Structural and functional analysis of GUN4 and ChlH subunits of the magnesium chelatase enzyme

by

Shabnam Tarahi Tabrizi

MACQUARIE University

This thesis is presented for the award of the degree of

Doctor of Philosophy

Department of Chemistry and Biomolecular Sciences

Faculty of Science and Engineering

Macquarie University, Sydney, New South Wales 2109, Australia

February 2016

Table of Contents

Declaration

I certify that this thesis titled '**Structural and functional analysis of GUN4 and ChlH subunits of the magnesium chelatase enzyme**' is my original research. Some parts of this research was achieved in collaboration with other researchers; any form of assistance received from others has been duly acknowledged and their contribution recognized. All sources of information and cited material have been referenced in this thesis. No part of this thesis has been submitted to any other institution as a component of a degree or award. I also certify that this thesis contains no information that has been formerly written or published by anyone else except where due reference is cited in the text.

Shabnam Tarahi Tabrizi

Department of Chemistry and Biomolecular Sciences

Macquarie University, Sydney, NSW, 2109

Australia

November 2015

Acknowledgements

The journey towards this PhD has been a remarkable time of learning and fun. It was a roller coaster ride of ups and downs, mixed emotions, failures and successes. It not only made me a better scientific researcher but also helped carving out my overall personality. The success and final outcome of this thesis required a lot of guidance and support from many people. First and foremost, thanks to my principal supervisor, Professor Robert Willows for trusting in me and selecting me for the PhD position and then guiding and supporting me throughout my PhD. It goes without saying I wouldn't be able to finish this journey successfully without his constant support. I am grateful to my adjunct supervisor, Dr Anthony Duff for his supports and comments on my project whenever I needed it. I would like to sincerely thank Dr David Langley who has greatly helped me on crystallography and inspired me to challenge myself and never give up on research even when all the other experts had already given up. I would also like to thank Dr Stephen Harrop for encouraging me on X-ray crystallography techniques and the art part of crystallography. I gratefully thank my co-supervisor Professor Ian Paulsen for his motivative talk during my PhD. Special thanks to Dr Artur Sawicki for his comments on shaping up my thesis. Thanks to Dr Mehdi Mirzaei and Samantha Emery for their help on proteomics analysis. I also thank all those whom I have co-authored with, for the opportunity to collaborate and also for their intellectual contribution. I thank Ms. Jane Yang for her kind assistance in regard to higher degree research matters. I thank Catherine Wong, Michelle Kang and Jinxia Ke for administrative assistance. Here, I thank Associate Professor Brian Atwell for graciously giving me the opportunity to teach a plant structure and function unit which was a great experience. Additional thanks to Negin Farzadian and Muhammad Masood for their help and friendly assistance during teaching practicals.

I extend my gratitude to all my colleagues and friends in the Department of Chemistry and Biomolecular Sciences (CBMS) and Department of Biological Science. Thanks to my lab mates in Willow's lab: Dr Roger Hiller, Artur, Andrew, Paul, Mehdi, and Irene whom I have learnt many things in experimentation. Also Paulsen's Lab members: Deepa, Ali, Anne, Sophie, Danie, Sheemal, and Hasinika. Special Thanks to Iniga and Sammi for sharing the office and therefore all my ups and downs, their friendship made the PhD journey more pleasant and memorable. Thanks to friends who gave me never ending support and heart full of joy. Thank you Mohamad Karthik, Jashan, Zari, Sahar, Nima, Tony, Francesca, Raha, and Azadeh. Thank you Negin for our friendly coffee club after work, Thank you Tony and Kartika for your wonderful photography and my soul mate Manveen for all her support and gym company.

I also gratefully acknowledge the help and motivation from my friends outside university, Charlotte, Onur, Sina, Milad, Mahboubeh, Misagh, Hamid, Azadeh, Hamed, Nooshin, Samneh, Alireza, and Ehsan.

The biggest gift from God is family and I am blessed to have a loving, endearing and immensely supportive family. Thanks to the love of my life, my husband (Masoud), who is an epitome of love, kindness and patience. Words cannot explain the gratitude I feel towards your love, because your love kept me going during my hard moments. Thanks to my mum (Nahid), for her love, and solidarity, you have sacrificed a lot, always supported me and have been like my greatest friend with whom I can discuss anything, my father (Mohsen) who is my role model and inspiration, my brother (Shahrooz) for all the fun in my life, Parents in-law, Mitra and Heidar for their prayers, concern and encouragement, and my husband's grandmother (dear Effat) who is one of a kind, and a very inspiring lady. I Love you All.

