cross-links between members of the ChlG complex; ChlG, HliD, HliC, YidC and Ycf39. Only low confidence cross-links were detected in the experiments, highlighting the requirement for specific enrichment of ChlG containing complexes upstream of nanoLC-MS/MS analysis to enhance signal-to-noise in the mass spectra.

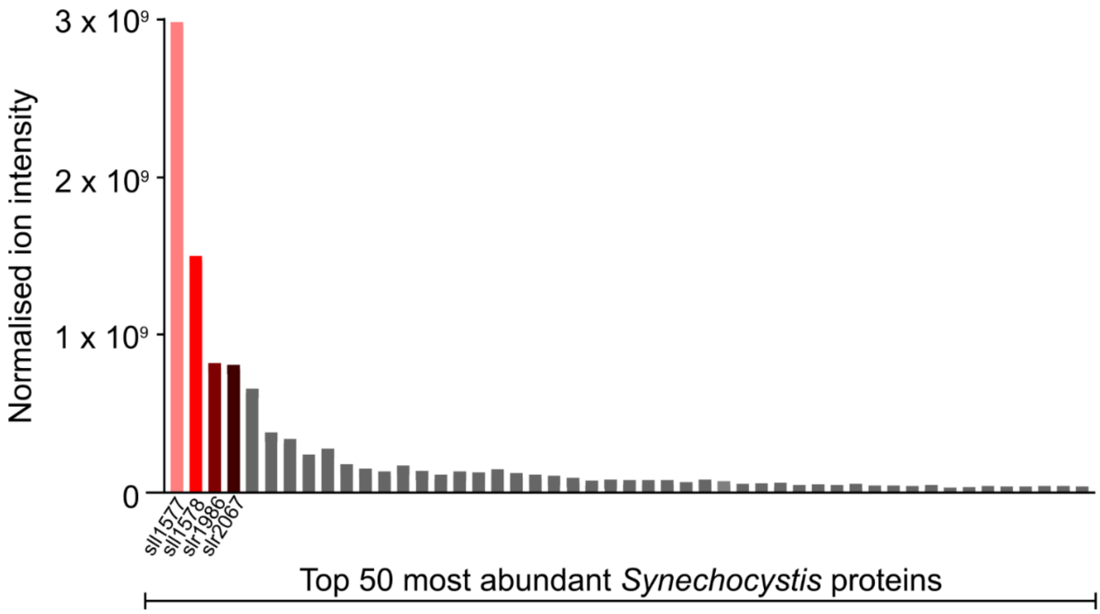


Figure 5.3: Identification of the top 50 most abundant *Synechocystis* proteins by quantative mass spectrometry. Proteins were extracted from *Synechocystis* cells from three independent cultures and analysed in triplicate by intensity based absolute quantification (IBAQ). Proteins quantification was achieved using MaxQuant (Cox and Mann, 2008). The top 4 most abundant proteins sll1578, sll1577, slr2067 and slr1986 are shown in red.

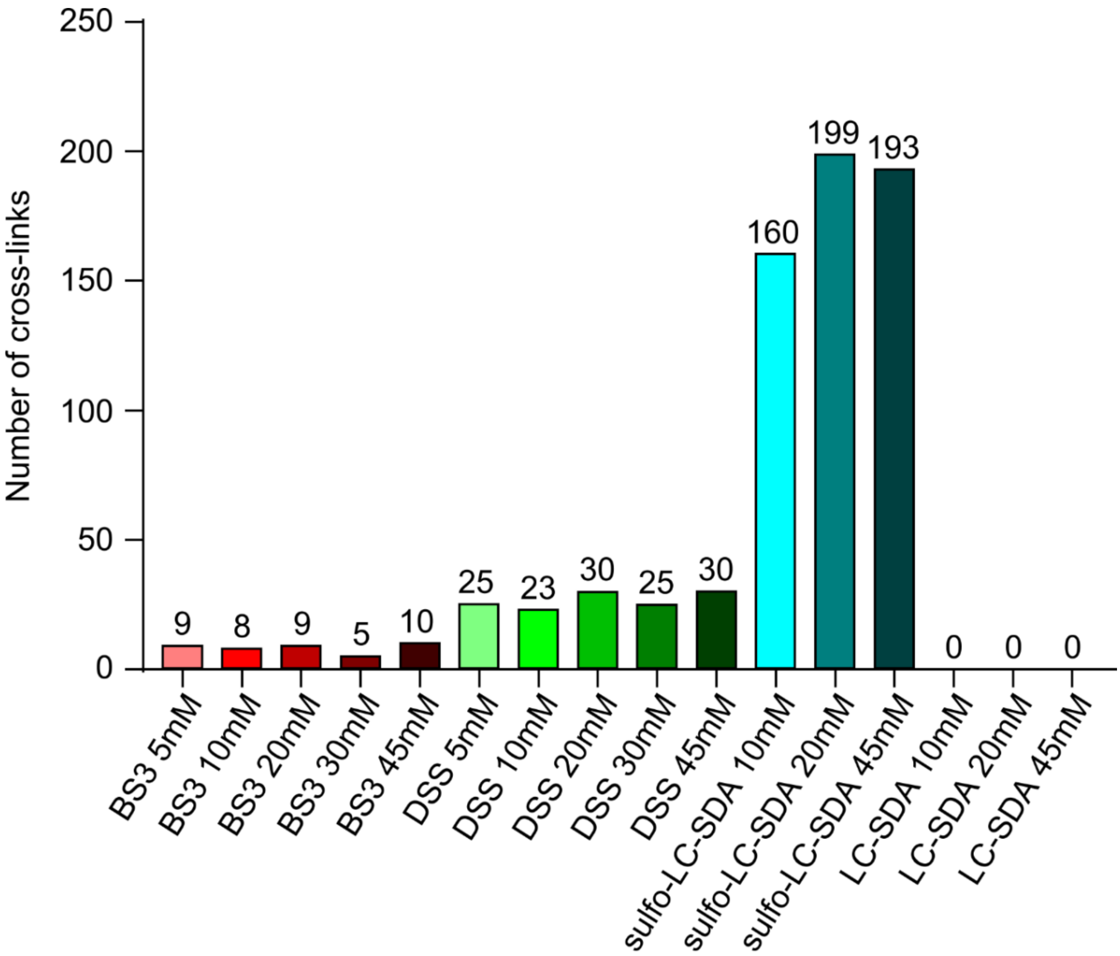


Figure 5.4: Number of cross-links detected *in vivo* between abundant *Synechocystis* proteins when varying cross-linking reagent species and concentration. *In vivo* cross-linking of intact *Synechocystis* cells was performed using DSS (red), BS3 (green), LC-SDA and sulfo-SDA (blue) at various concentrations. Cross-link modification to the 4 most abundant *Synechocystis* proteins (Figure 5.3) as well as photosystem apoproteins slr1834, slr1835, slr1311 and slr0906 were identified using the Byonic search engine (Protein Metrics). Peptides modified with a cross-link were summed for each experiment and used as a measure to determine the optimal conditions for *in vivo* cross-linking.

5.3.3 *In vivo* cross-linking of the FLAG-ChlG complex in intact *Synechocystis* cells using the heterobifunctional cross-linking reagents LC-SDA and sulfo-LC-SDA

In the experiments detailed above (Section 5.3.2), the heterobifunctional water soluble sulfo-LC-SDA outperformed BS3 and DSS in terms of the number of detectable cross-linked peptides. Although the number of cross-links generated at concentrations of 20 mM and 45 mM were similar, the higher concentration was selected here. The rationale for this decision was (a) 45 mM sulfo-LC-SDA did not over-saturate the proteins with cross-links otherwise there would have been a decrease in cross-link detection and (b) to increase the probability of detection of cross-links involving the low abundance ChlG complex in a background of highly abundant phycobiliproteins and photosystem apoproteins. Furthermore, the use of the non-polar LC-SDA in DMSO in combination with B-PER reagent lead to cell leakage therefore, this reagent was utilised here in DMSO alone.

In vivo cross-linking experiments were performed as described in Section 2.10.1. No pigmentation of the of cell supernatants was observed following treatment with the cross-linking reagents, indicating that no cell leakage had occurred. The cells were lysed and the thylakoid membrane fraction prepared by differential centrifugation. Following solubilisation of the thylakoid membrane fraction FLAG-ChlG complex was isolated by FLAG-immunoprecipitation and digested with a combination of endoproteinase Lys-C and trypsin (described in Section 2.9.4). The peptide fragments were analysed by nano-LC-MS/MS (Section 2.9.5) and putative cross-linked peptides identified using the Byonic search engine and visual curation.

The following cross-links were detected: ChlG-ChlG (Figure 5.5A), ChlG-Ycf39 (Figure 5.5B), ChlG-YidC (Figure 5.5C), Ycf39-ChlG (Figure 5.5D), YidC-ChlG (Figure 5.5E), Ycf39-Ycf39, Ycf39-HliD, Ycf39-HliC, Ycf39-YidC and YidC-YidC. As was the case for the *in vitro* cross-linking experiments (Section 5.3.1) specific amino acid side chains involved in cross-linking could not be identified as a result of (a) idiosyncratic dissociation of cross-linked peptides during mass spectrometry and (b) the non-specific nature of diazirine reactivity. Nevertheless, it was possible to narrow down the cross-linked sites to short sequences within the peptides (Table 4). Multiple cross-links between amino acid side chains located in the N-terminal FLAG extension of ChlG and the central region of the cytoplasmic protein Ycf39 were identified. These are shown in Figure 5.8 (dashed blue and black lines). This is in agreement with the results obtained from the *in vitro* cross-linking experiments described in Section 5.3.1 and lends further support to the prediction that that the N-terminal membrane extrinsic domain of ChlG is located in the cytoplasm. Further cross-links were detected within the natural N-terminal cytoplasmic domain of ChlG with Ycf39, indicating that the cross-linking of ChlG and Ycf39 is not an artefact due to the presence of the FLAG-tag. A single cross-link was observed between the C-terminus of Ycf39 and predicted cytoplasmic extra-membranous loop of ChlG located between TMH2 and TMH3 (black dashed line). Taken together, the results indicate that the central and C-terminal regions of Ycf39 are in close proximity to the cytoplasmic N-terminal and extra-membranous TMH2-TMH3 domains of ChlG.

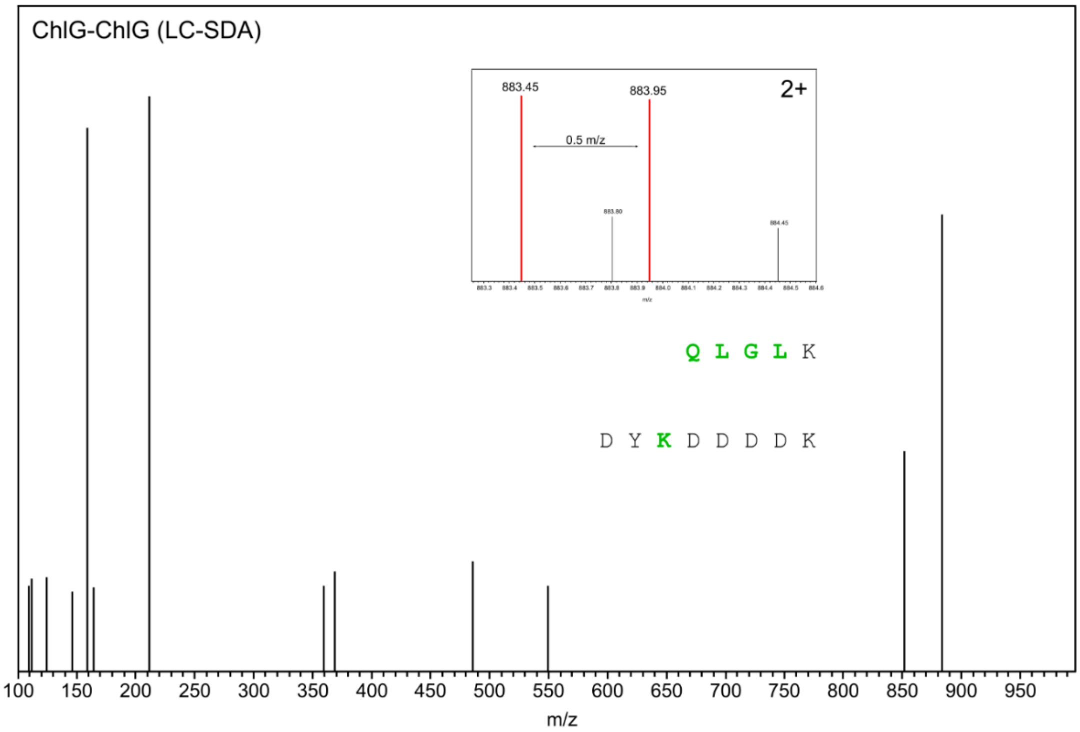
A cross-link was detected between Ycf39 and a domain of ChlG, TMH7-TMH8, predicted to be located in the thylakoid lumen (black dashed line). As Ycf39 is localised to the cytoplasmic side of the thylakoid lumen, this cross-link is improbable. A similar cross-link between the TMH7-TMH8 domain of ChlG and the predicted cytoplasmic domain of HliD (dashed green line) was also identified during *in vitro* cross-linking experiments, described above. This was explained by the potential for inversion of a sub-population of ChlG complexes during solubilisation in detergent. However, no detergent was present during the *in vivo* cross-linking experiment, highlighting the need for refinement of the structural model of ChlG.

Similarly, there was also evidence for an interaction between the ChlG cytoplasmic N-terminal FLAG extension and the predicted TMH5-TMH6 extra-membranous domain (dashed black line). This mirrored the results obtained during *in vitro* studies in which a cross-link was observed between the N-terminal cytoplasmic domain and the predicted TMH7-TMH8 extra-membranous region of ChlG (Section 5.3.1). As mentioned above, the improbability of this cross-link highlights the need for structural refinement of the ChlG model. As before, it was not possible to determine whether this cross-link occurred between two neighbouring ChlG proteins or within a single polypeptide. Taking the latter scenario, an alternative explanation can be considered in which the cross-link was formed during the co-translational assembly of the ChlG complex before the protein was inserted and orientated in the thylakoid membrane. It is possible that the prospective cytoplasmic and thylakoid luminal domains of the protein are located in close enough proximity during the process of co-translational membrane insertion for a cross-link to form between them. The possibility that FLAG immunoprecipitation can capture nascent complexes cannot be ruled out.

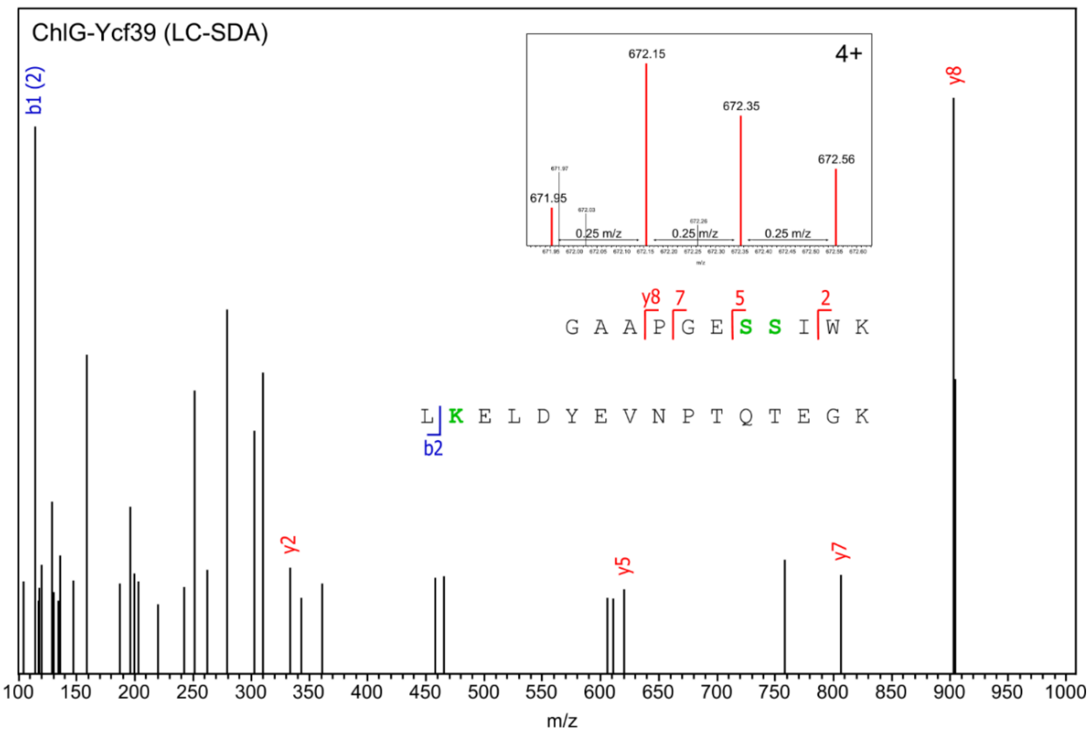
Table 4: *In vivo* LC-SDA and sulfo-LC-SDA cross-linked peptides. Analysis by mass spectrometry of a ChlG complex isolated from cells after in vivo chemical cross-linking with LC-SDA and sulfo-LC-SDA. Proteins 1-2 pertains to the left and right proteins in the pair respectively. Putative cross-linked residues are highlighted in bold red. PSM (peptide spectrum match) refers to the number of times a peptide was identified within the mass spectra. K indicates monolink modification of a K side chain. *Under normal circumstances, a cross-linked K would not be cleaved by trypsin. In the case of TGFLFI**K**, it is possible that the diazirine-derived carbene group added across a C-H bond in the K side chain instead of an N-H. Retention of the positive charge on the K epsilon amino group might allow trypsin cleavage at the residue. The evidence for this possibility is the clear presence of non-cross-linked a6 and b6 product ions, mapping to the T-I part of the sequence in the MS-MS spectra (see Figure 5.5E).

Cross-linked Proteins 1-2	Cross-linker	Peptide in protein 1	Peptide in protein 2	Calculated m/z	Observed m/z	PSM
ChlG-ChlG	LC-SDA	QLGLK	DYKDDDDK	883.4463	883.4473	1
ChlG-Ycf39	LC-SDA	DIDAINEPYR	KKK	601.6719	601.6731	1
ChlG-Ycf39	LC-SDA	GAAPGESSIWK	LKELDYEVNPTQTEGK	671.9520	671.9537	1
ChlG-YidC	LC-SDA	DYK	QEEIQKRYK	920.9938	920.9956	1
Ycf39-Ycf39	LC-SDA	KLK	AAFLKEWGATIVGGNICK	820.8064	820.8084	1
Ycf39-Ycf39	LC-SDA	SFTR	KK	490.3004	490.3010	1
Ycf39-Ycf39	LC-SDA	KKK	AAEYPKVPLMDIK	691.4111	691.4107	1
Ycf39-ChlG	LC-SDA	LNLIR	NPLENDVKYQASAQPFLVFGMLATGLALGHAG	1384.7486	1384.7525	1
Ycf39-ChlG	LC-SDA	IIK	MDYKDDDDKDYK	706.6819	706.6804	1
Ycf39-ChlG	LC-SDA	AGIK	MDYKDDDDKDYK	1173.5747	1173.5760	2
Ycf39-ChlG	LC-SDA	FAVR	AAASDTQNTGQNQAK	731.0416	731.0435	1
Ycf39-YidC	LC-SDA	KK	ATQGRESLPFEK	458.7624	458.7643	1
Ycf39-HliC	LC-SDA	SLR	MNNENSKFGFTAFAENWNGR	968.4680	968.4698	1
Ycf39-HliD	LC-SDA	KAAFLK	MSEELQPNQTPVQEDPKFGFNNYAEKLNGR	4547.3278	4547.3293	1
Ycf39-ChlG	sulfo-LC-SDA	IIK	DYKDDDDKDAAASDTQNTGQNQAK	1087.8755	1087.8743	1
Ycf39-ChlG	sulfo-LC-SDA	NCTEK	DDDDKDYKDDDDK	1302.0630	1320.0646	1
Ycf39-ChlG	sulfo-LC-SDA	AGIK	MDYKDDDDKDYK	1173.5747	1173.5767	2
Ycf39-ChlG	sulfo-LC-SDA	AGIKK	MDYKDDDK	618.9766	618.9753	1
Ycf39-ChlG	sulfo-LC-SDA	FVFFSILR	DDDDKAAAMSDTQNTGQNQAK	663.9284	663.9285	1
Ycf39-Ycf39	sulfo-LC-SDA	NCTEK	KAAFLK	756.4016	756.4005	1
Ycf39-YidC	sulfo-LC-SDA	SFTR	ATQGRESLPFEKK	439.8422	439.8428	1
YidC-ChlG	sulfo-LC-SDA	ITQPLMKER	ARQLLGMKGAAPGESSIWK	885.4926	885.4941	1
YidC-ChlG	sulfo-LC-SDA	SSK	MDYKDDDDKDYK	1131.0301	1131.0319	1
YidC-ChlG	sulfo-LC-SDA	TGFLFIK*	DDDDKDYK	678.3420	678.3463	1
YidC-ChlG	sulfo-LC-SDA	TGFLFIK*	NNENSKFGFTAFAENWNGR	645.3260	645.3273	1
YidC-YidC	sulfo-LC-SDA	LGLYPLSAGQIR	ESLPFEKKSSK	986.2301	986.2399	1
YidC-YidC	sulfo-LC-SDA	ITQPLMK	QQEEMAKVMK	749.4034	749.4073	2

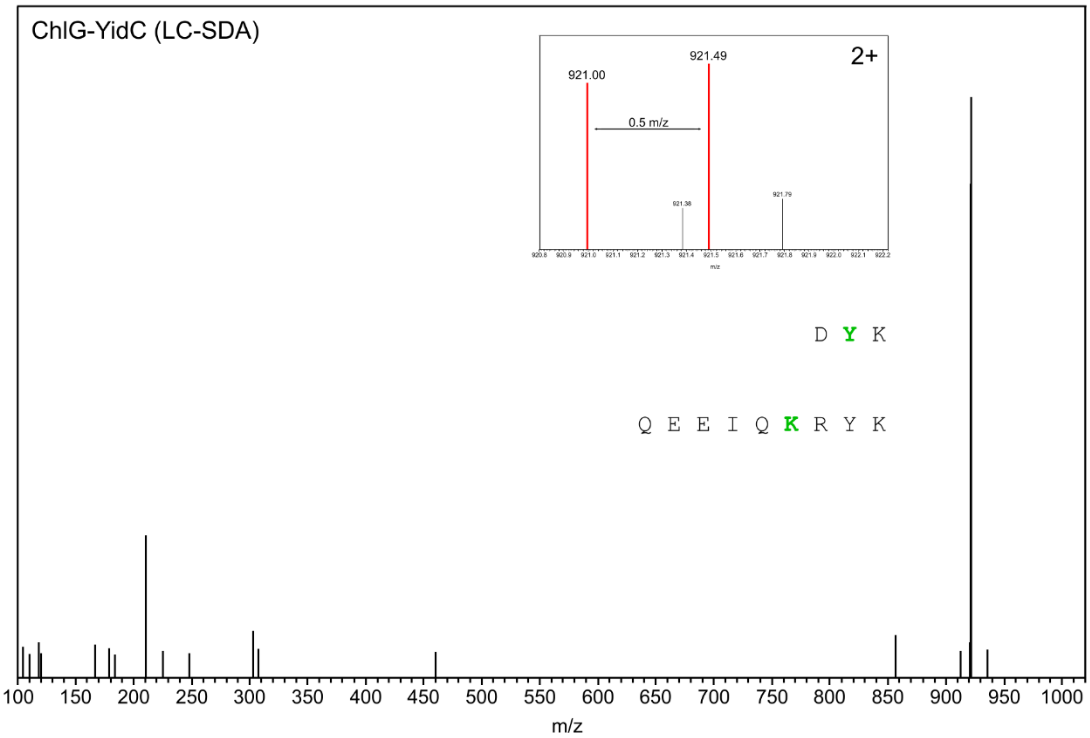
A


ChlG-ChlG (LC-SDA)
883.45
883.95
2+
0.5 m/z
Q L G L K
D Y K D D D D K
100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950
m/z

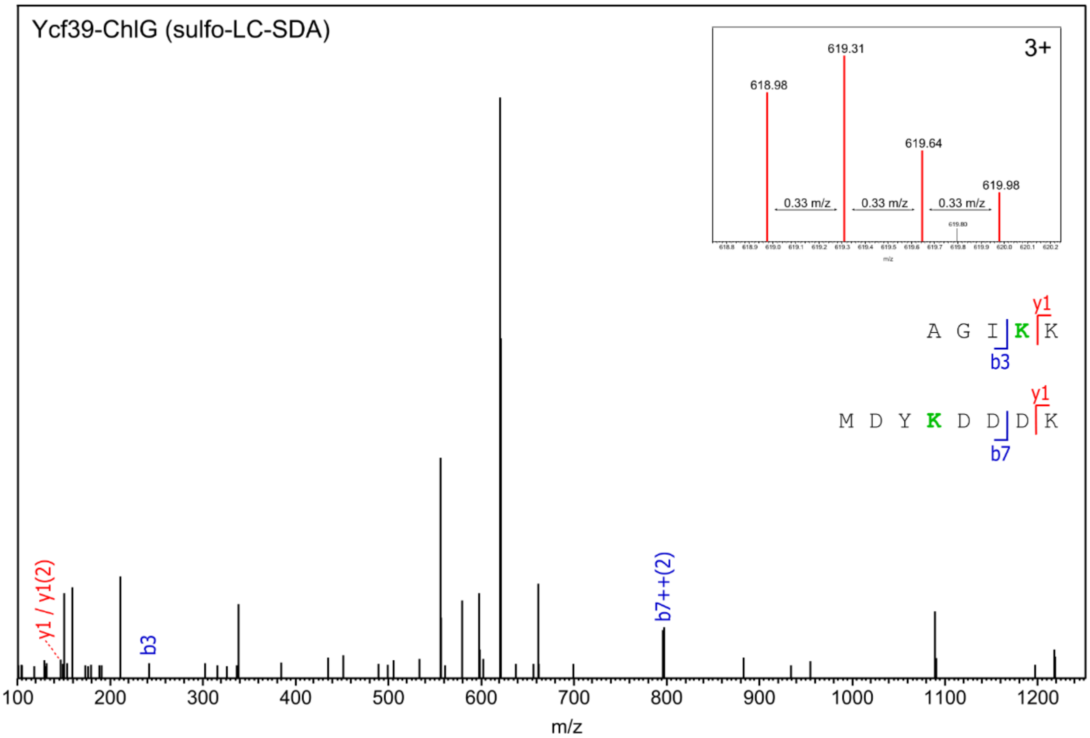
B


ChlG-Ycf39 (LC-SDA)
b1 (2)
672.15
672.35
672.56
671.95
4+
0.25 m/z
0.25 m/z
0.25 m/z
y8
G A A P G E S S I W K
y8 7 5 2
L K E L D Y E V N P T Q T E G K
b2
y2
y5
y7
100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000
m/z

C


ChlG-YidC (LC-SDA)
2+
921.00
921.49
0.5 m/z
D Y K
Q E E I Q K R Y K
m/z

D


Ycf39-ChlG (sulfo-LC-SDA)
3+
618.98
619.31
619.64
619.98
0.33 m/z
0.33 m/z
0.33 m/z
A G I K K
y1
b3
M D Y K D D D K
y1
b7
y1 / y1(2)
b3
b7++(2)
m/z

E

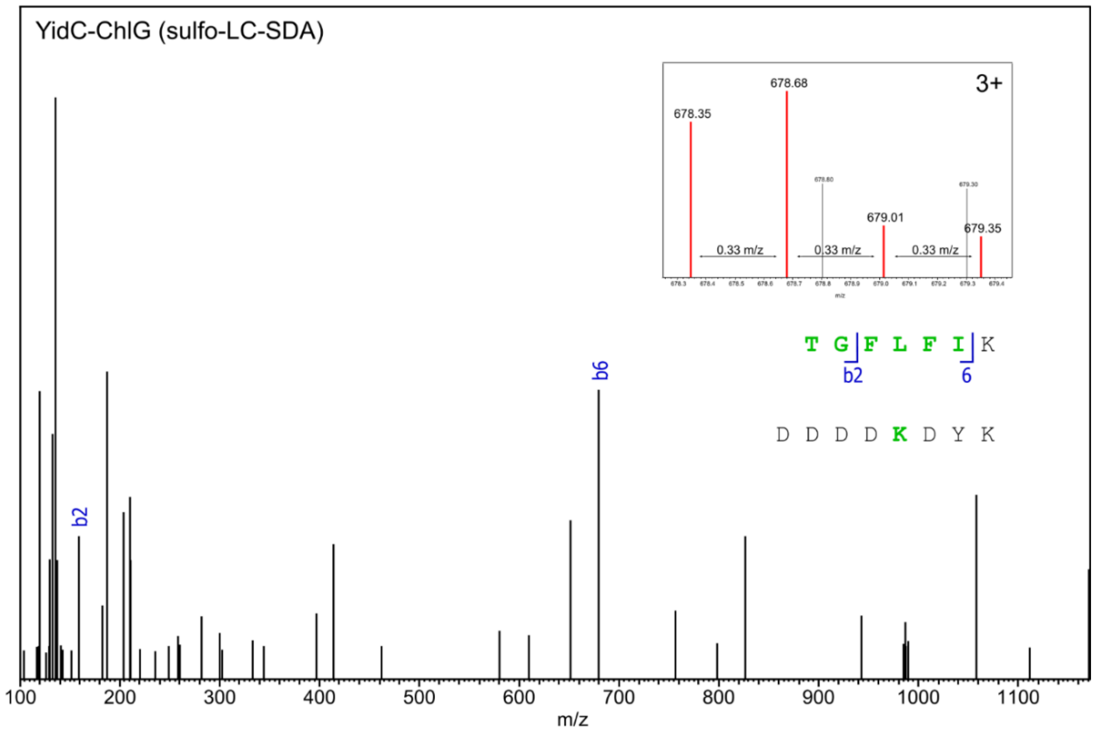


Figure 5.5: Analysis of FLAG-ChlG eluate cross-linked with LC-SDA and sulfo-LC-SDA *in vivo*, by mass spectrometry. Intact *Synechocystis* cells were treated with LC-SDA/DMSO and sulfo-LC-SDA/B-PER solutions before being analysed by nanoLC-MS/MS. Cross-linked peptides were identified using the search engine Byonic (Protein Metrics). MS/MS spectra representing (A) ChlG-ChlG, (B) ChlG-Ycf39, (C) ChlG-YidC, (D) Ycf39-ChlG and (E) YidC-ChlG among others (Table 4) were analysed to determine the sites of interaction between these proteins. Residues shown in bold green are potential cross-link sites. MS spectra (panels) were used to confirm the charge state and isotopomer peak distribution of the precursor ion selected for MS/MS.

5.3.4 *In vitro* cross-linking of the ChlG complex using the heterobifunctional cross-linking reagents LC-SDA and sulfo-LC-SDA

In vitro cross-linking experiments using the homobifunctional cross-linking reagents BS3 and DSS, and heterobifunctional LC-SDA at 0.88 mM are described in Section 5.3.1. In the light of the optimisation results for *in vivo* cross-linking (Section 5.3.2) the *in vitro* experiment was repeated using a higher concentration (45 mM) of both LC-SDA and sulfo-LC-SDA with an 8 L instead of 4 L cell culture, as described in Section 2.11.2.

In vitro cross-linking reactions, protein digestion and peptide identification were performed as described in Sections 2.10.2, 2.9.4 and 2.9.5. Cross-links were identified between the following: ChlG-ChlG (Figure 5.6A), ChlG-YidC (Figure 5.6B), Ycf39-ChlG (Figure 5.6C), Ycf39-Ycf39 and YidC-Ycf39, shown as blue and purple dashed lines in Figure 5.8. Amino acid side chains that were identified as cross-linked are shown in bold in the example spectra presented in Table 5.

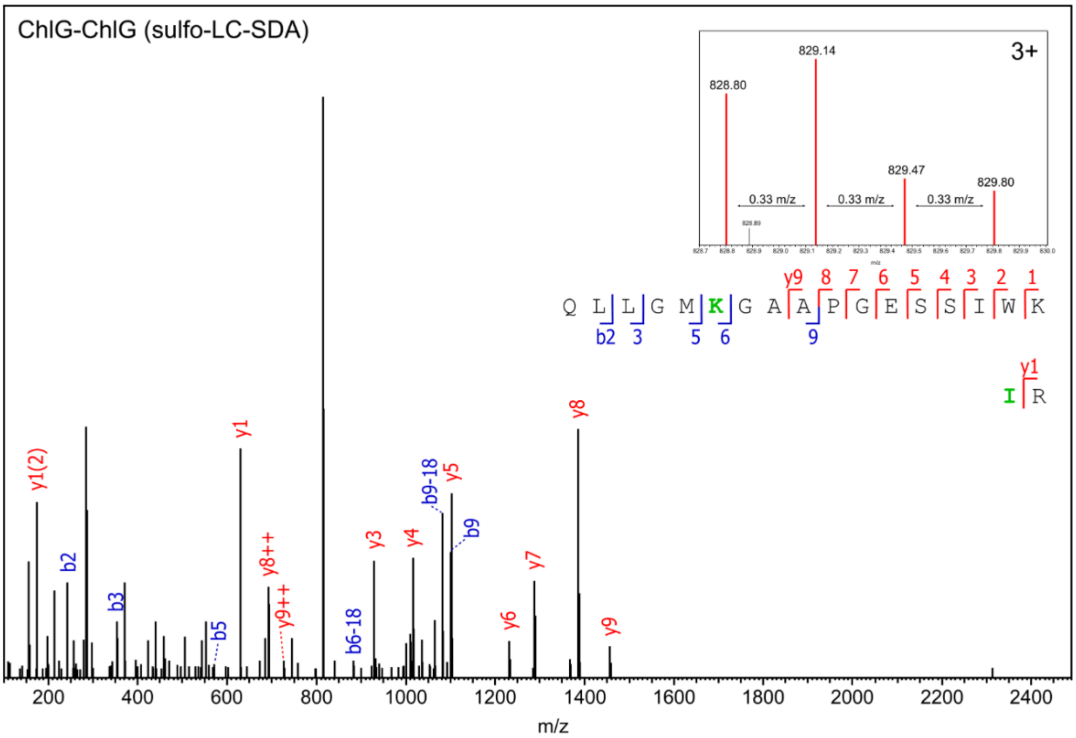
Two cross-links were identified between the central region of Ycf39 and the N-terminal FLAG extension of ChlG (blue dashed lines). These involved the same peptides previously identified in the *in vivo* LC-SDA/sulfo-LC-SDA cross-linking experiment described in Section 5.3.3, lending further support to the notion that Ycf39 is located in close proximity to the N-terminal cytoplasmic domain of ChlG. Likewise, a single cross-link was observed within this region of ChlG (purple dashed lines) which may be interpreted as either within one ChlG or between two neighbouring ChlGs, similar to the earlier results obtained with BS3 (Section 5.3.1). A cross-link was observed between the predicted TMH7-TMH8 thylakoid luminal domain of ChlG and the predicted cytoplasmic C-terminal domain of YidC (bold purple dashed line). Due to the spatial separation of these two domains proposed in the current model, this cross-link is unlikely to occur within a mature ChlG complex. However, as stated in Section 5.3.3, there is the possibility that this cross-link occurred during *de novo* insertion of ChlG into the thylakoid membrane. Alternatively, this cross-link may be an artefact of detergent solubilisation of the ChlG complex as described previously (Section 5.3.1).

A cross-link was identified between the central region of Ycf39 and the TMH2-TMH3 extra-membranous region of YidC predicted to be located in the thylakoid lumen (bold purple dashed line). Again this may be due to detergent solubilisation of the ChlG complex, generating a sub-population of molecules that are inverted within the detergent micelle and presenting an opportunity for cross-linking of protein domains that would be compartmentalised *in vi*vo.

Table 5: *In vitro* LC-SDA and sulfo-LC-SDA cross-linked peptides. Analysis by mass spectrometry of a ChlG complex isolated from 8 L of cells and after chemical cross-linking with LC-SDA and sulfo-LC-SDA. Proteins 1-2 pertains to the left and right proteins in the pair respectively Putative cross-linked residues are highlighted in bold red. PSM (peptide spectrum match) refers to the number of times a peptide was identified within the mass spectra. K indicates monolink modification of a K side chain. * Under normal circumstances, a cross-linked K would not be cleaved by trypsin. In the case of QLLGMKGAAPGESSIW**K**, it is possible that the diazirine-derived carbene group added across a C-H bond in the **K** side chain instead of an N-H. Retention of the positive charge on the K epsilon amino group might allow trypsin cleavage at the residue.

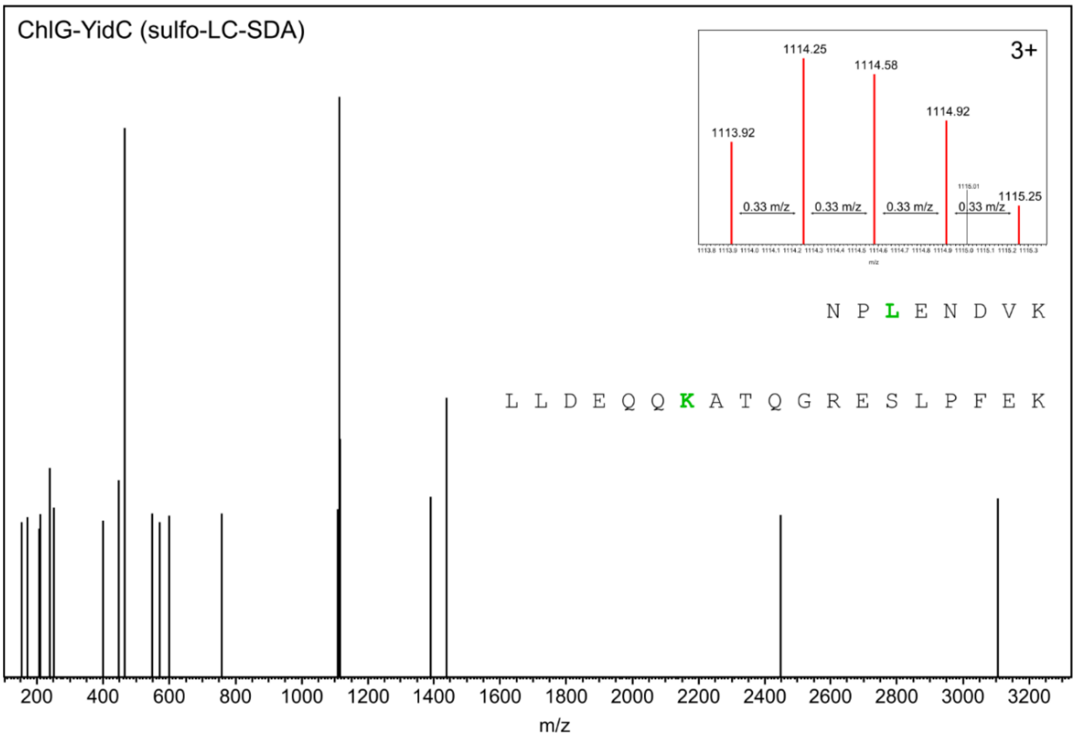
Cross-linked Proteins 1-2	Cross-linker	Peptide in protein 1	Peptide in protein 2	Calculated m/z	Observed m/z	PSM
Ycf39-ChlG	LC-SDA	**AGI**K	MDY**K**DDDD**K**DYK	1173.5747	1173.5766	2
Ycf39-ChlG	LC-SDA	**NCTE**K	DDDD**K**DY**K**DDDDK	1302.0630	1302.0660	1
Ycf39-Ycf39	LC-SDA	AA**EYP**K	AVQ**K**AGI**K**K	605.6961	605.6963	1
ChlG-ChlG	sulfo-LC-SDA	QLLGMKGAAPGESSIW**K***	**I**R	828.8013	828.8020	2
ChlG-YidC	sulfo-LC-SDA	**NPLENDV**K	LLDEQQ**K**ATQGRESLPFEK	1113.9192	1113.9182	1
Ycf39-Ycf39	sulfo-LC-SDA	L**NLI**R	**K**AAFL**K**EWGATIVGGNICK	707.9109	707.9119	2
Ycf39-Ycf39	sulfo-LC-SDA	**K**TYPVVGSR	**AV**ELDSVAR	2160.2023	2160.2051	1
Ycf39-YidC	sulfo-LC-SDA	**S**LR	EPLPENLQ**K**LLDEQQK	831.1304	831.1302	1
Ycf39-YidC	sulfo-LC-SDA	**FAV**R	NSQILTPTYSVT**K**GEDR	865.7967	865.7974	1
YidC-Ycf39	sulfo-LC-SDA	K**K**EK	**ELDYEVNPTQTEG**K	783.7443	783.7448	1

A



ChlG-ChlG (sulfo-LC-SDA)
3+
828.80
829.14
829.47
829.80
0.33 m/z
0.33 m/z
0.33 m/z
Q L L G M K G A A P G E S S I W K
y9 8 7 6 5 4 3 2 1
b2 3 5 6 9
I R
y1
y1(2)
b2
b3
b5
y1
y8++
y9++
b6-18
y3
y4
b9-18
y5
b9
y6
y7
y8
y9
200 400 600 800 1000 1200 1400 1600 1800 2000 2200 2400
m/z

B



ChlG-YidC (sulfo-LC-SDA)
3+
1113.92
1114.25
1114.58
1114.92
1115.25
0.33 m/z
0.33 m/z
0.33 m/z
0.33 m/z
N P L E N D V K
L L D E Q Q K A T Q G R E S L P F E K
200 400 600 800 1000 1200 1400 1600 1800 2000 2200 2400 2600 2800 3000 3200
m/z

C

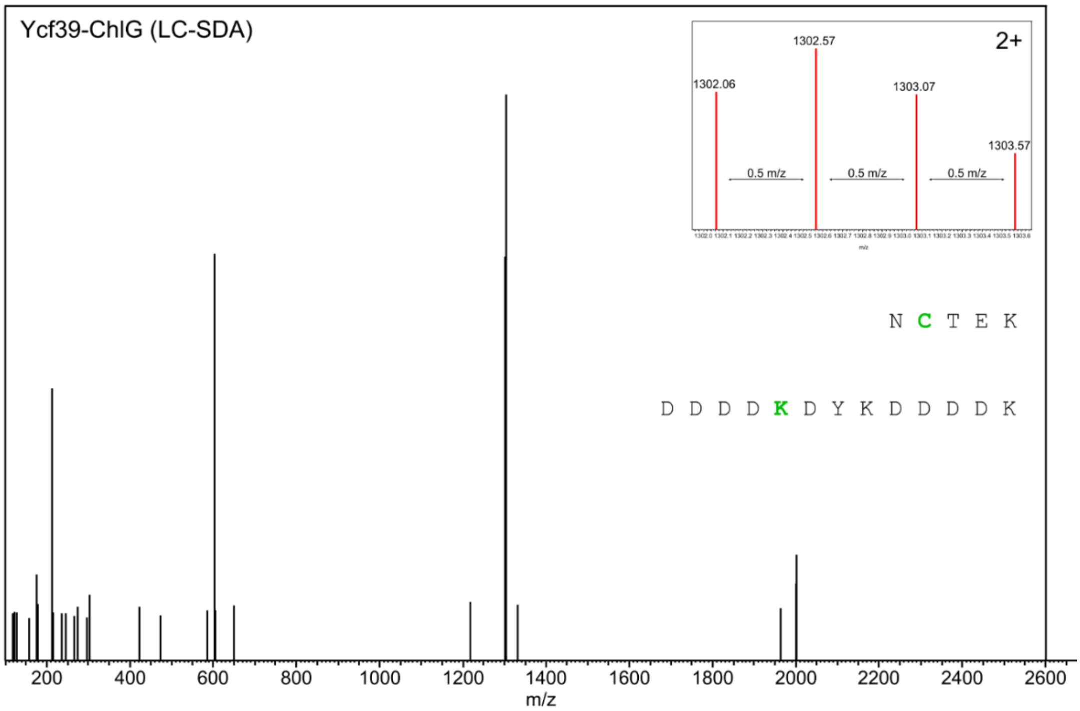


Figure 5.6: Analysis of FLAG-ChlG eluate, cross-linked with LC-SDA and sulfo-LC-SDA *in vitro*, by mass spectrometry. FLAG-ChlG was purified from 8 L of *Synechocystis* cells and treated with LC-SDA and sulfo-LC-SDA cross-linking reagents before being analysed by nanoLC-MS/MS. Cross-linked peptides were identified using the search engine Byonic (Protein Metrics). MS/MS spectra representing (A-B) ChlG-ChlG, (C-D) ChlG-YidC and (E) Ycf39-ChlG among others (Table 5) were analysed to determine the sites of interaction between these proteins. Residues shown in bold green are potential cross-link sites. MS spectra (panels) were used to confirm the charge state and isotopomer peak distribution of the precursor ion selected for MS/MS.