Above all and most importantly, I would like to thank God Almighty who gives me the strength to do things. With God, impossible is impossible.

Abstract

In plants and algae the Genomes Uncoupled 4 (GUN4) is a nuclear-encoded, chloroplast localized, porphyrin binding protein implicated in retrograde signaling between the chloroplast and nucleus. Functionally, it enhances Mg-chelatase activity, the enzyme that inserts magnesium into protoporphyrin IX (PPIX) in the chlorophyll biosynthesis pathway, possibly through the removal of Mg protoporphyrin IX (Mg-PPIX) the product of the reaction. Our *in vitro* results show that purified GUN4 increases the production of PPIX-generated singlet oxygen in the light by a factor of five compared with PPIX alone. Additionally the functional GUN4-PPIX-ChlH complex and GUN4-Mg-PPIX generate singlet oxygen at a reduced rate compared with GUN4-PPIX. Recent evidence points to singlet oxygen being a possible plastid to nucleus signal, possibly through second messengers. The light dependent singlet oxygen generation by GUN4-PPIX may be part of a signal transduction pathway from the chloroplast to the nucleus which senses the availability and flux of PPIX through the chlorophyll biosynthetic pathway in order to control the concentrations of enzymes and chlorophyll binding proteins in the chloroplast. Additionally, unlike the cyanobacterial GUN4, the chloroplastic orthologues have an extra C-terminal domain that is phosphorylated and is required for magnesium chelatase activity. The low resolution solution structure of full length GUN4 at ~20 Å was determined, by using small-angle X-ray scattering (SAXS) and report that the protein has a more elongated structure compared to the cyanobacterial protein. Also, the first crystal structure of truncated eukaryotic GUN4 from *Chlamydomonas reinhardtii* was solved. The structure is in broad agreement with those of previously solved cyanobacterial structures. Most interestingly, conformational divergence is restricted to several loops which cover the porphyrin-binding cleft. The conformational dynamics suggested by this ensemble of structures lend support to our understanding of how GUN4 binds PPIX or Mg-PPIX.

List of publications

Paper I: Müller AH, Sawicki A, Zhou S, **Tarahi Tabrizi S**, Luo M, Hansson M, Willows RD. 2014, Inducing the oxidative stress response in *Escherichia coli* improves the quality of a recombinant protein: Magnesium chelatase ChlH, Protein Expression and Purification, 101:61-67 **(Appendix 2)**

Paper II: **Tarahi Tabrizi S**, Langley DB, Harrop SJ, Duff AP, Willows RD. 2015. Structure of GUN4 from *Chlamydomonas reinhardtii*. Acta Crystallographica Section F Structural Biology communications. 71:1094-1099 **(Appendix 4)**

Paper III: **Tarahi Tabrizi**, S., Sawicki, A., Zhou, S., Luo, M. and Willows R.D. 2015. "GUN4 of *Chlamydomonas reinhardtii* enhances singlet oxygen production in the presence of light and protoporphyrin IX" (Prepared for submission) **(Chapter 5).**

I was first author on two publications (Publication 2 and 3) and co-author on one publication (Publication 1) presented in this thesis.

In publication I, I performed the initial culturing, and all the protein expression and purification of GUN4 and ChlH from *Chlamydomonas reinhardtii*. Also, I completed the immobilized metal ion affinity chromatography for Cr-ChlH.

In publication II, I performed the entire culturing, protein expression, purification, protein crystallization, crystal screening and crystal optimization methods. I also performed in-gel trypsin digestion, and I prepared the manuscript. Dr Anthony Duff provided feedback on protein crystallization methods and manuscript. Dr Stephen Harrop assisted and supervised me on protein crystallization, crystal optimization methods, and X-ray data collection at the Australian Synchrotron. Dr David Langley assisted me on crystal optimization and professionally conducted crystallography (data analysis, solving crystal structure, and structure depositing on

PDB), and was greatly involved in manuscript preparation. Prof Robert Willows provided feedback on protein purification steps and, edited and critically reviewed the manuscript.

In Publication III, I was involved in culturing, expression, and purification of all protein samples. Also, I performed the entire singlet oxygen detection experiments and data analysis and prepared the manuscript. A/Prof Meizhong Lou and Dr Shuaxiang Zhou provided the protein expression clones. Dr Artur Sawicki provided feedback and edited the manuscript. Prof Robert Willows was involved from the initial concept and design of the experiments, analysis of data, supervision and critically reviewing the manuscript.

Conference presentation (*Oral presentations)

1. Shabnam Tarahi Tabrizi, David Langley, Anthony Duff, Stephen J. Harrop, Robert D. Willows: Crystal structure of GUN4 protein and analysis of its role in singlet oxygen production, **Combio 2015 conference, Melbourne, Victoria, Australia.**

2. Shabnam Tarahi Tabrizi, Anthony Duff, Robert D. Willows: Structure characterization of the *Chlamydomonas reinhardtii* magnesium chelatase GUN4 and H subunits by small-angle X-ray scattering, **Australian Synchrotron User meeting, 2014, Melbourne, Victoria, Australia.**

3. Shabnam Tarahi Tabrizi, Anthony Duff, Stephen J. Harrop, Robert D. Willows: GUN4 increases the singlet oxygen production in plant chlorophyll biosynthesis, **MQ Biofocus Research Conference, 2014, Macquarie University, Sydney, Australia.**

4. Shabnam Tarahi Tabrizi, Anthony Duff, Artur Sawicki, Shuaixiang Zhou, Meizhong Luo, Robert D. Willows: Structure characterization of GUN4 and analysis of its role in activation of protoporphyrin IX, **The 6th Asia & Oceania Conference on Photobiology, 2013, Sydney, Australia.**

I also have 5 poster presentations in international and local conferences during my PhD candidature.

Awards

- 2015 Maslen Travel Scholarship, Society of Crystallographers in Australia and New Zealand (AUD 1000)

- 2014 Postgraduate Research Fund, Macquarie University (AUD 5000)

- 2014 Australian Society of Plant Scientists Student Travel Grant (AUD 200)

- 2014 Italian-Australian bilateral SAXS and SANS school student travel fund (AUD 1500)

- 2013 The student excellent oral presentation award in the 6th Asia & Oceania Conference on Photobiology Sydney, Australia (AUD 200)

- 2013 Postgraduate Research Fund, Macquarie University (AUD 2500)

- 2012 International Macquarie University Research Excellence Scholarship (iMQRES) covering a total tuition fee of AUD 117, 9801 (based on 2011 rate) with a living stipend of AUD 25, 392 per annum (tax exempted)

Abbreviations

Å	Angstrom
A 280	absorption at 280 nm
A. thaliana	*Arabidopsis thaliana*
AAA^+	ATPases Associated with diverse cellular Activities
ABA	Abscisic acid
ADP	Adenosine diphosphate
ALA	δ-Aminolevulinic acid
ATP	Adenosine triphosphate
Bchl	Bacteriochlorophyll
B-ME	β-mercaptoethanol
µM	Micromolar
µL	Microlitre
°C	Degrees Celsius
CCP4	Collaborative Computational Project Number 4
C. reinhardtii	*Chlamydomonas reinhardtii*
Chl	Chlorophyll
ChlM	magnesium-protoporphyrin IX methyltransferase
cDNA	Complementary deoxyribonucleic acid
COOT	Crystallographic Object-Oriented Toolkit
Da	Dalton
Dmax	Maximum dimension
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EM	Electron Microscopy
GUN	Genomes uncoupled
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
His	Histidine
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	Kilodalton
K_d	Dissociation constant
K_m	Substrate concentration at half maximal velocity, Michaelis-Menten constant
LB	Luria Bertani
Mg-chelatase	Magnesium-chelatase
Mg-PPIX	Magnesium-Protoporphyrin IX
mL	Millilitre
mM	Millimolar
mm	Millimetre

MR	Molecular Replacement
nm	Nano metre
$O_2(a^1\Delta_g)$	Singlet oxygen
PDB	Protein Data Bank
PEG	Polyethylene glycol
PhANG	photosynthesis associated nuclear-encoded genes
PPIX	Protoporphyrin IX
R. sphaeroides	*Rhodobacter sphaeroides*
R. capsulatus	*Rhodobacter capsulatus*
r.m.s.d	root mean square deviation
SAXS	Small angle x-ray scattering
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
T. elongatus	*Thermosynechococcus elongatus*
Tris	Tris(hydroxylmethyl)aminomethane
UV	Ultraviolet
Vmax	Maximal velocity, Michaelis-Menten constant

Chapter 1

Chapter 1. Introduction

1.1 Photosynthesis

Photosynthesis is directly or indirectly responsible for the survival of nearly all living organisms, with only chemolithotrophic bacteria not requiring this natural process (Nisbet et al., 1995, Raymond et al., 2002, Xiong and Bauer, 2002, Leslie, 2009). Photosynthetic organisms require chlorophyll (Chl) or bacteriochlorophyll (Bchl) for the collection and conversion of solar energy into chemical energy (Chew and Bryant, 2007, Reinbothe et al., 2010). In plants, algae and oxygenic bacteria (cyanobacteria) Chls are involved in aerobic photosynthesis, while for anaerobic photosynthetic bacteria Bchl molecules are responsible for light trapping (Zeilstra-Ryalls and Kaplan, 2004, Kiang et al., 2007).