5.3.5 Cross-linking of the *Synechocystis* thylakoid membrane fraction using sulfo-LC-SDA

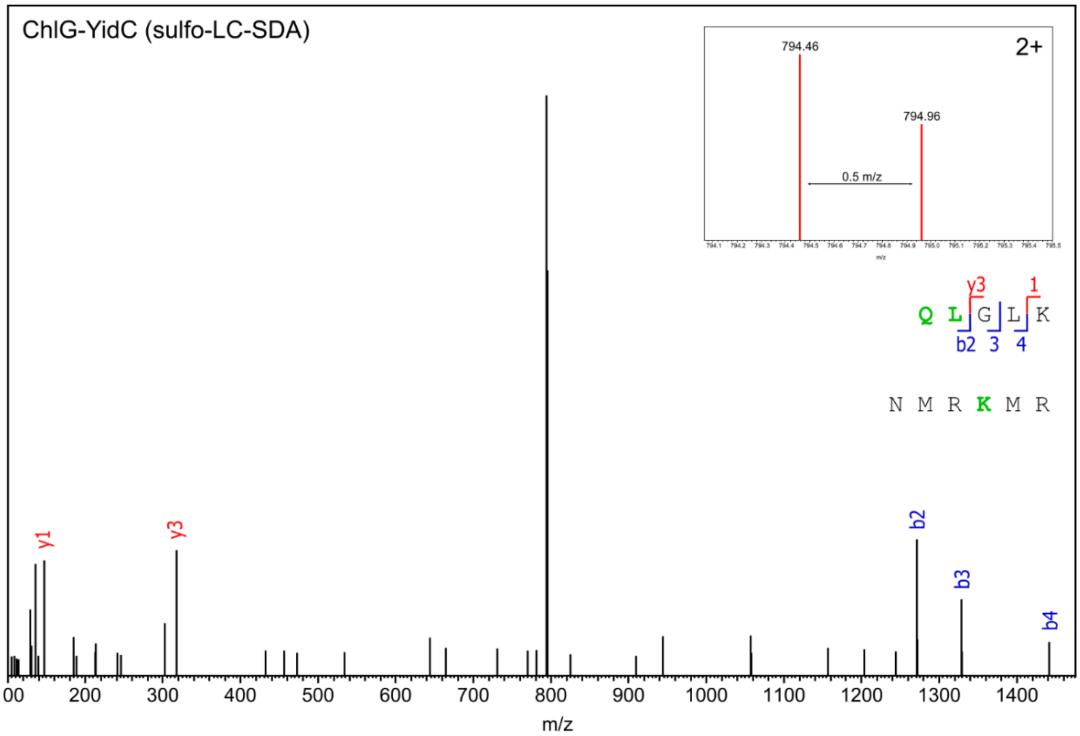
Throughout the series of cross-linking experiments performed so far, the most effective cross-linking reagent was sulfo-LC-SDA, which repeatedly produced the highest yield of putative cross-linked peptides. However, the number of cross-linked peptides mapping to the ChlG complex remained lower than expected in both *in vitro* and *in vivo* experiments. The yields of cross-linked peptides produced *in vivo* is most likely limited by the low abundance of the ChlG complex in comparison to other proteins which effectively dilute the cross-linker. A potential disadvantage of the *in vitro* cross-linking approach was the need for detergent solubilisation which may perturb the orientation of the complexes. A third option, which would eliminate the highly abundant phycobiliproteins and the use of detergent, was to isolate the thylakoid membrane fraction and add cross-linking reagent before detergent solubilisation and FLAG-immunoprecipition. The thylakoid membrane fraction was prepared from 4 L of *Synechocystis* FLAG-*chlG* Δ*chlG* cells (Sections 2.11.1) and treated with sulfo-LC-SDA according to the method described in Section 2.10.3. FLAG-ChlG was immunoprecipitated from the thylakoid membranes and the extracted proteins digested and analysed as detailed Sections 2.9.4 and 2.9.5.

Cross-links were identified between the following: ChlG-YidC (Figure 5.7A), Ycf39-ChlG (Figure 5.7B), YidC-ChlG (Figure 5.7C), Ycf39-Ycf39, Ycf39-YidC and YidC-YidC shown as yellow dashed lines in Figure 5.8. The peptides involved in these cross-linking events are shown in Table 6. A cross-link was detected between the TMH5-TMH6 thylakoid luminal domain of ChlG and the TMH1-TMH2 cytoplasmic domain of YidC. The conflicting nature of this cross-link may be due to reasons previously discussed: (a) that the current structural models require refinement or (b) interactions between prospective cytoplasmic and thylakoid luminal domains might be feasible in a nascent complex.

Table 6: Sulfo-LC-SDA cross-linked peptides within thylakoid membranes. Analysis by mass spectrometry of a ChlG complex after chemical cross-linking of purified thylakoid membranes with LC-SDA and sulfo-LC-SDA. Proteins 1-2 pertains to the left and right proteins in the pair respectively. Putative cross-linked residues are highlighted in bold red. PSM (peptide spectrum match) refers to the number of times a peptide was identified within the mass spectra.

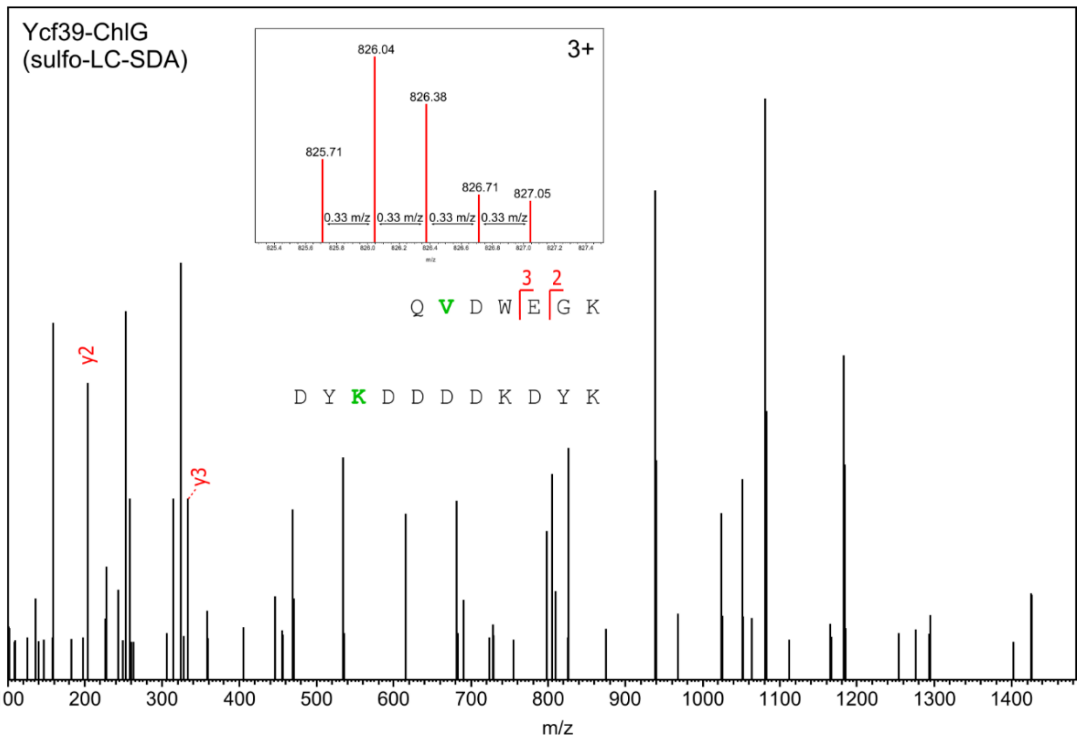
Cross-linked Proteins 1-2	Cross-linker	Peptide in protein 1	Peptide in protein 2	Calculated m/z	Observed m/z	PSM
ChlG-YidC	sulfo-LC-SDA	**QL**GLK	NMR**K**MR	794.4629	794.4590	1
Ycf39-ChlG	sulfo-LC-SDA	**QVDW**EGK	DY**K**DDDD**K**DYK	825.7083	825.7114	1
Ycf39-Ycf39	sulfo-LC-SDA	**K**LK	**AAFLKEWGATIVGGNIC**K	615.8566	615.8571	1
Ycf39-Ycf39	sulfo-LC-SDA	**K**TYPVVGSR	**LKELDYEVNPT**QTEGK	766.9105	766.9087	1
Ycf39-YidC	sulfo-LC-SDA	**S**LR	ESLPFE**K**K	516.2995	516.2999	1
YidC-ChlG	sulfo-LC-SDA	ITQ**PLM**K	GAAPGESSIW**K**IR	799.4505	799.4446	1
YidC-Ycf39	sulfo-LC-SDA	ITQ**PLM**K	LAF**S**EVLA**S**GK	1180.1814	1180.1792	1
YidC-Ycf39	sulfo-LC-SDA	ESLPFE**K**K	**ATDSLTI**R	683.3806	683.3780	1
YidC-YidC	sulfo-LC-SDA	ESLPF**E**K	ERQEEIQ**K**R	753.7375	753.7341	1

A



ChlG-YidC (sulfo-LC-SDA)
794.46
794.96
0.5 m/z
2+
Q L G L K
y3
1
b2 3 4
N M R K M R
y1
y3
b2
b3
b4
m/z

B



Ycf39-ChlG
(sulfo-LC-SDA)
826.04
826.38
825.71
826.71
827.05
0.33 m/z
3+
Q V D W E G K
3 2
D Y K D D D D K D Y K
y2
y3
m/z

C

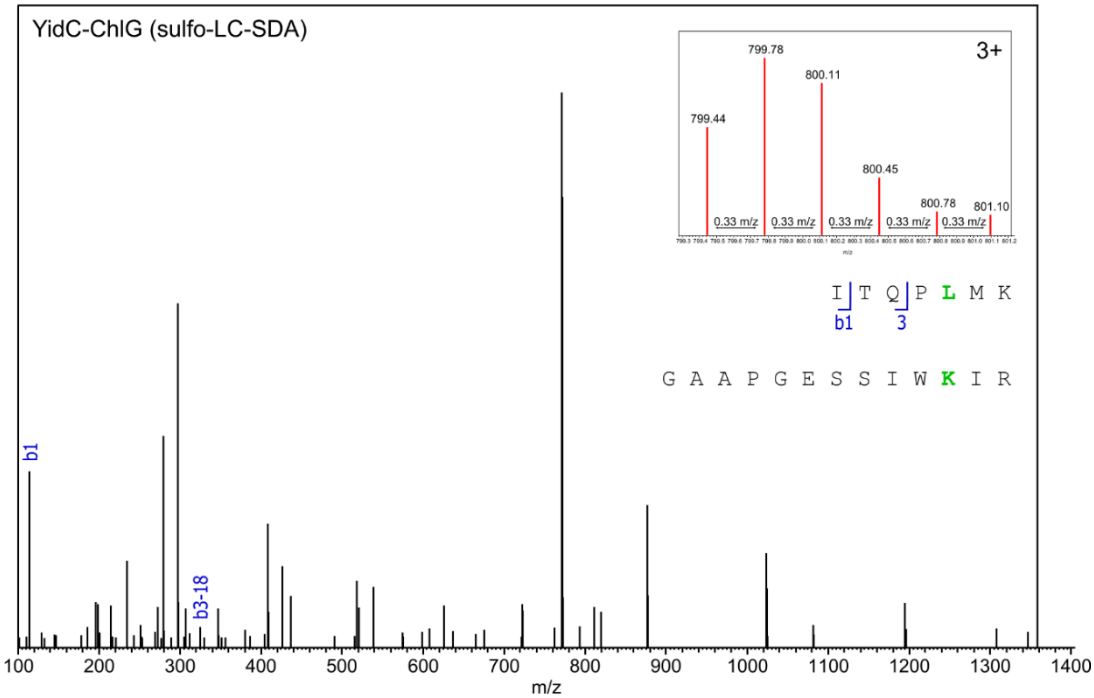


Figure 5.7: Analysis of FLAG-ChlG, isolated from thylakoid membranes cross-linked with sulfo-LC-SDA, by mass spectrometry. FLAG-ChlG was purified from 8 L of *Synechocystis* cells and treated with LC-SDA and sulfo-LC-SDA cross-linking reagents before being analysed by nanoLC-MS/MS. Cross-linked peptides were identified using the search engine Byonic (Protein Metrics). MS/MS spectra representing (A) ChlG-YidC, (B) Ycf39-ChlG and (D) YidC-ChlG among others (Table 6) were analysed to determine the sites of interaction between these proteins. Residues shown in bold green are potential cross-link sites. MS spectra (panels) were used to confirm the charge state and isotopomer peak distribution of the precursor ion selected for MS/MS.

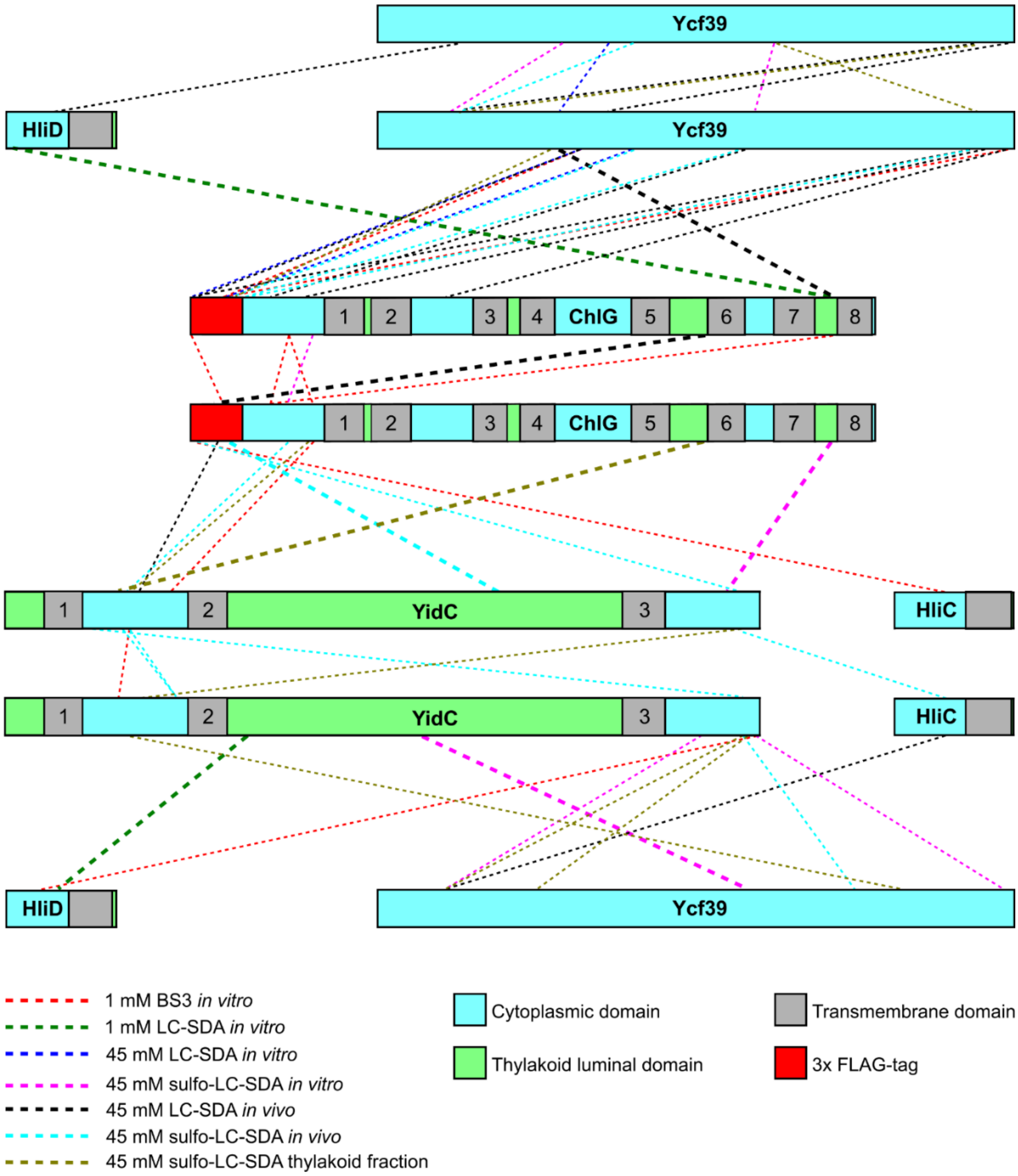


Figure 5.8: Diagram of the ChlG complex showing total cross-links identified in each experiment described in this study. ChlG, YidC, HliD and HliC were modelled using TMHMM to determine the number and location of transmembrane regions (grey) within each protein sequence. Green and blue represent thylakoid luminal and cytoplasmic domains respectively. Putative cross-links identified throughout the experiments described in this study are represented as coloured dashed lines. Bold lines represent cross-links which are inconsistent with the models shown and are explained in the text.

5.4 Discussion

5.4.1 Model of the ChlG complex

In each of the *in vivo* and *in vitro* cross-linking experiments presented in this study, 12 putative cross-links were detected between the N-terminal domain of ChlG, including the FLAG-tag extension, and the central and C-terminal regions of Ycf39 (Figure 5.8). Ycf39 is an atypical short-chain dehydrogenase associated with the cytoplasmic side of the thylakoid membrane in *Synechocystis* (Knoppová *et al*., 2014). The cross-linking results therefore suggest that the N-terminus of ChlG is located within the cytoplasm in close proximity to Ycf39. The function of Ycf39 within the ChlG complex has not yet been elucidated, however it has been shown to form a complex with HliD and HliC and bind to the early PSII assembly intermediate pD1, aiding its biosynthesis (Knoppová *et al*., 2014). In this study, one cross-link between the extra-membranous domain of HliD and Ycf39 and a second between the extra-membranous domain of HliC and Ycf39 were formed *in vivo* (Figure 5.8). These cross-links can be used to inform a model of HliC and HliD in which they are orientated within the thylakoid membrane such that their N-terminal membrane extrinsic domains are situated in the cytoplasm, enabling the interaction of these proteins with Ycf39. Interestingly, both HliC and HliD were found to cross-link to the same N-terminal region of Ycf39, whereas ChlG-Ycf39 cross-links appeared to be confined to the central and C-terminal regions of Ycf39, indicating that the interaction sites of the Hlips and ChlG with Ycf39 occur at different locations.

HliC and HliD also formed one cross-link each with ChlG in two different experiments. The predicted cytoplasmic domain of HliC cross-linked with the N-terminal cytoplasmic domain of ChlG on the FLAG-tag extension *in vivo* (Figure 5.8). The predicted N-terminal cytoplasmic domain of HliD, however, was shown to form a cross-link to the TMH7-TMH8 thylakoid luminal domain of ChlG, in direct conflict with the HliD-Ycf39 result (Figure 5.8). As the HliD-ChlG cross-link was formed *in vitro*, it was possible that either HliD or ChlG had become inverted in the detergent micelle, presenting an opportunity for a cross-link to from between two regions of each protein that would otherwise be compartmentalised *in vivo*. For this reason, the cross-link was not used to inform the model of the ChlG complex presented in Figure 5.9.

YidC has been hypothesised to mediate the partitioning of the transmembrane segments of chlorophyll-binding proteins into the thylakoid membrane lipid bilayer as *de novo* chlorophyll molecules are being delivered by nearby ChlG (Chidgey et al., 2014). In this work the N-terminal FLAG extension and natural N-terminal cytoplasmic domain of ChlG were shown to form cross-links with the predicted TMH1-TMH2 cytoplasmic domain of YidC, providing further evidence for the close association of these two proteins (Figure 5.8). One cross-link was detected within this domain of YidC and Ycf39, however the majority of the cross-links between these two proteins occurred at the predicted C-terminal cytoplasmic domain of YidC. The C-terminal domain of YidC was also shown to form a single cross-link to both HliD and HliC within the cytoplasmic domains of the latter proteins (Figure 5.8). Taken together, these cross-links can be used to inform the model of the ChlG complex, presented in Figure 5.9.

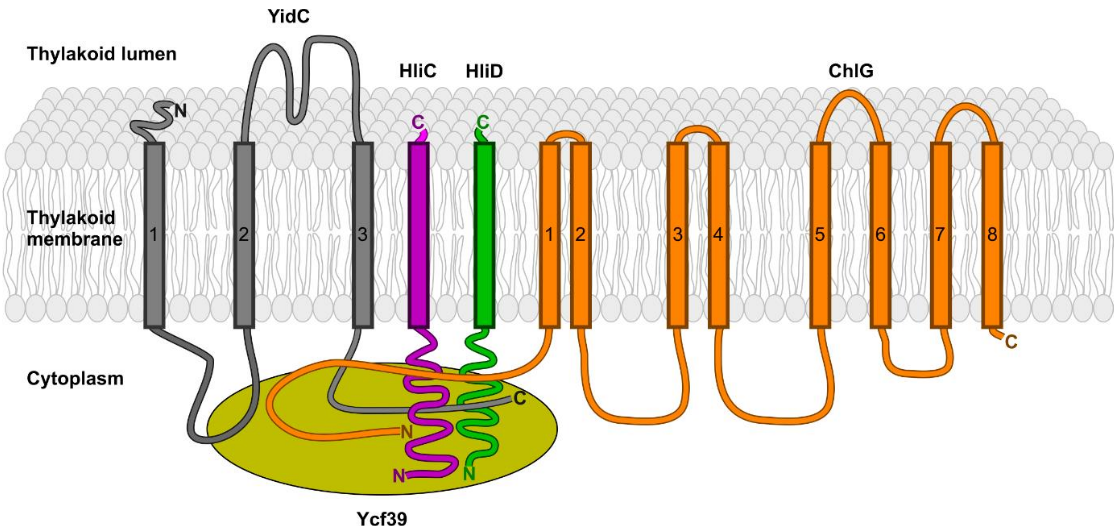


Figure 5.9: 2D model of the ChlG-HliD-HliC-YidC-Ycf39 complex. The total cross-linking data obtained from the experiments described in this study were used to identify regions of interaction between members of the ChlG complex. This information was used to inform the construction of a 2D model of the complex.

5.4.2 Possibilities for future refinement of the ChlG complex model

The cross-linking studies performed in this work were used to inform a model of the ChlG complex (Figure 5.9). One limitation was the relatively low abundance of the ChlG complex within *Synechocystis* which resulted in low cross-linked peptide yields for the target proteins in a background of highly abundant phycobiliproteins (Section 5.3.2). Simply increasing the quantity of material used for cross-linking is not a viable solution because the maximum quantity of total peptides that can be injected onto the nanoLC-MS/MS system is 500 ng. Enrichment of the ChlG complex by FLAG-immunoprecipitation both before and after treatment with cross-linking reagents did compensate for the limitation of low yields of ChlG complex by increasing the proportion of target peptides within the 500 ng analyses (Sections 5.3.3 and 5.3.4). Nevertheless, the cross-linked peptides were only a minor proportion of the total population of unmodified ChlG complex derived peptides. Therefore, further increases in the number and spectral intensities of ions representing cross-linked peptides are required in order to enhance the accuracy of this model. Several approaches have been developed to aid the MS detection of cross-linked peptides within complex mixtures, including specific enrichment of cross-linked peptides and use of isotope (Petrotchenko *et al*., 2005), fluorescent (Sinz and Wang, 2004) or mass-tag (Back *et al*., 2003) labelled cross-linking reagents. Affinity tagged cross-linkers are in development that would enable the specific purification of cross-linked peptides before analysis by MS. Prominent amongst these are biotinylated cross-linkers, of which there are several examples available within the literature e.g. (Trester-Zedlitz *et al*., 2003). Biotin is often selected as the affinity tag owing to its high affinity and small size. Affinity tags have also been used in tandem with other labelling techniques, such as isotope labelling (Chu *et al*., 2006), to enhance the detection and analysis of cross-linked peptides by MS. Although biotinylated cross-linking reagents are not yet commercially available at the time of this study, the strategies discussed above could be employed in future cross-linking experiments on the ChlG complex.

A second challenge encountered during this study was the inefficient and/or idiosyncratic dissociation of peptide ions in the mass spectrometer during product ion scanning, resulting in the inability to unambiguously determine the specific cross-link sites between peptides. Several novel cross-linker reagents have been developed to facilitate the interpretation of product ion spectra derived from cross-linked peptides. One such cross-linker, disuccinimidyl sulfoxide (DSSO), has recently become commercially available (Thermo Fisher Scientific). This reagent is membrane permeable in DMSO and contains an NHS ester at each end of a 10.3 Å spacer which, unlike DSS, is cleavable by collision induced dissociation (CID) during MS2 scanning. This results in distinctive ion doublets detectable in the product ion spectrum that can be used to confirm cross-link modification and the type of cross-link (monolink, intralink or interlink) based on characteristic dissociation patterns pertaining to each scenario. Additionally, DSSO interlinked peptides identified as doublet ions by MS2 are automatically selected for MS3 analysis, resulting in precise identification of the cross-linked residues (Kao *et al*., 2011). DSSO has been used to characterise both the yeast and human 20 S proteasome using *in vitro* and, in the latter case, *in vivo* approaches (Kao *et al*., 2011; Wang *et al*., 2017). Future studies involving cross-linking of the ChlG complex with DSSO in combination with MS3 analysis should enable more precise identification of cross-linked sites.

5.5 Future Work

Specific enrichment of cross-linked ChlG is required to increase the concentration of cross-linked proteins above the background. This could be achieved by treatment of isolated FLAG-ChlG complex with biotinylated cross-linking reagents followed by targeted isolation of cross-linked peptides using this affinity tag. Alternatively, cross-linked proteins could be separated by SDS-PAGE and ChlG containing complexes identified by immunoblotting. These bands could then be excised from the SDS gel and the proteins subjected to in-gel tryptic digestion before analysis by mass spectrometry.

Treatment of isolated FLAG-ChlG complex with MS cleavable cross-linking reagent DSSO (Thermo Fisher Scientific) in future experiments would enable MS3 analysis of the cross-linked ChlG complex. This would potentially yield more informative product ions that would reveal the precise amino acid residues involved in cross-linking.

In this study, many cross-links were detected that involved the FLAG-tag extension of ChlG, both *in vivo* and *in vitro*, demonstrating the potential for artificial association of proteins with this region (Figure 5.8). Although this seems unlikely given that the N-terminally FLAG-tagged *Chlamydomonas reinhardtii* and *Arabidopsis thaliana* ChlG proteins did not co-purify with HliD or Ycf39 when heterlogously produced in *Synechcosytis* (Proctor *et al*., 2018) (Chapter 3); this possibility could be further explored by immunoprecipitation of an untagged ChlG complex using antibodies raised against ChlG followed by chemical cross-linking.

Chapter 6: Truncations of *Synechocystis* chlorophyll synthase reveal that the N-terminus is important for enzyme activity but is not required for interaction with YidC, HliD or Ycf39

6.1 Summary

Chlorophyll synthase (ChlG) is the terminal enzyme of chlorophyll biosynthesis in oxygenic photosynthetic organisms. This enzyme forms a complex with high-light inducible protein D (HliD), Ycf39 and the membrane insertase YidC in order to coordinate Chl delivery with the co-translational insertion of Chl-binding proteins into the thylakoid membrane. From preliminary cross-linking data presented in Chapter 5, the extreme N-terminus of ChlG was predicted to be involved in the binding of these proteins, particularly Ycf39 due to the fact that a large number of cross-links were found between this protein and the N-terminus of ChlG across multiple cross-linking experiments. To investigate this prediction, mutant ChlG proteins with truncated N-termini were produced in *Synechocystis*. The truncated proteins were used as the basis for immunoprecipitation experiments which revealed that in each case the truncated enzyme was able to maintain an interaction with HliD, YidC and Ycf39, demonstrating that the N-terminus is not essential for mediating the formation of the ChlG complex. Deletion of the native *chlG*, an essential gene, was not possible from the strains containing FLAG-ChlG proteins lacking 32 or more residues from the N-terminus, indicating that these proteins were not active *in vivo*. This was confirmed by *in vitro* enzyme activity assays. A C-terminally FLAG-tagged ChlG (ChlG-FLAG) was generated as a control to see if a complex could be purified that resembled N-terminally tagged protein. Although the native *chlG* gene could not be deleted from this strain, indicating that the ChlG-FLAG enzyme was not able to complement the function of the native protein *in vivo*, the tagged enzyme produced Chl a_{GG} during *in vitro* enzyme assays. This protein did not co-purify with HliD or Ycf39 but retained an association with YidC and could form a dimer with a second ChlG molecule. The enzyme activity observed is likely due to co-purification of the native ChlG enzyme with the ChlG-FLAG variant.

6.2 Introduction

In oxygenic photosynthetic organisms, the chlorophyll synthase (ChlG) enzyme catalyses the esterification of chlorophyllide (Chlide) with geranylgeranyl pyrophosphate (GGPP) producing a mature chlorophyll (Chl) molecule. In the model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*), ChlG forms a complex with the high-light inducible proteins C and D (HliC and HliD) as well as the membrane insertase YidC and atypical alcohol dehydrogenase Ycf39 (Chidgey et al., 2014).

There is evidence to suggest that the ChlG-HliC/D-Ycf39-YidC complex is exclusive to cyanobacteria. Heterologous production and purification of plant and algal ChlG proteins in *Synechocystis* showed that only YidC remained bound to the eukaryotic homologs, whereas HliD and Ycf39 only co-purified with cyanobacterial ChlG (Proctor *et al*., 2018) (Chapter 3). These results demonstrate the important nature of the ChlG-YidC association and indicate that the interaction site could be conserved amongst eukaryotic and prokaryotic ChlG homologs.

Preliminary results from cross-linking experiments (discussed in Chapter 5) suggested that the ChlG-Ycf39, ChlG-YidC and ChlG-ChlG interactions involve the extreme N-terminus of ChlG. In particular, Ycf39, a thylakoid membrane associated protein that localises to the cytoplasmic side of the thylakoid membrane (Knoppová *et al*., 2014), was shown to form multiple cross-links with the N-terminus of ChlG across multiple cross-linking experiments that utilised various cross-linking reagents. This prompted selection of the ChlG N-terminus for the functional studies discussed in this chapter.

A structural model of *Synechocystis* ChlG was generated using the crystal structure of a related protein, ubiquinone synthase (UbiA) (Cheng and Li, 2014), as a template (Figure 6.1B). The N-terminus of the ChlG model appeared unstructured as it could not be mapped to regions of homology within UbiA. Sequence alignments of ChlG homologs with UbiA showed that the N-terminal tail of ChlG is extended in comparison to UbiA (Figure 6.1A). Modelling software (QUARK) was used to predict possible secondary structure configurations for the ChlG N-terminus (Figure 6.1C). The top four predictions all contained two alpha helices. From the cross-linking results presented in Section 5.3.1, it is apparent the N-terminal domain of ChlG is located within the cytoplasm, therefore these alpha helices may be situated on the cytoplasmic side of the thylakoid membrane and facilitate the interaction of ChlG with its partners. Unlike ChlG, UbiA has not been reported to interact with any other protein partners.

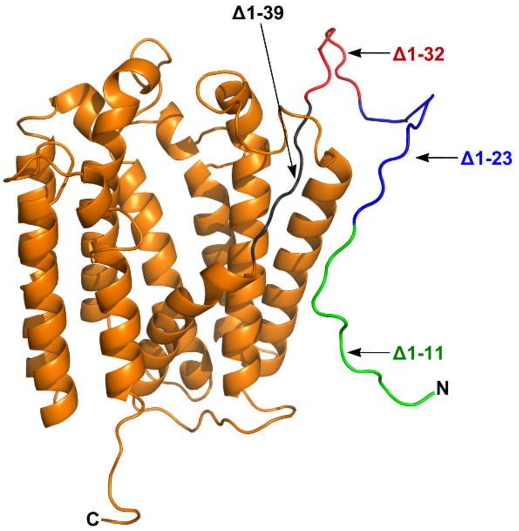
The N-terminus of the ChlG homologue in oat (*Avena sativa*) has been previously targeted for mutagenesis studies. The authors made sequential truncations to the N-terminus of the enzyme and heterologously produced them in *Escherichia coli*. *In vitro* enzyme assays using cell lysates revealed that removal of the first 89 residues abolished the activity of the protein (Schmid *et al*., 2001). However, the importance of the ChlG N-terminus *in vivo* has not been investigated.

In this study, sequential truncations were made to the N-terminus of a FLAG-tagged ChlG protein in *Synechocystis*. These truncated proteins were used as the basis for immunoprecipitation experiments and the resulting eluates analysed by immunoblot for the presence of HliD, Ycf39 and YidC. The activities of these mutant proteins were examined *in vivo*, by attempted deletion of the native *chlG* gene, and *in vitro* using the FLAG pulldown eluates for enzyme activity assays. The truncated ChlG proteins maintained an interaction with all components of the native ChlG complex although the activity of the mutants was impaired as the truncations became progressively larger.

A

```
ChlG        MSDTQNTGQNQAKARQLLGMKGAAPGESSIWKIRLQLM-KPITWIPLIWGVVCGAASSGG 59
UbiA        --------------------------------MRLVRIEHTIFSLPFAY--V-G-ALLSR 24
                                            :**  : : *  :*: :  * * *  .

ChlG        YIWSVEDFLKALTCMLLSGPLMTGYTQTLNDFYDRDIDAINE--PYRPIPSGAISVPQVV 117
UbiA        YPFTLADAIL--MAAAVVGLRMAGM--AYNNIADLDIDRLNPRTAKRPLVVGAVSLREAW 80
            * ::: * :    .  : *  *:*   : *:: * *** :*     **:  **:*: :.

ChlG        TQILILLVAGIGVAYG----LDVWAQHDFPIMMVLTLGGAFVAYIYSAPPLKLKQNGWLG 173
UbiA        ----ALVAAGSAIYFASAALLNTYALLLSPLVLAIALTYPHAKRLHPLPHLH--LGIVLG 134
                 *:.** .: :.    *:.:*    *:::.::*   ..  ::  * *:   .  **

ChlG        --------------NYALGASYIALPWWAGHALFGTLNPTIMVLTLIYSLAGLGIAVVND 219
UbiA        SVVFGGAVAASGDEASSLGEVLRSVPWLYVAAV---------------SLWVAGFDTIYS 179
                            :**    ::**    *:               **   *: .: .

ChlG        FKSVEGDRQLGLKSLPVMFGIGTAAWICV----IMIDVFQAGIAGYLIYVHQQLYATIVL 275
UbiA        IMDIDFDRSHGLGSIPALLGPKGALAASLAMHAAAVALFIAGVEAYGLGAIATVSTALTA 239
            : .:: **. ** *:*.::*   *   .:      : :* **: .* : .   : :::.

ChlG        LLLIPQITFQDMYFLRNPLENDVKYQASAQPFLVFGMLATGLALG--HAGI---       324
UbiA        LV---IILVQAMAWLGRVKE-SFNLNL-AVPII----IGAGIIVDMLHHMIRLL       284
            *:    * .* * :* .  * ..: :  * *::    :.:*: :.  *  *
```

B



C

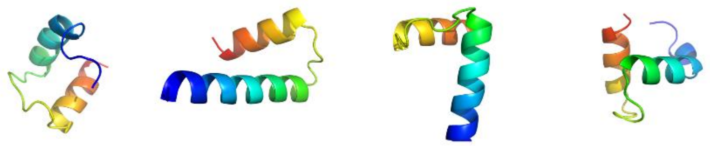


Figure 6.1: Sequence alignments and structural model of *Synechocystis* ChlG. (A) The amino acid sequences of ChlG from *Synechocystis* and UbiA from *Aeropyrum pernix were* aligned using Clustal Omega software. (B) A structural model of *Synechocystis* ChlG was generated using HHPred software and the published crystal structure of *Aeropyrum pernix* UbiA as a template. The extreme N-terminus remained unstructured in the model. N-terminal truncations were made to ChlG, removing 11 (green), 23 (blue), 32 (red) and 39 (black) residues. (C) Potential secondary structure configurations of the first 39 residues of ChlG. This region, unstructured in (B) due to lack of sequence homology between ChlG and UbiA, were modelled using QUARK software.

6.3 Results

6.3.1 Sequential truncation of the ChlG N-terminus

The importance of the N-terminal tail region of ChlG to the function of the enzyme was investigated by generation of four sequential truncations to the 5' end of the *chlG* gene, removing the first 33, 69, 96 and 117 nucleotides corresponding to the N-terminal 11, 23, 32 and 39 amino acids residues respectively of the protein (Figure 6.2A). These truncation points were chosen such that they roughly divided the extreme N-terminus into four equal segments ending just before the start of the first predicted transmembrane helix. The *Synechocystis chlG* gene was amplified from genomic DNA using 5' primers that complemented the regions of the gene immediately downstream of the desired truncation sites described above. These genes were cloned into the pPD-NFLAG plasmid in frame with sequence encoding a N-terminal 3xFLAG tag (Hollingshead *et al*., 2012) (Figure 6.2B). Each vector was introduced into WT *Synechocystis* by natural transformation and integrated into the *psbAII* locus, replacing *psbAII* and putting the heterologous *chlG* genes under the control of the *psbAII* promoter. Segregation of the mutant genes was confirmed by PCR; in all cases complete replacement of the *psbAII* gene with the gene encoding the tagged enzyme was achieved (Figure 6.2C). The resulting strains FLAG-*chlG* Δ*chlG* (FLAG-*chlG*), FLAG-*chlG* Δ1-11 (Δ1-11), FLAG-*chlG* Δ1-23 (Δ1-23), FLAG-*chlG* Δ1-32 (Δ1-32) and FLAG-*chlG* Δ1-39 (Δ1-39) were all capable of photoautotrophic growth under normal growth conditions.

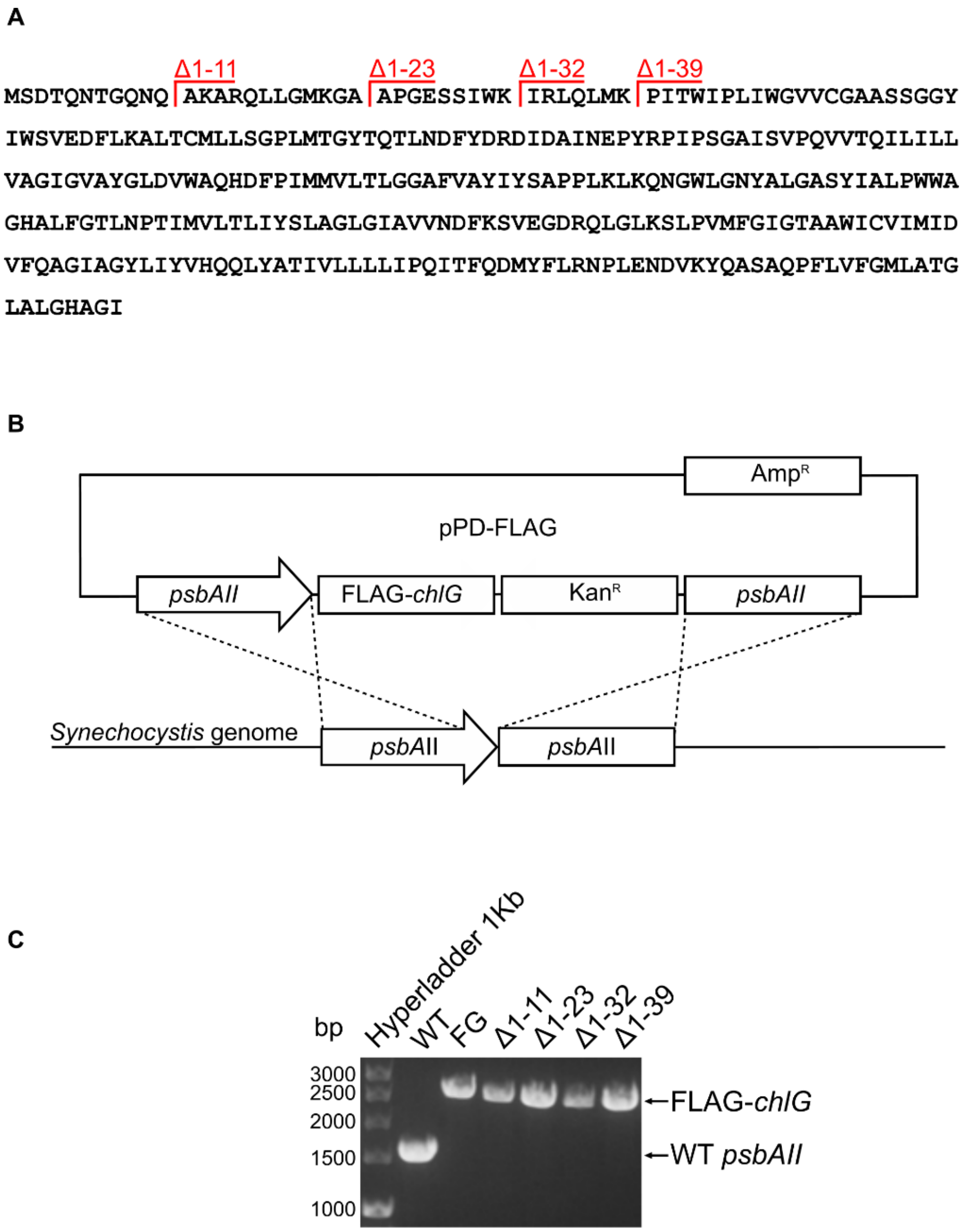


Figure 6.2: N-terminal truncations of chlorophyll synthase. (A) Four sequential truncations were made to the 5' end of the *Synechocystis chlG* gene so that the corresponding proteins lacked the first 11, 23, 32 and 39 residues from the N-terminus. (B) Genes encoding FLAG-tagged truncated ChlGs were inserted in place of the *psbAII* gene in the *Synechocystis* genome. (C) Agarose gel electrophoresis showing segregation of the mutant *chlG* proteins as well as a full length positive control (FLAG-*chlG*) in comparison a WT strain (WT).

6.3.2 The extreme N-terminus of ChlG is not required for binding HliD, Ycf39 and YidC

In order to examine whether the truncated FLAG-ChlG proteins could still interact with YidC, HliD and Ycf39, FLAG-immunoprecipitation experiments were conducted on each strain and the resulting eluates separated by SDS-PAGE (Figure 6.3A). In each case the bait protein was visible by silver staining and showed a clear reduction in size from Δ1-11 to Δ1-39 that correlates with the sequentially increasing truncation. There was also a clear reduction in the quantity of protein purified from Δ1-11 to Δ1-39. The Δ1-11 and Δ1-23 eluates contained the most FLAG-ChlG protein, the concentration decreased in the Δ1-32 sample and decreased further in the Δ1-39 eluate. The presence of FLAG-ChlG, YidC, HliD and Ycf39 was confirmed in all cases by immunoblot (Figure 6.3B). Ycf39 levels appeared to be reduced in the Δ1-39 sample, however the immunoprecipitation experiments were repeated for each strain and a positive signal for Ycf39 detected by immunblot in each case (data not shown). The interaction of Ycf39 with the ChlG complex has been shown to be dynamic depending on the light intensity under which the cells are cultured, which is the most likely explaination for the discrepancy in the data presented here (Proctor *et al*., 2018; Shukla *et al*., 2018b). As such, the results suggest that the first 39 residues of ChlG are not required for interaction with YidC, HliD or Ycf39.

Analysis of the pigment content of the eluates by absorbance spectroscopy and reverse-phase chromatography showed that the pigment profiles were comparable to a control immunoprecipitation eluate retrieved from the FLAG-6803 strain, containing the full length ChlG protein, indicating that carotenoid binding to the truncated FLAG-ChlG variants was unaffected by removal of the N-terminus (Figure 6.3C and Figure 6.4F-J). The mutant *Synechocystis* cells harbouring truncated ChlG proteins all exhibited a pigment profile that was comparable to that of the FLAG-*chlG* strain; each contained myxoxanthophyll, zeaxanthin, β-carotene, echinenone and Chl *a*, demonstrating that carotenoid and Chl biosynthesis was unaffected in these strains (Figure 6.4A-E).

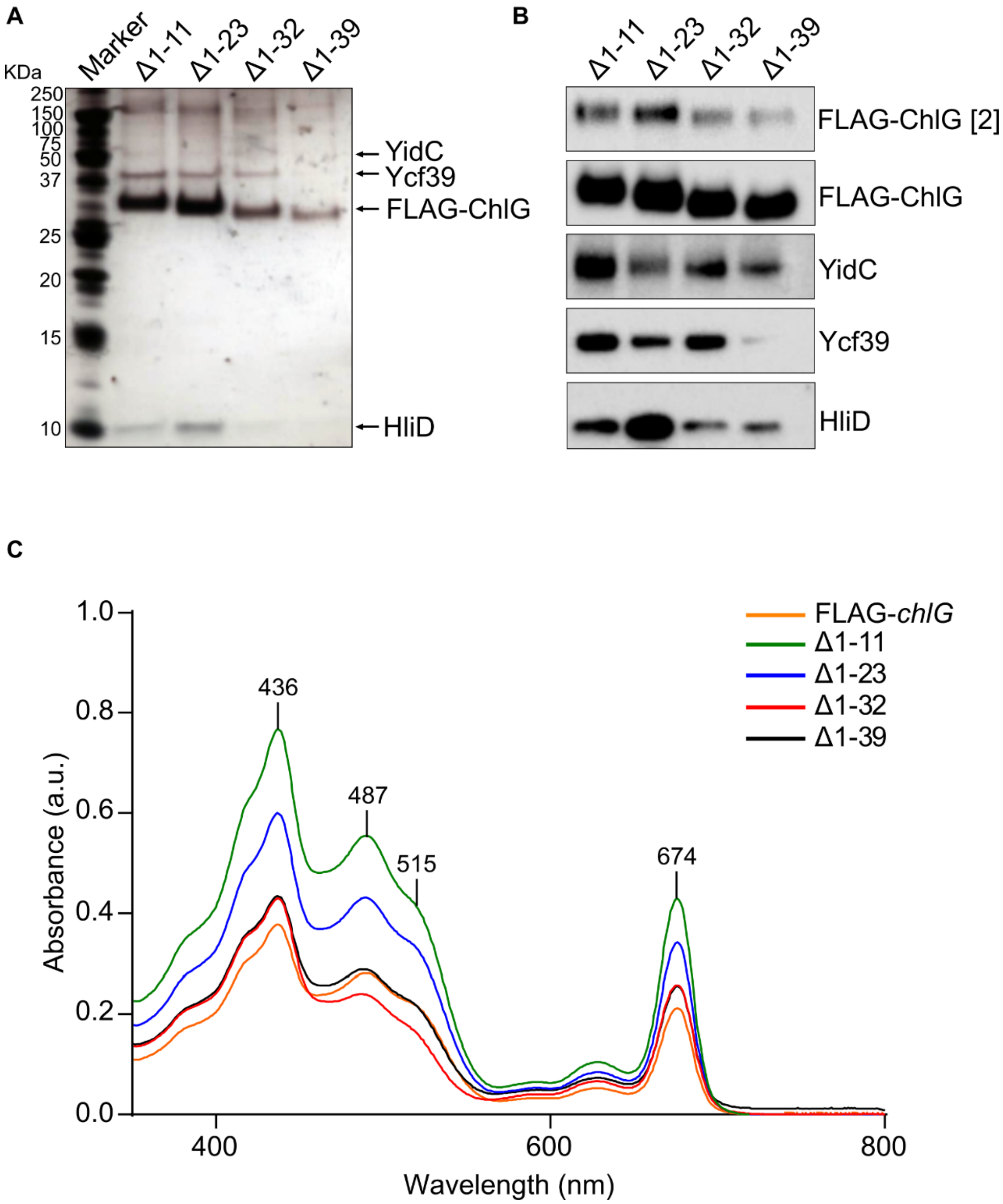


Figure 6.3: Purification of FLAG-ChlG from *Synechocystis* FLAG-*chlG* Δ*chlG*, FLAG-*chlG* Δ1-11, Δ1-23, Δ1-32 and Δ1-39 strains and identification of interacting proteins. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by silver staining. (B) Immunoblots using antibodies raised against 3xFLAG and ChlG interaction partners YidC, Ycf39, and HliD. FLAG-ChlG dimers (FLAG-ChlG [2]) were visible in each of the eluates (C) Absorption spectra of FLAG-immunoprecipitation eluates measured directly after elution.