1.1.1 Tetrapyrroles

The basic structure of a pyrrole is a 5-membered carbon-nitrogen heterocyclic compound with the formula C_4H_4NH. Tetrapyrroles consist of 4 covalently attached pyrroles and may be either linear (e.g. hydroxymethylbilane) or cyclic (e.g. chlorophyll). The most well-known tetrapyrroles, heme and chlorophyll play vital roles in higher plants in photosynthesis and metabolism (Tanaka and Tanaka, 2007). Heme is an iron-centered tetrapyrrole while Bchl/Chls are magnesium-tetrapyrroles, consisting of four pyrrole rings joined to form a tetrapyrrolic macrocycle with a magnesium atom in the center (Figure 1.1), and variations in the side chains and/or reduction states (Wettstein et al., 1995, Chen, 2014). Two other tetrapyrroles, siroheme, and phytochromobilin are vital, the former as a cofactor for sulphite and nitrite reductases and the latter as a photoreceptor (Tanaka and Tanaka, 2007, Tripathy et al., 2010). All four tetrapyrroles are derived from a common biosynthetic pathway (Figure. 1.2). The major site of tetrapyrrole biosynthesis is localized in plastids and only the last few steps of heme biosynthesis occurs in both mitochondria and plastids (Tanaka and Tanaka, 2007).

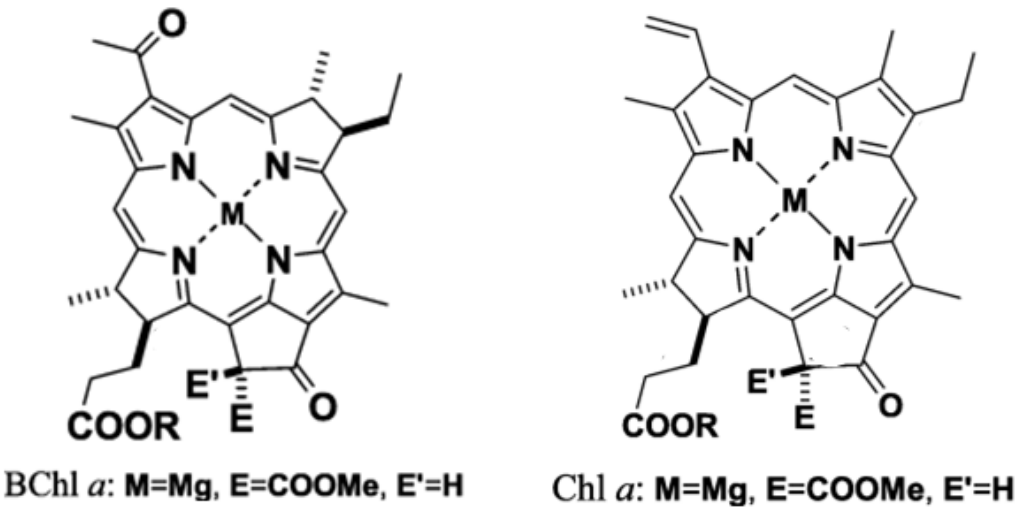


Figure 1. 1 Molecular structure of bacteriochlorophyll and chlorophyll

The two predominant light-capturing pigments found in nature (Tamiaki et al., 2007).