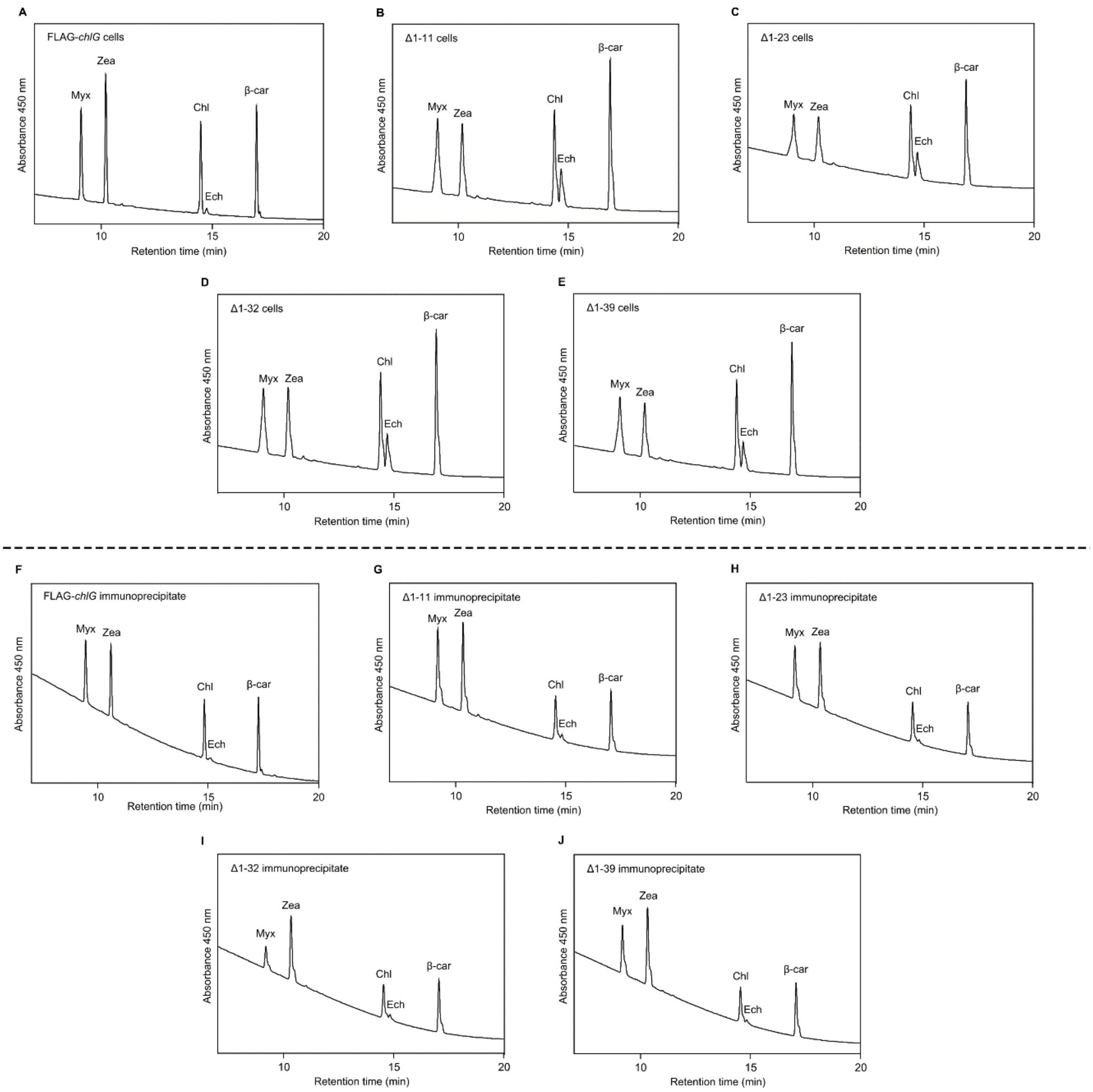


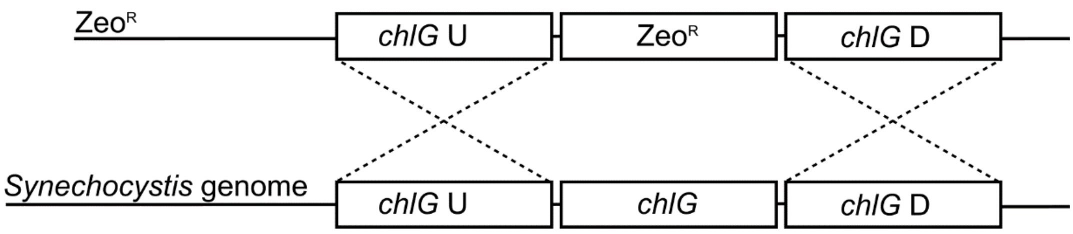
Figure 6.4: Analysis of the pigment content of whole cell and immunoprecipitation eluates of the FLAG-*chlG* Δ*chlG*, FLAG-*chlG* Δ1-11, Δ1-23, Δ1-32 and Δ1-39 strains. (A-E) Pigments were extracted from FLAG-*chlG* Δ*chlG*, FLAG-*chlG* Δ1-11, Δ1-23, Δ1-32 and Δ1-39 cells in methanol and separated by reverse phase HPLC. Myxoxanthophyll (Myx), zeaxanthin (Zea), β-carotene (β-car), echinenone (Ech) and chlorophyll (Chl) were all present in the mutant strains in levels comparable to FLAG-ChlG. (F-J) Pigments were extracted from FLAG-*chlG* and truncated FLAG-ChlG immunoprecipitation eluates and separated by reverse-phase HPLC. Pigment ratios were comparable between all of the eluates.

6.3.3 A ChlG enzyme lacking 32 residues from the N-terminus has impaired activity

ChlG is an essential enzyme in *Synechocystis*, so it is not possible to generate a fully segregated *chlG* deletion mutant in the WT background. To ascertain if the truncated enzymes were able to functionally complement the native enzyme, the deletion of the native *chlG* gene from the mutant strains was attempted, by replacement of the native *chlG* gene with a zeocin resistance cassette by linear mutagenesis as described in chapter 3 (Figure 6.5A). Segregation of the genome copies was checked by PCR screening of the *chlG* locus. Full segregation was only possible for the two least extensive truncations, Δ1-11 and Δ1-23, generating FLAG-*chlG* Δ1-11 Δ*chlG* (Δ1-11/Δ*chlG*) and FLAG-*chlG* Δ1-23 Δ*chlG* (Δ1-23/Δ*chlG*). The larger truncations, Δ1-32 and Δ1-39, still maintained wildtype copies of *chlG* even when plated onto the highest concentration of the selective antibiotic that was still permissive for growth (Figure 6.5B). This indicated that at least the extreme 23 N-terminal residues of ChlG are not required for enzyme function, but somewhere between residues 23 and 32 is the cut-off for producing a functional enzyme.

To test the activity of the truncated mutant ChlG enzymes *in vitro*, ChlG assays using immunoprecipitation eluates purified from the FLAG-*chlG* Δ1-11, Δ1-23, and Δ1-32 strains, still containing the native *chlG* gene, were performed in comparison to a full length FLAG-*chlG* eluate (Figure 6.6). Assays were stopped using excess methanol and the extract were analysed by reverse-phase HPLC. Chl a_{GG} and Chlide pigments were identified by their absorbance spectra. Consistent with the *in vivo* results, the Δ1-11 and Δ1-23 eluates catalysed the esterification of GGPP to Chlide. A small Chl a_{GG} peak was visible in the Δ1-32 ChlG assay. This enzyme was unable to complement deletion of the native chlG gene *in vivo* and the concentration of ChlG within the eluate was much lower than in the Δ1-11 and Δ1-23 samples. It is possible that the Δ1-32 ChlG protein dimerises and co-purifies with the native WT ChlG which would be expected to confer activity on the eluate. A ChlG dimer (FLAG-ChlG [2]) that survived the denaturing conditions used during SDS-PAGE is visible via immunoblotting using anti-FLAG antibodies (Figure 6.3B). The dimerisation of mutant and native ChlG may also enable the binding of ChlG interaction partners in the Δ1-11 and Δ1-23 complexes. This prospect is discussed further in the following section.

A



B

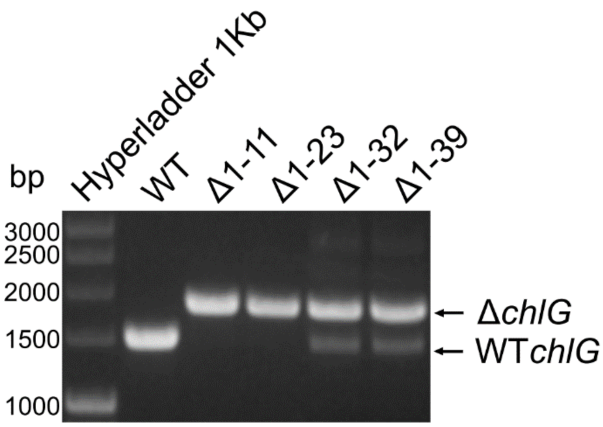


Figure 6 5: Deletion of native *chlG* from mutant strains harbouring truncated FLAG-*chlG* genes. (A) The native *chlG* gene was deleted from the strains producing truncated FLAG-tagged ChlGs by replacement with a zeocin resistance cassette (Zeo^R). (B) PCR using primers flanking the *chlG* locus revealed that only the two smaller truncations (Δ1-11 and Δ1-23) allow full deletion of the native *chlG* gene. The Δ1-32 and Δ1-39 ChlG mutants still retain copies of the WT *chlG* gene, indicating that full segregation of the knock out construct is not possible.

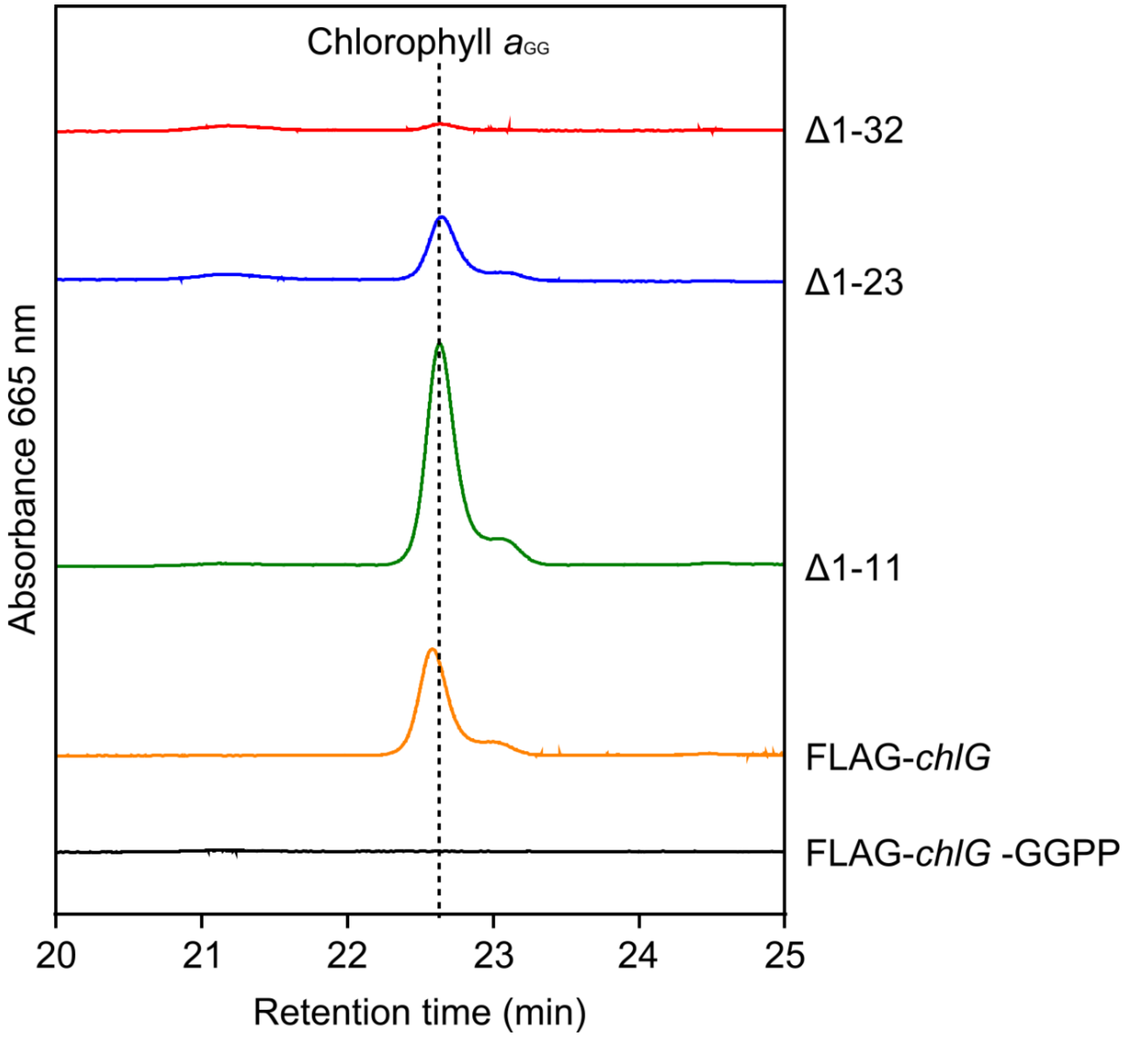


Figure 6.6: Reverse-phase HPLC separation of pigments extracted from ChlG assays. Absorbance profiles of pigments extracted from ChlG assays using immunoprecipitant eluates purified from FLAG-*chlG* Δ*chlG* (FLAG-*chlG*), a negative control lacking GGPP (FLAG-*chlG* – GGPP), FLAG-ChlG truncated by 11 (Δ1-11), 23 (Δ1-23) and 32 (Δ1-32) residues from the N-terminus.

6.3.4 The first 23 residues of ChlG are not required for activity

The fact that the truncated enzymes co-immunoprecipitate with the same interaction partners as the full-length enzymes indicates that the N-terminal 39 residues are not involved in the formation of the ChlG complex. However, it has previously been reported that in a strain producing FLAG-ChlG where the WT copy of the enzyme has not been deleted, dimers of FLAG-ChlG and the native un-tagged enzyme are formed (Shukla *et al*., 2018b), thus it is possible that the truncated enzyme may form a dimer with the native enzyme and that HliD-Ycf39-YidC may be interacting with the native enzyme. To investigate this possibility, the FLAG-ChlG protein was immunoprecipitated from the Δ1-11/Δ*chlG* and Δ1-23/Δ*chlG* strains which lacked the full length native *chlG* enzyme. The eluates were analysed by SDS-PAGE which revealed clear bands for FLAG-ChlG and HliD in both samples (Figure 6.7A). Immunoblotting confirmed the presence of FLAG-ChlG and HliD as well as both YidC and Ycf39 within the samples, indicating the truncated ChlG proteins were able to from the native complex in the absence of the full length native ChlG (Figure 6.7B). The eluates were both pigmented, with absorbance spectra that were comparable to an eluate retrieved from a strain containing the full length ChlG protein (Figure 6.7C). Further analysis of the pigment content of the eluates by reverse-phase HPLC showed that in both cases the truncated ChlG enzymes contained zeaxanthin, myxoxanthophyll and β-carotene in the absence of the full length ChlG (Figure 6.7D-E). These two eluates were also enzymatically active, although the Δ1-11/Δ*chlG* produced more Chl a_{GG} product despite being of lower concentration than Δ1-23/Δ*chlG* judging from the relative band intensities visible by SDS-PAGE (Figure 6.3A). This again demonstrates the importance of the N-terminus to the activity of the enzyme which appears to decrease as more of the N-terminus is removed (Figure 6.7F).

Taken together the results indicate that the first 32 residues of ChlG are not essential for the formation of the ChlG complex and that the first 23 residues are dispensable for enzyme activity. As the native WT *chlG* cannot be deleted from the Δ1-32 strain, it remains to be seen whether the truncated protein is marginally active and able to bind HliD, YidC and Ycf39, or whether this is due to formation of a dimer with the native ChlG enzyme. This is discussed further in Section 6.4.2.

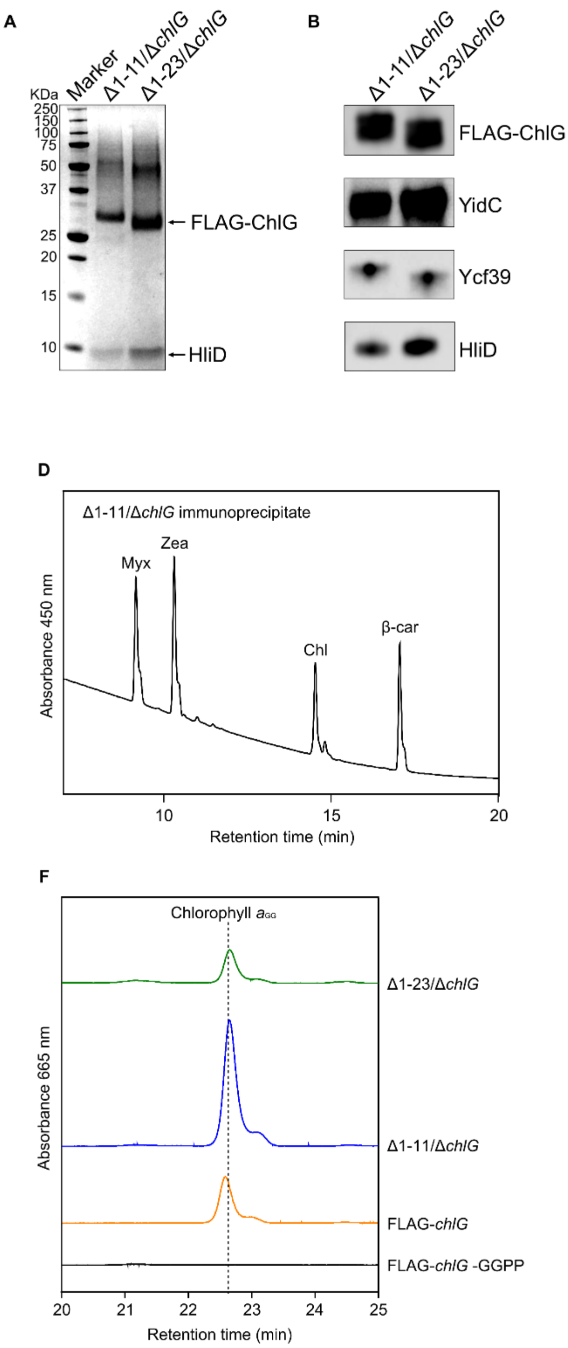


Figure 6.7: Purification of FLAG-ChlG from *Synechocystis* FLAG-*chlG* Δ1-11 Δ*chlG* and Δ1-23 Δ*chlG* strains and identification of interacting proteins. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG and the ChlG interaction partners YidC, Ycf39 and HliD. (C) Absorption spectra of FLAG-immunoprecipitation eluates. (D-E) Pigments were extracted in methanol and separated by reverse-phase HPLC. Myxoxanthophyll (Myx), zeaxanthin (Zea), β-carotene (β-car) and chlorophyll (Chl) were all present within both eluates. (F) Reverse-phase HPLC separation of pigments extracted from ChlG assays.

6.3.5 Deletion of 51 residues from the ChlG N-terminus destabilises the protein

Analysis of the FLAG-immunoprecipitation eluates obtained from the mutant *Synechocystis* strains harbouring N-terminally truncated ChlG proteins indicated that there was a reduction in the concentration of the ChlG protein as the truncations became more severe. To test if further truncations of N-terminus would abolish accumulation of the protein completely, two more truncated *chlG* genes were generated resulting in mutant proteins lacking 45 and 51 residues from the extreme N-terminus as described previously. The FLAG-immunoprecipitation eluate from both the FLAG-*chlG* Δ1-45 and FLAG-*chlG* Δ1-52 strains were not coloured and no ChlG bands were observed on a stained SDS-PAGE gel (Figure 6.8A). Although no ChlG protein was visible by SDS-PAGE, ChlG lacking 45 residues from the N-terminus was detectable by immunoblot. Immunoblots using the FLAG-ChlG Δ1-52 eluate produced a barely visible signal (Figure 6.8B). This result indicates that extensive truncation of the ChlG N-terminus significantly reduces accumulation of the protein in the cell.

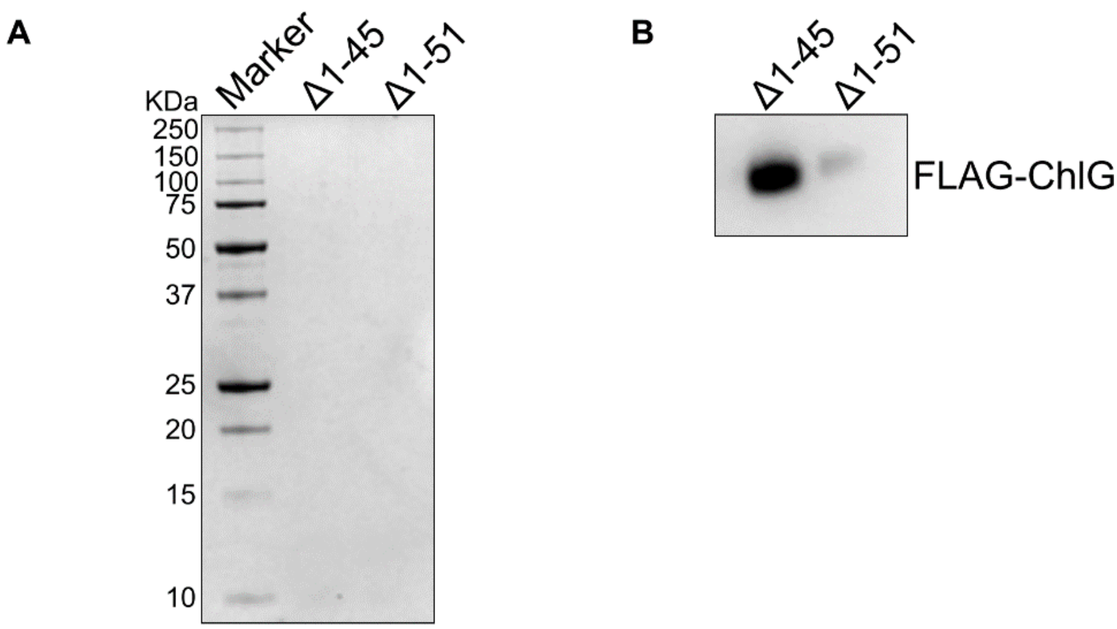


Figure 6.8: Purification of FLAG-ChlG from *Synechocystis* FLAG-*chlG* Δ1-45 and Δ1-51 strains. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG.

6.3.6 A C-terminally FLAG-tagged *Synechocystis* ChlG enzyme is inactive *in vivo*

Although truncation of the ChlG N-terminus by 39 residues did not prevent formation of the ChlG complex, it was possible that the N-terminal FLAG-tag was artificially extending the N-terminal region and facilitating ChlG complex formation. A strain containing a FLAG-tag on the C-terminus of the full length *chlG* gene was generated (*chlG*-FLAG) with the intention of using this strain as a control to test the effects of N-terminally truncating ChlG in the absence of the N-terminal FLAG-tag. The *chlG*-FLAG strain was used as the basis for immunoprecipitation experiments to first ascertain whether the same ChlG complex could be purified as the one isolated with the N-terminally FLAG-tagged enzyme. The *Synechocystis chlG* gene was amplified from genomic DNA and cloned into the pPD-CFLAG plasmid (Chidgey et al., 2014) in frame with sequence encoding a C-terminal 3xFLAG tag. The vector was introduced into WT *Synechocystis* by natural transformation and integrated into the genome at the *psbAII* locus by homologous recombination, placing the gene under the control of the *psbAII* promoter (Figure 6.9A) (Hollingshead *et al*., 2012). Segregation of genome copies was confirmed by PCR (Figure 6.9B).

As *chlG* is an essential gene in *Synechocystis*, the native gene can only be deleted from the *chlG*-FLAG strain if the tagged enzyme is functional. The native *chlG* gene was disrupted using the linear mutagenesis construct described by Chidgey *et al*. (2014) which partially replaced the *chlG* gene with a zeocin resistance cassette (*zeo*^{R}) (Figure 6.9C). Despite multiple attempts under different growth conditions, PCR screening using primers that flank the *chlG* locus revealed that full segregation of the strain was not possible (Figure 6.9D). This indicates that, unlike the FLAG-*chlG* variant, the ChlG-FLAG protein cannot functionally complement the activity of the native ChlG enzyme. The FLAG-tag on the C-terminus may impede the active site of the enzyme or structurally deform the protein and the surrounding thylakoid membrane, perturbing ChlG function.

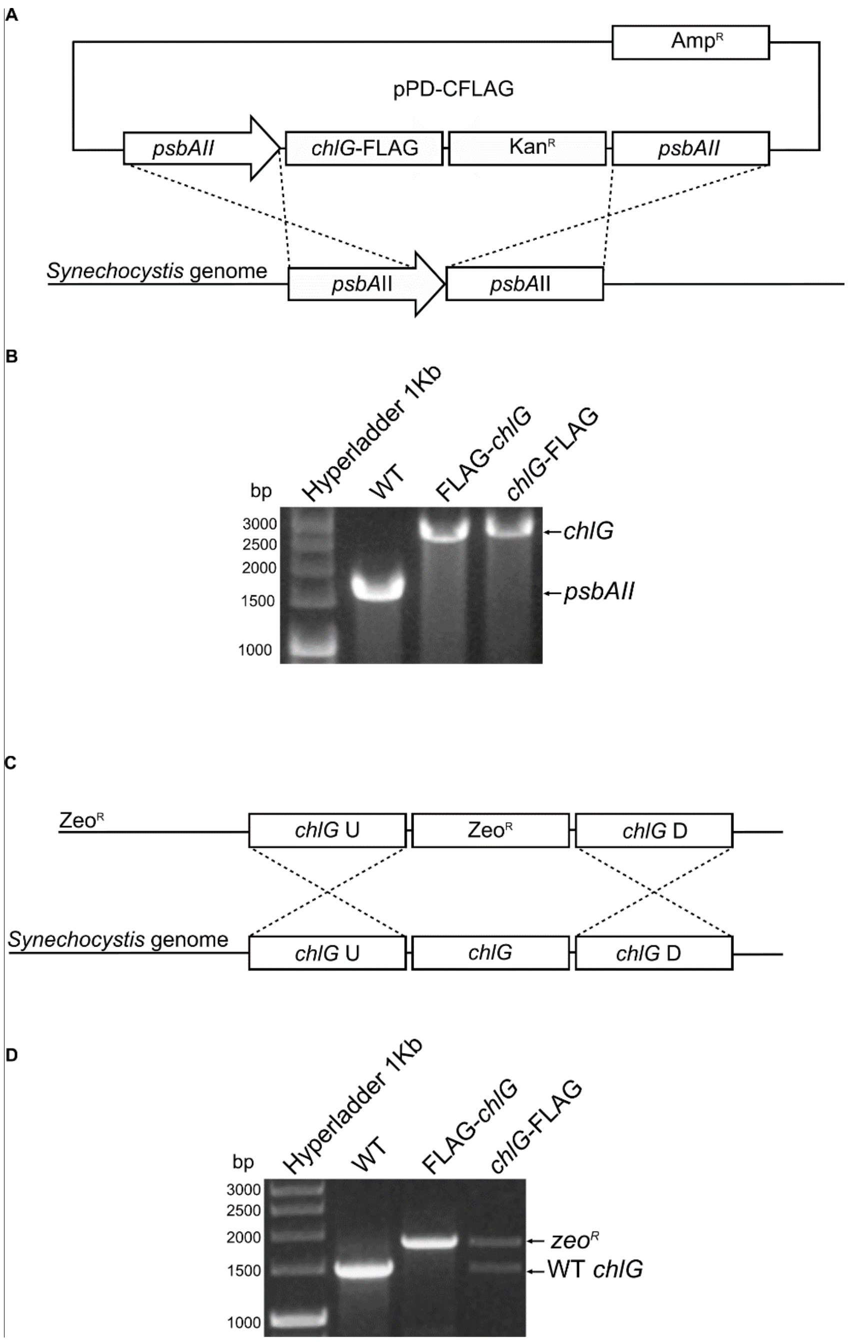


Figure 6.9: Generation of a *Synechocystis* strain expressing a C-terminally FLAG-tagged *chlG* gene and subsequent deletion of the native *chlG* gene. (A) A Construct encoding a C-terminally 3xFLAG tagged *chlG* was inserted in place of the *psbAII* gene in the *Synechocystis* genome. (B) PCR amplification of the *psbAII* locus of WT, FLAG-*chlG* Δ*chlG* and transformant *chlG*-FLAG *Synechocystis* strains. (C) The native *chlG* gene was deleted from the strains producing foreign FLAG-tagged ChlGs by replacement with a zeocin resistance cassette (Zeo^R). (D) Only partial segregation of the *chlG* locus was achievable in the *chlG*-FLAG strain as revealed by PCR screens using primers flanking the integration site.

6.3.7 C-terminally FLAG-tagged ChlG co-purifies with YidC but not HliD or Ycf39 and is active *in vitro*

To test whether the C-terminal FLAG-tagged ChlG protein is being produced and if so, whether it can bind to YidC, HliD and Ycf39, the ChlG-FLAG protein was immunoprecipitated from the *chlG*-FLAG strain. The strain containing the N-terminally FLAG-tagged ChlG was included as a control. The eluates were separated by SDS-PAGE (Figure 6.10A) and analysed for ChlG-FLAG, YidC, HliD and Ycf39 by immunoblot using the appropriate antibodies (Figure 6.10B). A band approximately 30 kDa in size was visible and confirmed to be FLAG-ChlG by immunoblot using antibodies raised against the FLAG-tag, indicating that the protein, despite being inactive, is still produced and maintained by the cell.

Immunoblot analysis of the resultant eluate showed the presence of ChlG-FLAG and YidC, but not HliD and Ycf39. This indicates that the C-terminally FLAG-tagged enzyme cannot interact with HliD or Ycf39 in contrast to the N-terminally FLAG-tagged variant (Figure 6.10B). The lack of HliD within the ChlG-FLAG eluate was also evident by spectrophotometric analysis. The eluate did not produce peaks at 487 and 515 nm, characteristic of the HliD bound carotenoids present within the native complex (Figure 6.10C). However, like the plant and algal versions of the complex described in chapter 3, the association of ChlG with YidC is maintained despite the abolished interactions with HliD and Ycf39. To rule out the possibility that the lack of Ycf39 and HliD observed in the ChlG-FLAG eluate is due to lack of accumulation of HliD and Ycf39 within this strain, immunoblot analysis of solubilised thylakoid membranes (prepared as described in Section 2.11.1) using antibodies raised against HliD and Ycf39 must be performed.

In vitro ChlG enzyme activity assays were performed using the ChlG-FLAG and FLAG-*chlG* eluates. Both eluates were able to produce Chl a_{GG} although ChlG-FLAG generated a smaller yield of product than the FLAG-*chlG* variant indicating that the C-terminal FLAG-tag is impeding ChlG activity (Figure 6.10D).

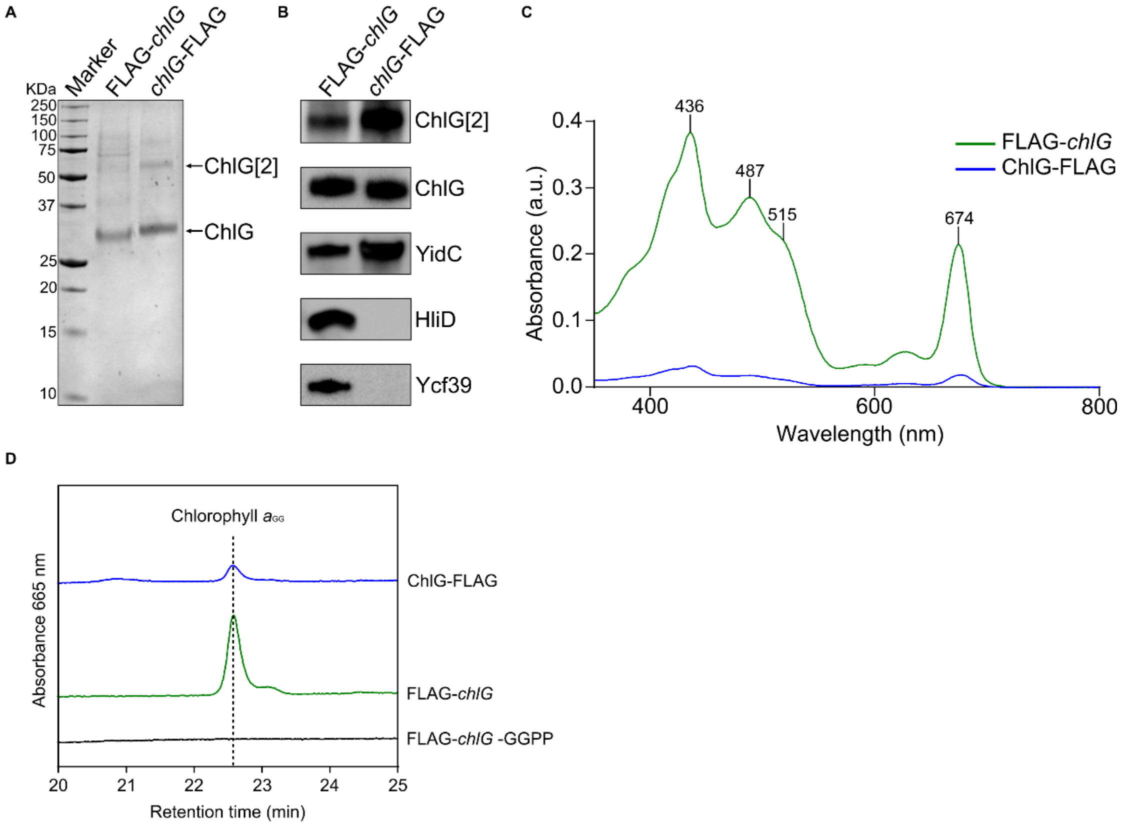


Figure 6.10: Purification and activity of ChlG-FLAG. (A) FLAG-immunoprecipitation eluates were separated by SDS-PAGE and analysed by staining with Coomassie Brilliant Blue. (B) Immunoblots using antibodies raised against 3xFLAG and ChlG interaction partners YidC, Ycf39, and HliD. ChlG[2] indicates a ChlG dimer. (C) Absorption spectra of FLAG-immunoprecipitation eluates. (D) Reverse-phase HPLC separation of pigments extracted from ChlG-FLAG and FLAG-*chlG* assays.

6.4 Discussion

6.4.1 The N-terminus of ChlG is not required for binding of YidC, HliD and Ycf39

Synechocystis ChlG binds to HliD, Ycf39 and YidC forming a complex that can be purified from the cell by FLAG-tagging the ChlG enzyme and performing FLAG-immunoprecipitation experiments (Chidgey et al., 2014). From the cross-linking data presented in Chapter 5, the first 39 residues of the ChlG N-terminus was predicted to be involved in the binding of these interaction partners. In this chapter, the N-terminus of a FLAG-tagged ChlG was sequentially truncated, removing 11, 23, 32 and 39 residues from the protein. In each case, the ChlG complex was still able to assemble in the cell, despite the absence of a maximum of 39 residues from the N-terminus. As ChlG has been shown to form a dimer in *Synechocystis* (Shukla *et al*., 2018b), the native *chlG* gene was deleted from the cells harbouring the truncated FLAG-tagged proteins. This was only possible in the Δ1-11 and Δ1-23 strains (see below), although FLAG-immunoprecipitation experiments using the resulting Δ1-11/Δ*chlG* and Δ1-23/Δ*chlG* strains demonstrated that the ChlG complex was still able to assemble.

The results from this chapter appear to conflict with the results described in Chapter 5. In Chapter 5, the N-terminal domain of ChlG was shown to form cross-links with Ycf39 (indicating that the N-terminus of ChlG is located in the cytoplasm) in addition to YidC and HliD. Cross-links between YidC and Ycf39/HliD were also identified, as well as cross-links between Ycf39 and HliD. The removal of the N-terminus of ChlG may not abolish the formation of the ChlG complex because (a) the cross-links identified may represent the close proximity of two polypeptides but not necessarily their interaction, (b) the N-terminal domain of ChlG does interact with YidC/Ycf39/HliD, however, the interactions between the latter proteins may suffice to facilitate formation of the ChlG complex, (c) other, so far unobserved, domains of ChlG could also interact with its partners and stabilise the ChlG complex. Expanding on the latter scenario, ChlG, HliD and YidC are all integral membrane proteins, predicted to feature 8, 1 and 3 transmembrane helices respectively (see Chapter 5 Section 5.3.1). Despite the fact that no cross-link modifications of amino acid residues predicted to be located within the transmembrane helices were identified, it is possible that interactions between these three proteins occur within the lipid bilayer. The cross-linking reagents may simply not have been able to permeate the membrane and/or access these regions of the proteins. The results from this chapter highlight the need for further structural characterisation of the ChlG complex, as outlined in Chapter 5 Section 5.4.2.

6.4.2 The first 32 residues of *Synechocystis* ChlG are essential for enzyme function

N-terminal truncation of ChlG by 32 amino acids almost entirely abolished the enzyme activity, as demonstrated *in vivo* by the failed attempt to delete the native *chlG* gene from the FLAG-*chlG* Δ1-32 background, and *in vitro* by negligible production of Chl a_{GG} during enzyme activity assays. However, ChlG variants lacking 11 and 23 residues from the N-terminus were both active *in vivo* as the native *chlG* gene was successfully deleted from these backgrounds indicting that the truncated variants were able to functionally complement the WT protein. Both proteins were also active *in vitro*, although the Δ1-23 to a lesser extent than the Δ1-11 variant despite the fact that the FLAG-ChlG concentration within the undiluted Δ1-23 eluate used for *in vitro* enzyme assays was significantly higher than in the Δ1-11 eluate (Figure 6.3A). This suggests that the removal of 23 residues from the N-terminus of ChlG perturbs activity without abolishing it completely. The length of the N-terminus therefore appears to correlate with ChlG activity and one or more amino acids essential to the activity of ChlG is located somewhere between residues 23 and 32. It may also be the case that removal of the amino acids immediately proximal to the predicted start of the first transmembrane helix could destabilise the protein, abolishing its activity.

In order to check whether the observed activity of the Δ1-11 and Δ1-23 mutants arose from co-immunoprecipitation of the WT enzyme, the native chlG gene was removed from the strains producing these proteins and FLAG-ChlG immunoprecipitated from the resulting strains. The Δ1-11/Δ*chlG* and Δ1-23/ Δ*chlG* eluates remained active *in vitro*, demonstrating that these enzymes are active in the absence of native ChlG. However, deletion of the native *chlG* gene was not possible from the Δ1-32 mutant. Furthermore, the Δ1-32 immunoprecipitation eluate contained ChlG dimers that survived denaturation during SDS-PAGE (Figure 6.3B). Whether or not this was a homodimer consisting of two Δ1-32 ChlG proteins, a heterodimer of Δ1-32 and WT ChlG or a mixture of both of these could not be ascertained. Therefore the possibility remains that the small amount of *in vitro* activity of this eluate could be due to the presence of a small amount of native ChlG. A method by which the dimerization of Δ1-32 with WT ChlG could be detected is outlined in Section 6.5. Alternatively the Δ1-32

ChlG variant may well have some residual activity; however the protein concentrations within this eluate were generally much lower than the Δ1-11 and Δ1-23 samples and so the smaller amount of product produced by Δ1-32 may be due to there being less FLAG-ChlG present within this assay.

In a previous study, truncations of recombinant oat (*Avena sativa*) ChlG, which shares 62% identity with the *Synechocystis* homologue, revealed that a core protein comprised of residues 88 to 377 is enzymatically active *in vitro* (Schmid *et al*., 2001). The first 46 residues of the oat ChlG enzyme correspond to a chloroplast transit peptide. If this region is discounted then residue Trp88 is equivalent to residue Trp31 in the *Synechocystis* ChlG, thus this result is in agreement with the finding that the Δ1-23 truncation is active and Δ1-32 has only marginal activity.

6.4.3 A C-terminally FLAG-tagged ChlG protein is inactive and cannot bind to HliD or Ycf39

It is possible that the FLAG-tag on the N-terminus of the truncated ChlG proteins can non-specifically facilitate the interaction of the enzyme with any of HliD, Ycf39 and YidC, essentially replacing the missing N-terminal regions. Cross-linking experiments using the full length FLAG-ChlG complex indicated that the FLAG-tag is in close enough proximity to YidC to form a cross-link between them (Chapter 5). To test this hypothesis, the full length ChlG protein was C-terminally FLAG-tagged (ChlG-FLAG) and used as bait during FLAG pulldown experiments. Analysis of the resulting eluate for interaction partners of ChlG-FLAG by immunoblot revealed that the presence of a C-terminal FLAG-tag lead to the abolishment of the HliD and Ycf39 interactions with ChlG, although an association with the YidC protein was maintained. As the interaction of HliD and Ycf39 was not maintained when the FLAG-tag was added to the C-terminus of ChlG, this approach was not viable as a means to examine the effects of truncating the ChlG N-terminus in the absence of the N-terminal FLAG-tag. Despite this, the activity of this protein was tested *in vivo* by attempting to delete the native *chlG* gene. This proved to be impossible, indicating that the C-terminal FLAG-tag inhibited enzyme activtiy, preventing it from functionally complementing the native protein. In contrast to this, the ChlG-FLAG eluate was able to produce Chl a_{GG} product during *in vitro* enzyme assays, although to a lesser extent than a FLAG-*chlG* control. As discussed with regards to the Δ1-32 truncated ChlG, the *in vitro* activity of ChlG-FLAG eluate may be conveyed by the co-purification of the tagged protein with the WT ChlG enzyme, causing the discrepancy in results between the *in vivo* and *in vitro* activities of the tagged enzyme. A ChlG dimer (ChlG[2]) was observed by SDS-PAGE and immunoblot (Figure 6.10A-B), indicating that this dimer is strong enough to survive the relatively denaturing conditions in which SDS-PAGE was performed. There is the possibility, however, that the dimerisation of ChlG is an artefact of treatment with SDS. Alternatively, ChlG-FLAG may be marginally active but is not competent enough to completely complement the WT enzyme. Another possibility is that the presence of a FLAG-tag on the C-terminus of ChlG retards the translocation of the protein across the thylakoid membrane, resulting in destabilisation of the thylakoid membrane and preventing the formation of the ChlG complex. Likewise, this may also perturb the activity of the enzyme *in vivo* by preventing the channelling of substrates to the active site.

6.5 Future work

Only the N-terminus of *Synechocystis* ChlG was targeted for truncation in this study. Schmid *et al*. (2001) showed that the final His378 could be removed from the C-terminal end of the Oat ChlG homologue without affecting activity, although any further truncation of the C-terminus resulted in the complete loss of activity. Interestingly, although deletion of the penultimate Ser377 abolished enzyme activity, substituting it for an Ala did not, indicating that the length of the C-terminus is the important factor (Schmid *et al*., 2001). This makes sense given that the final transmembrane helix is predicted to run from Ala359-Ser377 (Schmid *et al*., 2001). Sequence alignments of the *Synechocystis* ChlG with the Oat homologue show that the C-terminus of *Synechocytis* ChlG extends beyond the equivalent His378 by an additional 3 residues (Ala-Gly-Ile). Additionally, the results presented in this chapter demonstrate that the addition of 3xFLAG-tag to the C-terminus of ChlG abolished binding of HliD and Ycf39 as well as synthase activity *in vivo*, suggesting that ChlG may be more sensitive to modification of the C-terminus than the N-terminus. Truncations should be made to the C-terminus of *Synechocystis* FLAG-*chlG*, causing the loss of the Ala, Gly and Ile residues before sequential removal of further C-terminal residues one at a time. The activity of the resulting ChlG variants will be tested *in vivo* by attempted removal of the native *chlG* gene from the host strain. The FLAG-tagged proteins will be purified by FLAG-immunoprecipitation and tested for activity *in vitro* as well as analysed for interaction partners by immunoblot.

The Δ1-32 and C-terminally FLAG-tagged ChlG variants were both poorly functional *in vivo*, surmised from the fact that it was not possible to delete the native *chlG* gene from these backgrounds, but exhibited activity *in vitro* during enzyme activity assays. As described previously, it is possible that a dimer is able to form between the native and tagged proteins, the former of which would co-purify with the tagged protein and confer ChlG activity on the eluate. Support for this notion will require a demonstration of the presence of 'native' ChlG, perhaps His-tagged, in purified Δ1-32 and ChlG-FLAG eluates. Analysis of the resulting eluates by immunoblot using antibodies raised against the His-tag would enable the presence or absence of the His-ChlG to be determined.

Chapter 7: Heterologous production and *in vitro* mutagenesis of *Arabidopsis thaliana* chlorophyll synthase

7.1 Summary

Chlorophyll synthase (ChlG) and Bacteriochlorophyll synthase (BchG) catalyse the attachment of the tail moiety to the chlorin ring of (Bacterio)chlorophyllide, forming (Bacterio)chlorophyll in oxygenic and anoxygenic phototrophs respectively. These enzymes are structurally related but each is unable to utilise the other's natural substrate. To better understand the importance of conserved ChlG residues predicted to be involved in substrate specificity or substrate binding to the enzyme, N-terminally His-tagged *Arabidopsis thaliana* ChlG (AtChlG) was heterologously produced in *E. coli* and several point mutants were tested for activity by *in vitro* enzyme assays. Production of chlorophyllide *a* (Chlide *a*), the substrate of ChlG, was achieved by removing the tail moiety from chlorophyll *a* (Chl *a*) using the *Arabidopsis* enzyme chlorophyllase (CLH-1) also produced recombinantly in *E. coli*. ChlG activity was demonstrated in *E. coli* lysates containing recombinant AtChlG but AtChlG solubilised in detergent and purified by Ni^{2+} IMAC did not produce any detectable product. Three point mutations were generated, targeting residues conserved in ChlG and replacing them with their conserved counterparts in BchG; Q46E, L56P, V60Y. Q46E and L56P were active but to a lesser extent than WT ChlG, demonstrating a non-critical role of these residues in the activity of ChlG. V60Y was almost completely devoid of activity. Additionally, point mutations were made to residues N99 and A225, predicted to be located in the active site of AtChlG. Substitution of N99 for Ala and A225 for Met inhibited AtChlG catalysis completely, demonstrating the importance of these residues to ChlG function. Finally, residue P54, predicted to be important for determining the substrate specificity of AtChlG, was substituted by Phe. P54F was able to utilise Chlide *a* as its substrate to a lesser extent than WT AtChlG, suggesting that mutation of this residue had reduced the affinity of AtChlG for Chlide. This system can be used to quickly and easily test the viability of any AtChlG mutant *in vitro*.

7.2 Introduction

Chlorophyll synthase (ChlG) catalyses the esterification of the tetraprenyl tail to ring D of Chlorophyllide *a* (Chlide), which is the last step of the chlorophyll biosynthesis pathway. Activity of this enzyme was first detected in the etioplast membranes of oat (*Avena sativa*) (Rüdiger *et al*., 1980). The enzyme in purple bacteria, bacteriochlorophyll synthase (BchG) (Bollivar *et al*., 1994b) and the ChlG homologue in the cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*) (Kaneko *et al*., 1996) were subsequently identified. The activity of both BchG and ChlG was confirmed by production in *E. coli* (Oster *et al*., 1997). Further studies were performed on these enzymes with respect to their substrate specificity (Kim and Lee, 2010) and their mechanism of catalysis (Schmid *et al*., 2002) (see Section 1.8.11 for discussion).

There are only two examples of ChlG mutagenesis studies reported in the literature. Schmid *et al.* (2001) performed *in vitro* mutagenesis of the oat (*Avena sativa*) ChlG. Deleting residues 1-87 from the N-terminus resulted in a "core" protein that retained 46% activity of the full length ChlG. Deletion of residues 1-88 abolished all enzyme activity. At the C-terminus, only one His residue (H378) could be removed before the protein lost all activity upon deletion of the Ser residue directly upstream (S377). However, substitution of S377 to Ala (S377A) resulted in a protein that was 63% active in comparison to the WT enzyme. The authors concluded that it was the overall length of the sequence, rather than the specific amino acid, that was critical to producing a functioning enzyme. The same study also targeted four Arg and five Cys residues, identified as important for ChlG activity, by first testing the activity of the enzyme in the presence of specific Arg and Cys inhibitors. They subsequently altered these residues to Ala and tested the enzyme activity with and without the appropriate inhibitor. Of the four Arg residues, R91 and R161 were essential to activity. For the Cys residues, C109, conserved in all of the known ChlG homologs at the time, was also critical to activity whereas C130 showed reduced activity.

Kim, Kim and Lee, (2016) identified residue I44 of the *Synechocystis* ChlG and the equivalent F28 of *Rba. sphaeroides* BchG as critical to the substrate specificity of the enzyme. They found that expression of *Synechocystis chlG* in a *Rba. sphaeroides bchG* null mutant resulted in multiple suppressor strains that all harboured a ChlG I44F point mutation. ChlG I44F was able to esterify BChlide, although with reduced substrate affinity, and was thus able to restore photosynthetic growth to these strains, albeit at greatly reduced growth rates. Subsequently, the authors performed *in vitro* mutagenesis experiments where they generated ChlG I44F and the corresponding BchG F28I variant proteins and produced them in *E. coli*. They showed that ChlG I44F was able to esterify BChlide and that BchG F28I was able to esterify Chlide, albeit with a 10 fold reduction in efficiency relative to their respective native substrates (see Section 1.8.11 for further discussion).

The discovery of chlorophyll synthase gene in the higher plant, *Arabidopsis thaliana* (*A. thaliana*), came in 1995 when the gene, annotated G4, was identified during the sequencing of the *A. thaliana* genome and was found to encode a protein with significant sequence homolog to the *Rba. capsulatus* BchG (Gaubier *et al*., 1995). The enzyme was produced in *E. coli* and was able to esterify Chlide with geranylgeranyl pyrophosphate (GGPP) in preference to utilising phytyl pyrophosphate as a substrate (Oster and Rüdiger, 1997). However, there have been no further reports of kinetics, *in vitro* mutagenesis or substrate specificity studies. This chapter reports optimisation of AtChlG production in *E. coli* and generation of several point mutations that change conserved ChlG residues to their BchG counterparts. A method for the production of Chlide substrate was developed, which was tested for activity with mutant AtChlG proteins. The aim was to establish a recombinant system to characterise *A. thaliana* ChlG *in vitro*, to enable comparisons with (B)Chl synthases from *Avena sativa*, *Synechocystis* and purple bacteria.

7.3 Results

7.3.1 Heterologous production of *Arabidopsis thaliana* ChlG in *E. coli*

In order to produce point-mutant variants of the *A. thaliana* ChlG (hereafter AtChlG) using the QuickChange II Site-Directed Mutagenesis Kit (Agilent Technologies), it was first essential to establish a viable over-production system in *E. coli*. Following trials with various expression vectors; the pET28a vector was found to be suitable for AtChlG production in *E. coli* cell line BL21 (DE3). The At_*chlG* gene was sub-cloned from the *Synechocystis* strain discussed in Chapter 3. The gene was amplified by PCR so that it contained NdeI and EcoRI restriction enzyme digest sites and ligated into the corresponding sites of pET28a, in frame with an N-terminal 6x His tag (His_6).

The recombinant pET28a::AtChlG plasmid was introduced into *E. coli* BL21(DE3), which was cultured in 1 L of LB medium at 37 °C with shaking at 180 rpm until the OD_{600} of the cells reached 0.6, at which point expression was induced with IPTG. Production of AtChlG was confirmed by separation of whole cell lysates by SDS-PAGE followed by immunoblots using antibodies raised against ChlG (already shown to recognise the AtChlG enzyme in Chapter 3) (Figure 7.1A) and the His-tag epitope (Novagen®) (Figure 7.1B). Both antibodies gave a strong signal at 25 kDa, smaller than would be expected for His_6-AtChlG which is calculated to be 37 kDa. It is probable that ChlG, a hydrophobic integral membrane protein which likely binds more SDS than an equivalent soluble protein, may migrate faster than expected during SDS-PAGE (Rath *et al*., 2009). AtChlG purified from *Synechocystis,* and analysed via SDS-PAGE and immunoblot, also migrated as approximately 25 kDa, as discussed in Chapter 3. A control sample, prepared from BL21 harbouring an empty pET28a vector, did not cross-react with either antibody.

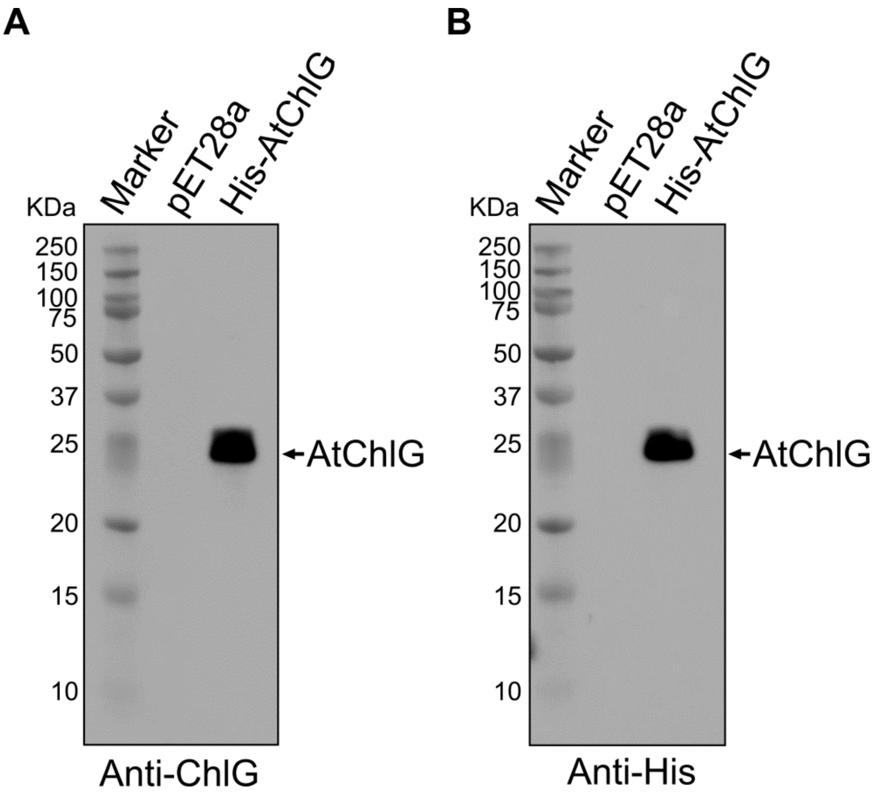


Figure 7.1: Production of AtChlG in *E. coli*. Immunodetection of AtChlG in cell lysates of *E. coli* using antibodies raised against ChlG (A) and the His-tag epitope (B).

7.3.2 Purification of AtChlG by Ni^{2+} immobilised metal affinity chromatography

Attempts were made to purify AtChlG, after confirming its production in *E. coli*. The AtChlG membrane fraction was prepared from 2 L of cells (Section 2.11.4), cultured as described above, solubilised with 1.5% (w/v) β-DDM, then the clarified detergent extract was applied to a Ni^{2+} NTA IMAC column. The AtChlG protein eluted from the column in 400 mM imidazole and the eluate was analysed by SDS-PAGE. There was no prominent protein band detectable upon staining with Commassie Brialliant Blue (Figure 7.2A) although a band at 30 kDa was tentatively assigned to AtChlG. However, AtChlG production was confirmed by immunoblot using antibodies raised against ChlG (Figure 7.2B). AtChlG appeared to aggregate in solution as indicated by multiple signals of varied MW. There also appeared to be some degradation of the protein and so AtChlG may not be very stable in solution or is partially degraded in the host cell.

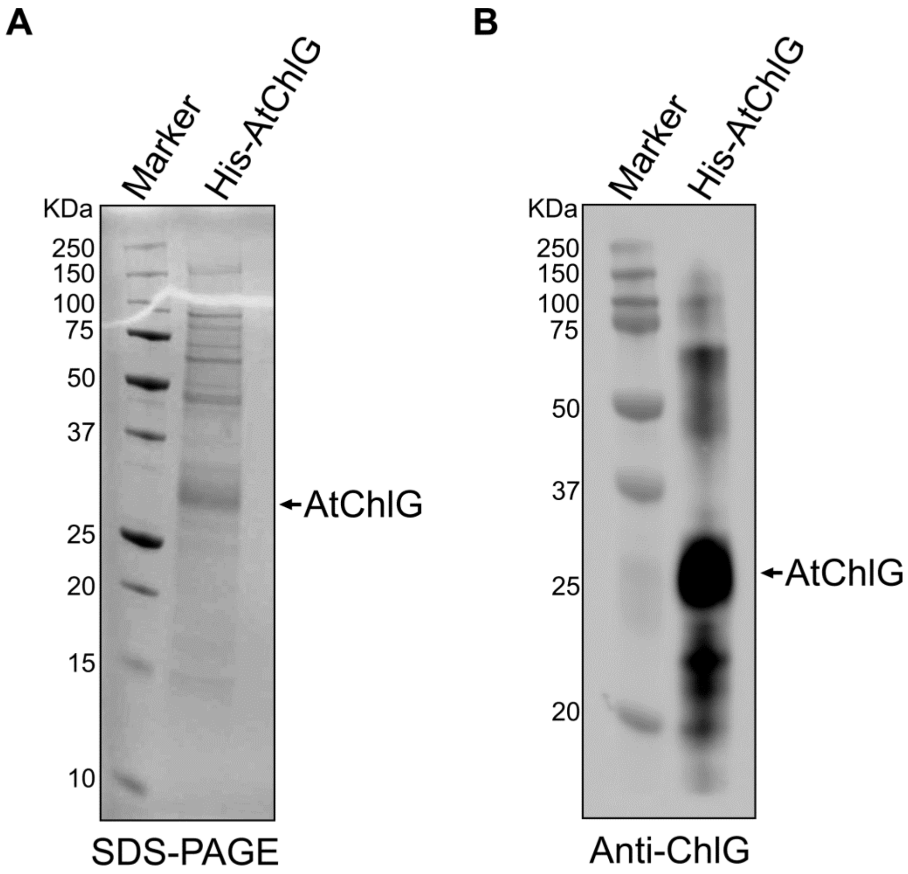


Figure 7.2: Purification of His-AtChlG by Ni^{2+} IMAC. Separation of AtChlG eluate by SDS-PAGE, followed by staining with Commassie Brilliant Blue (A) and immunoblotting using anti-ChlG antibodies (B).

7.3.3 Optimisation of AtChlG production

An attempt was made to increase the yield of AtChlG obtained per litre of culture by altering the growth temperature and medium. Lowering the temperature after induction has been reported to aid recombinant protein expression (San-Miguel *et al.*, 2013). A 1 L LB AtChlG culture was grown as described above except that the temperature was lowered from 37 °C to 18 °C after induction. AtChlG was purified from the cultures as described above and analysed by SDS-PAGE in comparison to a control sample maintained at 37 °C. The culture grown at 18 °C exhibited a significant increase in AtChlG production as represented by an intense band that contrasts with the comparably weak band visible in the 37 °C sample (Figure 7.3). Secondly, the culture medium, LB, was switched for LB autoinduction medium (Formedium™), which can increase cell densities and the yield of correctly folded recombinant protein (Sivashanmugam *et al*., 2009; Studier, 2005). Autoinduction medium places the induction of the target gene under the control of the host's metabolism at a level tolerable to the cell, avoiding the inhibition of cell growth often seen in IPTG induced cultures. The *lac* operon is initially repressed by glucose, but as the glucose is depleted, the cells switch to consumption of lactose and glycerol. Lactose breakdown produces allolactose, the native inducer of the *lac* operon, which relieves repression of the *lac* operon and enables target gene expression (Blommel *et al*., 2007). AtChlG was purified from an autoinduced 1 L LB culture, grown at 37 °C until the cell density reached an optical density of 0.6 when measured at 600 nm, at which point the temperature was lowered to 18 °C and the culture incubated overnight. The AtChlG eluate was analysed by SDS-PAGE (Figure 7.3). The yield of AtChlG increased slightly in autoinduction medium in comparison to the LB culture, so this method of AtChlG production was adopted for this work.

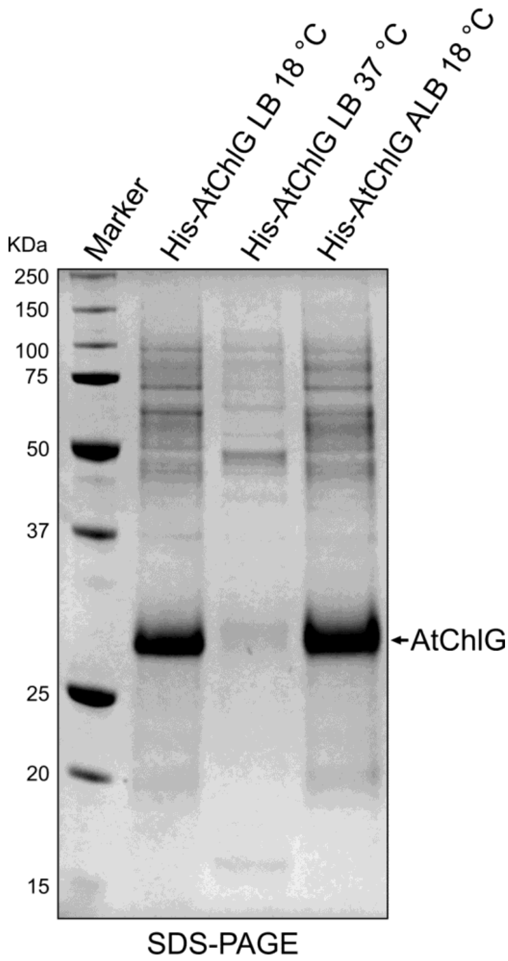


Figure 7.3: Purification of AtChlG produced in *E. coli* grown in varying media and temperatures. Recombinant AtChlG *E. coli* cells were grown in LB medium at 37 °C and 18 °C after IPTG induction and a third culture was grown in autoinduction LB medium (ALB) at 18 °C after the culture had become turbid. His-AtChlG was purified by IMAC and separated by SDS-PAGE before staining with Coomassie Brilliant Blue.

7.3.4 Generating a Chlide *a* producing *Rba. sphaeroides* mutant

A source of Chlide, the substrate of ChlG, was required in order to perform enzyme activity assays with the AtChlG enzyme. A *Rhodobacter* (*Rba.*) *sphaeroides* mutant lacking *bchC*, *bchX* and *bchF* (Δ*bchCXF*) is perturbed in the bacteriochlorophyll *a* (Bchl) specific steps of the Bchl biosynthesis pathway, i.e. modification of Chlide to BChlide (Figure 7.4A). This strain accumulates Chlide, which is excreted from the cells when the strain is cultured in growth medium supplemented with 0.02% (v/v) Tween 20

(Chidgey, 2014). However, this strain also produces an excess of Pchlide, necessitating the separation of Pchlide and Chlide by reverse-phase HPLC, and so it was beneficial to attempt to further improve the extent of Pchlide to Chlide conversion in this strain. Deletion of *bchF* in the original Δ*bchCXF* strain did not take into account the overlap between the 3' end of *bchF* and the start of the downstream *bchN* gene, which encodes protochlorophyllide reductase (DPOR), the enzyme responsible for converting Pchlide to Chlide (Figure 7.4B). Deletion of *bchF* could have perturbed the expression of *bchN*, for example by impeding RNA polymerase from binding to the truncated *bchN* promoter region which may be located within the *bchF* coding region. As such, a new Δ*bchCXF* strain (Δ*bchCXF**) was generated in which only the 5' end of *bchF* was deleted and a frame shift was introduced into the 3' end of the gene. Production of BchF in Δ*bchCXF** was therefore abolished whilst the upstream *bchN* region remained intact.

Truncation of *bchF* was achieved using the pK18mob*sacB* suicide vector system (Schäfer *et al*., 1994). This plasmid contains a kanamycin-resistance cassette and the gene *sacB*, encoding the enzyme levansucrase that catalyses the conversion of sucrose into toxic sugars, fructooligosaccharides, and thereby acts as a second selection marker. DNA complementary to the regions upstream of, and to the centre of, *bchF*, were cloned into pK18mob*sacB*, which was then transformed into the *E. coli* strain S17-1. Subsequently, the plasmid was transferred to a *Rba. sphaeroides* strain already lacking *bchC* and *bchX* (Δ*bchCX*) by bacterial conjugation (Chidgey, 2014). Selection of conjugants on kanamycin initiated the first homologous recombination event which incorporated the vector into the host genome (Figure 7.5A). These were then exposed to sucrose, which prompted a second homologous recombination event, resulting in either the desired Δ*bchCXF** mutant (Figure 7.5B), or, if the second recombination event occurs at the same site as the first, colonies that have reverted back to Δ*bchCX* (Figure 7.5C). Δ*bchCXF** mutants were identified by PCR screening (Figure 7.5D).

To test whether Δ*bchCXF** produced more Chlide than Δ*bchCXF*, both strains were cultured in 0.02% (v/v) Tween 20 and the pigments extracted from the growth medium as described in Section 2.14.3. Pigments were analysed by reverse-phase HPLC (Section 2.14.5), monitoring absorbance at 632 nm and 665 nm (data not shown) for PChlide and Chlide respectively and fluorescence at 670 nm (excited at 440 nm). Chlide eluted from the column after 20 minutes and Pchlide after 27 minutes as determined by their absorbance spectra. Unfortunately, Δ*bchCXF** produced a higher Pchlide to Chlide ratio (Figure 7.6B) than the original Δ*bchCXF* strain (Figure 7.6A). Calculating the area of each peak; Δ*bchCXF* produced a Pchlide:Chlide ratio of 1:1.5, whereas Δ*bchCXF** produced these pigments in a ratio of 2.5:1. There values were approximate, as the exact ratios of the pigments vary between independent cultures, however it can be concluded that the Δ*bchCXF** strain is less efficient in producing Chlide than Δ*bchCXF*. The experiment was pursued no further.

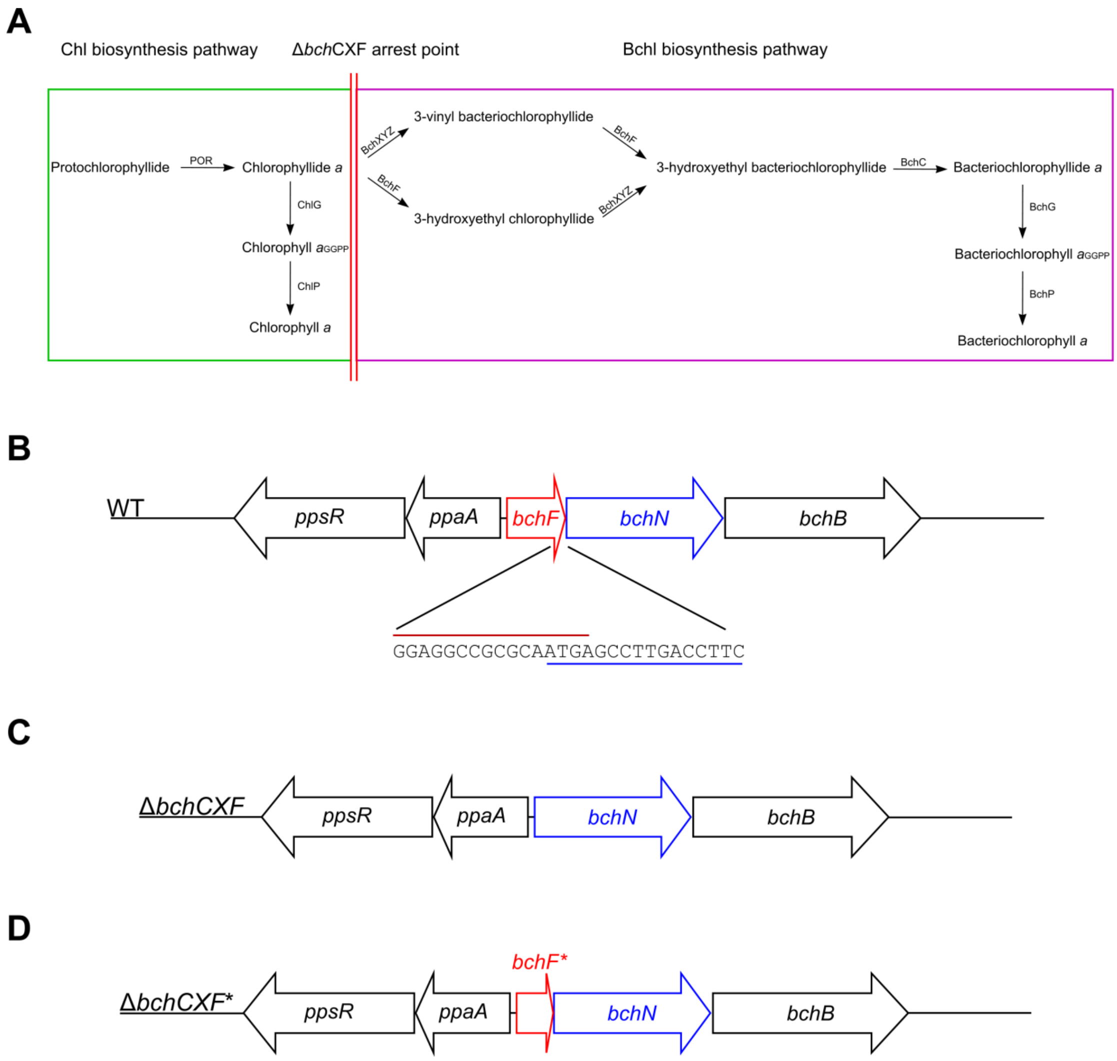


Figure 7.4: Bchl a biosynthesis is perturbed in a *Rba. sphaeroides* Δ*bchCXF* mutant. (A) The Bchl-specific steps (purple box) of BChl biosynthesis; *bchF* encodes 3-vinyl bacteriochlorophyll hydratase; BchXYZ are the three subunits of Chlide *a* reductase (COR) and *bchC* encodes 3-hydroxyethyl bacteriochlorophyllide dehydrogenase. *Rba. sphaeroides* Δ*bchCXF* and Δ*bchCXF** mutants are arrested in BChl biosynthesis (red lines). (B) Overlap of *bchF* (red) and *bchN* (blue) genes within the *Rba. sphaeroides* genome. (C) The *bchF* gene was completely deleted from the Δ*bchCX* mutant resulting in the original *Rba. sphaeroides* Δ*bchCXF* strain. (D) The 5' end of *bchF* was deleted from a Δ*bchCX* mutant resulting in the *Rba. sphaeroides* Δ*bchCXF** mutant.

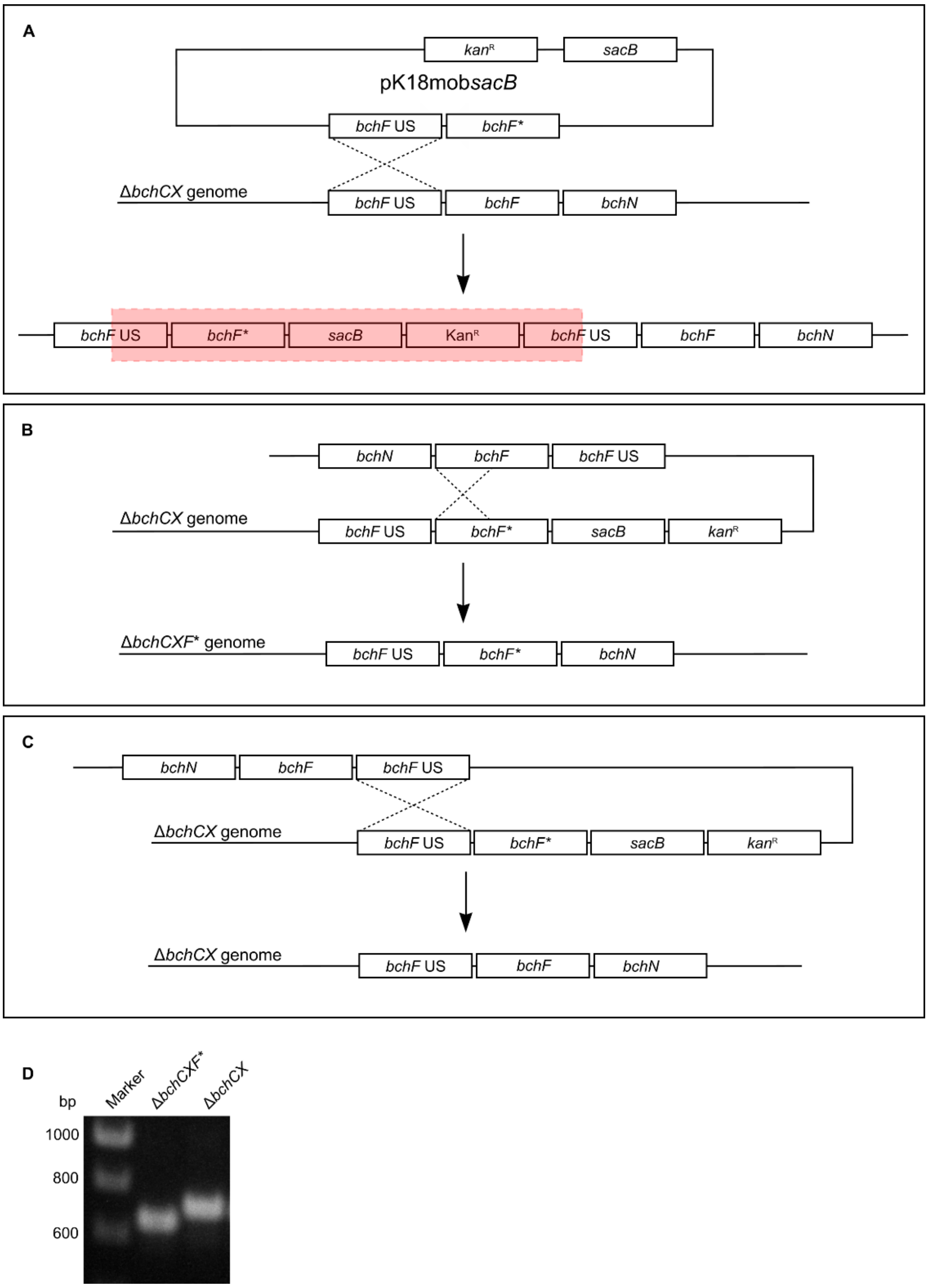


Figure 7.5: Truncation of *bchF* from the mutant *Rba. sphaeroides* Δ*bchCX* genome using the pK18mob*sacB* suicide vector. (A) The upstream region of *bchF* (*bchF* US) and N-terminal half of *bchF* (*bchF**) was ligated into the pK18mobsacB vector, followed by incorporation into the genome of *Rba. sphaeroides* Δ*bchCX* by kanamycin induced homologous recombination. (B) A second recombination event, induced by sucrose, removed the plasmid DNA and N-terminal part of *bchF* from the genome, resulting in mutant Δ*bchCXF**. (C) A secondary recombination event occurring at the same site as the first results in a reversion back to Δ*bchCX*. (D) Successful recombinants were confirmed by PCR screening.

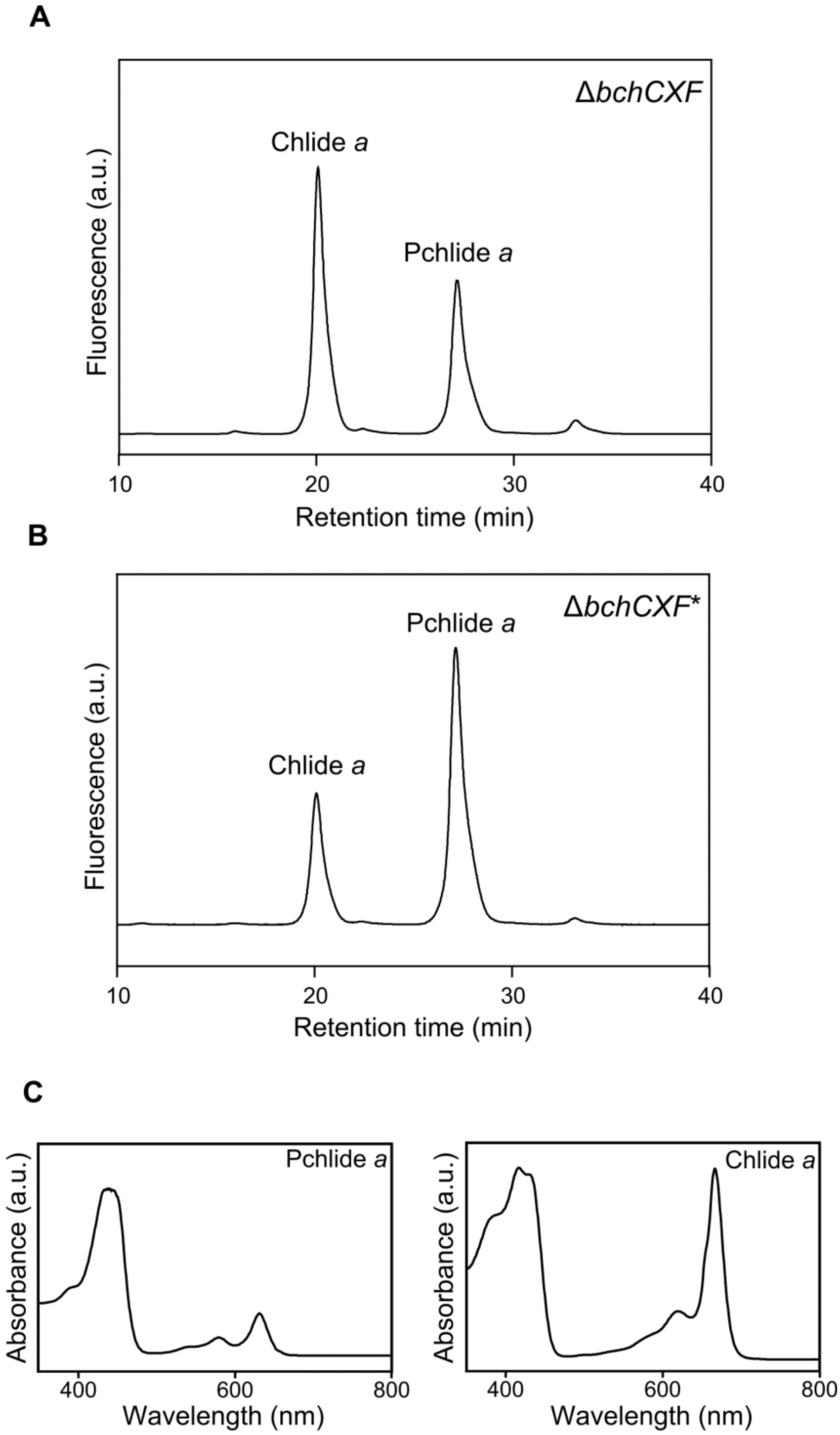


Figure 7.6: Reverse-phase HPLC separation of pigments extracted from mutant Δ*bchCXF* and Δ*bchCXF *Rba. sphaeroides* strains.** (A-B) Fluorescence profiles of Chlide *a* and Pchlide *a* pigments extracted from the Δ*bchCXF* (A) and Δ*bchCXF** (B) mutants. (C) Pchlide *a* (left panel) and Chlide *a* (right panel) were identified by their characteristic absorbance spectra.

7.3.5 Heterologous expression of *A. thaliana* chlorophyllase (CLH-1) in *E. coli*

Due to the difficulty in purifying Chlide *a* from the Δ*bchCXF Rba. sphaeroides* mutant, a potentially more efficient and less laborious approach was conceived whereby the phytyl tail is removed from Chl *a* using the plant enzyme chlorophyllase (CLH-1). A similar approach has been reported before where chorophyllase and Chl *a* were both prepared from the leaves of *A. altissima* (Oster *et al*., 1997). The chlorophyllase gene from *A. thaliana* (*CLH-1*) was codon optimised for expression in *E. coli* and artificially synthesised by Integrated DNA Technologies®**.** *CLH-1* was sub-cloned into pET21a and produced in *E. coli* BL21 (DE3) following the method described by Tsuchiya *et al.* (1999). Although masked by a native *E. coli* protein of a similar size, the intensity of a band on stained SDS-PAGE at approximately 35 kDa, the predicted molecular mass of CLH-1, was increased in the CLH-1 producing cells in comparison to a control sample (Figure 7.7A). CLH-1 production was confirmed by immunodetection of the His-tagged version of the protein (Figure 7.7B).

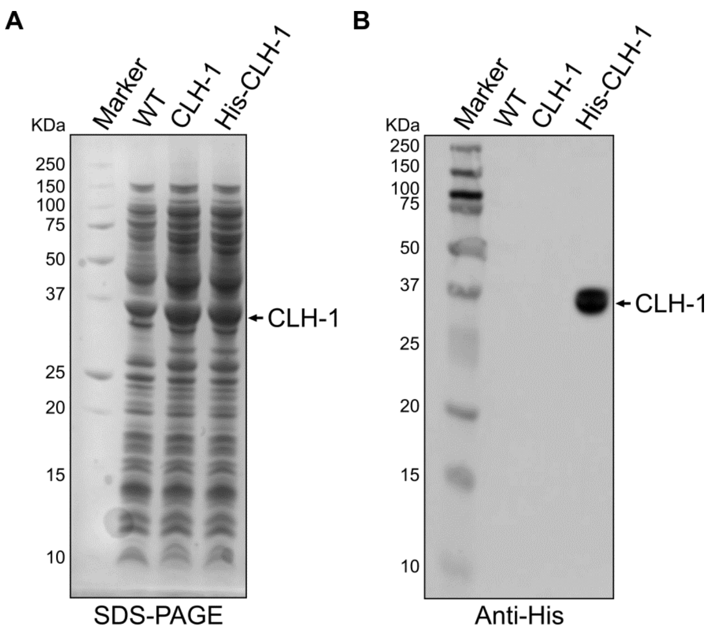


Figure 7.7: Production of *A. thaliana* chlorophyllase-1 in *E. coli*. (A) SDS-PAGE separation and Coomassie Brilliant Blue staining of recombinant *E. coli* lysates producing His-tagged and non His-tagged CLH-1. (B) CLH-1 was production was confirmed by immunoblot with anti-His antibodies.

7.3.6 Production of Chlide *a* from Chl *a* by CLH-1 containing *E. coli* lysates

Attempts to purify CLH-1 to homogeneity were unsuccessful as the protein yield was not sufficient (data not shown), however, significant dephytylase activity was present in total cell lysates. Clarified lysates were prepared from CLH-1 producing *E. coli* cells and added to Chl *a*, isolated from *Synechocystis*, as described in Section 2.14.1. At the end of the assay, pigments were extracted and analysed by reverse-phase HPLC (Section 2.14.4). A major peak eluted after 12 minutes (Figure 7.8B) and was confirmed to be Chlide by its absorbance spectra and its absence in control assays lacking CLH-1, in which only Chl *a* was present (Figure 7.8A). The results show complete conversion of Chl *a* to Chlide *a* by CLH-1, demonstrating the promise of this method as a means to produce Chlide for use in AtChlG assays. However, Chlide produced this way was not stable and degraded extensively during storage, as described later.

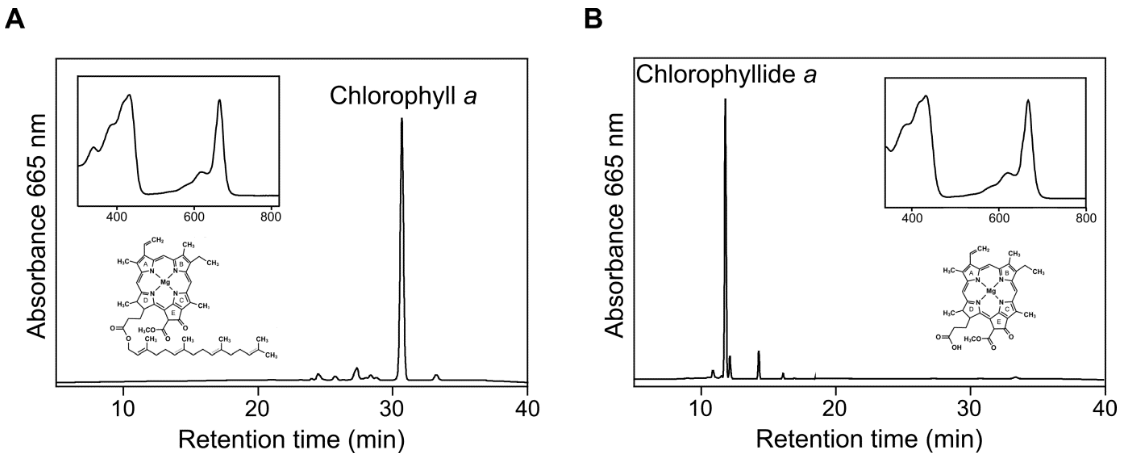


Figure 7.8: Reverse-phase HPLC separation of pigments extracted from *Synechocystis* and CLH-1 enzyme assays. (A) Chl *a* was extracted from *Synechocystis* and analysed by reverse-phase HPLC. (B) Chl *a* was converted to Chlide *a* by reaction with CLH-1. The absorbance spectra of the two pigments are shown boxed in their respective panels.

7.3.7 Enzyme activity of AtChlG solubilised in detergent

In order to test whether the recombinant AtChlG is active in detergent micelles, the *E. coli* membranes were solubilised in β-DDM and AtChlG purified by Ni^{2+} IMAC. The eluate was buffer exchanged into FLAG buffer (25 mM Na-phosphate, 50 mM NaCl, 10 mM MgCl, 10% glycerol, pH 7.4), and tested for ChlG activity. Enzyme assays were performed in triplicate and consisted of 20 µM of AtChlG, 30 µM Chlide (produced by CLH-1 assay), 100 µM GGPP and 0.01% (w/v) βDDM diluted to 30 µL in FLAG buffer as described by Chidgey, 2014. Assays were started by addition of Chlide and incubated for 60 minutes in the dark at 30 °C before being stopped by addition of 60 µL of MeOH. The pigments were extracted from the protein precipitate in excess MeOH and immediately analysed by reverse phase HPLC. A positive control, consisting of purified FLAG-ChlG complex from *Synechocystis*, and a negative AtChlG control assay lacking GGPP were performed.

The FLAG-ChlG complex produced two peaks attributed to Chl a_{GG}, which was formed during the assay, and Chl *a*, which was bound to the HliD component of the ChlG complex and therefore is already present in the assay mixture. The abosorbance spectra of Chl *a* and Chl a_{GG} are identical, however, the pigments can be distuingshed from one another by the difference in elution time between the two from the HPLC column. Chl a_{GG} eluted after 22.5 minutes and Chl *a*, the more hydrophobic of the two pigments, after 30 minutes (Figure 7.9A). The identity of these two peaks was confirmed using the same HPLC program to analyse Chl *a*, purified from *Synechocystis*, and Chl a_{GG}, extracted from a *Synechocystis* Δ*chlP* mutant lacking geranyl geranyl diphosphate reductase (GGPP) (Hitchcock *et al*., 2016). Chl *a* eluted after 30 minutes, as was the case in the FLAG-ChlG assay. Chl a_{GG} eluted slightly later than expected, after 24 minutes (Figure 7.9B); there could be slight differences in the chemical structure of the Chl a_{GG} species produced in the two systems, although the exact nature of these changes has yet to be resolved.

AtChlG produced no Chl a_{GG} indicating that the enzyme is not active when solubilised in detergent (Figure 7.9A). It was also confirmed that WT *E. coli* lysate has no enzyme activity that can esterify Chlide *a* with GGPP.

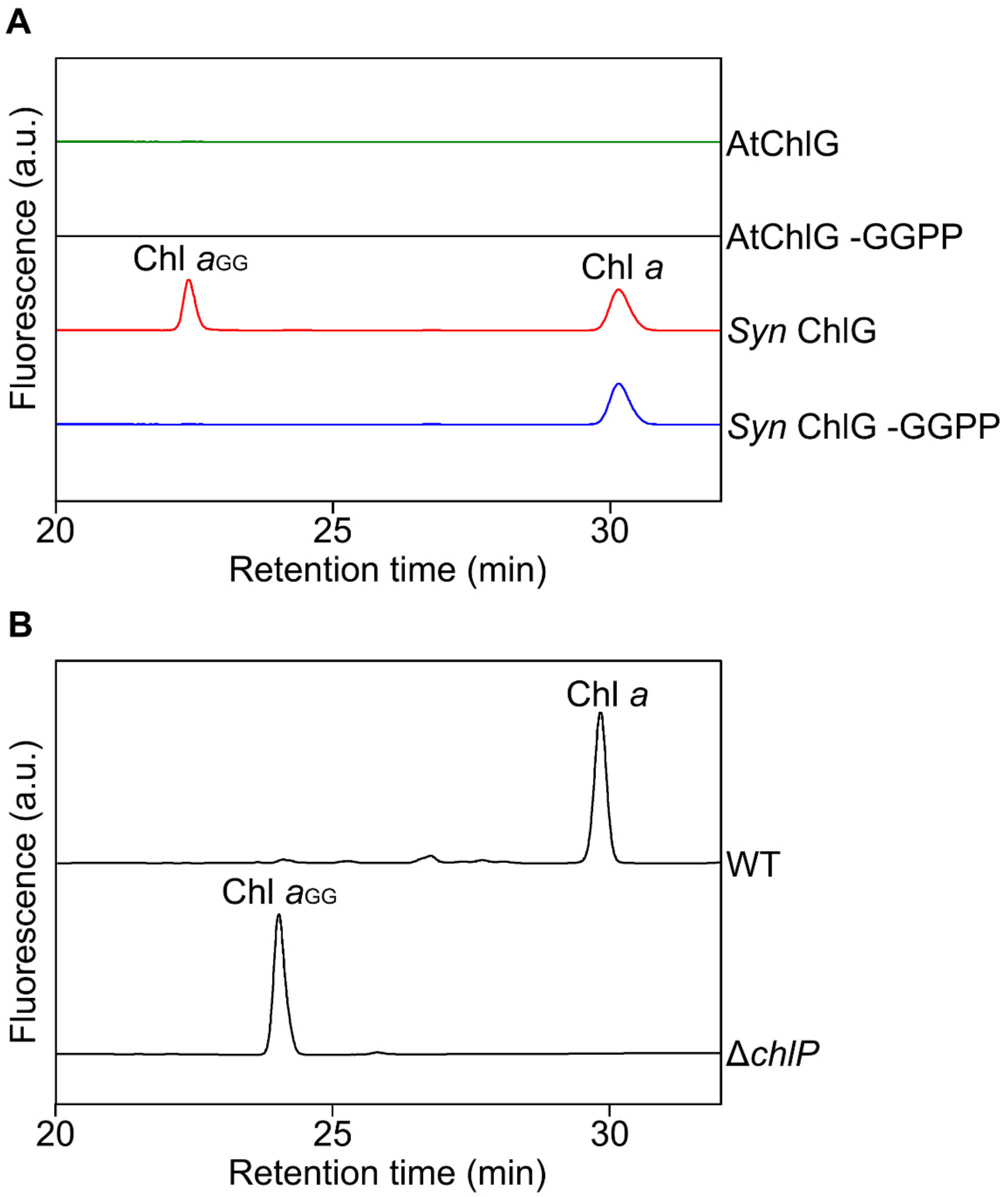
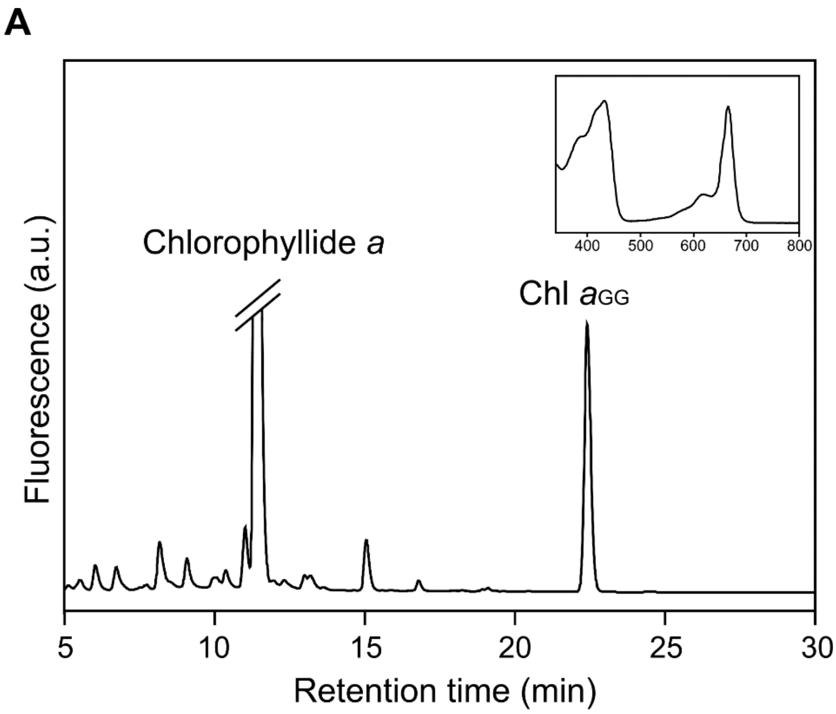


Figure 7.9: Reverse-phase HPLC separation of pigments extracted from purified ChlG assays. (A) Pigments extracted from enzyme assays using purified AtChlG eluate (AtChlG), *Synechocystis* FLAG-ChlG pulldown eluate assays (*Syn*ChlG) and from *Synechocystis* ChlG (*Syn* ChlG -GGPP) and AtChlG (AtChlG -GGPP) negative control assays lacking GGPP. (B) Pigments extracted from WT *Synechocystis* (WT) and *Synechocystis* Δ*chlP* (Δ*chlP*) were used as standards.