1.1.2 Tetrapyrrole biosynthetic pathway

The first six steps of siroheme, heme, phytochromobilin and Chl biosynthesis have a shared pathway (Figure 1.2). The first step, catalysed by glutamyl tRNA synthetase, activates glutamate by ligating it to its cognate tRNA to form tRNAGlu (Schon et al., 1986). Following this, glutamyl tRNA reductase (GluTR) reduces glutamyl-tRNA, producing glutamate-1-semialdehyde (GSA) with this step coupled to oxidation of NADPH (Wang et al., 1984). GSA is then converted into 5-aminolevulinic acid (ALA) by glutamate 1-semialdehyde aminotransferase (GSA-AT) (Hoober et al., 1988) in the third step of the pathway. The fourth enzyme, ALA dehydratase, converts two molecules of ALA to porphobilinogen, a pyrrole molecule (Jaffe, 1993). Subsequently, four porphobilinogen molecules are joined to form a linear tetrapyrrole, 1-hydroxymethylbilane. The fifth enzyme, uroporphyrinogen III synthase, converts hydroxymethylbilane to uroporphyrinogen III which is the first closed tetrapyrrole in the pathway (Beale, 1999). This is the first branchpoint of the four tetrapyrrolic end products. In the first step of the siroheme branch of the pathway, S-adenosyl-L-methionine-dependent

methyltransferase methylates uroporphyrinogen III to form dihydrosirohydrochlorin. Subsequently, siroheme ferrochelatase chelates an iron atom into dihydrosirohydrochlorin to form siroheme in an ATP-independent manner (Tanaka and Tanaka, 2007). The remaining steps are still shared with uroporphyrinogen III decarboxylase converting uroporphyrinogen III to coproporphyrinogen III, which in turn is decarboxylated to form protoporphyrinogen IX by an eighth enzyme, coproporphyrinogen III oxidase (Beale, 1999). The ninth step in the pathway is accomplished by protoporphyrinogen IX oxidase that generates PPIX, a substrate common to both the heme and Chl branches (Lermontova et al., 1997).

This is the second branchpoint in the tetrapyrrole biosynthetic pathway and is differentiated by chelation of PPIX by two different metals Fe^{2+} or Mg^{2+}. The iron (Fe) branch produces heme, into which Fe is inserted into the center of the PPIX macrocycle by Fe-chelatase (Ferrochelatase) enzyme, while the Mg branch leads to chlorophyll as the end product, which starts with the insertion of Mg^{2+} into PPIX (Xiong and Bauer, 2002, Chen, 2014).

1.1.3 Tetrapyrroles as photosensitizers

From uroporphyrinogen III onwards, pathway intermediates can be oxidized by molecular oxygen to form the corresponding porphyrins, which are capable of absorbing light energy forming highly reactive singlet oxygen ($O_2(a^1\Delta_g)$) (Tanaka and Tanaka, 2007) and damaging the plastid membrane, and up-regulating detoxifying enzymes mRNA in a process called photo-oxidative damage (Mock et al., 1998). Thus, flux through this pathway, specifically at the branch point of the Chl and heme branches, must be carefully controlled. Therefore, the regulation of the flux of the tetrapyrrole pathway is necessary as well as limiting the concentration of free tetrapyrroles *in vivo*. Recent evidence suggests sensor proteins such as GUN4 are responsible for photo-protection of tetrapyrroles (Brzezowski et al., 2014). However,

our finding about $O_2(a^1\Delta_g)$ production role of GUN4 contradicts the current supposition that GUN4 has a photo-protective role (Chapter 5).