7.3.8 Enzyme activity of recombinant AtChlG *E. coli* lysate

As AtChlG was not active when extracted from the lipid bilayer of *E. coli* membranes, clarified lysates were tested for enzyme activity. The cells were lysed in FLAG buffer and centrifuged at 8000 x*g* for 20 minutes to remove insoluble material. Chlide (30 µM) and GGPP (100 µM) were added to 27 µL of lysate and incubated for 60 mins at 30 °C. The pigment content of the assays were analysed as described previously (Figure 7.10A) and a negative control lacking GGPP (-GGPP) was included (Figure 7.10B).

The majority of the pigment within the assay was Chlide *a*, as indicated by the peak at 12 minutes in both the -GGPP control and AtChlG sample. However the AtChlG sample also contained a pigment that eluted after 22.5 minutes that was not present in the – GGPP control, attributed to be Chl a_{GG} as surmised from the absorbance spectra. The Chl a_{GG} peak was relatively small compared to the Chlide *a* peak. Calculating the integrals of the Chlide *a* and Chla_{GG} absorbance peaks at 665 nm across three separate assays, just 1.4% of the Chlide *a* was converted to Chl *a* by AtChlG. Attempts were made to improve Chl a_{GG} production by incremental increases in the assay solvent (MeOH) concentration, halving the Chlide *a* concentration to 15 µM, purifying AtChlG membranes on sucrose gradients and by assaying AtChlG membranes solubilised in 1.5% (w/v) βDDM. None of these changes increased Chl a_{GG} production and some, such as solubilising the AtChlG membranes in detergent, abolished the enzyme activity completely (data not shown).



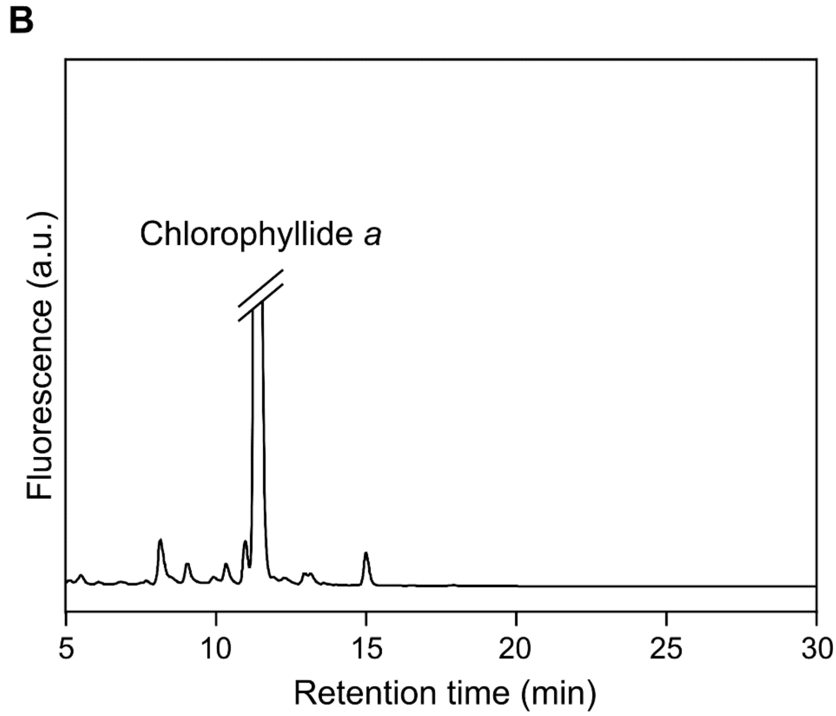


Figure 7.10: Reverse-phase HPLC separation of pigments extracted from AtChlG assay. Pigments extracted from enzyme assays using recombinant AtChlG *E. coli* lysates (A) and a negative control assay (B) lacking the substrate GGPP were analysed by reverse-phase HPLC. The absorbance spectra of Chl a_{GG} is shown boxed in (A).

7.3.9 Mutation of conserved ChlG residues to their BchG equivalents

Despite the fact that purified AtChlG is inactive in detergent micelles, the demonstration that AtChlG activity is measurable within clarified *E. coli* lysates presented an opportunity to generate point mutant variants of this enzyme and test them for activity *in vitro*. Residues that are conserved in homologs of ChlG, but are replaced with a different conserved residue in BchG (Figure 7.11 and 7.12), were targeted for mutation to test the hypothesis that these residues are either important for determining the substrate specificity of ChlG for Chlide, and/or are important for the general activity of the enzyme. Although there are many such examples of residues that fit these criteria, three amino acids within AtChlG were selected due to their predicted location within the first transmembrane helix of the enzyme. These residues were substituted by their counterparts in BchG. The AtChlG residues, and the equivalent BchG residue, were: glutamine 46 by glutamic acid (Q46E), leucine 56 by proline (L56P) and valine 60 by tyrosine (V60Y) (Figure 7.11).

In addition, the proline residue at position 54 in AtChlG was substituted by a phenylalanine residue (P54F). The equivalent residue in *Synechocystis* ChlG (I44) and BchG (F28) has been reported to be involved in the substrate specificity of ChlG and BchG for Chlide *a* and BChlide *a* respectively (Kim *et al*., 2016), as discussed previously in Section 1.8.11 and 7.2.

Two more AtChlG residues were targeted based on sequence alignments of ChlG and BchG with a related protein, ubiquinone synthase (UbiA) (Figure 7.13). UbiA catalyses the condensation of isoprenylpyrophosphate (IPP) with the aromatic molecule *p*-hydroxybenzoate (PHB). This reaction is analogous to reaction catalysed by ChlG, the esterification of Chlide with GGPP. The structure of this enzyme has been solved by overproduction of UbiA in *E. coli* followed by purification in detergent micelles and X-ray crystallography (Figure 7.13B) (Cheng and Li, 2014). Using this structure, the authors identified many UbiA residues predicted to be essential for activity due to their location within the enzyme active site (Figure 7.13). Substitution of each of these residues for Ala, followed by enzyme assays using *E. coli* lysates, showed that all mutants exhibited perturbed enzyme activity to varying degrees. Among these, N50A abolished UbiA activity completely and reduced binding of PHB and geranyl thiolopyrophosphate (GSPP), an analogue of IPP, to approximately 40% and 3% the levels of the WT respectively. The equivalent residue in AtChlG, N99, is conserved in ChlG and BchG homologs and, like N50 of UbiA, was predicted to be involved in facilitating the binding of GGPP to AtChlG (Figure 7.13B). This residue was subsequently substituted by an Ala residue (N99A). D175 of UbiA is located close to the binding site of GSPP. Substitution of this residue for Ala also resulted in an inactive protein and reduced binding of GSPP to UbiA to approximately 5% WT levels (Cheng and Li, 2014). The equivalent residue in AtChlG is A225, which is conserved in ChlG homologs but is replaced by a conserved Met residue in BchG homologs (Figure 7.11, 7.12 and 7.14). A model of AtChlG was constructed using the crystal structure of UbiA as a template (Figure 7.14). The A225 residue was located in the putative GGPP binding pocket of the model, akin to D175 of UbiA, and was therefore substituted by the equivalent Met residue of BchG (A225M).

Plasmids encoding the 6 variant enzymes; Q46E, P54F, L56P, V60Y, N99A and A225M, were generated using the pET28a::AtChlG plasmid as template with the Quickchange II site directed mutagenesis kit from Agilent. The sequence-verified plasmids were transformed into BL21 and clarified cell lysates were prepared as described for the 'wild-type' AtChlG enzyme. For each variant, the lysates were diluted to the same total protein concentration in FLAG buffer and production of enzyme was confirmed by immunoblot (Figure 7.15). Several immunoblots were performed in an attempt to check the parity of the AtChlG concentration between the lysates. However, there was variation in the AtChlG signal intensities produced by the same samples between the different immunoblots, making it difficult to estimate the relative levels of AtChlG. Although AtChlG levels were approximately consistent between the mutants a more accurate method of normalising AtChlG concentrations must be developed.

```
Rs      -------------------------------------------MSVNLSL---HPRSVPE
Cr      --------------------------MAMNQQATEEKSDTNSAARQMLGMKGAA-LETDI
At      -------------------MAAETDTDKVKSQTPDKAPAGGSSINQLLGIKGAS-QETNK
Am      MANSDPSQ----VPASADNTEATATPSEPSAVEASEATEQGSAARQLLGMKGADTGDTNI
Te      MEESE---RTIIIEVLLSTTSPMTETPDSTTT-STSSESTTAAARQLLGMKGAKSGETNI
6803    ------------------------------MSDTQNTGQNQAKARQLLGMKGAAPGESSI
7002    MPNDEWFSFTLFFCTKYSNPDAMTETPNPDTKPATAPEEQGSKARQLLGMKGAAGGETSI
                                                      *.:      .

Rs      PRALLELIQPITWFPPIWAYLCGTVSVGI--WPGE-KWPLVLLGMVLAGPLVCGMSQAAN
Cr      WKIRVQLTKPVTWIPLIWGVACGAAASGHYQWNNPTQIAQLLTCMMMSGPFLTGYTQTIN
At      WKIRLQLTKPVTWPPLVWGVVCGAAASGNFHWTPE-DVAKSILCMMMSGPCLTGYTQTIN
Am      WKIRVQLMKPITWVPLVWGVICGAAASGHFTWTLE-HVLMSAVCMVMSGPLLTGYTQTIN
Te      WKIRLQLMKPITWIPLIWGVVCGAASSGGFTWSLE-DILKAATCMLLSGPLMAGYTQTLN
6803    WKIRLQLMKPITWIPLIWGVVCGAASSGGYIWSVE-DFLKALTCMLLSGPLMTGYTQTLN
7002    WKIRLQLMKPITWIPLIWGVVCGAASSGGYVWGVE-DFLKAMTCMLLSGPLLTGYTQTLN
        :  ::* :*:** * :*.  **:.: *   *    .       *:::** : * :*: *

Rs      NWCDRHVDAVNEPDRPIPSGRIPGRWGLYIALLMTVLSLAVGWMLGPWG-------FGAT
Cr      DWYDREIDAINEPYRPIPSGRISERDVIVQIWVLLLGGIGLAYTLDQWAGHTTPVMLQLT
At      DWYDRDIDAINEPYRPIPSGAISEPEVITQVWVLLLGGLGIAGILDVWAGHTTPTVFYLA
Am      DFYDREIDAINEPYRPIPSGAISIPQVVTQIFVLLGAGIGVAYGLDRWAGHEFPTLTVLA
Te      DYYDREIDAINEPYRPIPSGAISLNQVRAQIIFLLVAGLTLAVLLDLWSDHATFPVTKIA
6803    DFYDRDIDAINEPYRPIPSGAISVPQVVTQILILLVAGIGVAYGLDVWAQHDFPIMMVLT
7002    DFYDREIDAINEPYRPIPSGAISVPQVVTQILVLLGSGIGLSYLLDLWAGHDFPVMLVLT
        :: **.:**:*** ****** *          .:   .: :.  *. *.          :

Rs      VFGVLAAWAYSVEPIRLKRSGWWGPGLVALCYEGLPWFTGAAVLSAGAPSFFIVTVALLY
Cr      IFGSFISYIYSAPPLKLKQSGWAGNYALGSSYIALPWWAGQALFGTLTLDVM--ALTIAY
At      LGGSLLSYIYSAPPLKLKQNGWVGNFALGASYISLPWWAGQALFGTLTPDVV--VLTLLY
Am      IFGSFISFIYSAPPIKLKQNGWTGNFALGASYIALPWWAGQALFGTLTPKVM--VLTLAY
Te      LLGGFLAYIYSAPPLKLKKNGWLGNYALGASYIALPWWAGHALFGELTPTIV--ILTLIY
6803    LGGAFVAYIYSAPPLKLKQNGWLGNYALGASYIALPWWAGHALFGTLNPTIM—VLTLIY
7002    VVGCFIAYIYSAPPLKLKQNGWLGNYALGASYIALPWWAGHALFGTLTPTVM--VVTLIY
        : * : :: **. *::**:.** *   :. .* .***::* *::.     ..   ::: *

Rs      AFGAHGIMTLNDFKALEGDRQHGVRSLPVMLGPEVAAKLACTVMAMAQILVITLLVIWG-
Cr      SLAGLGIAIVNDFKSIEGDRQMGLQSLPVAFGVDTAKWICVSTIDVTQLGV-AAYLAWGL
At      SIAGLGIAIVNDFKSVEGDRALGLQSLPVAFGTETAKWICVGAIDITQLSV-AGYLL-AS
Am      SLSGLGIAIINDFKAVEGDRELGLKSLPVVFGIEKAAWICVLMIDVFQIGM-ALFLI-SI
Te      SLAGLGIAIVNDFKSVEGDRQLGLASLPVMFGITTAAWICVLMIDIFQLGI-AGYLM-AI
6803    SLAGLGIAVVNDFKSVEGDRQLGLKSLPVMFGIGTAAWICVIMIDVFQAGI-AGYLI-YV
7002    SFAGLGIAVVNDFKSVEGDRQLGLKSLPVMFGVGTAAWICVLMIDIFQVGI-AGYLV-SI
        ::.. **  :****::****  *: **** :*   *  :.   : : *  : :  :

Rs      -KPIHAGIITALLVAQ-LFAMRVLLRDPAGKCPWYNGTGVTLYVLGMMVAAFAIRGLEVL
Cr      HEELYGAVLLALILPQIYFQYKYFLPDPIANDVKYQASAQPFLVFGLLTAGLACGHHVNA
At      GKPYYALALVALIIPQIVFQFKYFLKDPVKYDVKYQASAQPFLVLGIFVTALASQH----
Am      HENLYAVILFLLVIPQITFQDMYFLRDPVKNDVKYQASAQPFLVLGTLVAGLAIGHAGL-
Te      HANLYAVLLILLIIPQIVFQDMYFLRDPLKNDVKYQASAQPFLVLGMLVVGLALGHTLV-
6803    HQQLYATIVLLLLIPQITFQDMYFLRNPLENDVKYQASAQPFLVFGMLATGLALGHAGI-
7002    HEQLYATILLLLVIPQITFQDMYFLRDPLKNDVKYQASAQPFLVLGMLVAGLAMGHAGIS
         :.  :  *:: *  *    :* :*      *:.:.  : *:* :...:*

Rs      P----------
Cr      VAAAASAAGAL
At      -----------
Am      -----------
Te      -----------
6803    -----------
7002    -----------
```

Figure 7.11: Alignments of ChlG homologs with BchG from *Rba. sphaeroides*. ChlG sequences from *Chlamydomonas reinhardtii* (Cr), *Arabidopsis thaliana* (At), *Acaryochloris marina* (Am), *Thermosynechococcus elongatus* BP-1 (Te), *Synechocystis* sp. PCC 6803 (6803), *Synechococcus* sp. PCC 7002 (7002) and BchG from *Rhodobacter sphaeroides* were aligned using Clustal Omega. The amino acids substituted in the variant AtChlG proteins are colour coded as follows; Q46E (yellow), P54F (grey), L56P (blue), V60Y (purple), N99A (red) and A225M (green). The Rs sequence is representative of BchGs; the residues are conserved in numerous BchG proteins (see Figure 7.12).

```
Ca      MSDMSDQTRLSSPPSLPLHKQPQSRYAWLVRSIQLMKPVTWFAPTWAFMCGAIASGALGW
Rs      ----MSVNLSLHPRSV-----P-----EPRALLELIQPITWFPPIWAYLCGTVSVGIWPG
Rc      ---MSAQDLSPSRRSI-----P-----EPRAMLELIKPVTWFPPMWAYLCGAVSSNVPIW
Rp      ---------MSNSVAVR-----P-----APSAVLEVLHPITWFPPMWAFGCGVVSSGVPIS
Rr      --------------MQR-----T-----AVLPYVQLLKPITWFAPMWAFGCGLISSGLPVW
                                        :::::*:*** * **: ** :: .

Ca      NESIGRLLLGMFMAGPILCGLSQVVNDYADREVDAINEPHRLIPSGQVSLRHVYILTAVL
Rs      EK-WPLVLLGMVLAGPLVCGMSQAANNWCDRHVDAVNEPDRPIPSGRIPGRWGLYIALLM
Rc      EN-KGVVVLGIVLAGPIVCGMSQAANDWCDRHVDAINEPHRPIPSGRIPGLWGLYIAIAM
Rp      SR-WVEVIAGIVLCGPLLVATSQVVNDWFDRDVDAINEPNRPIPSGRIPGRWGLYLSFLW
Rr      DR-WPVIALGVLLCGPLVCGTSQAVNDWFDRHVDAINEPDRPIPSGRIPGRVGLYIAIGW
        ..    :  *:.:.**:: . **..*:: **.***:***.* ****::       ::

Ca      TWIGASIALFLGRQVAFFVALGLVFALAYSLRPIRGKRNGWIGNALVAISYEGLAWMAGH
Rs      TVLSLAVGWMLGPWGFGATVFGVLAAWAYSVEPIRLKRSGWWGPGLVALCYEGLPWFTGA
Rc      SLLSLVVGWQLGSWGFVATLLGVAAAWAYSVEPIRLKRSGWWGPGLVGLAYEGLPWITGA
Rp      TAASLLVASQLGAWVFGAAVLGLVLAWMYSMPPFRLKQNGWLGNGACAITYEGFAWFTGA
Rr      TVLSLAVAWVLGPWVFGAAIFGLALAWAYSAPPFRLKGNGWWGNSAVGLCYEGLPWFTGA
        :  .  :.  **      . :*:  *  **  *:* * .** * .  .: ***: *::*

Ca      AAFAPLT-----GESVTIALLYSLGAHGIMTVNDFKSIRGDTIMGIRSIPVQYGKVMAAR
Rs      AVLSA---GAPSFFIVTVALLYAFGAHGIMTLNDFKALEGDRQHGVRSLPVMLGPEVAAK
Rc      AVLLATADTSPGFPIVMMATLYALGAHGIMTINDFKAIEGDRKLGIKSLPAVYGPEVAAK
Rp      AVMLG---GLPPWWIVTLALLYSAGAHGIMTLNDFKSIEGDIKTGVGSLPVKLGVDNAAR
Rr      AVIAG---ALPDTRIVLLAVLYSIGAHGIMTLNDFKAVEGDIRMGVRSLPVQLGVAPAAR
        *.:            * :* **: *******:****::.**   *: *:*.  *   **:

Ca      MVVTTMGVAQIAVIGLLFHWGHPVAATVVAILLAAQSIPNARFIRDPENNEVFFNATAIM
Rs      LACTVMAMAQILVITLLVIWGKPIHAGIITALLVAQLFAMRVLLRDPAGKCPWYNGTGVT
Rc      IACTVMGLAQALVITMLYLFSKPYHATAVLVLLCGQFWAMSVWMRDPEGKAPWYNGTGVV
Rp      VLCAVMAIPQVIVVALLLSWDRPIQAGIVGFVLAVQLALMVRFLRSPVERATWFSGLGVA
Rr      LACAVMAVPQMVVIGLVASWDRPYHAGAVGLLLLAQFVLMARLLKRPRELAPWYNATGTT
        :  :.*.: *  *: ::  :.:*  *  :  :*  *       :: *     ::.. .

Ca      LYVWGMLAAAIGLAA---------------
Rs      LYVLGMMVAAFAIRGLEVLP----------
Rc      MYVSGMMITAFAIRGFTV------------
Rp      LYVSGMMASAVAVSSFGAA-----------
Rr      LYVIGMMVSAFALAPLIGAVELAPLIGAAP
```

Figure 7.12: Alignments of various BchG homologs from purple bacteria. BchG sequences from *Chloroflexus aurantiacus* (Ca), *Rba. sphaeroides* (RS), *Rba. capsulatus* (Rc), *Rhodopseudomonas palustris* (Rp) and *Rhodospirillum rubrum* (Rr) were aligned using Clustal Omega software. Conserved residues targeted for mutation in the corresponding AtChlG sequence are coloured as described in Figure 7.11.

A

```
UbiA      ------------------------------------------MRLVRI-EHTIFSLPFAY
At        MAAETDTDKVKSQTPDKAPAGGSSINQLLGIKGASQETNKWKIRLQLTKPVTWPPLVWGV
Rs        ----------------------MSVN----LSLHPRSVPEPRALLELIQPITWFPPIWAY
                                                      *      *     :.

UbiA      VGALLSRYP--FTLAD--AILMAAAVVGLRMAG--MAYNNIADLDIDRLNPRTAKRPLVV
At        VCGAAASGNFHWTPEDVAKSILCMMMSGPCLTGYTQTINDWYDRDIDAINEP--YRPIPS
Rs        LCGTVSVGI--WPGEKWPLVLLGMVLAGPLVCGMSQAANNWCDRHVDAVNEP--DRPIPS
          : .  :     :   .    ::   : *  : *   : *:  * .:* :*     **:

UbiA      GAVSLRE----AWALVAAGSAIYFASAALLNTYALLLSPLVLAIALTYPHA---KRLHPL
At        GAISEPEVITQVWVLLLGGLGI----AGILDVWAGHTTPTVFYLALGGSLLSYIYSAPPL
Rs        GRIPGRWGLYIALLMTVLSLAV----GWMLGPWG-------FGATVFGVLAAWAYSVEPI
          * :        .  :   . .:    . :*. :.       :  ::            *:

UbiA      PHLHLGIVLGSVVFGGAVAASGDEASSLGEVLRSVPWLYVAAV-----------------
At        KLKQNGW---------------VGNFALGASYISLPWWAGQALFGTLTPDVV--VLTLLY
Rs        RLKRSGW---------------WGPGLVALCYEGLPWFTGAAVLSAGAPSFFIVTVALLY
             : *                     :.    .:**    *:

UbiA      SLWVAGFDTIYSIMDIDFDRSHGLGSIPALLGPKGALAASLAMHAAAVALFIAGVEAYGL
At        SIAGLGIAIVNDFKSVEGDRALGLQSLPVAFGTETAKWICVGAIDITQLSVAGYLLASGK
Rs        AFGAHGIMTLNDFKALEGDRQHGVRSLPVMLGPEVAAKLACTVMAMAQILVITLLVIWGK
          ::   *:  : .:  :: **  *: *:*. :* : *   .      :   .   :   *

UbiA      GAIATVSTALTALVII-LVQAMAWLGRVKESFN--LN-LAVPIIIGAGIIVDMLH-HMIR
At        PYYAL---ALVALIIPQIVFQFKYFLKDPVKYDVKYQASAQPFL-VLGIFVTALASQH--
Rs        PIHAG---IITALLVA-QLFAMRVLLRDPAGKCPWYNGTGVTLY-VLGMMVAAFAIRGLE
            *     :.**::   :  :  : :         :  .  :    *::*  :  :

UbiA      LL-
At        ---
Rs        VLP
```

B

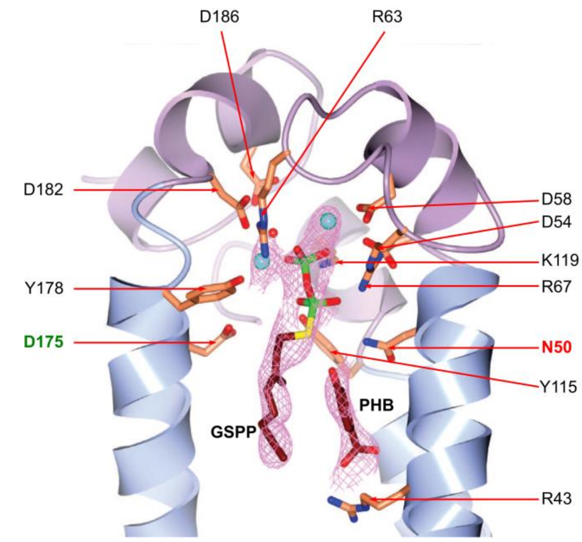


Figure 7.13: Sequence alignments of UbiA, AtChlG and BchG. (A) The amino acid sequences of UbiA from *Aeropyrum pernix, AtChlG from A. thaliana (At) and BchG from Rba. sphaeroides* (Rs) were aligned using Clustal Omega software. The residues corresponding to AtChlG N99 (red) and A225 (green) are shown. UbiA residues targeted for mutation by Cheng and Li (2014) are in bold. (B) The crystal structure of UbiA containing substrates *p*-hydroxybenzoate (PHB) and geranyl thiolopyrophosphate (GSPP) ((Cheng and Li, 2014). Red arrows highlight residues within the active site targeted for mutation by Cheng and Li (2014); D175 (green) and N50 (red) are highlighted in bold.

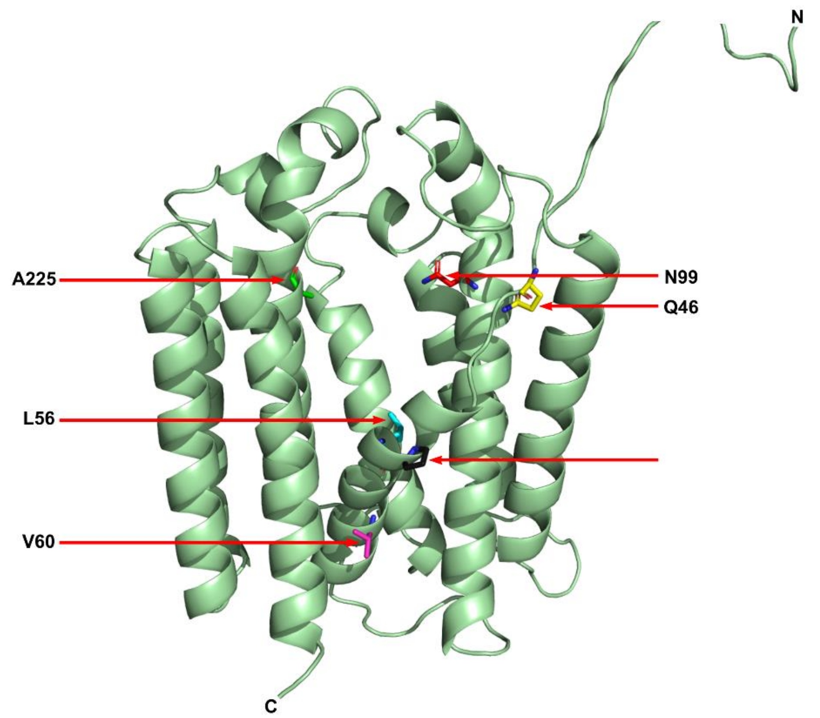


Figure 7.14: Structural model of AtChlG. A structural model of Arabidopsis thaliana ChlG (AtChlG) was generated using HHPred software and the published crystal structure of Aeropyrum pernix UbiA as a template. The residues targeted for substitution in this study are indicated (red arrows).

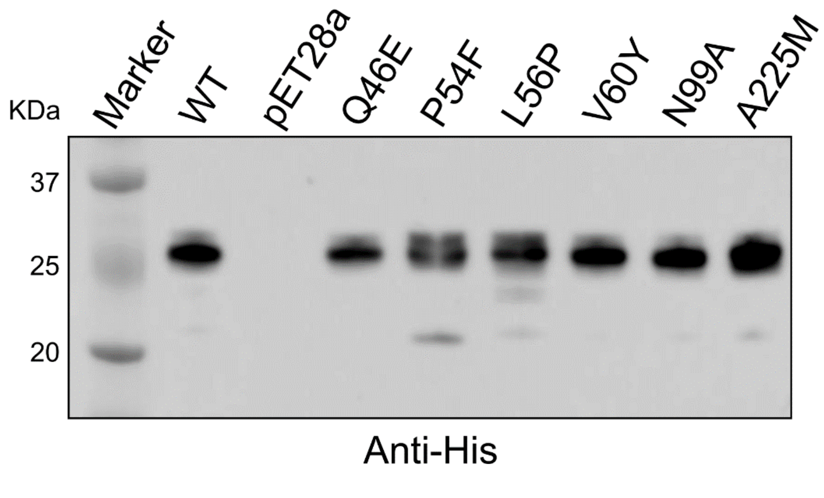


Figure 7.15: Immunoblot of variant AtChlG proteins. Immunoblot analysis of equal loadings of cell lysates prepared from *E. coli* strains producing recombinant AtChlG variants probed with antibodies raised against the His-tag epitope.

7.3.10 Enzyme activities of AtChlG point mutants

AtChlG lysates were assayed in triplicate and analysed by reverse-phase HPLC as described above. The results showed that WT AtChlG was the most active enzyme, producing the largest Chl a_{GG} peaks (Figure 7.16A). The Chl a_{GG} produced in all of the functional assays eluted as several distinct peaks, the intensity of which varied between the mutants (Figure 7.16A, C, D, E). This is in contrast to the earlier WT AtChlG assay where ChlG a_{GG} eluted mostly as a single peak after 22.5 minutes (Figure 7.10A). The spectra of the peaks in these assays are all characteristic of Chl *a*; however, the hydrophobicities of these Chl *a* species are slightly different, indicated by elution from the column after different periods of time. The usual peak at 22.5 minutes was produced in each of the functional assays, but additional peaks at approximately 23.5, 24 and 25 minutes were also produced (Figure 7.16A, C, D, E). This is likely due to the degradation of the Chlide substrate which also eluted as several peaks, besides the main peak at 12 minutes (Figure 7.17A). The Chlide *a* species that elutes at 14.3 minutes in the negative control (Figure 7.17A) was completely consumed by the enzyme in the WT assay (Figure 7.17B).

All of the various Chl a_{GG} peaks for each of the samples were integrated and the values summed to give a value for total Chlide esterification produced by each mutant. The value for Chl a_{GG} species produced by the WT AtChlG was artificially set to 100%, and the values for the variant AtChlG proteins are presented as a percentage of the WT activity (Figure 7.18). The most active mutant was P54F, predicted to be important in determining the substrate specificity of AtChlG, which evolved 64.7% the Chl a_{GG} product produced by the WT. The activities of Q46E, L56P and V60Y, which featured substitution of residues that are conserved in ChlG homolog, were significantly impaired in activity, producing Chl a_{GG} at 12.7%, 7.5% and 0.1% WT levels respectively. A225M, predicted to be located within the GGPP binding domain of AtChlG, had negligible activity, producing just 0.1% the Chl a_{GG} yield generated by the WT protein. N99A, predicted to be located within the active site of AtChlG and to facilitate the binding of GGPP, was devoid of activity.

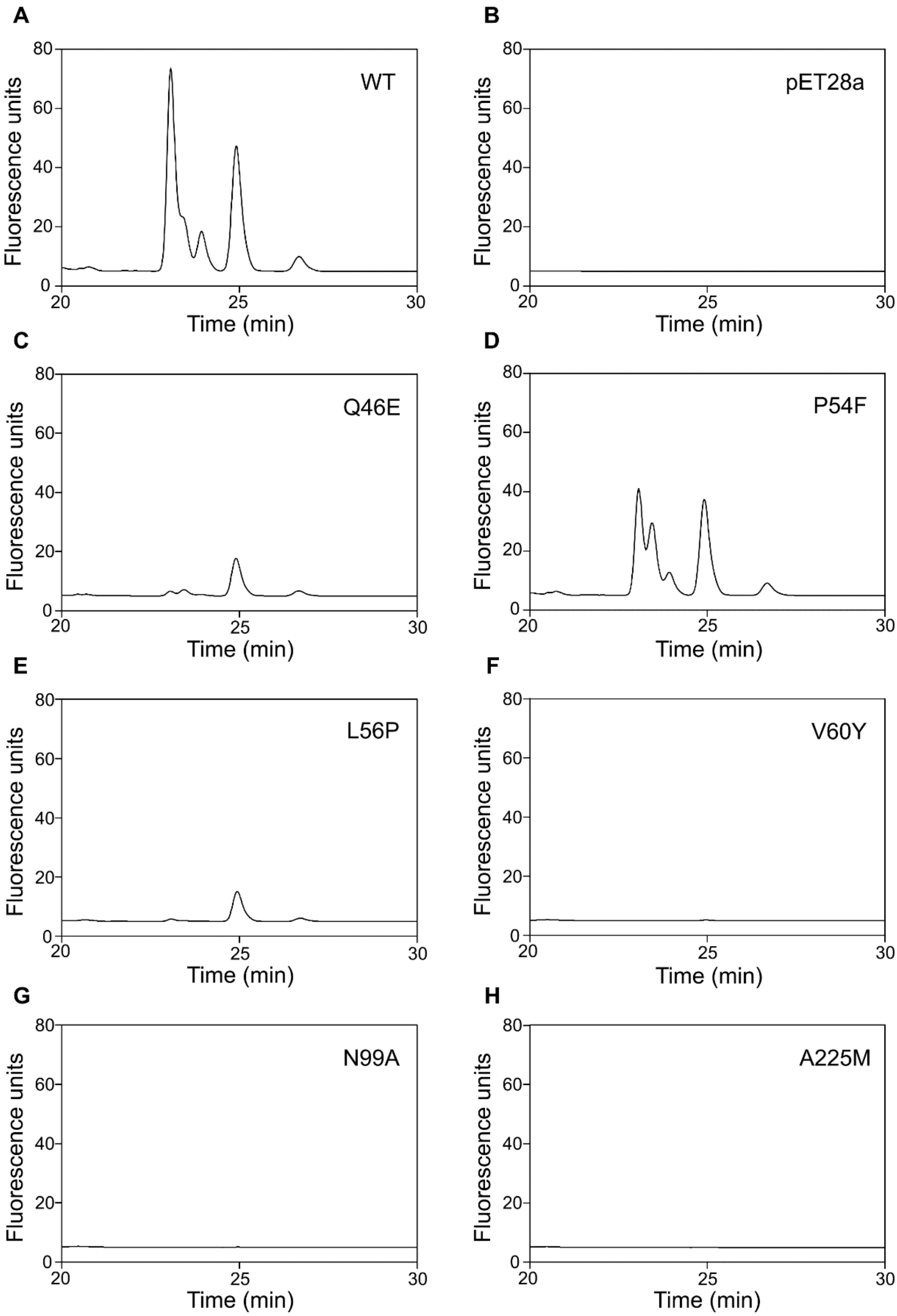


Figure 7.16: Reverse-phase HPLC separation of pigments extracted from 'wildtype' and variant AtChlG assays. Fluorescence profile of pigments extracted from enzyme assays using lysates from the following *E. coli* AtChlG variants: WT (A), empty pET28 vector negative control (B), Q46E (C), P54F (D), L56P (E), V60Y (F), N99A (G) and A225M (H).

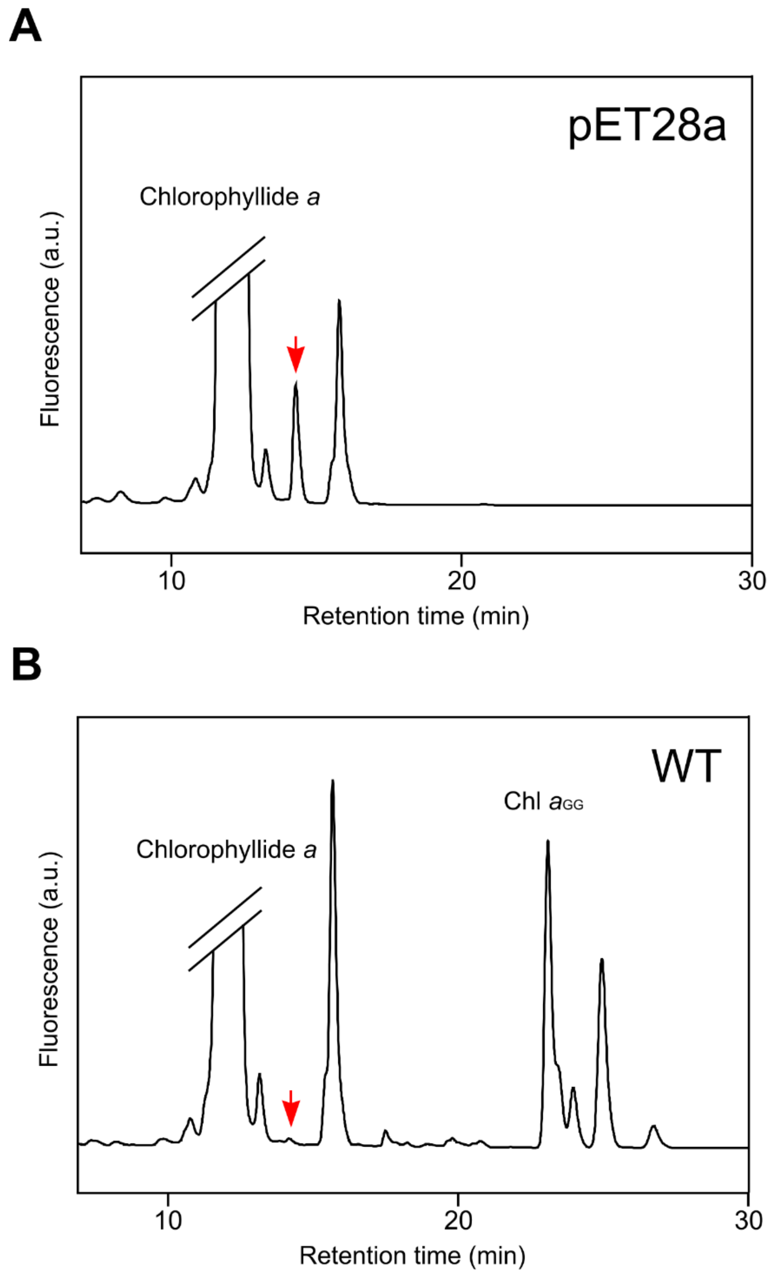


Figure 7.17: Reverse-phase HPLC separation of pigments extracted from AtChlG assays. (A) Fluorescence profile of pigments extracted from a pET28a negative control (pET28a) assay, showing the degradation of the Chlide substrate, and the AtChlG (WT) activity assays, showing the generation of multiple Chl a_{GG} species. (B) One of the Chlide *a* species (red arrow) was completely consumed throughout the course of the enzyme assays using the WT AtChlG protein.

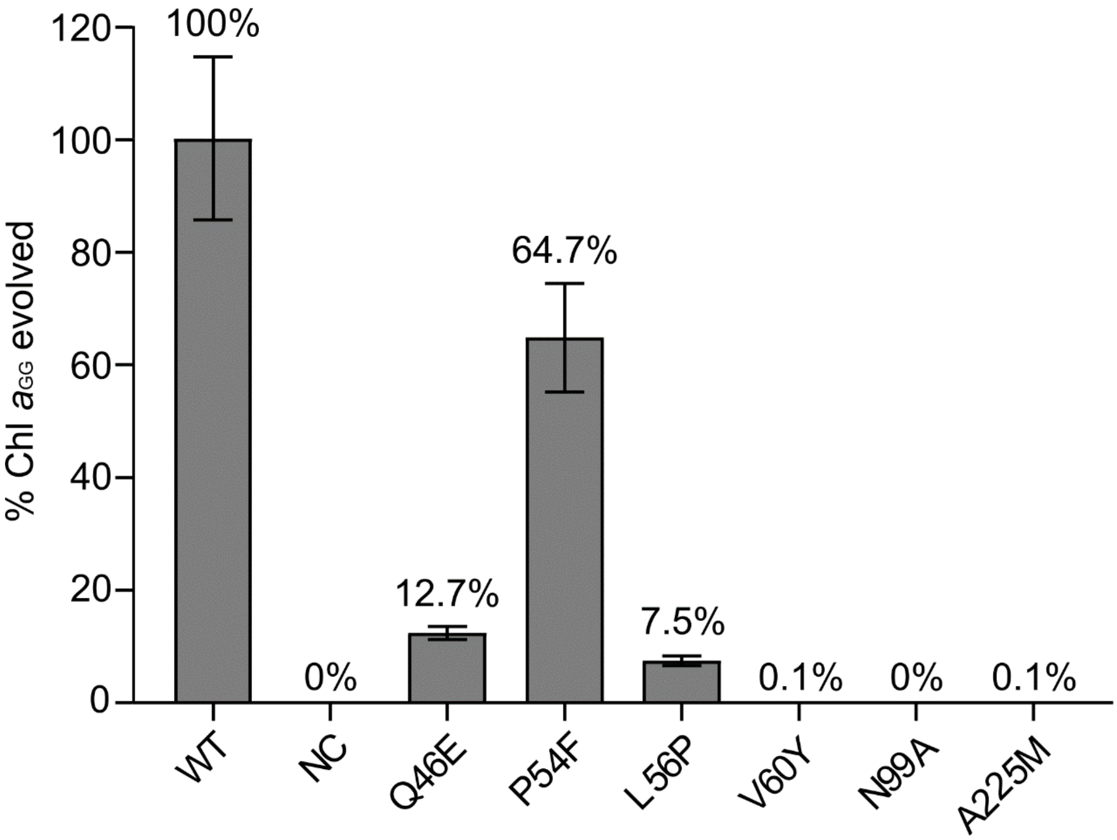


Figure 7.18: Histogram showing the percentage activity of variant AtChlG mutants. Reverse-phase HPLC peaks corresponding to Chl a_{GG} were integrated and summed. These values were used to calculate the percentage activity of each AtChlG variant in comparison to the WT enzyme, which was set as 100%. AtChlG assays were repeated in triplicate and the standard error is presented.

7.4 Discussion

7.4.1 AtChlG activity is abolished in detergent micelles

AtChlG was purified from *E. coli* cells by solubilisation in the detergent βDDM followed by Ni^{2+} IMAC. The eluate was tested for activity by *in vitro* enzyme assays with substrates Chlide and GGPP. Analysis of the pigments extracted from the assays showed that no Chlide was esterified with GGPP, thus AtChlG is not active in detergent micelles. It has been reported previously that solubilisation of ChlG from etiolated plant extracts resulted in complete loss of enzyme activity (Rüdiger *et al*., 1980), and so previous *in vitro* activity assays of ChlG and BchG homolog produced in *E. coli* have always been performed using cell lysates without removing the recombinant enzyme from the lipid bilayer (Kim and Lee, 2010; Kim *et al*., 2016; Oster and Rüdiger, 1997; Oster *et al*., 1997; Schmid *et al*., 2001, 2002). It may be that the enzyme loses structural stability, due to the loss of essential lipids from around the hydrophobic transmembrane helices. It follows that the enzyme should be purified from the lipid bilayer in the presence of exogenously added lipids to see if this results in an active enzyme. Alternatively, the detergent solubilised enzyme could be reconstituted into liposomes. It is also possible that a detergent belt surrounding the core of AtChlG is preventing the diffusion of substrates and product to and from the active site. A phylogenetically related protein, ubiquinone synthase (UbiA), features a lateral portal which opens into the lipid bilayer, through which hydrophobic substrates enter the catalytic site of the protein (Cheng and Li, 2014). The presence of a detergent micelle encompassing the protein may effectively block this opening and inhibit the enzyme. If AtChlG possess a similar lateral portal, through which hydrophobic substrates GGPP and Chlide enter the active site, this could explain why the detergent solubilised enzyme is non-functional. As is the case with ChlG homologs, *in vitro* activity of a detergent solubilised UbiA protein has never been demonstrated, with assays instead performed on the membrane fraction of recombinant *E. coli* cells (Melzer and Heide, 1994; Wessjohann and Sontag, 1996).

7.4.2 Production of Chl *a* in recombinant AtChlG *E. coli* lysates is low

Although AtChlG is active within clarified *E. coli* lysate, the quantity of esterified pigment produced is small in comparison to the total Chlide available. AtChlG esterified just 1.4% of the total Chlide pool (30 µM). This may be due to one or more of several factors. The concentration of AtChlG in the lysates could be low, thus there is only very limited capacity for esterification of Chlide within the assays. The reaction rate of the enzyme could be low due to removal of the protein from its native environment within plant thylakoid membranes. The diffusion of the substrates to the active site of AtChlG could be impeded or, similarly, the enzyme is inhibited by an excess of substrate. Finally, the diffusion of the hydrophobic Chl *a* product away from the active site may not be possible, thus each enzyme is only capable of a single turnover.

E. coli lysates containing ChlG from *Synechocystis* assayed in a similar manner but using 1-1.5 mL of lysate in comparison to the 27 µL used in this study, still only esterified 52% of the substrate (Oster *et al*., 1997). In light of this, it is probable that the efficiency of Chlide *a* to Chl a_{GG} conversion is limited either by the concentration of AtChlG and/or the enzyme is inhibited by Chl a_{GG} which, as a hydrophobic molecule, cannot diffuse away from the active site into the relatively aqueous environment in which the assays are performed. This should be tested by performing stop flow assays with increasing concentrations of AtChlG enzyme. If Chl a_{GG} production ceases after a consistent period of time at all AtChlG concentrations, yet the quantity of product increases with increasing AtChlG concentrations, it would suggest that the enzyme is only able to catalyse one esterification event before it is inhibited by lack of product release.

7.4.3 AtChlG variant P54F is able to esterify Chlide with GGPP

The isoleucine 44 (I44) residue of *Synechocystis* ChlG is important in determining the substrate specificity of the enzyme. This residue was identified by replacement of *bchG* in *Rba. sphaeroides* with *chlG* from *Synechocystis*. Multiple suppressors arose, all of which harboured the same point mutation, resulting in the substitution of I44 for a Phe residue, which corresponds to residue F28 in the *Rba. sphaeroides* BchG. The variant ChlG enzyme was capable of restoring photoautotrophic growth to the *Rba. sphaeroides* mutant, indicating that the mutant ChlG enzyme was able to utilise BChlide as a substrate and produce Bchl. *In vitro*, the ChlG I44F mutant had the same affinity for Chlide *a* (Km = 0.09 mM +/- 0.01 mM) as the WT enzyme (Kim *et al*., 2016) although the enzymes affinity for BChlide was much lower (Km = 1.42 mM +/- 0.21 mM). Similarly, the equivalent mutation was made to BchG, substituting F28 for Iso, resulting in a BchG F28I variant that could utilise Chlide to produce Chl *a* (Km = 2.87

mM +/- 0.35 mM) albeit to a much lesser extent than its native BChlide substrate (Km = 0.16 mM +/- 0.03).

Sequence alignments of multiple ChlG and BchG homolog allowed the identification of P54 of AtChlG as the residue corresponding to I44 in *Synechocystis* ChlG and F28 in purple bacterial BchGs. Subsequently, P54 was substituted by Phe and the activity of the variant AtChlG lysates assayed for activity. The exact AtChlG concentrations within each assay were unknown, as concentrations were estimated by the relative AtChlG signals detected by immunoblots. As such, the total enzyme activity of the AtChlG variants could not be determined. Instead, the total amount of Chl a_{GG} produced by each mutant was summed and calculated as a percentage of the total Chl a_{GG} produced by the WT protein. The P54F AtChlG variant exhibited 64.7% the activity of the WT AtChlG, demonstrating that this protein was still capable of utilising its native substrate but with reduced activity. This is in partial contrast to the results published by Kim, Kim and Lee, (2016) who showed that ChlG I44F retained WT levels of activity when tested with Chlide *in vitro*. AtChlG and *Synechocystis* ChlG share 66% sequence identity and both share 35% sequence identity with BchG. However, this residue is not conserved between AtChlG (P54) and *Synechocystis* ChlG (I44), and so the effects of substituting residues at this position to the Phe residue of BchG are not directly comparable. The activity of AtChlG P54F should be tested with BChlide to examine whether this mutation confers some level of BchG activity, as was the case with ChlG I44F, which was capable of esterifying BChlide, but to a lesser extent than Chlide (Kim *et al*., 2016). However, BChlide is difficult to make and store due to its inherent instability in solution. Removal of the BChl tail using CLH-1 under anaerobic conditions may be a viable strategy to produce BChlide for future experiments.

7.4.4 AtChlG residues Q46, L56 and V60 are important for enzyme activity

Three AtChlG residues were substituted by their BchG counterparts generating variants Q46E, L56P and V60Y. These residues were chosen based on their conservation in ChlG homolog, and the equivalent residues which they are substituted by are conserved in BchG homologs. There are many more residues that fulfil these selection criteria; however these three were chosen for proof of concept experiments based on their location within the first transmembrane helix of ChlG/BchG, which is predicted to be crucial for activity and substrate specificity. All of the mutants were significantly perturbed in ChlG activity with Q46E and L56P producing just 12.7% and 7.5% the Chl a_{GG} evolved by the WT protein. V60Y was essentially inactive, producing Chl a_{GG} that was barely detectable by fluorescence, at 0.1% WT activity. As stated above, the exact concentration of each AtChlG variant was not calculable, however, immunoblot analysis indicates that the levels are approximately consistent between the assays. In the case where a significant reduction in enzyme activity is observed, it can tentatively be concluded that the substituted residues are important, but not essential, for AtChlG activity. Where the activity is essentially abolished, a crucial role of the specific amino acid in enzyme function can be more confidently concluded. The nature of the enzyme inhibition caused by these substitutions has yet to be investigated, however one can speculate that the replacement of conserved amino acids with ones that are different in size and/or charge causes a change in the tertiary structure of the enzyme. A change in structure can inhibit enzyme activity in a number of ways, for example, by preventing access of the substrates to the active site, or impeding the catalytic mechanism of the enzyme.

7.4.5 N99A and A225M are involved in substrate binding to AtChlG

UbiA is structurally and functionally related to ChlG/BchG (Bonitz *et al*., 2011). The elucidation of the UbiA structure enabled the identification of 12 residues within the active site of the protein that were predicted to be important in coordinating the binding of the substrates to the protein. These residues were substituted by Ala and tested for activity *in vitro* (Cheng and Li, 2014). Among them, N50A and D175A variants were devoid of activity and were significantly impaired in binding the substrate GSPP. The equivalent residue of UbiA N50 in AtChlG is N99 (after removal of the AtChlG chloroplast targeting sequence), which is conserved in all ChlG and BchG homologs. N99 was predicted to be important for binding of GGPP to AtChlG, akin to the binding of GSPP to UbiA. It was therefore predicted that substitution of this residue for Ala (N99A) in AtChlG would disrupt the binding of GGPP to AtChlG and result in a non-functional protein, as was the case for N50A of UbiA. The N99A protein was devoid of activity, as predicted.

The UbiA residue D175 was also predicted to be important for coordinating the binding of GSPP to the enzyme. The equivalent residue of UbiA D175 in AtChlG is A225, which is conserved in all ChlG homologs. A model of AtChlG was constructed using the published structure of UbiA as a template (Figure 7.14). From the model, A225 was predicted to be located within the GGPP binding pocket of AtChlG and is involved in coordinating the binding of GGPP to the enzyme, akin to the binding of GSPP to UbiA. A225 was substituted by the residue at the equivalent position in BchG, Met, which is also conserved in BchG homologs. The activity of this AtChlG variant, A225M, was severely inhibited, producing just 0.1% the activity of the WT protein.

These results not only demonstrate the essential nature of residues N99 and A225 to the function of AtChlG, but suggest a high degree of structural similarity between UbiA and ChlG. In this respect, these results are a step towards validating the AtChlG model generated in this study as a viable means to predict further residues that are likely to be important to the function of the enzyme.

The active site of ChlG/BchG appears to be very sensitive to small alterations in the substrate, shown by the inability of either ChlG or BchG to utilise the other enzyme's natural substrate, Chlide and BChlide respectively (Kim and Lee, 2009), despite the minor structural differences between these two pigments. However, a single amino acid change in ChlG, from I44 to the equivalent residue F28 of BchG (I44F), enabled the variant ChlG I44F to utilise BChlide as a substrate (Kim *et al*., 2016) in addition to maintaining its native affinity for Chlide. The results from this study are in agreement with Cheng and Li (2014), as the substitution of two residues predicted to be located within the active site of AtChlG, N99 and A225, abolished the activity of the enzyme. It would however be interesting to test the activity of the N99A and A225M AtChlG variants for activity with BChlide.

7.5 Future work

The results from this chapter report the construction of a quick and easy system for the generation and testing of the activity of AtChlG point mutants *in vitro*. In addition to the six mutations described here, there are many more examples of conserved residues between ChlG homologs that have different, but also conserved, counterparts in BchG. These can be targeted for mutation in future experiments. The existing AtChlG mutants should also be tested for activity with BChlide, to see if, like the reported *Synechocystis* I44F ChlG mutant, they are now able to esterify BChlide with GGPP. In particular, the P54F AtChlG mutant generated in this study is a direct analogue of I44F and would be predicted to exhibit some BchG activity. In addition, the results from a study in which sequential truncations were made to the N-terminus of *Avena sativa* ChlG indicated that the first 87 residues were dispensable for protein activity. In this same study, the protein was made inactive by the removal of just two residues from the C-terminus (Schmid *et al*., 2001). The equivalent truncations could be made to AtChlG and the variant proteins tested for activity. Furthermore, sequence alignments of UbiA and AtChlG, in combination with the UbiA crystal structure, allowed the identification of the N99 and A225 residues of AtChlG, which were correctly predicted to be essential for function. Subsequently, a model of AtChlG was generated using the published structure of UbiA as a template (Figure 7.14). This model can be used to predict other residues that may be important for AtChlG function.

Despite the fact that all attempts to solubilise recombinant ChlG homologs in the past have resulted in loss of the proteins activity, the reconstitution of the enzyme with lipids within detergent micelles to try and restore activity has never been attempted. As AtChlG is active within *E. coli* lysates, a simple mix of exogenously added *E. coli* lipids may suffice to restore activity to the enzyme, although lipids native to the thylakoid membrane of *A. thaliana* could also be tested. Alternatively, the detergent solubilised protein could be reconstituted into liposomes before being used in enzyme activity assays.

Throughout this study, a major hindrance to the AtChlG activity assays performed was the inherent instability of the Chlide *a* substrate, generated via the methods discussed previously (Sections 7.3.4 and 7.3.6). The Chlide would degrade rapidly once prepared, despite attempts to store the pigment in various solvents at low temperatures and in the dark. Although AtChlG was able to utilise the degraded Chlide as a substrate, the pigment produced multiple peaks upon analysis by reverse-phase HPLC, both alone and when esterified to GGPP, making the results more difficult to interpret. The various Chlide and Chl a_{GG} species generated by the degradation of the pigment should be analysed by mass spectrometry to determine the differences between them. It was postulated that the degradation of the Chlide could be due to the formation of reactive oxygen species during preparation. As such, performing all experimental steps during Chlide production and purification in an anaerobic chamber should be attempted to see if this improves the stability of the pigment.

Chapter 8: General discussion

The research findings presented in this thesis relate to the functional and structural characterisation of the chlorophyll synthase (ChlG) proteins from cyanobacteria, plants and algae. Previous research identified a novel protein-pigment ChlG complex in the model cyanobacteria *Synechocysti*s composed of ChlG, high light inducible proteins (HliPs) C and D (HliC/HliD), the PSII assembly factor Ycf39 and the membrane translocase YidC (Chidgey et al., 2014). ChlG catalyses the final reaction of the chlorophyll (Chl) biosynthesis pathway, producing a mature Chl molecule that is then inserted co-translationally into Chl-binding proteins during photosystem assembly. The YidC component of the ChlG complex was hypothesised to stabilise membrane segments of these proteins during Chl insertion, whilst Ycf39 is involved in rearrangement of the complex under different environmental conditions and HliD/HliC provide photoprotection (Chidgey et al., 2014; Niedzwiedzki et al., 2016).

Following the discovery of the ChlG complex in cyanobacteria (Chidgey et al., 2014); the research presented in Chapter 3 of this thesis aimed to establish whether or not a similar complex exists in plants and algae. Plant and algal ChlG homologs, produced in *Synechocystis* and isolated from solubilised thylakoid membranes, whilst able to functionally complement the deletion of the essential *Synechocystis chlG* gene, did not form a complex with HliD or Ycf39. This indicates that a ChlG-HliD-Ycf39-like complex does not form in these organisms, despite the fact that the ChlG homologs are similar enough to functionally complement one another. The same experiments performed using a ChlG homolog from another cyanobacteria resulted in the isolation of the full complex, thus the formation of a ChlG complex is most likely confined to cyanobacteria. These results are in agreement with the existing body of research. Related proteins to Ycf39 and HliPs in the plant *Arabidopsis thaliana*, HCF244 and OHPs respectively, were shown to interact with each other but did not co-purify with ChlG when isolated from the native organism (Hey and Grimm, 2018; Myouga *et al*., 2018). Furthermore, under normal and high light growth conditions, the *Synechocystis* strains harbouring plant and algal *chlG* genes did not exhibit any phenotypes, and so photoprotection of ChlG may not be necessary in plants and algae, or is conveyed through other means. YidC remained associated with both the plant and algal ChlG proteins, indicating that this interaction is likely conserved in higher oxygenic phototrophs.

The pigments associated with the ChlG complex are Chl, chlorophyllide and carotenoids β-carotene, zeaxanthin and myxoxanthophyll (Chidgey et al., 2014; Shukla et al., 2018a). The Chl and β-carotene pigments are bound to HliD and HliC and have been shown to able to quench light absorbed by the complex, whereas the zeaxanthin and myxoxanthophyll were speculated to promote association of HliD with ChlG by acting at the interface between these two proteins (Llansola-Portoles *et al*., 2017; Niedzwiedzki *et al*., 2016). The experiments presented in Chapter 4 tests this hypothesis by abolishing the synthesis of myxoxanthophyll, deoxy-myxoxanthophyll (a precursor to myxoxanthophyll) and zeaxanthin in a *Synechocystis* strain harbouring a FLAG-tagged ChlG protein. Retrieval of ChlG from this strain by FLAG pulldown showed that the interaction of ChlG with HliD is lost in the absence of these carotenoids. This is the first evidence of a structural role of carotenoids within an enzyme complex of the chlorophyll biosynthesis pathway; expanding their repertoire of functions outside of their usual roles as structural and light-harvesting components of the photosystems, as well as photoprotective agents.

Chemical cross-linking of the ChlG complex *in vitro* and *in vivo* enabled the generation of a model showing its orientation and arrangement within the thylakoid membrane and confirms the close spatial proximity between the protein members of this complex. Although cross-linking has previously been used to characterise interactions between other proteins involved in photosynthesis, the cross-linking targets of these studies have been abundant, often soluble, molecules. As ChlG is a low abundance integral membrane protein, the cross-linking methods were optimised in this thesis for the structural characterisation of the ChlG complex and may therefore be applicable to other challenging targets.

Further structural information was gathered by truncation of the N-terminus of ChlG, predicted to be important for formation of the ChlG complex from the results presented in Chapter 5. Although the formation of the ChlG complex was unaffected by the removal of the N-terminus, the enzyme activity of ChlG was shown to reduce both *in vivo* and *in vitro*. This is in agreement with the findings of Schmid *et al*. 2001, in which similar truncations were made to the N-terminus of the oat (*Avena sativa*) ChlG protein, abolishing enzyme activity. The similarity between the results presented in this thesis and the studies on the oat ChlG enzyme alludes to the structural similarity of ChlG homologs that have long since evolutionarily diverged.

The heterologous production of a desired protein in *E. coli* is the most widely used system for generating the quantities of material needed for biochemical characterisation. ChlG from *Synechocystis* and *Arabidopsis* have been successfully produced in *E. coli* and lysates prepared from the cells for use in enzyme activity assays, however, neither enzyme have been purified in high yields. In this work, the production of a His-tagged *Arabidopsis* ChlG enzyme in *E. coli* has been optimised and the recombinant protein purified. Although not active in detergent micelles, in agreement with earlier reports (Rüdiger *et al*., 1980), the protein was active within cell-free lysates. In combination with a method for the production of the enzymes substrate, chlorophyllide *a*, these two systems were demonstrated to be useful for the production and functional characterisation of ChlG mutants. In addition, a model of the protein was generated using the crystal structure of a related protein as a template. Using this model, two residues were predicted to be important for enzyme activity; confirmed by protein mutagenesis and enzyme assays. The results lend credibility to the accuracy of the ChlG model which, in the absence of a crystal or Cryo EM structure, can be further tested by the mutagenesis of other residues predicted from the model to be important for enzyme activity. The high yields of ChlG produced using the methods outlined in this thesis may also enable the structural characterisation of the enzyme by protein crystallography or Cryo EM.

References

Acharya, K., Zazubovich, V., Reppert, M., and Jankowiak, R. (2012). Primary Electron Donor(s) in Isolated Reaction Center of Photosystem II from Chlamydomonas reinhardtii. J. Phys. Chem. B *116*, 4860–4870.

Adams, P.G., and Hunter, C.N. (2012). Adaptation of intracytoplasmic membranes to altered light intensity in Rhodobacter sphaeroides. Biochim. Biophys. Acta - Bioenerg. *1817*, 1616–1627.

Adams, N.B.P., Marklew, C.J., Brindley, A.A., Hunter, C.N., and Reid, J.D. (2014). Characterization of the magnesium chelatase from *Thermosynechococcus elongatus*. Biochem. J. *457*, 163–170.

Adams, N.B.P., Vasilev, C., Brindley, A.A., and Hunter, C.N. (2016a). Nanomechanical and Thermophoretic Analyses of the Nucleotide-Dependent Interactions between the AAA+Subunits of Magnesium Chelatase. J. Am. Chem. Soc. *138*, 6591–6597.

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

Adamska, I., Roobol-Bóza, M., Lindahl, M., and Andersson, B. (1999). Isolation of pigment-binding early light-inducible proteins from pea. Eur. J. Biochem. *260*, 453–460.

Addlesee, H.A., and Hunter, C.N. (1999). Physical mapping and functional assignment of the geranylgeranyl-bacteriochlorophyll reductase gene, bchP, of Rhodobacter sphaeroides. J. Bacteriol. *181*, 7248–7255.

Addlesee, H.A., Gibson, L.C.D., Jensen, P.E., and 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.

Akopian, D., Shen, K., Zhang, X., and Shan, S. (2013). Signal Recognition Particle: An Essential Protein-Targeting Machine. Annu. Rev. Biochem. *82*, 693–721.

Akulinkina, D. V., Bolychevtseva, Y. V., Elanskaya, I. V., Karapetyan, N. V., and Yurina, N.P. (2015). Association of high light-inducible HliA/HliB stress proteins with photosystem 1 trimers and monomers of the cyanobacterium Synechocystis PCC 6803. Biochem. *80*, 1254–1261.

Alawady, A.E., and Grimm, B. (2004). Tobacco Mg protoporphyrin IX methyltransferase is involved in inverse activation of Mg porphyrin and protoheme synthesis. Plant J. *41*, 282–290.

Alawady, A., Reski, R., Yaronskaya, E., and 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.

Albus, C.A., Ruf, S., Schöttler, M.A., Lein, W., Kehr, J., and Bock, R. (2010). Y3IP1, a nucleus-encoded thylakoid protein, cooperates with the plastid-encoded Ycf3 protein in photosystem I assembly of tobacco and Arabidopsis. Plant Cell *22*, 2838–2855.

Albus, C.A., Salinas, A., Czarnecki, O., Kahlau, S., Rothbart, M., Thiele, W., Lein, W., Bock, R., Grimm, B., and Schöttler, 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.

Allakhverdiev, S.I., and Murata, N. (2004). Environmental stress inhibits the synthesis de novo of proteins involved in the photodamage–repair cycle of Photosystem II in Synechocystis sp. PCC 6803. Biochim. Biophys. Acta - Bioenerg. *1657*, 23–32.

Anbudurai, P.R., Mor, T.S., Ohad, I., Shestakov, S. V, and Pakrasi, H.B. (1994). The ctpA gene encodes the C-terminal processing protease for the D1 protein of the photosystem II reaction center complex. Proc. Natl. Acad. Sci. U. S. A. *91*, 8082–8086.

and, C.F., and Vermaas, W. (1999). A Cyanobacterial Gene Family Coding for Single-Helix Proteins Resembling Part of the Light-Harvesting Proteins from Higher Plants†.

Andersson, U., Heddad, M., and Adamska, I. (2003). Light stress-induced one-helix protein of the chlorophyll a/b-binding family associated with photosystem I. Plant Physiol. *132*, 811–820.

Armbrust, T.S., Chitnis, P.R., and Guikema, J.A. (1996). Organization of photosystem I polypeptides examined by chemical cross-linking. Plant Physiol. *111*, 1307–1312.

Aro, E.M., Virgin, I., and Andersson, B. (1993). Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. Biochim. Biophys. Acta *1143*, 113–134.

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

Back, J.W., de Jong, L., Muijsers, A.O., and de Koster, C.G. (2003). Chemical cross-linking and mass spectrometry for protein structural modeling. J. Mol. Biol. *331*, 303–313.

Ballottari, M., Girardon, J., Dall'Osto, L., and Bassi, R. (2012). Evolution and functional properties of Photosystem II light harvesting complexes in eukaryotes. Biochim. Biophys. Acta - Bioenerg. *1817*, 143–157.

Baniulis, D., Zhang, H., Zakharova, T., Hasan, S.S., and Cramer, W.A. (2011). Purification and Crystallization of the Cyanobacterial Cytochrome b 6 f Complex. In Methods in Molecular Biology

(Clifton, N.J.), pp. 65–77.

Barber, J. (2006). Photosystem II: an enzyme of global significance. Biochem. Soc. Trans. *34*, 619–631.

Barber, J. (2016). Mn_4Ca Cluster of Photosynthetic Oxygen-Evolving Center: Structure, Function and Evolution. Biochemistry *55*, 5901–5906.

Barber, J., and Andersson, B. (1992). Too much of a good thing: light can be bad for photosynthesis. Trends Biochem. Sci. *17*, 61–66.

Barker, M., de Vries, R., Nield, J., Komenda, J., and Nixon, P.J. (2006). The deg proteases protect Synechocystis sp. PCC 6803 during heat and light stresses but are not essential for removal of damaged D1 protein during the photosystem two repair cycle. J. Biol. Chem. *281*, 30347–30355.

Barker, M., Boehm, M., Nixon, P.J., and Nield, J. (2008). Structural Analysis of an FtsH2/FtsH3 Complex Isolated from Synechocystis sp. PCC 6803. In Photosynthesis. Energy from the Sun, (Dordrecht: Springer Netherlands), pp. 737–740.

Barros, T., and Kühlbrandt, W. (2009). Crystallisation, structure and function of plant light-harvesting Complex II. Biochim. Biophys. Acta - Bioenerg. *1787*, 753–772.

Beale, S.I. (2006). Biosynthesis of 5-Aminolevulinic Acid. In Chlorophylls and Bacteriochlorophylls, (Dordrecht: Springer Netherlands), pp. 147–158.

Beale, S.I., and Castelfranco, P.A. (1974). The Biosynthesis of delta-Aminolevulinic Acid in Higher Plants: I. Accumulation of delta-Aminolevulinic Acid in Greening Plant Tissues. Plant Physiol. *53*, 291–296.

Beck, J., Lohscheider, J.N., Albert, S., Andersson, U., Mendgen, K.W., Rojas-Stütz, M.C., Adamska, I., and Funck, D. (2017). Small One-Helix Proteins Are Essential for Photosynthesis in Arabidopsis. Front. Plant Sci. *8*, 1–14.

Beck, K., Eisner, G., Trescher, D., Dalbey, R.E., Brunner, J., and Müller, M. (2001). YidC, an assembly site for polytopic Escherichia coli membrane proteins located in immediate proximity to the SecYE translocon and lipids. EMBO Rep. *2*, 709–714.

Becker, K., Cormann, K.U., and Nowaczyk, M.M. (2011). Assembly of the water-oxidizing complex in photosystem II. J. Photochem. Photobiol. B Biol. *104*, 204–211.

Bečková, M., Yu, J., Krynická, V., Kozlo, A., Shao, S., Koník, P., Komenda, J., Murray, J.W., and Nixon, P.J. (2017). Structure of Psb29/Thf1 and its association with the FtsH protease complex involved in photosystem II repair in cyanobacteria. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *372*.

Bédard, J., Trösch, R., Wu, F., Ling, Q., Flores-Pérez, Ú., Töpel, M., Nawaz, F., and Jarvis, P. (2017).

Suppressors of the Chloroplast Protein Import Mutant *tic40* Reveal a Genetic Link between Protein Import and Thylakoid Biogenesis. Plant Cell *29*, 1726–1747.

Bellafiore, S., Ferris, P., Naver, H., Göhre, V., and Rochaix, J.-D. (2002). Loss of Albino3 leads to the specific depletion of the light-harvesting system. Plant Cell *14*, 2303–2314.

Ben-Shem, A., Frolow, F., and Nelson, N. (2003). Crystal structure of plant photosystem I. Nature *426*, 630–635.

Benz, M., Bals, T., Gügel, I.L., Piotrowski, M., Kuhn, A., Schünemann, D., Soll, J., and Ankele, E. (2009). Alb4 of Arabidopsis Promotes Assembly and Stabilization of a Non Chlorophyll-Binding Photosynthetic Complex, the CF1CF0–ATP Synthase. Mol. Plant *2*, 1410–1424.

Bhaya, D., Dufresne, A., Vaulot, D., and Grossman, A. (2002). Analysis of the hli gene family in marine and freshwater cyanobacteria. FEMS Microbiol. Lett. *215*, 209–219.

Bhuiyan, N.H., Friso, G., Poliakov, A., Ponnala, L., and van Wijk, K.J. (2015). MET1 is a thylakoid-associated TPR protein involved in photosystem II supercomplex formation and repair in Arabidopsis. Plant Cell *27*, 262–285.

Blankenship, R.E. (2014). Molecular Mechanisms of Photosynthesis, 2nd Edition (Wiley-Blackwell).

Block, M.A., Joyard, J., and Douce, R. (1980). Site of synthesis of geranylgeraniol derivatives in intact spinach chloroplasts. Biochim. Biophys. Acta - Gen. Subj. *631*, 210–219.

Boehm, M., Romero, E., Reisinger, V., Yu, J., Komenda, J., Eichacker, L.A., Dekker, J.P., and Nixon, P.J. (2011). Investigating the Early Stages of Photosystem II Assembly in *Synechocystis* sp. PCC 6803. J. Biol. Chem. *286*, 14812–14819.

Boehm, M., Yu, J., Reisinger, V., Beckova, M., Eichacker, L.A., Schlodder, E., Komenda, J., and Nixon, P.J. (2012). Subunit composition of CP43-less photosystem II complexes of Synechocystis sp. PCC 6803: implications for the assembly and repair of photosystem II. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *367*, 3444–3454.

Boekema, E.J., Hankamer, B., Bald, D., Kruip, J., Nield, J., Boonstra, A.F., Barber, J., and Rögner, M. (1995). Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. Proc. Natl. Acad. Sci. U. S. A. *92*, 175–179.

Böger, P. (1965). Chlorophyllase of Chlorella vulgaris Beijerinck. Phytochemistry *4*, 435–443.

Bollivar, D., Braumann, I., Berendt, K., Gough, S.P., and Hansson, M. (2014). The Ycf54 protein is part of the membrane component of Mg-protoporphyrin IX monomethyl ester cyclase from barley (*Hordeum vulgare* L.). FEBS J. *281*, 2377–2386.

Bollivar, D.W., Jiang, Z.Y., Bauer, C.E., and 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., Suzuki, J.Y., Beatty, J.T., Dobrowolski, J.M., and Bauer, C.E. (1994b). Directed Mutational Analysis of Bacteriochlorophyll a Biosynthesis in Rhodobacter capsulatus. J. Mol. Biol. *237*, 622–640.

Bollivar, D.W., Wang, S., Allen, J.P., and Bauer, C.E. (1994c). 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.

Bonitz, T., Alva, V., Saleh, O., Lupas, A.N., and Heide, L. (2011). Evolutionary relationships of microbial aromatic prenyltransferases. PLoS One *6*, e27336.

Boudreau, E., Takahashi, Y., Lemieux, C., Turmel, M., and Rochaix, J.D. (1997). The chloroplast ycf3 and ycf4 open reading frames of Chlamydomonas reinhardtii are required for the accumulation of the photosystem I complex. EMBO J. *16*, 6095–6104.

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

Breitenbach, J., Fernández-González, B., Vioque, A., and Sandmann, G. (1998). A higher-plant type ζ-carotene desaturase in the cyanobacterium Synechocystis PCC6803. Plant Mol. Biol. *36*, 725–732.

Brettel, K., and Leibl, W. (2001). Electron transfer in photosystem I. Biochim. Biophys. Acta - Bioenerg. *1507*, 100–114.

Bricker, T.M., Roose, J.L., Fagerlund, R.D., Frankel, L.K., and Eaton-Rye, J.J. (2012). The extrinsic proteins of Photosystem II. Biochim. Biophys. Acta - Bioenerg. *1817*, 121–142.

Brindley, A.A., Adams, N.B.P., Hunter, C.N., and Reid, J.D. (2015). Five Glutamic Acid Residues in the C-Terminal Domain of the ChlD Subunit Play a Major Role in Conferring Mg^{2+} Cooperativity upon Magnesium Chelatase. Biochemistry *54*, 6659–6662.

Bröcker, M.J., Virus, S., Ganskow, S., Heathcote, P., Heinz, D.W., Schubert, W.-D., Jahn, D., and Moser, J. (2008a). ATP-driven Reduction by Dark-operative Protochlorophyllide Oxidoreductase from *Chlorobium tepidum* Mechanistically Resembles Nitrogenase Catalysis. J. Biol. Chem. *283*, 10559–10567.

Bröcker, M.J., Wätzlich, D., Uliczka, F., Virus, S., Saggu, M., Lendzian, F., Scheer, H., Rüdiger, W., Moser, J., and Jahn, D. (2008b). Substrate recognition of nitrogenase-like dark operative protochlorophyllide oxidoreductase from Prochlorococcus marinus. J. Biol. Chem. *283*, 29873–29881.

Bröcker, M.J., Wätzlich, D., Saggu, M., Lendzian, F., Moser, J., and Jahn, D. (2010a). Biosynthesis of (Bacterio)chlorophylls: ATP-dependent transient subunit interaction and electron transfer of dark operative protochlorophyllide oxidoreductase. J. Biol. Chem. *285*, 8268–8277.

Bröcker, M.J., Schomburg, S., Heinz, D.W., Jahn, D., Schubert, W.-D., and Moser, J. (2010b). Crystal Structure of the Nitrogenase-like Dark Operative Protochlorophyllide Oxidoreductase Catalytic Complex (ChlN/ChlB) $_2$. J. Biol. Chem. *285*, 27336–27345.

Bryant, D.A. (1994). The Molecular Biology of Cyanobacteria (Springer Netherlands).

Bučinská, L., Kiss, É., Koník, P., Knoppová, J., Komenda, J., and Sobotka, R. (2018). The Ribosome-Bound Protein Pam68 Promotes Insertion of Chlorophyll into the CP47 Subunit of Photosystem II. Plant Physiol. *176*, 2931–2942.

Burke, D.H., Alberti, M., and Hearst, J.E. (1993a). 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.

Burke, D.H., Alberti, M., and Hearst, J.E. (1993b). The Rhodobacter capsulatus chlorin reductase-encoding locus, bchA, consists of three genes, bchX, bchY, and bchZ. J. Bacteriol. *175*, 2407–2413.

Cai, W., Ma, J., Chi, W., Zou, M., Guo, J., Lu, C., and Zhang, L. (2010). Cooperation of LPA3 and LPA2 is essential for photosystem II assembly in Arabidopsis. Plant Physiol. *154*, 109–120.

Callahan, F.E., Wergin, W.P., Nelson, N., Edelman, M., and Mattoo, A.K. (1989). Distribution of Thylakoid Proteins between Stromal and Granal Lamellae in Spirodela : Dual Location of Photosystem II Components. Plant Physiol. *91*, 629–635.

Camara, B. (1984). Terpenoid metabolism in plastids : sites of phytoene synthetase activity and synthesis in plant cells. Plant Physiol. *74*, 112–116.

Canniffe, D.P., Jackson, P.J., Hollingshead, S., Dickman, M.J., and Hunter, C.N. (2013). Identification of an 8-vinyl reductase involved in bacteriochlorophyll biosynthesis in *Rhodobacter sphaeroides* and evidence for the existence of a third distinct class of the enzyme. Biochem. J. *450*, 397–405.

Cardona, T., Sedoud, A., Cox, N., and Rutherford, A.W. (2012). Charge separation in Photosystem II: A comparative and evolutionary overview. Biochim. Biophys. Acta - Bioenerg. *1817*, 26–43.

Cartron, M.L., Olsen, J.D., Sener, M., Jackson, P.J., Brindley, A.A., Qian, P., Dickman, M.J., Leggett, G.J., Schulten, K., and Neil Hunter, C. (2014). Integration of energy and electron transfer processes in the photosynthetic membrane of Rhodobacter sphaeroides. Biochim. Biophys. Acta - Bioenerg. *1837*, 1769–

1780.

Caspy, I., and Nelson, N. (2018). Structure of the plant photosystem I. Biochem. Soc. Trans. *46*, 285–294.

Cazzaniga, S., Li, Z., Niyogi, K.K., Bassi, R., and Dall'Osto, L. (2012). The Arabidopsis szl1 mutant reveals a critical role of β-carotene in photosystem I photoprotection. Plant Physiol. *159*, 1745–1758.

Cazzonelli, C.I. (2011). Goldacre Review: Carotenoids in nature: insights from plants and beyond. Funct. Plant Biol. *38*, 833.

Cereda, A., Hitchcock, A., Symes, M.D., Cronin, L., Bibby, T.S., and Jones, A.K. (2014). A Bioelectrochemical Approach to Characterize Extracellular Electron Transfer by Synechocystis sp. PCC6803. PLoS One *9*, e91484.

Charuvi, D., Kiss, V., Nevo, R., Shimoni, E., Adam, Z., and Reich, Z. (2012). Gain and loss of photosynthetic membranes during plastid differentiation in the shoot apex of Arabidopsis. Plant Cell *24*, 1143–1157.

Chen, G.E., Canniffe, D.P., Martin, E.C., and Hunter, C.N. (2016). Absence of the cbb3 Terminal Oxidase Reveals an Active Oxygen-Dependent Cyclase Involved in Bacteriochlorophyll Biosynthesis in Rhodobacter sphaeroides. J. Bacteriol. *198*, 2056–2063.

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

Chen, M., Samuelson, J.C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R.E. (2002). Direct Interaction of YidC with the Sec-independent Pf3 Coat Protein during Its Membrane Protein Insertion. J. Biol. Chem. *277*, 7670–7675.

Chen, X., Wang, X., Feng, J., Chen, Y., Fang, Y., Zhao, S., Zhao, A., Zhang, M., and Liu, L. (2014). Structural Insights into the Catalytic Mechanism of *Synechocystis* Magnesium Protoporphyrin IX *O* - Methyltransferase (ChlM). J. Biol. Chem. *289*, 25690–25698.

Chen, Y., Li, F., and Wurtzel, E.T. (2010). Isolation and Characterization of the Z-ISO Gene Encoding a Missing Component of Carotenoid Biosynthesis in Plants. PLANT Physiol. *153*, 66–79.

Cheng, W., and Li, W. (2014). Structural insights into ubiquinone biosynthesis in membranes. Science *343*, 878–881.

Cheregi, O., Sicora, C., Kós, P.B., Barker, M., Nixon, P.J., and Vass, I. (2007). The role of the FtsH and Deg proteases in the repair of UV-B radiation-damaged Photosystem II in the cyanobacterium Synechocystis PCC 6803. Biochim. Biophys. Acta - Bioenerg. *1767*, 820–828.

Chew, A.G.M., and Bryant, D.A. (2007). Characterization of a Plant-like Protochlorophyllide *a* Divinyl Reductase in Green Sulfur Bacteria. J. Biol. Chem. *282*, 2967–2975.

Chew, A.G.M., Frigaard, N.-U., and Bryant, D.A. (2007). Identification of the bchP gene, encoding geranylgeranyl reductase in Chlorobaculum tepidum. J. Bacteriol. *190*, 747–749.

Chidgey, J.W. (2014). Investigations into Synechocystis chlorophyll synthase. University of Sheffield.

Chidgey, J.W., Linhartová, M., Komenda, J., Jackson, P.J., Dickman, M.J., Canniffe, D.P., Koník, P., Pilný, J., Hunter, C.N., and Sobotka, R. (2014). A Cyanobacterial Chlorophyll Synthase-HliD Complex Associates with the Ycf39 Protein and the YidC/Alb3 Insertase. Plant Cell *26*, 1267–1279.

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

Chitnis, P.R. (1996). Photosystem I. Plant Physiol. *111*, 661–669.

Chitnis, V.P., Xu, Q., Yu, L., Golbeck, J.H., Nakamoto, H., Xie, D.L., and Chitnis, P.R. (1993). Targeted inactivation of the gene psaL encoding a subunit of photosystem I of the cyanobacterium Synechocystis sp. PCC 6803. J. Biol. Chem. *268*, 11678–11684.

Chorus, I., and Bartram, J. (1999). Toxic cyanobacteria in water : a guide to their public health consequences, monitoring, and management (E & FN Spon).

Chu, F., Mahrus, S., Craik, C.S., and Burlingame, A.L. (2006). Isotope-Coded and Affinity-Tagged Cross-Linking (ICATXL): An Efficient Strategy to Probe Protein Interaction Surfaces. J. Am. Chem. Soc. *128*, 10362–10363.

Cohen-Bazire, G., Sistrom, W.R., and Stanier, R.Y. (1957). Kinetic studies of pigment synthesis by non-sulfur purple bacteria. J. Cell. Comp. Physiol. *49*, 25–68.

Coomber, S.A., and Hunter, C.N. (1989). Construction of a physical map of the 45 kb photosynthetic gene cluster of Rhodobacter sphaeroides. Arch. Microbiol. *151*, 454–458.

Coomber, S.A., Chaudhri, M., Connor, A., Britton, G., and Hunter, C.N. (1990). Localized transposon Tn5 mutagenesis of the photosynthetic gene cluster of Rhodobacter sphaeroides. Mol. Microbiol. *4*, 977–989.

Cooper, D.E., Rands, M.B., and Woo, C.P. (1975). Sulfide reduction in fellmongery effluent by red sulfur bacteria. J. Water Pollut. Control Fed. *47*, 2088–2100.

Corradi, H.R., Corrigall, A. V, Boix, E., Mohan, C.G., Sturrock, E.D., Meissner, P.N., and Acharya, K.R.

(2006). Crystal structure of protoporphyrinogen oxidase from Myxococcus xanthus and its complex with the inhibitor acifluorfen. J. Biol. Chem. *281*, 38625–38633.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Croce, R., and van Amerongen, H. (2014). Natural strategies for photosynthetic light harvesting. Nat. Chem. Biol. *10*, 492–501.

Dalbey, R.E., Kuhn, A., Zhu, L., and Kiefer, D. (2014). The membrane insertase YidC. Biochim. Biophys. Acta - Mol. Cell Res. *1843*, 1489–1496.

Danielsson, R., Suorsa, M., Paakkarinen, V., Albertsson, P.-A., Styring, S., Aro, E.-M., and Mamedov, F. (2006). Dimeric and monomeric organization of photosystem II. Distribution of five distinct complexes in the different domains of the thylakoid membrane. J. Biol. Chem. *281*, 14241–14249.

Dasgupta, J., Ananyev, G.M., and Dismukes, G.C. (2008). Photoassembly of the Water-Oxidizing Complex in Photosystem II. Coord. Chem. Rev. *252*, 347–360.

Daum, B., Nicastro, D., Austin, J., McIntosh, J.R., and Kühlbrandt, W. (2010). Arrangement of Photosystem II and ATP Synthase in Chloroplast Membranes of Spinach and Pea. Plant Cell *22*, 1299–1312.

Davison, P.A., Schubert, H.L., Reid, J.D., Iorg, C.D., Heroux, A., Hill, C.P., and Hunter, C.N. (2005). Structural and Biochemical Characterization of Gun4 Suggests a Mechanism for Its Role in Chlorophyll Biosynthesis [†, ‡]. Biochemistry *44*, 7603–7612.

Debus, R.J., Barry, B.A., Sithole, I., Babcock, G.T., and McIntosh, L. (1988). Directed mutagenesis indicates that the donor to P+680 in photosystem II is tyrosine-161 of the D1 polypeptide. Biochemistry *27*, 9071–9074.

Demé, B., Cataye, C., Block, M.A., Maréchal, E., and Jouhet, J. (2014). Contribution of galactoglycerolipids to the 3-dimensional architecture of thylakoids. FASEB J. *28*, 3373–3383.

Dewez, D., Park, S., García-Cerdán, J.G., Lindberg, P., and Melis, A. (2009). Mechanism of REP27 protein action in the D1 protein turnover and photosystem II repair from photodamage. Plant Physiol. *151*, 88–99.

Diner, B.A., Ries, D.F., Cohen, B.N., and Metz, J.G. (1988). COOH-terminal processing of polypeptide D1 of the photosystem II reaction center of Scenedesmus obliquus is necessary for the assembly of the oxygen-evolving complex. J. Biol. Chem. *263*, 8972–8980.

Dobáková, M., Tichy, M., and Komenda, J. (2007). Role of the PsbI protein in photosystem II assembly and repair in the cyanobacterium Synechocystis sp. PCC 6803. Plant Physiol. *145*, 1681–1691.

Dobáková, M., Sobotka, R., Tichý, M., and Komenda, J. (2009). Psb28 protein is involved in the biogenesis of the photosystem II inner antenna CP47 (PsbB) in the cyanobacterium Synechocystis sp. PCC 6803. Plant Physiol. *149*, 1076–1086.

Dolganov, N.A., Bhaya, D., and Grossman, A.R. (1995). Cyanobacterial protein with similarity to the chlorophyll a/b binding proteins of higher plants: evolution and regulation. Proc. Natl. Acad. Sci. U. S. A. *92*, 636–640.

Domanskii, V.P., and Rüdiger, W. (2001). On the nature of the two pathways in chlorophyll formation from protochlorophyllide. Photosynth. Res. *68*, 131–139.

Domanskii, V., Rassadina, V., Gus-Mayer, S., Wanner, G., Schoch, S., and Rüdiger, W. (2002). Characterization of two phases of chlorophyll formation during greening of etiolated barley leaves. Planta *216*, 475–483.

Dorne, A.J., Joyard, J., and Douce, R. (1990). Do thylakoids really contain phosphatidylcholine? Proc. Natl. Acad. Sci. U. S. A. *87*, 71–74.

Drath, M., Kloft, N., Batschauer, A., Marin, K., Novak, J., and Forchhammer, K. (2008). Ammonia triggers photodamage of photosystem II in the cyanobacterium Synechocystis sp. strain PCC 6803. Plant Physiol. *147*, 206–215.

Drews, G., Peters, J., and Dierstein, R. (1983). Molecular organization and biosynthesis of pigment-protein complexes of Rhodopseudomonas capsulata. Ann. l'Institut Pasteur / Microbiol. *134*, 151–158.

Dühring, U., Irrgang, K.-D., Lünser, K., Kehr, J., and Wilde, A. (2006). Analysis of photosynthetic complexes from a cyanobacterial ycf37 mutant. Biochim. Biophys. Acta - Bioenerg. *1757*, 3–11.

Dühring, U., Ossenbühl, F., and Wilde, A. (2007). Late assembly steps and dynamics of the cyanobacterial photosystem I. J. Biol. Chem. *282*, 10915–10921.

Eichacker, L.A., Soll, J., Lauterbach, P., Rüdiger, W., Klein, R.R., and Mullet, J.E. (1990). In vitro synthesis of chlorophyll a in the dark triggers accumulation of chlorophyll a apoproteins in barley etioplasts. J. Biol. Chem. *265*, 13566–13571.

Eichacker, L.A., Helfrich, M., Rüdiger, W., and Müller, B. (1996). Stabilization of chlorophyll a-binding apoproteins P700, CP47, CP43, D2, and D1 by chlorophyll a or Zn-pheophytin a. J. Biol. Chem. *271*, 32174–32179.

Elder, G.H., and Roberts, A.G. (1995). Uroporphyrinogen decarboxylase. J. Bioenerg. Biomembr. *27*, 207–214.

Ellsworth, R.K. (1971). Studies on chlorophyllase.I.Hydrolytic and esterification activities of chlorophyllase from wheat seedlings. Photosynthetica *5*, 226–232.

Emanuelsson, O., Nielsen, H., and Heijne, G. Von (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. *8*, 978–984.

Engelken, J., Brinkmann, H., and Adamska, I. (2010). Taxonomic distribution and origins of the extended LHC (light-harvesting complex) antenna protein superfamily. BMC Evol. Biol. *10*, 233.

Ermakova-Gerdes, S., and Vermaas, W. (1999). Inactivation of the Open Reading Frame slr 0399 in Synechocystis sp. PCC 6803 Functionally Complements Mutations near the Q A Niche of Photosystem II. J. Biol. Chem. *274*, 30540–30549.

Ernst, S., Schönbauer, A.-K., Bär, G., Börsch, M., and Kuhn, A. (2011). YidC-Driven Membrane Insertion of Single Fluorescent Pf3 Coat Proteins. J. Mol. Biol. *412*, 165–175.

Fernández-González, B., Sandmann, G., and Vioque, A. (1997). A new type of asymmetrically acting beta-carotene ketolase is required for the synthesis of echinenone in the cyanobacterium Synechocystis sp. PCC 6803. J. Biol. Chem. *272*, 9728–9733.

Ferreira, G.C., Andrew, T.L., Karr, S.W., and Dailey, H.A. (1988). Organization of the terminal two enzymes of the heme biosynthetic pathway. Orientation of protoporphyrinogen oxidase and evidence for a membrane complex. J. Biol. Chem. *263*, 3835–3839.

Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J., and Iwata, S. (2004). Architecture of the photosynthetic oxygen-evolving center. Science *303*, 1831–1838.

Frain, K.M., Gangl, D., Jones, A., Zedler, J.A.Z., and Robinson, C. (2016). Protein translocation and thylakoid biogenesis in cyanobacteria. Biochim. Biophys. Acta - Bioenerg. *1857*, 266–273.

Frank, H.A., and Cogdell, R.J. (1996). Carotenoids in photosynthesis. Photochem. Photobiol. *63*, 257–264.

Friedel Drepper, *, Pierre Dorlet, and, and Mathis, P. (1997). Cross-Linked Electron Transfer Complex between Cytochrome c2 and the Photosynthetic Reaction Center of Rhodobacter sphaeroides†.

Fu, A., He, Z., Cho, H.S., Lima, A., Buchanan, B.B., and Luan, S. (2007). A chloroplast cyclophilin functions in the assembly and maintenance of photosystem II in Arabidopsis thaliana. Proc. Natl. Acad. Sci. *104*, 15947–15952.

Funk, C. (2001). The PsbS Protein: A Cab-protein with a Function of Its Own. In Regulation of Photosynthesis, (Dordrecht: Kluwer Academic Publishers), pp. 453–467.

Gabruk, M., Stecka, A., Strzałka, W., Kruk, J., Strzałka, K., and Mysliwa-Kurdziel, B. (2015). Photoactive protochlorophyllide-enzyme complexes reconstituted with PORA, PORB and PORC proteins of A. thaliana: fluorescence and catalytic properties. PLoS One *10*, e0116990.

Gao, J., Wang, H., Yuan, Q., and Feng, Y. (2018). Structure and Function of the Photosystem Supercomplexes. Front. Plant Sci. *9*, 357.

Gaubier, P., Wu, H.J., Laudié, M., Delseny, M., and Grellet, F. (1995). A chlorophyll synthetase gene from Arabidopsis thaliana. MGG Mol. Gen. Genet. *249*, 58–64.

Gerdes, L., Bals, T., Klostermann, E., Karl, M., Philippar, K., Hünken, M., Soll, J., and Schünemann, D. (2006). A second thylakoid membrane-localized Alb3/OxaI/YidC homologue is involved in proper chloroplast biogenesis in Arabidopsis thaliana. J. Biol. Chem. *281*, 16632–16642.

Giannino, D., Condello, E., Bruno, L., Testone, G., Tartarini, A., Cozza, R., Innocenti, A.M., Bitonti, M.B., and Mariotti, D. (2004). The gene geranylgeranyl reductase of peach (Prunus persica [L.] Batsch) is regulated during leaf development and responds differentially to distinct stress factors. J. Exp. Bot. *55*, 2063–2073.

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, K.D., Neuberger, A., and Tait, G.H. (1963). Studies on the biosynthesis of porphyrin and bacteriochlorophyll by Rhodopseudomonas spheroides. 4. S-adenosylmethioninemagnesium protoporphyrin methyltransferase. Biochem. J. *88*, 325–334.

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

Gibson, L.C., Jensen, P.E., and 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.

Gill, S.C., and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. *182*, 319–326.

Göhre, V., Ossenbühl, F., Crèvecoeur, M., Eichacker, L.A., and Rochaix, J.-D. (2006). One of two alb3 proteins is essential for the assembly of the photosystems and for cell survival in Chlamydomonas. Plant Cell *18*, 1454–1466.

González-Leiza, S.M., de Pedro, M.A., and Ayala, J.A. (2011). AmpH, a bifunctional DD-endopeptidase and DD-carboxypeptidase of Escherichia coli. J. Bacteriol. *193*, 6887–6894.

Goto, T., Aoki, R., Minamizaki, K., and 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.

Graham, J.E. (1998). Carotenoid biosynthesis in Synechococcus sp. PCC 7002: Identification of the enzymes and the carotenoids.

Graham, J.E., and Bryant, D.A. (2008). The Biosynthetic Pathway for Synechoxanthin, an Aromatic Carotenoid Synthesized by the Euryhaline, Unicellular Cyanobacterium Synechococcus sp. Strain PCC 7002. J. Bacteriol. *190*, 7966–7974.

Graham, J.E., and Bryant, D.A. (2009). The Biosynthetic Pathway for Myxol-2' Fucoside (Myxoxanthophyll) in the Cyanobacterium Synechococcus sp. Strain PCC 7002. J. Bacteriol. *191*, 4485–4485.

Grasses, T., Grimm, B., Koroleva, O., and Jahns, P. (2001). Loss of alpha-tocopherol in tobacco plants with decreased geranylgeranyl reductase activity does not modify photosynthesis in optimal growth conditions but increases sensitivity to high-light stress. Planta *213*, 620–628.

Griffin, B.M., Schott, J., and Schink, B. (2007). Nitrite, an Electron Donor for Anoxygenic Photosynthesis. Science (80-.). *316*, 1870–1870.

Grossman, A.R. (1990). Chromatic adaptation and the events involved in phycobilisome biosynthesis. Plant, Cell Environ. *13*, 651–666.

Grundmeier, A., and Dau, H. (2012). Structural models of the manganese complex of photosystem II and mechanistic implications. Biochim. Biophys. Acta - Bioenerg. *1817*, 88–105.

Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A., and Saenger, W. (2009). Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. Nat. Struct. Mol. Biol. *16*, 334–342.

Gutu, A., Chang, F., and O'Shea, E.K. (2018). Dynamical localization of a thylakoid membrane binding protein is required for acquisition of photosynthetic competency. Mol. Microbiol. *108*, 16–31.

Hahn, A., Vonck, J., Mills, D.J., Meier, T., and Kühlbrandt, W. (2018). Structure, mechanism, and regulation of the chloroplast ATP synthase. Science (80-.). *360*, eaat4318.

Hansson, M., and Hederstedt, L. (1994). Bacillus subtilis HemY is a peripheral membrane protein essential for protoheme IX synthesis which can oxidize coproporphyrinogen III and protoporphyrinogen

IX. J. Bacteriol. *176*, 5962–5970.

Harada, J., Nagashima, K.V.P., Takaichi, S., Misawa, N., Matsuura, K., and Shimada, K. (2001). Phytoene Desaturase, CrtI, of the Purple Photosynthetic Bacterium, Rubrivivax gelatinosus, Produces both Neurosporene and Lycopene. Plant Cell Physiol. *42*, 1112–1118.

Harada, J., Miyago, S., Mizoguchi, T., Azai, C., Inoue, K., Tamiaki, H., and Oh-oka, H. (2008). Accumulation of chlorophyllous pigments esterified with the geranylgeranyl group and photosynthetic competence in the CT2256-deleted mutant of the green sulfur bacterium Chlorobium tepidum. Photochem. Photobiol. Sci. *7*, 1179.

Havaux, M., Guedeney, G., He, Q., and Grossman, A.R. (2003). Elimination of high-light-inducible polypeptides related to eukaryotic chlorophyll a/b-binding proteins results in aberrant photoacclimation in Synechocystis PCC6803. Biochim. Biophys. Acta - Bioenerg. *1557*, 21–33.

He, Q., and Vermaas, W. (1998). Chlorophyll a availability affects psbA translation and D1 precursor processing in vivo in Synechocystis sp. PCC 6803. Proc. Natl. Acad. Sci. U. S. A. *95*, 5830–5835.

He, Q., Dolganov, N., Björkman, O., and Grossman, A.R. (2001). The high light-inducible polypeptides in Synechocystis PCC6803. Expression and function in high light. J. Biol. Chem. *276*, 306–314.

Heidrich, J., Thurotte, A., and Schneider, D. (2017). Specific interaction of IM30/Vipp1 with cyanobacterial and chloroplast membranes results in membrane remodeling and eventually in membrane fusion. Biochim. Biophys. Acta - Biomembr. *1859*, 537–549.

Heinz, S., Liauw, P., Nickelsen, J., and Nowaczyk, M. (2016). Analysis of photosystem II biogenesis in cyanobacteria. Biochim. Biophys. Acta - Bioenerg. *1857*, 274–287.

Hennig, R., Heidrich, J., Saur, M., Schmüser, L., Roeters, S.J., Hellmann, N., Woutersen, S., Bonn, M., Weidner, T., Markl, J., et al. (2015). IM30 triggers membrane fusion in cyanobacteria and chloroplasts. Nat. Commun. *6*, 7018.

Herbst, J., Girke, A., Hajirezaei, M.R., Hanke, G., and Grimm, B. (2018). Potential roles of YCF54 and ferredoxin-NADPH reductase for magnesium protoporphyrin monomethylester cyclase. Plant J. *94*, 485–496.

Hernandez-Prieto, M.A., Tibiletti, T., Abasova, L., Kirilovsky, D., Vass, I., and Funk, C. (2011). The small CAB-like proteins of the cyanobacterium Synechocystis sp. PCC 6803: Their involvement in chlorophyll biogenesis for Photosystem II. Biochim. Biophys. Acta - Bioenerg. *1807*, 1143–1151.

Hey, D., and Grimm, B. (2018). ONE-HELIX PROTEIN2 (OHP2) Is Required for the Stability of OHP1 and Assembly Factor HCF244 and Is Functionally Linked to PSII Biogenesis. Plant Physiol. *177*, 1453–1472.

Hey, D., Rothbart, M., Herbst, J., Wang, P., Müller, J., Wittmann, D., Gruhl, K., and Grimm, B. (2017). LIL3, a Light-Harvesting Complex Protein, Links Terpenoid and Tetrapyrrole Biosynthesis in Arabidopsis thaliana. Plant Physiol. *174*, 1037–1050.

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

Heyes, D.J., Heathcote, P., Rigby, S.E.J., Palacios, M.A., van Grondelle, R., and Hunter, C.N. (2006). The first catalytic step of the light-driven enzyme protochlorophyllide oxidoreductase proceeds via a charge transfer complex. J. Biol. Chem. *281*, 26847–26853.

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

Heyes, D.J., Menon, B.R.K., Sakuma, M., and 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.

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

Hinchigeri, S.B., Hundle, B., and 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.

Hippler, M., Drepper, F., Haehnel, W., and Rochaix, J.D. (1998). The N-terminal domain of PsaF: precise recognition site for binding and fast electron transfer from cytochrome c6 and plastocyanin to photosystem I of Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. U. S. A. *95*, 7339–7344.

Hippler, M., Drepper, F., Rochaix, J.D., and Mühlenhoff, U. (1999). Insertion of the N-terminal part of PsaF from Chlamydomonas reinhardtii into photosystem I from Synechococcus elongatus enables efficient binding of algal plastocyanin and cytochrome c6. J. Biol. Chem. *274*, 4180–4188.

Hitchcock, A., Jackson, P.J., Chidgey, J.W., Dickman, M.J., Hunter, C.N., and Canniffe, D.P. (2016). Biosynthesis of Chlorophyll *a* in a Purple Bacterial Phototroph and Assembly into a Plant Chlorophyll–Protein Complex. ACS Synth. Biol. *5*, 948–954.

Hollingshead, S., Kopečná, J., Jackson, P.J., Canniffe, D.P., Davison, P.A., Dickman, M.J., Sobotka, R., and 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.

Hollingshead, S., Bliss, S., Baker, P.J., and Neil Hunter, C. (2017). Conserved residues in Ycf54 are required for protochlorophyllide formation in Synechocystis sp. PCC 6803. Biochem. J. *474*, 667–681.

Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Pat Cerretti, D., Urdal, D.L., and Conlon, P.J. (1988). A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. Bio/Technology *6*, 1204–1210.

Hossain, M.M., and Nakamoto, H. (2002). HtpG Plays a Role in Cold Acclimation in Cyanobacteria. Curr. Microbiol. *44*, 291–296.

Hossain, M.M., and Nakamoto, H. (2003). Role for the Cyanobacterial HtpG in Protection from Oxidative Stress. Curr. Microbiol. *46*, 70–76.

Hunter, C.N., and Coomber, S.A. (1988). Cloning and Oxygen-regulated Expression of the Bacteriochlorophyll Biosynthesis Genes bch E, B, A and C of Rhodobacter sphaeroides. Microbiology *134*, 1491–1497.

Hutin, C., Nussaume, L., Moise, N., Moya, I., Kloppstech, K., and Havaux, M. (2003). Early light-induced proteins protect Arabidopsis from photooxidative stress. Proc. Natl. Acad. Sci. U. S. A. *100*, 4921–4926.

Hwang, H.J., Dilbeck, P., Debus, R.J., and Burnap, R.L. (2007). Mutation of Arginine 357 of the CP43 Protein of Photosystem II Severely Impairs the Catalytic S-State Cycle of the H_2O Oxidation Complex [†]. Biochemistry *46*, 11987–11997.

Ikeuchi, M. (1996). Complete genome sequence of a cyanobacterium Synechocystis sp. PCC 6803, the oxygenic photosynthetic prokaryote. Tanpakushitsu Kakusan Koso. *41*, 2579–2583.

Inagaki, N., Yamamoto, Y., and Satoh, K. (2001). A sequential two-step proteolytic process in the carboxyl-terminal truncation of precursor D1 protein in *Synechocystis* sp. PCC6803 [1]. FEBS Lett. *509*, 197–201.

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

Ito, H., Yokono, M., Tanaka, R., and 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. (1993). Porphyrin Accumulation and Export by Isolated Barley (Hordeum vulgare) Plastids (Effect of Diphenyl Ether Herbicides). Plant Physiol. *101*, 1181–1187.

Jansson, S. (1999). A guide to the Lhc genes and their relatives in Arabidopsis. Trends Plant Sci. *4*, 236–240.

Jansson, S. (2008). A Protein Family Saga: From Photoprotection to Light-Harvesting (and Back?). In Photoprotection, Photoinhibition, Gene Regulation, and Environment, (Dordrecht: Springer Netherlands), pp. 145–153.

Jansson, S., Andersson, J., Kim, S.J., and Jackowski, G. (2000). An Arabidopsis thaliana protein homologous to cyanobacterial high-light-inducible proteins. Plant Mol. Biol. *42*, 345–351.

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

Jensen, P.E., Gibson, L.C., and Hunter, C.N. (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.

Jensen, P.E., Haldrup, A., Zhang, S., and Scheller, H.V. (2004). The PSI-O Subunit of Plant Photosystem I Is Involved in Balancing the Excitation Pressure between the Two Photosystems. J. Biol. Chem. *279*, 24212–24217.

Jia, L., Dienhart, M., Schramp, M., McCauley, M., Hell, K., and Stuart, R.A. (2003). Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1. EMBO J. *22*, 6438–6447.

Jiang, F., Chen, M., Yi, L., de Gier, J.-W., Kuhn, A., and Dalbey, R.E. (2003). Defining the regions of Escherichia coli YidC that contribute to activity. J. Biol. Chem. *278*, 48965–48972.

Johnson, E.T., and Schmidt-Dannert, C. (2008). 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.

Jordan, P., Fromme, P., Witt, H.T., Klukas, O., Saenger, W., and Krausz, N. (2001). Three-dimensional structure of cyanobacterial photosystem I at 2.5[thinsp][angst] resolution. Nature *411*, 909–917.

Junglas, B., and Schneider, D. (2018). What is Vipp1 good for? Mol. Microbiol. *108*, 1–5.

Kalkhof, S., and Sinz, A. (2008). Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. Anal. Bioanal. Chem. *392*, 305–312.

Kallberg, Y., Oppermann, U., Jörnvall, H., and Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). Coenzyme-based functional assignments in completed genomes. Eur. J. Biochem. *269*, 4409–4417.

Kamata, T., Hiramoto, H., Morita, N., Shen, J.-R., Mann, N.H., and Yamamoto, Y. (2005). Quality control of Photosystem II: an FtsH protease plays an essential role in the turnover of the reaction center D1 protein in Synechocystis PCC 6803 under heat stress as well as light stress conditions. Photochem. Photobiol. Sci. *4*, 983.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., et al. (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.

Kao, A., Chiu, C., Vellucci, D., Yang, Y., Patel, V.R., Guan, S., Randall, A., Baldi, P., Rychnovsky, S.D., and Huang, L. (2011). Development of a Novel Cross-linking Strategy for Fast and Accurate Identification of Cross-linked Peptides of Protein Complexes. Mol. Cell. Proteomics *10*, M110.002212.

Kapri-Pardes, E., Naveh, L., and Adam, Z. (2007). The thylakoid lumen protease Deg1 is involved in the repair of photosystem II from photoinhibition in Arabidopsis. Plant Cell *19*, 1039–1047.

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

Kashino, Y., Lauber, W.M., Carroll, J.A., Wang, Q., Whitmarsh, J., Satoh, K., and Himadri, P.B. (2002). Proteomic Analysis of a Highly Active Photosystem II Preparation from the Cyanobacterium Synechocystis sp. PCC 6803 Reveals the Presence of Novel Polypeptides†.

Kato, Y., and Sakamoto, W. (2009). Protein Quality Control in Chloroplasts: A Current Model of D1 Protein Degradation in the Photosystem II Repair Cycle. J. Biochem. *146*, 463–469.

Kato, Y., and Sakamoto, W. (2014). Phosphorylation of photosystem II core proteins prevents undesirable cleavage of D1 and contributes to the fine-tuned repair of photosystem II. Plant J. *79*, 312–321.

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

Kawakami, K., Umena, Y., Iwai, M., Kawabata, Y., Ikeuchi, M., Kamiya, N., and Shen, J.-R. (2011). Roles of PsbI and PsbM in photosystem II dimer formation and stability studied by deletion mutagenesis and

X-ray crystallography. Biochim. Biophys. Acta - Bioenerg. *1807*, 319–325.

Kedrov, A., Sustarsic, M., de Keyzer, J., Caumanns, J.J., Wu, Z.C., and Driessen, A.J.M. (2013). Elucidating the Native Architecture of the YidC: Ribosome Complex. J. Mol. Biol. *425*, 4112–4124.

Keller, Y., Bouvier, F., d'Harlingue, A., and Camara, B. (1998). Metabolic compartmentation of plastid prenyllipid biosynthesis--evidence for the involvement of a multifunctional geranylgeranyl reductase. Eur. J. Biochem. *251*, 413–417.

Keren, N., Liberton, M., and Pakrasi, H.B. (2005). Photochemical competence of assembled photosystem II core complex in cyanobacterial plasma membrane. J. Biol. Chem. *280*, 6548–6553.

Kerfeld, C.A., Melnicki, M.R., Sutter, M., and Dominguez-Martin, M.A. (2017). Structure, function and evolution of the cyanobacterial orange carotenoid protein and its homologs. New Phytol. *215*, 937–951.

Khrouchtchova, A., Hansson, M., Paakkarinen, V., Vainonen, J.P., Zhang, S., Jensen, P.E., Scheller, H.V., Vener, A. V., Aro, E.-M., and Haldrup, A. (2005). A previously found thylakoid membrane protein of 14 kDa (TMP14) is a novel subunit of plant photosystem I and is designated PSI-P. FEBS Lett. *579*, 4808–4812.

Kiel, J.A.K.W., Ten Berge, A.M., and Venema, G. (1992). Nucleotide sequence of the Synechococcus sp. PCC7942 Heme gene encoding the homologue of mammalian uroporphyrinogen decarboxylase. DNA Seq. *2*, 415–418.

Kiley, P.J., and Kaplan, S. (1988). Molecular genetics of photosynthetic membrane biosynthesis in Rhodobacter sphaeroides. Microbiol. Rev. *52*, 50–69.

Kim, E.J., and Lee, J.K. (2010). Competitive Inhibitions of the Chlorophyll Synthase of Synechocystis sp. Strain PCC 6803 by Bacteriochlorophyllide a and the Bacteriochlorophyll Synthase of Rhodobacter sphaeroides by Chlorophyllide a. J. Bacteriol. *192*, 198–207.

Kim, E.-J., Kim, H., and Lee, J.K. (2016). The Photoheterotrophic Growth of Bacteriochlorophyll Synthase-Deficient Mutant of Rhodobacter sphaeroides Is Restored by I44F Mutant Chlorophyll Synthase of Synechocystis sp. PCC 6803. J. Microbiol. Biotechnol. *26*, 959–966.

Kim, J., Eichacker, L.A., Rudiger, W., and Mullet, J.E. (1994a). Chlorophyll regulates accumulation of the plastid-encoded chlorophyll proteins P700 and D1 by increasing apoprotein stability. Plant Physiol. *104*, 907–916.

Kim, J., Klein, P.G., and Mullet, J.E. (1994b). Synthesis and turnover of photosystem II reaction center protein D1. Ribosome pausing increases during chloroplast development. J. Biol. Chem. *269*, 17918–17923.

Klimmek, F., Sjodin, A., Noutsos, C., Leister, D., and Jansson, S. (2006). Abundantly and Rarely Expressed Lhc Protein Genes Exhibit Distinct Regulation Patterns in Plants. PLANT Physiol. *140*, 793–804.

Klinkert, B., Ossenbühl, F., Sikorski, M., Berry, S., Eichacker, L., and Nickelsen, J. (2004). PratA, a Periplasmic Tetratricopeptide Repeat Protein Involved in Biogenesis of Photosystem II in *Synechocystis* sp. PCC 6803. J. Biol. Chem. *279*, 44639–44644.

Klostermann, E., Droste Gen Helling, I., Carde, J.-P., and Schünemann, D. (2002). The thylakoid membrane protein ALB3 associates with the cpSecY-translocase in Arabidopsis thaliana. Biochem. J. *368*, 777–781.

Knoppová, J., Sobotka, R., Tichý, M., Yu, J., Konik, P., Halada, P., Nixon, P.J., and Komenda, J. (2014). Discovery of a Chlorophyll Binding Protein Complex Involved in the Early Steps of Photosystem II Assembly in Synechocystis. Plant Cell *26*, 1200 LP-1212.

Kobayashi, K. (2016). Role of membrane glycerolipids in photosynthesis, thylakoid biogenesis and chloroplast development. J. Plant Res. *129*, 565–580.

Kobayashi, K., Masuda, T., Tajima, N., Wada, H., and Sato, N. (2014). Molecular phylogeny and intricate evolutionary history of the three isofunctional enzymes involved in the oxidation of protoporphyrinogen IX. Genome Biol. Evol. *6*, 2141–2155.

Koch, M., Breithaupt, C., Kiefersauer, R., Freigang, J., Huber, R., and Messerschmidt, A. (2004). Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. EMBO J. *23*, 1720–1728.

Koivuniemi, A., Aro, E.-M., and Andersson, B. (1995). Degradation of the D1- and D2-Proteins of Photosystem II in Higher Plants Is Regulated by Reversible Phosphorylation. Biochemistry *34*, 16022–16029.

Kol, S., Nouwen, N., and Driessen, A.J.M. (2008). Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes. J. Biol. Chem. *283*, 31269–31273.

Komenda, J., and Barber, J. (1995). Comparison of psbO and psbH deletion mutants of Synechocystis PCC 6803 indicates that degradation of D1 protein is regulated by the QB site and dependent on protein synthesis. Biochemistry *34*, 9625–9631.

Komenda, J., and Sobotka, R. (2016). Cyanobacterial high-light-inducible proteins — Protectors of chlorophyll–protein synthesis and assembly. Biochim. Biophys. Acta - Bioenerg. *1857*, 288–295.

Komenda, J., Hassan, H.A.G., Diner, B.A., Debus, R.J., Barber, J., and Nixon, P.J. (2000). Degradation of the Photosystem II D1 and D2 proteins in different strains of the cyanobacterium Synechocystis PCC

6803 varying with respect to the type and level of psbA transcript. Plant Mol. Biol. *42*, 635–645.

Komenda, J., Reisinger, V., Müller, B.C., Dobáková, M., Granvogl, B., and Eichacker, L.A. (2004). Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in Synechocystis PCC 6803. J. Biol. Chem. *279*, 48620–48629.

Komenda, J., Barker, M., Kuviková, S., de Vries, R., Mullineaux, C.W., Tichy, M., and Nixon, P.J. (2006). The FtsH protease slr0228 is important for quality control of photosystem II in the thylakoid membrane of Synechocystis sp. PCC 6803. J. Biol. Chem. *281*, 1145–1151.

Komenda, J., Kuviková, S., Granvogl, B., Eichacker, L.A., Diner, B.A., and Nixon, P.J. (2007a). Cleavage after residue Ala352 in the C-terminal extension is an early step in the maturation of the D1 subunit of Photosystem II in Synechocystis PCC 6803. Biochim. Biophys. Acta - Bioenerg. *1767*, 829–837.

Komenda, J., Tichy, M., Prásil, O., Knoppová, J., Kuviková, S., de Vries, R., and Nixon, P.J. (2007b). The exposed N-terminal tail of the D1 subunit is required for rapid D1 degradation during photosystem II repair in Synechocystis sp PCC 6803. Plant Cell *19*, 2839–2854.

Komenda, J., Nickelsen, J., Tichý, M., Prášil, O., Eichacker, L.A., and Nixon, P.J. (2008). The Cyanobacterial Homologue of HCF136/YCF48 Is a Component of an Early Photosystem II Assembly Complex and Is Important for Both the Efficient Assembly and Repair of Photosystem II in Synechocystis sp. PCC 6803. J. Biol. Chem. *283*, 22390–22399.

Komenda, J., Knoppová, J., Krynická, V., Nixon, P.J., and Tichý, M. (2010). Role of FtsH2 in the repair of Photosystem II in mutants of the cyanobacterium Synechocystis PCC 6803 with impaired assembly or stability of the CaMn4 cluster. Biochim. Biophys. Acta - Bioenerg. *1797*, 566–575.

Komenda, J., Sobotka, R., and Nixon, P.J. (2012). Assembling and maintaining the Photosystem II complex in chloroplasts and cyanobacteria. Curr. Opin. Plant Biol. *15*, 245–251.

Kopečná, J., Sobotka, R., and 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.

Kopf, M., Klähn, S., Scholz, I., Matthiessen, J.K.F., Hess, W.R., and Voß, B. (2014). Comparative Analysis of the Primary Transcriptome of Synechocystis sp. PCC 6803. DNA Res. *21*, 527–539.

Kouřil, R., Dekker, J.P., and Boekema, E.J. (2012). Supramolecular organization of photosystem II in green plants. Biochim. Biophys. Acta - Bioenerg. *1817*, 2–12.

Kreuz, K., and Kleinig, H. (1981). On the compartmentation of isopentenyl diphosphate synthesis and utilization in plant cells. Planta *153*, 578–581.

Krieger-Liszkay, A., Fufezan, C., and Trebst, A. (2008). Singlet oxygen production in photosystem II and related protection mechanism. Photosynth. Res. *98*, 551–564.

Kroll, D., Meierhoff, K., Bechtold, N., Kinoshita, M., Westphal, S., Vothknecht, U.C., Soll, J., and Westhoff, P. (2001). VIPP1, a nuclear gene of Arabidopsis thaliana essential for thylakoid membrane formation. Proc. Natl. Acad. Sci. U. S. A. *98*, 4238–4242.

Kruse, O., Rupprecht, J., Mussgnug, J.H., Dismukes, G.C., and Hankamer, B. (2005). Photosynthesis: a blueprint for solar energy capture and biohydrogen production technologies. Photochem. Photobiol. Sci. *4*, 957.

Krynická, V., Shao, S., Nixon, P.J., and Komenda, J. (2015). Accessibility controls selective degradation of photosystem II subunits by FtsH protease. Nat. Plants *1*, 15168.

Kufryk, G.I., and Vermaas, W.F.J. (2003). Slr2013 is a novel protein regulating functional assembly of photosystem II in Synechocystis sp. strain PCC 6803. J. Bacteriol. *185*, 6615–6623.

Kufryk, G., Hernandez-Prieto, M.A., Kieselbach, T., Miranda, H., Vermaas, W., and Funk, C. (2008). Association of small CAB-like proteins (SCPs) of Synechocystis sp. PCC 6803 with Photosystem II. Photosynth. Res. *95*, 135–145.

Kunkel, D.D. (1982). Thylakoid centers: Structures associated with the cyanobacterial photosynthetic membrane system. Arch. Microbiol. *133*, 97–99.

Kurisu, G., Zhang, H., Smith, J.L., and Cramer, W.A. (2003). Structure of the Cytochrome b6f Complex of Oxygenic Photosynthesis: Tuning the Cavity. Science (80-.). *302*, 1009–1014.

Kusama, Y., Inoue, S., Jimbo, H., Takaichi, S., Sonoike, K., Hihara, Y., and Nishiyama, Y. (2015). Zeaxanthin and Echinenone Protect the Repair of Photosystem II from Inhibition by Singlet Oxygen in Synechocystis sp. PCC 6803. Plant Cell Physiol. *56*, 906–916.

Kwon, C.-T., Kim, S.-H., Song, G., Kim, D., and Paek, N.-C. (2017). Two NADPH: Protochlorophyllide Oxidoreductase (POR) Isoforms Play Distinct Roles in Environmental Adaptation in Rice. Rice (N. Y). *10*, 1.

Lagarde, D., and Vermaas, W. (1999). The zeaxanthin biosynthesis enzyme beta-carotene hydroxylase is involved in myxoxanthophyll synthesis in Synechocystis sp. PCC 6803. FEBS Lett. *454*, 247–251.

Larkin, R.M., Alonso, J.M., Ecker, J.R., and Chory, J. (2003). GUN4, a Regulator of Chlorophyll Synthesis and Intracellular Signaling. Science (80-.). *299*, 902–906.

Lash, T.D. (2005). The enigma of coproporphyrinogen oxidase: How does this unusual enzyme carry out oxidative decarboxylations to afford vinyl groups? Bioorg. Med. Chem. Lett. *15*, 4506–4509.

Layer, G., Verfürth, K., Mahlitz, E., and Jahn, D. (2002). Oxygen-independent Coproporphyrinogen-III Oxidase HemN from *Escherichia coli*. J. Biol. Chem. *277*, 34136–34142.

Lee, H.J., Duke, M. V., and Duke, S.O. (1993). Cellular Localization of Protoporphyrinogen-Oxidizing Activities of Etiolated Barley (Hordeum vulgare L.) Leaves (Relationship to Mechanism of Action of Protoporphyrinogen Oxidase-Inhibiting Herbicides). Plant Physiol. *102*, 881–889.

Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Förster, F., and Aebersold, R. (2012). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. Mol. Cell. Proteomics *11*, M111.014126.

Li, H.M., Kaneko, Y., and Keegstra, K. (1994). Molecular cloning of a chloroplastic protein associated with both the envelope and thylakoid membranes. Plant Mol. Biol. *25*, 619–632.

Liang, C., Zhao, F., Wei, W., Wen, Z., and Qin, S. (2006). Carotenoid biosynthesis in cyanobacteria: structural and evolutionary scenarios based on comparative genomics. Int. J. Biol. Sci. *2*, 197–207.

Lindsten, A., Welch, C.J., Schoch, S., Ryberg, M., Rudiger, W., and Sundqvist, C. (1990). Chlorophyll synthetase is latent in well preserved prolamellar bodies of etiolated wheat. Physiol. Plant. *80*, 277–285.

Lindsten, A., Wiktorsson, B., Ryberg, M., and Sundqvist, C. (1993). Chlorophyll synthetase activity is relocated from transforming prolamellar bodies to developing thylakoids during irradiation of dark-grown wheat. Physiol. Plant. *88*, 29–36.

Link, S., Engelmann, K., Meierhoff, K., and Westhoff, P. (2012). The atypical short-chain dehydrogenases HCF173 and HCF244 are jointly involved in translational initiation of the psbA mRNA of Arabidopsis. Plant Physiol. *160*, 2202–2218.

Liu, H., Roose, J.L., Cameron, J.C., and Pakrasi, H.B. (2011b). A Genetically Tagged Psb27 Protein Allows Purification of Two Consecutive Photosystem II (PSII) Assembly Intermediates in *Synechocystis* 6803, a Cyanobacterium. J. Biol. Chem. *286*, 24865–24871.

Liu, H., Huang, R.Y.-C., Chen, J., Gross, M.L., and Pakrasi, H.B. (2011a). Psb27, a transiently associated protein, binds to the chlorophyll binding protein CP43 in photosystem II assembly intermediates. Proc. Natl. Acad. Sci. U. S. A. *108*, 18536–18541.

Liu, H., Zhang, H., Niedzwiedzki, D.M., Prado, M., He, G., Gross, M.L., and Blankenship, R.E. (2013). Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. Science *342*, 1104–1107.

Liu, J., Yang, H., Lu, Q., Wen, X., Chen, F., Peng, L., Zhang, L., and Lu, C. (2012). PsbP-domain protein1, a nuclear-encoded thylakoid lumenal protein, is essential for photosystem I assembly in Arabidopsis. Plant Cell *24*, 4992–5006.

Llansola-Portoles, M.J., Sobotka, R., Kish, E., Shukla, M.K., Pascal, A.A., Polívka, T., and Robert, B. (2017). Twisting a β-Carotene, an Adaptive Trick from Nature for Dissipating Energy during Photoprotection. J. Biol. Chem. *292*, 1396–1403.

Lohr, M., Im, C.-S., and Grossman, A.R. (2005). Genome-Based Examination of Chlorophyll and Carotenoid Biosynthesis in Chlamydomonas reinhardtii Plant Physiol. *138*, 490 LP-515.

Lohscheider, J.N., Rojas-Stütz, M.C., Rothbart, M., Andersson, U., Funck, D., Mendgen, K., Grimm, B., and Adamska, I. (2015). Altered levels of LIL3 isoforms in *Arabidopsis* lead to disturbed pigment-protein assembly and chlorophyll synthesis, chlorotic phenotype and impaired photosynthetic performance. Plant. Cell Environ. *38*, 2115–2127.

Lopez, J.C., Ryan, S., and Blankenship, R.E. (1996). Sequence of the bchG gene from Chloroflexus aurantiacus: relationship between chlorophyll synthase and other polyprenyltransferases. J. Bacteriol. *178*, 3369–3373.

Lütz, C., Benz, J., and Rüdiger, W. (1981). Esterification of Chlorophyllide in Prolamellar Body (PLB) and Prothylakoid (PT) Fractions from Avena sativa Etioplasts. Zeitschrift Für Naturforsch. C *36*, 58–61.

MacGregor-Chatwin, C., Sener, M., Barnett, S.F.H., Hitchcock, A., Barnhart-Dailey, M.C., Maghlaoui, K., Barber, J., Timlin, J.A., Schulten, K., and Hunter, C.N. (2017). Lateral Segregation of Photosystem I in Cyanobacterial Thylakoids. Plant Cell *29*, 1119–1136.

Mackenzie, C., Eraso, J.M., Choudhary, M., Roh, J.H., Zeng, X., Bruscella, P., Puskás, Á., and Kaplan, S. (2007). Postgenomic Adventures with *Rhodobacter sphaeroides*. Annu. Rev. Microbiol. *61*, 283–307.

Malavath, T., Caspy, I., Netzer-El, S.Y., Klaiman, D., and Nelson, N. (2018). Structure and function of wild-type and subunit-depleted photosystem I in Synechocystis. Biochim. Biophys. Acta - Bioenerg. *1859*, 645–654.

Maresca, J.A., Graham, J.E., Wu, M., Eisen, J.A., and Bryant, D.A. (2007). Identification of a fourth family of lycopene cyclases in photosynthetic bacteria. Proc. Natl. Acad. Sci. *104*, 11784–11789.

Martínez-Férez, I.M., and Vioque, A. (1992). Nucleotide sequence of the phytoene desaturase gene from Synechocystis sp. PCC 6803 and characterization of a new mutation which confers resistance to the herbicide norflurazon. Plant Mol. Biol. *18*, 981–983.

Martínez-Férez, I., Fernández-González, B., Sandmann, G., and Vioque, A. (1994). Cloning and expression in Escherichia coli of the gene coding for phytoene synthase from the cyanobacterium Synechocystis sp. PCC6803. Biochim. Biophys. Acta - Gene Struct. Expr. *1218*, 145–152.

Martins, B.M., Grimm, B., Mock, H.P., Huber, R., and 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.

Masamoto, K., Misawa, N., Kaneko, T., Kikuno, R., and Toh, H. (1998). Beta-carotene hydroxylase gene from the cyanobacterium Synechocystis sp. PCC6803. Plant Cell Physiol. *39*, 560–564.

Masamoto, K., Zsiros, O., and Gombos, Z. (1999). Accumulation of Zeaxanthin in Cytoplasmic Membranes of the Cyanobacterium Synechococcus sp. Strain PCC7942 Grown under High Light Condition. J. Plant Physiol. *155*, 136–138.

Masamoto, K., Wada, H., Kaneko, T., and Takaichi, S. (2001). Identification of a Gene Required for cis-to-trans Carotene Isomerization in Carotenogenesis of the Cyanobacterium Synechocystis sp. PCC 6803. Plant Cell Physiol. *42*, 1398–1402.

Masoumi, A., Heinemann, I.U., Rohde, M., Koch, M., Jahn, M., and Jahn, D. (2008). Complex formation between protoporphyrinogen IX oxidase and ferrochelatase during haem biosynthesis in Thermosynechococcus elongatus. Microbiology *154*, 3707–3714.

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.

Mazor, Y., Borovikova, A., Caspy, I., and Nelson, N. (2015). Structure of the plant photosystem I supercomplex at 2.6 Å resolution. Nat. Plants *3*, 17014.

McFadden, G.I. (2001). Chloroplast origin and integration. Plant Physiol. *125*, 50–53.

McGlynn, P., and Hunter, C.N. (1993). Genetic analysis of the bchC and bchA genes of Rhodobacter sphaeroides. Mol. Gen. Genet. *236*, 227–234.

Van De Meene, A.M.L., Hohmann-Marriott, M.F., Vermaas, W.F.J., and Roberson, R.W. (2006). The three-dimensional structure of the cyanobacterium Synechocystis sp. PCC 6803. Arch. Microbiol. *184*, 259–270.

Melzer, M., and Heide, L. (1994). Characterization of Polyprenyldiphosphate: 4-Hydroxybenzoate Polyprenyltransferase from Escherichia coli. Biochim. Biophys. Acta - Lipids Lipid Metab. *1212*, 93–102.

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

Menon, B.R.K., Davison, P.A., Hunter, C.N., Scrutton, N.S., and Heyes, D.J. (2010). Mutagenesis alters the catalytic mechanism of the light-driven enzyme protochlorophyllide oxidoreductase. J. Biol. Chem. *285*, 2113–2119.

Merati, G., and Zanetti, G. (1987). Chemical cross-linking of ferredoxin to spinach thylakoids Evidence for two independent binding sites of ferredoxin to the membrane. FEB *215*, 37–40.

Mikami, K., Kanesaki, Y., Suzuki, I., and Murata, N. (2002). The histidine kinase Hik33 perceives osmotic stress and cold stress in Synechocystis sp PCC 6803. Mol. Microbiol. *46*, 905–915.

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

Moellering, E.R., and Benning, C. (2011). Galactoglycerolipid metabolism under stress: A time for remodeling. Trends Plant Sci. *16*, 98–107.

Mohamed, H.E., van de Meene, A.M.L., Roberson, R.W., and Vermaas, W.F.J. (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.

Moore, M., Harrison, M.S., Peterson, E.C., and Henry, R. (2000a). Chloroplast Oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. J. Biol. Chem. *275*, 1529–1532.

Moore, M., Harrison, M.S., Peterson, E.C., and Henry, R. (2000b). Albino3 Is Required for Post-translational Integration of the Light Harvesting Chlorophyll-binding Protein into Thylakoid Membranes *. 1529–1533.

Moore, M., Goforth, R.L., Mori, H., and Henry, R. (2003). Functional interaction of chloroplast SRP/FtsY with the ALB3 translocase in thylakoids. J. Cell Biol. *162*, 1245 LP-1254.

Mork-Jansson, A., Bue, A.K., Gargano, D., Furnes, C., Reisinger, V., Arnold, J., Kmiec, K., and Eichacker, L.A. (2015). Lil3 Assembles with Proteins Regulating Chlorophyll Synthesis in Barley. PLoS One *10*, e0133145.

Mothersole, D.J., Jackson, P.J., Vasilev, C., Tucker, J.D., Brindley, A.A., Dickman, M.J., and Hunter, C.N. (2016). PucC and LhaA direct efficient assembly of the light-harvesting complexes in *Rhodobacter sphaeroides*. Mol. Microbiol. *99*, 307–327.

Muhlenhoff, U., Zhao, J., and Bryant, D.A. (1996). Interaction between Photosystem I and Flavodoxin from the Cyanobacterium Synechococcus sp. PCC 7002 as Revealed by Chemical Cross-Linking. Eur. J. Biochem. *235*, 324–331.

Müller, B., and Eichacker, L.A. (1999). Assembly of the D1 precursor in monomeric photosystem II reaction center precomplexes precedes chlorophyll a-triggered accumulation of reaction center II in barley etioplasts. Plant Cell *11*, 2365–2377.

Muneer, S., Kim, E.J., Park, J.S., and Lee, J.H. (2014). Influence of green, red and blue light emitting diodes on multiprotein complex proteins and photosynthetic activity under different light intensities in lettuce leaves (Lactuca sativa L.). Int. J. Mol. Sci. *15*, 4657–4670.

Muranaka, L.S., Rütgers, M., Bujaldon, S., Heublein, A., Geimer, S., Wollman, F.-A., and Schroda, M. (2016). TEF30 Interacts with Photosystem II Monomers and Is Involved in the Repair of Photodamaged Photosystem II in *Chlamydomonas reinhardtii*. Plant Physiol. *170*, 821–840.

Murata, N., and Nishiyama, Y. (2018). ATP is a driving force in the repair of photosystem II during photoinhibition. Plant. Cell Environ. *41*, 285–299.

Myouga, F., Takahashi, K., Tanaka, R., Nagata, N., Kiss, A.Z., Funk, C., Nomura, Y., Nakagami, H., Jansson, S., and Shinozaki, K. (2018). Stable Accumulation of Photosystem II Requires ONE-HELIX PROTEIN1 (OHP1) of the Light Harvesting-Like Family. Plant Physiol. *176*, 2277–2291.

Nagamori, S., Smirnova, I.N., and Kaback, H.R. (2004). Role of YidC in folding of polytopic membrane proteins. J. Cell Biol. *165*, 53–62.

Nagata, N., Tanaka, R., Satoh, S., and Tanaka, A. (2005). Identification of a Vinyl Reductase Gene for Chlorophyll Synthesis in Arabidopsis thaliana and Implications for the Evolution of Prochlorococcus Species. PLANT CELL ONLINE *17*, 233–240.

Nagata, N., Tanaka, R., and Tanaka, A. (2007). The Major Route for Chlorophyll Synthesis Includes [3,8-divinyl]-chlorophyllide a Reduction in Arabidopsis thaliana. Plant Cell Physiol. *48*, 1803–1808.

Nakanishi, H., Nozue, H., Suzuki, K., Kaneko, Y., Taguchi, G., and Hayashida, N. (2005). Characterization of the Arabidopsis thaliana Mutant pcb2 which Accumulates Divinyl Chlorophylls. Plant Cell Physiol. *46*, 467–473.

Naver, H., Boudreau, E., and Rochaix, J.D. (2001). Functional studies of Ycf3: its role in assembly of photosystem I and interactions with some of its subunits. Plant Cell *13*, 2731–2745.

Naylor, G.W., Addlesee, H.A., Gibson, L.C.D., and Hunter, C.N. (1999). The photosynthesis gene cluster of Rhodobacter sphaeroides. Photosynth. Res. *62*, 121–139.

Neilson, J.A.D., and Durnford, D.G. (2010a). Evolutionary distribution of light-harvesting complex-like proteins in photosynthetic eukaryotes. Genome *53*, 68–78.

Neilson, J.A.D., and Durnford, D.G. (2010b). Structural and functional diversification of the light-harvesting complexes in photosynthetic eukaryotes. Photosynth. Res. *106*, 57–71.

Nellaepalli, S., Ozawa, S.-I., Kuroda, H., and Takahashi, Y. (2018). The photosystem I assembly apparatus consisting of Ycf3–Y3IP1 and Ycf4 modules. Nat. Commun. *9*, 2439.

Nickelsen, J., and Rengstl, B. (2013). Photosystem II assembly: from cyanobacteria to plants. Annu. Rev. Plant Biol. *64*, 609–635.

Nickelsen, J., and Zerges, W. (2013). Thylakoid biogenesis has joined the new era of bacterial cell biology. Front. Plant Sci. *4*, 458.

Nickelsen, J., Rengstl, B., Stengel, A., Schottkowski, M., Soll, J., and Ankele, E. (2011). Biogenesis of the cyanobacterial thylakoid membrane system - an update. FEMS Microbiol. Lett. *315*, 1–5.

Niedzwiedzki, D.M., Tronina, T., Liu, H., Staleva, H., Komenda, J., Sobotka, R., Blankenship, R.E., and Polívka, T. (2016). Carotenoid-induced non-photochemical quenching in the cyanobacterial chlorophyll synthase–HliC/D complex. Biochim. Biophys. Acta - Bioenerg. *1857*, 1430–1439.

Van Niel, C.B. (1962). The present status of the comparative study of photosynthesis. Annu. Rev. Plant Physiol. *13*, 1–25.

Nilsson, R., and van Wijk, K.J. (2002). Transient interaction of cpSRP54 with elongating nascent chains of the chloroplast-encoded D1 protein; “cpSRP54 caught in the act”. FEBS Lett. *524*, 127–133.

Nilsson, R., Brunner, J., Hoffman, N.E., and Wijk, K.J. van (1999). Interactions of ribosome nascent chain complexes of the chloroplast-encoded D1 thylakoid membrane protein with cpSRP54. EMBO J. *18*, 733–742.

Nixon, P.J., and Diner, B.A. (1994). Analysis of water-oxidation mutants constructed in the cyanobacterium Synechocystis sp. PCC 6803. Biochem. Soc. Trans. *22*, 338–343.

Nixon, P.J., Barker, M., Boehm, M., de Vries, R., and Komenda, J. (2004). FtsH-mediated repair of the photosystem II complex in response to light stress. J. Exp. Bot. *56*, 357–363.

Nixon, P.J., Michoux, F., Yu, J., Boehm, M., and Komenda, J. (2010). Recent advances in understanding the assembly and repair of photosystem II. Ann. Bot. *106*, 1–16.

Niyogi, K.K., Li, X.-P., Rosenberg, V., and Jung, H.-S. (2004). Is PsbS the site of non-photochemical quenching in photosynthesis? J. Exp. Bot. *56*, 375–382.

Nomata, J., Mizoguchi, T., Tamiaki, H., and Fujita, Y. (2006). A Second Nitrogenase-like Enzyme for Bacteriochlorophyll Biosynthesis. J. Biol. Chem. *281*, 15021–15028.

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

Nowaczyk, M.M., Hebeler, R., Schlodder, E., Meyer, H.E., Warscheid, B., and Rogner, M. (2006). Psb27, a Cyanobacterial Lipoprotein, Is Involved in the Repair Cycle of Photosystem II. PLANT CELL ONLINE *18*, 3121–3131.

O'Brian, M.R., and Panek, H. (2002). A whole genome view of prokaryotic haem biosynthesis. Microbiology *148*, 2273–2282.

Oborník, M., and Green, B.R. (2005). Mosaic Origin of the Heme Biosynthesis Pathway in Photosynthetic Eukaryotes. Mol. Biol. Evol. *22*, 2343–2353.

Ogawa, T., Bovey, F., and Shibata, K. (1975). An intermediate in the phytylation of chlorophyllide a in vivo. Plant Cell Physiol. *16*, 199–202.

Ohad, I., Kyle, D.J., and Arntzen, C.J. (1984). Membrane protein damage and repair: removal and replacement of inactivated 32-kilodalton polypeptides in chloroplast membranes. J. Cell Biol. *99*, 481–485.

Ortega, J., Iwanczyk, J., and Jomaa, A. (2009). Escherichia coli DegP: a structure-driven functional model. J. Bacteriol. *191*, 4705–4713.

Ossenbühl, F., Göhre, V., Meurer, J., Krieger-Liszkay, A., Rochaix, J.-D., and Eichacker, L.A. (2004). Efficient assembly of photosystem II in Chlamydomonas reinhardtii requires Alb3.1p, a homolog of Arabidopsis ALBINO3. Plant Cell *16*, 1790–1800.

Ossenbühl, F., Inaba-Sulpice, M., Meurer, J., Soll, J., and Eichacker, L.A. (2006). The synechocystis sp PCC 6803 oxa1 homolog is essential for membrane integration of reaction center precursor protein pD1. Plant Cell *18*, 2236–2246.

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., and Rüdiger, W. (1997). 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., and Astier, C. (2004). Aerobic and Anaerobic Mg-Protoporphyrin Monomethyl Ester Cyclases in Purple Bacteria. J. Biol. Chem. *279*, 6385–6394.

Pan, X., Ma, J., Su, X., Cao, P., Chang, W., Liu, Z., Zhang, X., and Li, M. (2018). Structure of the maize photosystem I supercomplex with light-harvesting complexes I and II. Science *360*, 1109–1113.

Parham, R., and Rebeiz, C.A. (1995). Chloroplast Biogenesis 72: A [4-Vinyl]Chlorophyllide a Reductase Assay Using Divinyl Chlorophyllide a as an Exogenous Substrate. Anal. Biochem. *231*, 164–169.

Pasch, J.C., Nickelsen, J., and Schünemann, D. (2005). The yeast split-ubiquitin system to study chloroplast membrane protein interactions. Appl. Microbiol. Biotechnol. *69*, 440–447.

Peers, G., Truong, T.B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A.R., Hippler, M., and Niyogi, K.K. (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. Nature *462*, 518–521.

Peng, L., Ma, J., Chi, W., Guo, J., Zhu, S., Lu, Q., Lu, C., and Zhang, L. (2006). LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in Arabidopsis thaliana. Plant Cell *18*, 955–969.

Peter, G.F., and Thornber, J.P. (1991). Biochemical composition and organization of higher plant photosystem II light-harvesting pigment-proteins. J. Biol. Chem. *266*, 16745–16754.

Petrotchenko, E. V., Olkhovik, V.K., and Borchers, C.H. (2005). Isotopically Coded Cleavable Cross-linker for Studying Protein-Protein Interaction and Protein Complexes. Mol. Cell. Proteomics *4*, 1167–1179.

Pfennig, N. (1978). General physiology and ecology of photosynthetic bacteria. In The Photosynthetic Bacteria, R.K. Clayton, and W.R. Sistrom, eds. (New York: Plenum Press), pp. 3–18.

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

Pinta, V., Picaud, M., Reiss-Husson, F., and 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.

Pisciotta, J.M., Zou, Y., and Baskakov, I. V. (2010). Light-Dependent Electrogenic Activity of Cyanobacteria. PLoS One *5*, e10821.

Plücken, H., Müller, B., Grohmann, D., Westhoff, P., and Eichacker, L.A. (2002). The HCF136 protein is essential for assembly of the photosystem II reaction center in Arabidopsis thaliana. FEBS Lett. *532*, 85–90.

Porra, R.J. (1997). Recent Progress in Porphyrin and Chlorophyll Biosynthesis. Photochem. Photobiol. *65*, 492–516.

Porra, R.J., Schafer, W., Gad'on, N., Katheder, I., Drews, G., and 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., and 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.

Pospíšil, P. (2011). Enzymatic function of cytochrome b559 in photosystem II. J. Photochem. Photobiol. B Biol. *104*, 341–347.

Powles, S.B. (1984). Photoinhibition of Photosynthesis Induced by Visible Light. Annu. Rev. Plant Physiol. *35*, 15–44.

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

Promnares, K., Komenda, J., Bumba, L., Nebesarova, J., Vacha, F., and Tichy, M. (2006). Cyanobacterial Small Chlorophyll-binding Protein ScpD (HliB) Is Located on the Periphery of Photosystem II in the Vicinity of PsbH and CP47 Subunits. J. Biol. Chem. *281*, 32705–32713.

Puthiyaveetil, S., Woodiwiss, T., Knoerdel, R., Zia, A., Wood, M., Hoehner, R., and Kirchhoff, H. (2014). Significance of the Photosystem II Core Phosphatase PBCP for Plant Viability and Protein Repair in Thylakoid Membranes. Plant Cell Physiol. *55*, 1245–1254.

Qian, P., Papiz, M.Z., Jackson, P.J., Brindley, A.A., Ng, I.W., Olsen, J.D., Dickman, M.J., Bullough, P.A., and Hunter, C.N. (2013). Three-Dimensional Structure of the *Rhodobacter sphaeroides* RC-LH1-PufX Complex: Dimerization and Quinone Channels Promoted by PufX. Biochemistry *52*, 7575–7585.

Qin, X., Sun, L., Wen, X., Yang, X., Tan, Y., Jin, H., Cao, Q., Zhou, W., Xi, Z., and Shen, Y. (2010). Structural insight into unique properties of protoporphyrinogen oxidase from Bacillus subtilis. J. Struct. Biol. *170*, 76–82.

Qin, X., Suga, M., Kuang, T., and Shen, J.-R. (2015). Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. Science (80-.). *348*, 989–995.

Raines, C.A. (2003). The Calvin cycle revisited. Photosynth. Res. *75*, 1–10.

Rappaport, F., Boussac, A., Force, D.A., Peloquin, J., Brynda, M., Sugiura, M., Un, S., Britt, R.D., and Diner, B.A. (2009). Probing the Coupling between Proton and Electron Transfer in Photosystem II Core Complexes Containing a 3-Fluorotyrosine. J. Am. Chem. Soc. *131*, 4425–4433.

Rasmussen, R.E., Erstad, S.M., Ramos-Martinez, E.M., Fimognari, L., De Porcellinis, A.J., and Sakuragi, Y. (2016). An easy and efficient permeabilization protocol for in vivo enzyme activity assays in cyanobacteria. Microb. Cell Fact. *15*, 186.

Rast, A., Heinz, S., and Nickelsen, J. (2015). Biogenesis of thylakoid membranes. Biochim. Biophys. Acta - Bioenerg. *1847*, 821–830.

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc. Natl. Acad. Sci. U. S. A. *106*, 1760–1765.

Rebeiz, C.A., Ioannides, I.M., Kolossov, V., and Kopetz, K.J. (1999). Chloroplast biogenesis 80. Proposal of a unified multibranched chlorophyll a/b biosynthetic pathway. Photosynthetica *36*, 117–128.

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.

Renger, G. (2011). Light induced oxidative water splitting in photosynthesis: Energetics, kinetics and mechanism. J. Photochem. Photobiol. B Biol. *104*, 35–43.

Rengstl, B., Oster, U., Stengel, A., and 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.

Rengstl, B., Knoppová, J., Komenda, J., and Nickelsen, J. (2013). Characterization of a Synechocystis double mutant lacking the photosystem II assembly factors YCF48 and Sll0933. Planta *237*, 471–480.

Richter, A.S., Peter, E., Rothbart, M., Schlicke, H., Toivola, J., Rintamäki, E., and Grimm, B. (2013). Posttranslational influence of NADPH-dependent thioredoxin reductase C on enzymes in tetrapyrrole synthesis. Plant Physiol. *162*, 63–73.

Richter, A.S., Wang, P., and Grimm, B. (2016). Arabidopsis Mg-Protoporphyrin IX Methyltransferase Activity and Redox Regulation Depend on Conserved Cysteines. Plant Cell Physiol. *57*, 519–527.

Rintamäki, E., Kettunen, R., and Aro, E.M. (1996). Differential D1 dephosphorylation in functional and photodamaged photosystem II centers. Dephosphorylation is a prerequisite for degradation of damaged D1. J. Biol. Chem. *271*, 14870–14875.

Roose, J.L., and Pakrasi, H.B. (2004). Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into photosystem II. J. Biol. Chem. *279*, 45417–45422.

Roose, J.L., and Pakrasi, H.B. (2008). The Psb27 Protein Facilitates Manganese Cluster Assembly in Photosystem II. J. Biol. Chem. *283*, 4044–4050.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2014). The PsbP domain protein 1 functions in the assembly of lumenal domains in photosystem I. J. Biol. Chem. *289*, 23776–23785.

Rossini, S., Casazza, A.P., Engelmann, E.C.M., Havaux, M., Jennings, R.C., and Soave, C. (2006). Suppression of both ELIP1 and ELIP2 in Arabidopsis does not affect tolerance to photoinhibition and photooxidative stress. Plant Physiol. *141*, 1264–1273.

Rudiger, W. (1993). Esterification of chlorophyllide and its implication for thylakoid development. In Pigment-Protein Complexes in Plastids : Synthesis and Assembly, C. Sundqvist, and M. Ryberg, eds. (Academic Press), pp. 219–237.

Rüdiger, W. (1992). Last Steps in Chlorophyll Biosynthesis: Esterification and Insertion into the Membrane. In Regulation of Chloroplast Biogenesis, (Boston, MA: Springer US), pp. 183–190.

Rüdiger, W. (1997). Chlorophyll metabolism: From outer space down to the molecular level. Phytochemistry *46*, 1151–1167.

Rüdiger, W., Hedden, P., Köst, H.-P., and Chapman, D.J. (1977). Esterification of chlorophyllide by geranylgeranyl pyrophosphate in a cell-free system from maize shoots. Biochem. Biophys. Res. Commun. *74*, 1268–1272.

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

Ruf, S., Kössel, H., and Bock, R. (1997). Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. J. Cell Biol. *139*, 95–102.

Ruffle, S. V, Mustafa, A.O., Kitmitto, A., Holzenburg, A., and Ford, R.C. (2000). The location of the mobile electron carrier ferredoxin in vascular plant photosystem I. J. Biol. Chem. *275*, 36250–36255.

Ruiz-Sola, M.Á., and Rodríguez-Concepción, M. (2012). Carotenoid Biosynthesis in Arabidopsis: A Colorful Pathway. Arab. B. *10*, e0158.

Sääf, A., Monné, M., de Gier, J.W., and von Heijne, G. (1998). Membrane topology of the 60-kDa Oxa1p homologue from Escherichia coli. J. Biol. Chem. *273*, 30415–30418.

Sachelaru, I., Petriman, N.A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., Drepper, F., Warscheid, B., and Koch, H.-G. (2013). YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein. J. Biol. Chem. *288*, 16295–16307.

Saito, M., Watanabe, S., Yoshikawa, H., and Nakamoto, H. (2008). Interaction of the Molecular Chaperone HtpG with Uroporphyrinogen Decarboxylase in the Cyanobacterium Synechococcus elongatus PCC 7942. Biosci. Biotechnol. Biochem. *72*, 1394–1397.

Samuelson, J.C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. (2000). YidC mediates membrane protein insertion in bacteria. Nature *406*, 637–641.

San-Miguel, T., Pérez-Bermúdez, P., and Gavidia, I. (2013). Production of soluble eukaryotic recombinant proteins in E. coli is favoured in early log-phase cultures induced at low temperature. Springerplus *2*, 89.

Sano, S. (1966). 2,4-Bis-(beta-hydroxypropionic acid) deuteroporphyrinogen IX, a possible intermediate between coproporphyrinogen 3 and protoporphyrin IX. J. Biol. Chem. *241*, 5276–5283.

Sano, S., and Granick, S. (1961). Mitochondrial coproporphyrinogen oxidase and protoporphyrin formation. J. Biol. Chem. *236*, 1173–1180.

Santabarbara, S., Casazza, A.P., Ali, K., Economou, C.K., Wannathong, T., Zito, F., Redding, K.E., Rappaport, F., and Purton, S. (2013). The requirement for carotenoids in the assembly and function of the photosynthetic complexes in Chlamydomonas reinhardtii. Plant Physiol. *161*, 535–546.

Sasarman, A., Letowski, J., Czaika, G., Ramirez, V., Nead, M.A., Jacobs, J.M., and Morais, R. (1993). Nucleotide sequence of the hemG gene involved in the protoporphyrinogen oxidase activity of Escherichia coli K12. Can. J. Microbiol. *39*, 1155–1161.

Satoh, K., and Yamamoto, Y. (2007). The carboxyl-terminal processing of precursor D1 protein of the photosystem II reaction center. Photosynth. Res. *94*, 203–215.

Sawicki, A., and Willows, R.D. (2010). BchJ and BchM interact in a 1 : 1 ratio with the magnesium chelatase BchH subunit of Rhodobacter capsulatus. FEBS J. *277*, 4709–4721.

Sawicki, A., Zhou, S., Kwiatkowski, K., Luo, M., and Willows, R.D. (2017). 1- *N* -histidine phosphorylation of ChlD by the AAA $^{+}$ ChlI2 stimulates magnesium chelatase activity in chlorophyll synthesis. Biochem. J. *474*, 2095–2105.

Schafer, L., Sandmann, M., Woitsch, S., and Sandmann, G. (2006). Coordinate up-regulation of carotenoid biosynthesis as a response to light stress in Synechococcus PCC7942. Plant, Cell Environ. *29*, 1349–1356.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum. Gene *145*, 69–73.

Scheller, H.V., Jensen, P.E., Haldrup, A., Lunde, C., and Knoetzel, J. (2001). Role of subunits in eukaryotic Photosystem I. Biochim. Biophys. Acta - Bioenerg. *1507*, 41–60.

Schmid, H.C., Oster, U., Kögel, J., Lenz, S., and Rüdiger, W. (2001). Cloning and Characterisation of Chlorophyll Synthase from Avena sativa. Biol. Chem. *382*, 903–911.

Schmid, H.C., Rassadina, V., Oster, U., Schoch, S., and Rüdiger, W. (2002). Pre-Loading of Chlorophyll Synthase with Tetraprenyl Diphosphate Is an Obligatory Step in Chlorophyll Biosynthesis. Biol. Chem. *383*, 1769–1778.

Schneider, A., Steinberger, I., Strissel, H., Kunz, H.-H., Manavski, N., Meurer, J., Burkhard, G., Jarzombski, S., Schünemann, D., Geimer, S., et al. (2014). The Arabidopsis Tellurite resistance C protein together with ALB3 is involved in photosystem II protein synthesis. Plant J. *78*, 344–356.

Schoch, S., Oster, U., Mayer, K., Feick, R., and Rüdiger, W. (1999). Subtrate Specificity of Overexpressed Bacteriochlorophyll Synthase from Chloroflexus Aurantiacus. In The Chloroplast: From Molecular Biology to Biotechnology, (Dordrecht: Springer Netherlands), pp. 213–216.

Schottkowski, M., Gkalympoudis, S., Tzekova, N., Stelljes, C., Schünemann, D., Ankele, E., and Nickelsen, J. (2009). Interaction of the periplasmic pratA factor and the PsbA (D1) protein during biogenesis of photosystem II in Synechocystis sp. PCC 6803. J. Biol. Chem. *284*, 1813–1819.

Schöttler, M.A., Albus, C.A., and Bock, R. (2011). Photosystem I: Its biogenesis and function in higher plants.

Schwabe, T.M.E., Gloddek, K., Schluesener, D., and Kruip, J. (2003). Purification of recombinant BtpA and Ycf3, proteins involved in membrane protein biogenesis in Synechocystis PCC 6803. J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. *786*, 45–59.

Seitl, I., Wickles, S., Beckmann, R., Kuhn, A., and Kiefer, D. (2014). The C-terminal regions of YidC from *Rhodopirellula baltica* and *Oceanicaulis alexandrii* bind to ribosomes and partially substitute for SRP receptor function in *Escherichia coli*. Mol. Microbiol. *91*, 408–421.

Sener, M., Strumpfer, J., Singharoy, A., Hunter, C.N., and Schulten, K. (2016). Overall energy conversion efficiency of a photosynthetic vesicle. Elife *5*, e09541.

Service, R.J., Yano, J., McConnell, I., Hwang, H.J., Niks, D., Hille, R., Wydrzynski, T., Burnap, R.L., Hillier, W., and Debus, R.J. (2011). Participation of glutamate-354 of the CP43 polypeptide in the ligation of manganese and the binding of substrate water in photosystem II. Biochemistry *50*, 63–81.

Shalygo, N., Czarnecki, O., Peter, E., and Grimm, B. (2009). Expression of chlorophyll synthase is also involved in feedback-control of chlorophyll biosynthesis. Plant Mol. Biol. *71*, 425–436.

Shen, J., Williams-Carrier, R., and Barkan, A. (2017). PSA3, a Protein on the Stromal Face of the Thylakoid Membrane, Promotes Photosystem I Accumulation in Cooperation with the Assembly Factor PYG7. Plant Physiol. *174*, 1850–1862.

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., Reid, J.D., and Hunter, C.N. (2003). Purification and kinetic characterization of the magnesium protoporphyrin IX methyltransferase from Synechocystis PCC6803. Biochem. J. *371*, 351–360.

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

Shi, L.-X., Hall, M., Funk, C., and Schröder, W.P. (2012). Photosystem II, a growing complex: Updates on newly discovered components and low molecular mass proteins. Biochim. Biophys. Acta - Bioenerg. *1817*, 13–25.

Shibata, M., Mikota, T., Yoshimura, A., Iwata, N., Tsuyama, M., and Kobayashi, Y. (2004). Chlorophyll formation and photosynthetic activity in rice mutants with alterations in hydrogenation of the chlorophyll alcohol side chain. Plant Sci. *166*, 593–600.

Shih, P.M. (2015). Cyanobacterial evolution: Fresh insight into ancient questions. Curr. Biol. *25*, R192–R193.

Shimizu, S., and Tamaki, E. (1962). Chlorophyllase of Tobacco Plants I. Preparation and Properties of Water Soluble Enzyme. Shokubutsugaku Zasshi *75*, 462–467.

Shimizu, S., and Tamaki, E. (1963). Chlorophyllase of tobacco plants. II. Enzymic phytylation of chlorophyllide and pheophorbide in vitro. Arch. Biochem. Biophys. *102*, 152–158.

Shpilyov, A. V., Zinchenko, V. V., Shestakov, S. V., Grimm, B., and Lokstein, H. (2005). Inactivation of the geranylgeranyl reductase (ChlP) gene in the cyanobacterium Synechocystis sp. PCC 6803. Biochim. Biophys. Acta - Bioenerg. *1706*, 195–203.

Shpilyov, A. V., Zinchenko, V. V., Grimm, B., and Lokstein, H. (2013). Chlorophyll *a* phytylation is required for the stability of photosystems I and II in the cyanobacterium *Synechocystis* sp. PCC 6803. Plant J. *73*, 336–346.

Shukla, M.K., Llansola-Portoles, M.J., Tichý, M., Pascal, A.A., Robert, B., and Sobotka, R. (2018a). Binding of pigments to the cyanobacterial high-light-inducible protein HliC. Photosynth. Res. *137*, 29–39.

Shukla, M.K., Jackson, P.J., Moravcová, L., Zdvihalová, B., Proctor, M.S., Brindley, A.A., Dickman, M.J., Hunter, C.N., and Sobotka, R. (2018b). Cyanobacterial LHC-like proteins control formation of the chlorophyll-synthase-Ycf39 complex. Mol. Plant. Under Rev.

Siegbahn, P.E.M. (2009). Structures and Energetics for O_2 Formation in Photosystem II. Acc. Chem. Res. *42*, 1871–1880.

Silva, P., Thompson, E., Bailey, S., Kruse, O., Mullineaux, C.W., Robinson, C., Mann, N.H., and Nixon, P.J. (2003). FtsH is involved in the early stages of repair of photosystem II in Synechocystis sp PCC 6803. Plant Cell *15*, 2152–2164.

Sinha, R.K., Komenda, J., Knoppová, J., Sedlářová, M., and Pospíšil, P. (2012). Small CAB-like proteins prevent formation of singlet oxygen in the damaged photosystem II complex of the cyanobacterium Synechocystis sp. PCC 6803. Plant, Cell Environ. *35*, 806–818.

Sinz, A., and Wang, K. (2004). Mapping spatial proximities of sulfhydryl groups in proteins using a fluorogenic cross-linker and mass spectrometry. Anal. Biochem. *331*, 27–32.

Sirijovski, N., Olsson, U., Lundqvist, J., Al-Karadaghi, S., Willows, R.D., and Hansson, M. (2006). ATPase activity associated with the magnesium chelatase H-subunit of the chlorophyll biosynthetic pathway is an artefact. Biochem. J. *400*, 477–484.

Sirpiö, S., Khrouchtchova, A., Allahverdiyeva, Y., Hansson, M., Fristedt, R., Vener, A. V., Scheller, H.V., Jensen, P.E., Haldrup, A., and Aro, E.-M. (2008). AtCYP38 ensures early biogenesis, correct assembly and sustenance of photosystem II. Plant J. *55*, 639–651.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. Protein Sci. *18*, 936–948.

Smith, C.A., Suzuki, J.Y., and 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.

Sobotka, R. (2014). Making proteins green; biosynthesis of chlorophyll-binding proteins in cyanobacteria. Photosynth. Res. *119*, 223–232.

Sobotka, R., McLean, S., Zuberova, M., Hunter, C.N., and Tichy, M. (2008). 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.

Sofia, H.J., Chen, G., Hetzler, B.G., Reyes-Spindola, J.F., and 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., and Schultz, G. (1981). Phytol synthesis from geranylgeraniol in spinach chloroplasts. Biochem. Biophys. Res. Commun. *99*, 907–912.

Soll, J., Schultz, G., Rüdiger, W., and Benz, J. (1983). Hydrogenation of geranylgeraniol : two pathways exist in spinach chloroplasts. Plant Physiol. *71*, 849–854.

Sozer, O., Kis, M., Gombos, Z., and Ughy, B. (2011). Proteins, glycerolipids and carotenoids in the functional photosystem II architecture. Front. Biosci. (Landmark Ed. *16*, 619–643.

Spence, E., Bailey, S., Nenninger, A., Møller, S.G., and Robinson, C. (2004). A homolog of Albino3/OxaI is essential for thylakoid biogenesis in the gyanobacterium Synechocystis sp. PCC6803. J. Biol. Chem. *279*, 55792–55800.

Srivastava, R., Battchikova, N., Norling, B., and Aro, E.-M. (2006). Plasma membrane of Synechocystis PCC 6803: a heterogeneous distribution of membrane proteins. Arch. Microbiol. *185*, 238–243.

Staleva, H., Komenda, J., Shukla, M.K., Šlouf, V., Kaňa, R., Polívka, T., and Sobotka, R. (2015). Mechanism of photoprotection in the cyanobacterial ancestor of plant antenna proteins. Nat Chem Biol *11*, 287–291.

Stamatakis, K., Tsimilli-Michael, M., and Papageorgiou, G.C. (2014). On the question of the light-harvesting role of β-carotene in photosystem II and photosystem I core complexes. Plant Physiol. Biochem. *81*, 121–127.

Standfuss, J., and Kühlbrandt, W. (2004). The Three Isoforms of the Light-harvesting Complex II. J. Biol. Chem. *279*, 36884–36891.

Steiger, S., Schäfer, L., and Sandmann, G. (1999). High-light-dependent upregulation of carotenoids and their antioxidative properties in the cyanobacterium Synechocystis PCC 6803. J. Photochem. Photobiol. B Biol. *52*, 14–18.

Stengel, A., Gugel, I.L., Hilger, D., Rengstl, B., Jung, H., and Nickelsen, J. (2012). Initial Steps of Photosystem II de Novo Assembly and Preloading with Manganese Take Place in Biogenesis Centers in Synechocystis. PLANT CELL ONLINE *24*, 660–675.

Stöckel, J., Bennewitz, S., Hein, P., and Oelmüller, R. (2006). The evolutionarily conserved tetratrico peptide repeat protein pale yellow green7 is required for photosystem I accumulation in Arabidopsis and copurifies with the complex. Plant Physiol. *141*, 870–878.

Storbeck, S., Rolfes, S., Raux-Deery, E., Warren, M.J., Jahn, D., and Layer, G. (2010). A novel pathway for the biosynthesis of heme in Archaea: genome-based bioinformatic predictions and experimental evidence. Archaea *2010*, 175050.

Storm, P., Hernandez-Prieto, M.A., Eggink, L.L., Hoober, J.K., and Funk, C. (2008). The small CAB-like proteins of Synechocystis sp. PCC 6803 bind chlorophyll. Photosynth. Res. *98*, 479–488.

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

Su, Q., Frick, G., Armstrong, G., and Apel, K. (2001). POR C of Arabidopsis thaliana: a third light- and NADPH-dependent protochlorophyllide oxidoreductase that is differentially regulated by light. Plant Mol. Biol. *47*, 805–813.

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H., et al. (2015). Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. Nature *517*, 99–103.

Sun, X., Peng, L., Guo, J., Chi, W., Ma, J., Lu, C., and Zhang, L. (2007a). Formation of DEG5 and DEG8 Complexes and Their Involvement in the Degradation of Photodamaged Photosystem II Reaction Center D1 Protein in Arabidopsis. PLANT CELL ONLINE *19*, 1347–1361.

Sun, X., Wang, L., and Zhang, L. (2007b). Involvement of DEG5 and DEG8 proteases in the turnover of the photosystem II reaction center D1 protein under heat stress in Arabidopsis thaliana. Chinese Sci. Bull. *52*, 1742–1745.

Sun, X., Fu, T., Chen, N., Guo, J., Ma, J., Zou, M., Lu, C., and Zhang, L. (2010). The stromal chloroplast Deg7 protease participates in the repair of photosystem II after photoinhibition in Arabidopsis. Plant Physiol. *152*, 1263–1273.

Sundberg, E., Slagter, J.G., Fridborg, I., Cleary, S.P., Robinson, C., and Coupland, G. (1997). ALBINO3, an Arabidopsis nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. Plant Cell *9*, 717–730.

Suzuki, J.Y., and Bauer, C.E. (1995). A prokaryotic origin for light-dependent chlorophyll biosynthesis of plants. Proc. Natl. Acad. Sci. U. S. A. *92*, 3749–3753.

Suzuki, I., Kanesaki, Y., Mikami, K., Kanehisa, M., and Murata, N. (2001). Cold-regulated genes under control of the cold sensor Hik33 in Synechocystis. Mol. Microbiol. *40*, 235–244.

Takahashi, S., and Murata, N. (2008). How do environmental stresses accelerate photoinhibition? Trends Plant Sci. *13*, 178–182.

Takahashi, K., Takabayashi, A., Tanaka, A., and Tanaka, R. (2014). Functional Analysis of Light-harvesting-like Protein 3 (LIL3) and Its Light-harvesting Chlorophyll-binding Motif in *Arabidopsis*. J. Biol. Chem. *289*, 987–999.

Takaichi, S., Maoka, T., and Masamoto, K. (2001). Myxoxanthophyll in Synechocystis sp. PCC 6803 is myxol 2'-dimethyl-fucoside, (3R,2'S)-myxol 2'-(2,4-di-O-methyl-alpha-L-fucoside), not rhamnoside. Plant Cell Physiol. *42*, 756–762.

Takemoto, J.Y., Peters, J., and Drews, G. (1982). Crosslinking of photosynthetic membrane polypeptides of Rhodopseudomonas capsulata. FEBS Lett. *142*, 227–230.

Tanaka, N., and Nakamoto, H. (1999). HtpG is essential for the thermal stress management in cyanobacteria. FEBS Lett. *458*, 117–123.

Tanaka, R., and Tanaka, A. (2007). Tetrapyrrole Biosynthesis in Higher Plants. Annu. Rev. Plant Biol. *58*, 321–346.

Tanaka, R., Oster, U., Kruse, E., Rudiger, W., and Grimm, B. (1999). Reduced activity of geranylgeranyl reductase leads to loss of chlorophyll and tocopherol and to partially geranylgeranylated chlorophyll in transgenic tobacco plants expressing antisense RNA for geranylgeranyl reductase. Plant Physiol. *120*, 695–704.

Tanaka, R., Rothbart, M., Oka, S., Takabayashi, A., Takahashi, K., Shibata, M., Myouga, F., Motohashi, R., Shinozaki, K., Grimm, B., et al. (2010). LIL3, a light-harvesting-like protein, plays an essential role in chlorophyll and tocopherol biosynthesis. Proc. Natl. Acad. Sci. U. S. A. *107*, 16721–16725.

Taylor, D.P., Cohen, S.N., Clark, W.G., and Marrs, B.L. (1983). Alignment of genetic and restriction maps of the photosynthesis region of the Rhodopseudomonas capsulata chromosome by a conjugation-mediated marker rescue technique. J. Bacteriol. *154*, 580–590.

Theis, J., and Schroda, M. (2016). Revisiting the photosystem II repair cycle. Plant Signal. Behav. *11*, e1218587.

Thornton, L.E., Ohkawa, H., Roose, J.L., Kashino, Y., Keren, N., and Pakrasi, H.B. (2004). Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem ii activity in the cyanobacterium Synechocystis 6803. Plant Cell *16*, 2164–2175.

Tichý, M., Bečková, M., Kopečná, J., Noda, J., Sobotka, R., and Komenda, J. (2016). Strain of Synechocystis PCC 6803 with Aberrant Assembly of Photosystem II Contains Tandem Duplication of a Large Chromosomal Region. Front. Plant Sci. *7*, 648.

Torabi, S., Umate, P., Manavski, N., Plöchinger, M., Kleinknecht, L., Bogireddi, H., Herrmann, R.G., Wanner, G., Schröder, W.P., and Meurer, J. (2014). PsbN Is Required for Assembly of the Photosystem

II Reaction Center in *Nicotiana tabacum*. Plant Cell *26*, 1183–1199.

Tóth, T.N., Chukhutsina, V., Domonkos, I., Knoppová, J., Komenda, J., Kis, M., Lénárt, Z., Garab, G., Kovács, L., Gombos, Z., et al. (2015). Carotenoids are essential for the assembly of cyanobacterial photosynthetic complexes. Biochim. Biophys. Acta - Bioenerg. *1847*, 1153–1165.

Trester-Zedlitz, M., Kamada, K., Burley, S.K., Fenyö, D., Chait, B.T., and Muir, T.W. (2003). A Modular Cross-Linking Approach for Exploring Protein Interactions. J. Am. Chem. Soc. *125*, 2416–2425.

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

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

Tsuchiya, T., Ohta, H., Masuda, T., Mikami, B., Kita, N., Shioi, Y., and Takamiya, K. -i. (1997). Purification and Characterization of Two Isozymes of Chlorophyllase from Mature Leaves of Chenopodium album. Plant Cell Physiol. *38*, 1026–1031.

Tsuchiya, T., Ohta, H., Okawa, K., Iwamatsu, A., Shimada, H., Masuda, T., and Takamiya, K. (1999). Cloning of chlorophyllase, the key enzyme in chlorophyll degradation: finding of a lipase motif and the induction by methyl jasmonate. Proc. Natl. Acad. Sci. U. S. A. *96*, 15362–15367.

Tsukatani, Y., Yamamoto, H., Harada, J., Yoshitomi, T., Nomata, J., Kasahara, M., Mizoguchi, T., Fujita, Y., and Tamiaki, H. (2013). An unexpectedly branched biosynthetic pathway for bacteriochlorophyll b capable of absorbing near-infrared light. Sci. Rep. *3*, 1217.

Tucker, J.D., Siebert, C.A., Escalante, M., Adams, P.G., Olsen, J.D., Otto, C., Stokes, D.L., and Hunter, C.N. (2010). Membrane invagination in Rhodobacter sphaeroides is initiated at curved regions of the cytoplasmic membrane, then forms both budded and fully detached spherical vesicles. Mol. Microbiol. *76*, 833–847.

Umena, Y., Kawakami, K., Shen, J.-R., and Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. Nature *473*, 55–60.

Uniacke, J., and Zerges, W. (2007). Photosystem II Assembly and Repair Are Differentially Localized in Chlamydomonas. PLANT CELL ONLINE *19*, 3640–3654.

Vavilin, D., and Vermaas, W. (2007). Continuous chlorophyll degradation accompanied by chlorophyllide and phytol reutilization for chlorophyll synthesis in Synechocystis sp. PCC 6803. Biochim. Biophys. Acta - Bioenerg. *1767*, 920–929.

Vavilin, D., Yao, D., and Vermaas, W. (2007). Small Cab-like proteins retard degradation of photosystem II-associated chlorophyll in Synechocystis sp. PCC 6803: kinetic analysis of pigment labeling with 15N and 13C. J. Biol. Chem. *282*, 37660–37668.

Vershinin, A. (1999). Biological functions of carotenoids--diversity and evolution. Biofactors *10*, 99–104.

Walter, B., Pieta, T., and Schünemann, D. (2015a). Arabidopsis thaliana mutants lacking cpFtsY or cpSRP54 exhibit different defects in photosystem II repair. Front. Plant Sci. *6*, 250.

Walter, B., Hristou, A., Nowaczyk, M.M., and Schünemann, D. (2015b). *In vitro* reconstitution of co-translational D1 insertion reveals a role of the cpSec–Alb3 translocase and Vipp1 in photosystem II biogenesis. Biochem. J. *468*, 315–324.

Wang, P., Gao, J., Wan, C., Zhang, F., Xu, Z., Huang, X., Sun, X., and Deng, X. (2010). Divinyl chlorophyll(ide) a can be converted to monovinyl chlorophyll(ide) a by a divinyl reductase in rice. Plant Physiol. *153*, 994–1003.

Wang, P., Li, C., Wang, Y., Huang, R., Sun, C., Xu, Z., Zhu, J., Gao, X., Deng, X., and Wang, P. (2014). Identification of a Geranylgeranyl reductase gene for chlorophyll synthesis in rice. Springerplus *3*, 201.

Wang, Q., Jantaro, S., Lu, B., Majeed, W., Bailey, M., and He, Q. (2008). The high light-inducible polypeptides stabilize trimeric photosystem I complex under high light conditions in Synechocystis PCC 6803. Plant Physiol. *147*, 1239–1250.

Wang, X., Cimermancic, P., Yu, C., Schweitzer, A., Chopra, N., Engel, J.L., Greenberg, C., Huszagh, A.S., Beck, F., Sakata, E., et al. (2017). Molecular Details Underlying Dynamic Structures and Regulation of the Human 26S Proteasome. Mol. Cell. Proteomics *16*, 840–854.

Wang, Y., Mao, L., and Hu, X. (2004). Insight into the structural role of carotenoids in the photosystem I: a quantum chemical analysis. Biophys. J. *86*, 3097–3111.

Watanabe, M., Iwai, M., Narikawa, R., and Ikeuchi, M. (2009). Is the Photosystem II Complex a Monomer or a Dimer? Plant Cell Physiol. *50*, 1674–1680.

Watanabe, S., Kobayashi, T., Saito, M., Sato, M., Nimura-Matsune, K., Chibazakura, T., Taketani, S., Nakamoto, H., and Yoshikawa, H. (2007). Studies on the role of HtpG in the tetrapyrrole biosynthesis pathway of the cyanobacterium Synechococcus elongatus PCC 7942. Biochem. Biophys. Res. Commun. *352*, 36–41.

Wei, L., Guo, J., Ouyang, M., Sun, X., Ma, J., Chi, W., Lu, C., and Zhang, L. (2010). LPA19, a Psb27 Homolog in *Arabidopsis thaliana* , Facilitates D1 Protein Precursor Processing during PSII Biogenesis. J. Biol. Chem. *285*, 21391–21398.

Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X., and Liu, Z. (2016). Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. Nature *534*, 69–74.

Wessjohann, L., and Sontag, B. (1996). Prenylation of Benzoic Acid Derivatives Catalyzed by a Transferase fromEscherichia coli Overproduction: Method Development and Substrate Specificity. Angew. Chemie Int. Ed. English *35*, 1697–1699.

Westphal, S., Heins, L., Soll, J., and Vothknecht, U.C. (2001). Vipp1 deletion mutant of Synechocystis: A connection between bacterial phage shock and thylakoid biogenesis? Proc. Natl. Acad. Sci. *98*, 4243–4248.

Whitby, F.G., Phillips, J.D., Kushner, J.P., and Hill, C.P. (1998). Crystal structure of human uroporphyrinogen decarboxylase. EMBO J. *17*, 2463–2471.

Williams, J.G.K. (1988). Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in Synechocystis 6803. Methods Enzymol. *167*, 766–778.

Willstatter, Richard ; Stoll, A. (1928). Investigations On Chlorophyll : Methods And Results (The Science Press Printing Company ; Lancaster).

Winterfeld, S., Imhof, N., Roos, T., Bär, G., Kuhn, A., and Gerken, U. (2009). Substrate-Induced Conformational Change of the *Escherichia coli* Membrane Insertase YidC. Biochemistry *48*, 6684–6691.

Woese, C.R., Stackebrandt, E., Weisburg, W.G., Paster, B.J., Madigan, M.T., Fowler, V.J., Hahn, C.M., Blanz, P., Gupta, R., Nealson, K.H., et al. (1984). The phylogeny of purple bacteria: The alpha subdivision. Syst. Appl. Microbiol. *5*, 315–326.

Wong, Y.S., Castelfranco, P.A., Goff, D.A., and Smith, K.M. (1985). Intermediates in the formation of the chlorophyll isocyclic ring. Plant Physiol. *79*, 725–729.

Woolhead, C.A., Thompson, S.J., Moore, M., Tissier, C., Mant, A., Rodger, A., Henry, R., and Robinson, C. (2001). Distinct Albino3-dependent and -independent pathways for thylakoid membrane protein insertion. J. Biol. Chem. *276*, 40841–40846.

Wu, Z., Zhang, X., He, B., Diao, L., Sheng, S., Wang, J., Guo, X., Su, N., Wang, L., Jiang, L., et al. (2007). A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. Plant Physiol. *145*, 29–40.

Xiong, W., Shen, G., and Bryant, D.A. (2017). Synechocystis sp. PCC 6803 CruA (sll0147) encodes lycopene cyclase and requires bound chlorophyll a for activity. Photosynth. Res. *131*, 267–280.

Xu, H., Vavilin, D., Funk, C., and Vermaas, W. (2002). Small Cab-like proteins regulating tetrapyrrole biosynthesis in the cyanobacterium Synechocystis sp. PCC 6803. Plant Mol. Biol. *49*, 149–160.

Xu, H., Vavilin, D., Funk, C., and Vermaas, W. (2004). Multiple deletions of small cab-like proteins in the cyanobacterium Synechocystis sp. PCC 6803. Consequences for pigment biosynthesis and accumulation. J. Biol. Chem. *279*, 27971–27979.

Xu, Q., Guikema, J.A., and Chitnis, P.R. (1994). Identification of surface-exposed domains on the reducing side of photosystem I. Plant Physiol. *106*, 617–624.

Xu, Y., Alvey, R.M., Byrne, P.O., Graham, J.E., Shen, G., and Bryant, D.A. (2011). Expression of Genes in Cyanobacteria: Adaptation of Endogenous Plasmids as Platforms for High-Level Gene Expression in Synechococcus sp. PCC 7002. In Methods in Molecular Biology (Clifton, N.J.), pp. 273–293.

Yamamoto, T., Burke, J., Autz, G., and Jagendorf, A.T. (1981). Bound Ribosomes of Pea Chloroplast Thylakoid Membranes: Location and Release in Vitro by High Salt, Puromycin, and RNase. Plant Physiol. *67*, 940–949.

Yamazaki, S., Nomata, J., and Fujita, Y. (2006). Differential operation of dual protochlorophyllide reductases for chlorophyll biosynthesis in response to environmental oxygen levels in the cyanobacterium Leptolyngbya boryana. Plant Physiol. *142*, 911–922.

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.

Yang, H., Liao, L., Bo, T., Zhao, L., Sun, X., Lu, X., Norling, B., and Huang, F. (2014). Slr0151 in *Synechocystis* sp. PCC 6803 is required for efficient repair of photosystem II under high-light condition. J. Integr. Plant Biol. *56*, 1136–1150.

Yang, H., Liu, J., Wen, X., and Lu, C. (2015). Molecular mechanism of photosystem I assembly in oxygenic organisms. Biochim. Biophys. Acta - Bioenerg. *1847*, 838–848.

Yang, H., Li, P., Zhang, A., Wen, X., Zhang, L., and Lu, C. (2017). Tetratricopeptide repeat protein Pyg7 is essential for photosystem I assembly by interacting with PsaC in Arabidopsis. Plant J. *91*, 950–961.

Yao, D., Kieselbach, T., Komenda, J., Promnares, K., Prieto, M.A.H., Tichy, M., Vermaas, W., and Funk, C. (2007). Localization of the Small CAB-like Proteins in Photosystem II. J. Biol. Chem. *282*, 267–276.

Yu, Y., You, L., Liu, D., Hollinshead, W., Tang, Y., and Zhang, F. (2013). Development of Synechocystis sp. PCC 6803 as a Phototrophic Cell Factory. Mar. Drugs *11*, 2894–2916.

Yurina, N.P., Mokerova, D. V., and Odintsova, M.S. (2013). Light-inducible stress plastid proteins of phototrophs. Russ. J. Plant Physiol. *60*, 577–588.

Zak, E., Norling, B., Maitra, R., Huang, F., Andersson, B., and Pakrasi, H.B. (2001). The initial steps of biogenesis of cyanobacterial photosystems occur in plasma membranes. Proc. Natl. Acad. Sci. U. S. A.

98, 13443–13448.

Zang, X., Liu, B., Liu, S., Arunakumara, K.K.I.U., and Zhang, X. (2007). Optimum conditions for transformation of Synechocystis sp. PCC 6803. J. Microbiol. *45*, 241–245.

Zeilstra-Ryalls, J., Gomelsky, M., Eraso, J.M., Yeliseev, A., O'Gara, J., and Kaplan, S. (1998). Control of photosystem formation in Rhodobacter sphaeroides. J. Bacteriol. *180*, 2801–2809.

Zhang, L., Paakkarinen, V., van Wijk, K.J., and Aro, E.M. (1999). Co-translational assembly of the D1 protein into photosystem II. J. Biol. Chem. *274*, 16062–16067.

Zhang, L., Paakkarinen, V., Suorsa, M., and Aro, E.M. (2001). A SecY homologue is involved in chloroplast-encoded D1 protein biogenesis. J. Biol. Chem. *276*, 37809–37814.

Zhang, S., Frankel, L.K., and Bricker, T.M. (2010). The Sll0606 protein is required for photosystem II assembly/stability in the cyanobacterium Synechocystis sp. PCC 6803. J. Biol. Chem. *285*, 32047–32054.

Zhang, S., Shen, G., Li, Z., Golbeck, J.H., and Bryant, D.A. (2014). Vipp1 Is Essential for the Biogenesis of Photosystem I but Not Thylakoid Membranes in *Synechococcus* sp. PCC 7002. J. Biol. Chem. *289*, 15904–15914.

Zhou, S., Sawicki, A., Willows, R.D., and 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.

Zhu, F., Suzuki, K., Okada, K., Tanaka, K., Nakagawa, T., Kawamukai, M., and Matsuda, H. (1997). Cloning and Functional Expression of a Novel Geranylgeranyl Pyrophosphate Synthase Gene from Arabidopsis thaliana in Escherichia coli. Plant Cell Physiol. *38*, 357–361.

Zhu, L., Wasey, A., White, S.H., and Dalbey, R.E. (2013). Charge composition features of model single-span membrane proteins that determine selection of YidC and SecYEG translocase pathways in Escherichia coli. J. Biol. Chem. *288*, 7704–7716.

Zhu, Y., Graham, J.E., Ludwig, M., Xiong, W., Alvey, R.M., Shen, G., and Bryant, D.A. (2010). Roles of xanthophyll carotenoids in protection against photoinhibition and oxidative stress in the cyanobacterium Synechococcus sp. strain PCC 7002. Arch. Biochem. Biophys. *504*, 86–99.

Zilber, A.L., and Malkin, R. (1988). Ferredoxin Cross-Links to a 22 kD Subunit of Photosystem I. Plant Physiol. *88*, 810–814.

Zsebo, K.M., and Hearst, J.E. (1984). Genetic-physical mapping of a photosynthetic gene cluster from R. capsulata. Cell *37*, 937–947.

(2007). Class 2 · Transferases VI (Berlin, Heidelberg: Springer Berlin Heidelberg).