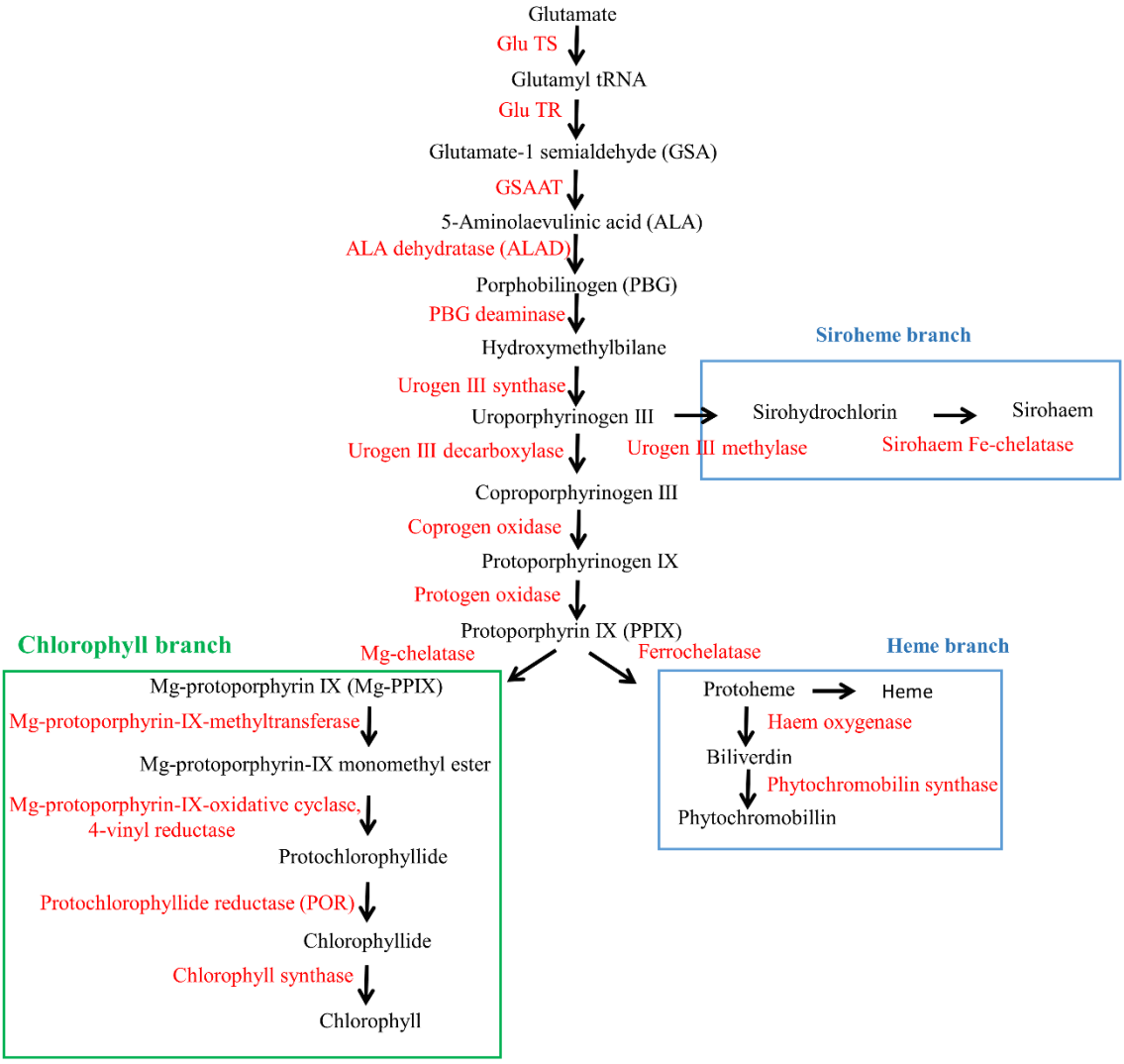


Figure 1.2 The tetrapyrrole biosynthetic pathway in plants

The enzymatic steps of tetrapyrrole production in higher plants. Tetrapyrrole biosynthetic pathway, with outlined heme, siroheme branches presented in blue boxes, and the Chl branch presented in a green box.

1.2 Mg-chelatase enzyme

The Mg-chelatase enzyme is a key regulator of Chl biosynthesis, as it catalyses the adenosine triphosphate (ATP) dependent insertion of magnesium into PPIX to form Mg-PPIX (Figure 1.2) in the branch point between Chl and heme synthesis (Walker and Weinstein, 1994).

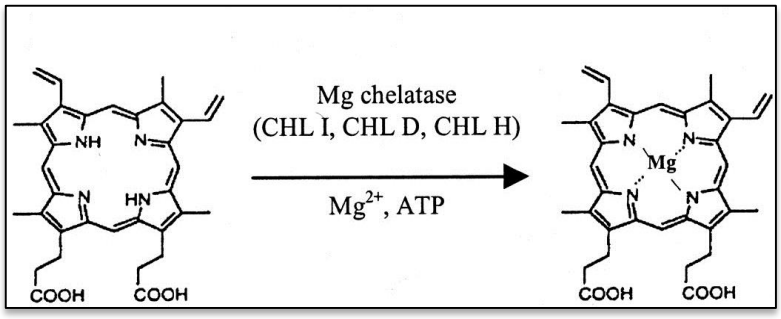


Figure 1.3 Mg-chelatase enzymatic activity

Mg-chelatase inserts a divalent magnesium ion into the substrate PPIX, generating Mg-PPIX (Lake et al., 2004).

This enzyme has three genes called *bchI*/*chlI*, *bchD*/*chlD*, and *bchH*/*chlH*, which encode three protein subunits: ChlI/BchI (~45 kDa), ChlD/BchD (~80 kDa) and ChlH/BchH (~145 kDa). These subunits are conserved from bacteria to higher plants, and are responsible for different functions. ChlI/BchI are members of the AAA^+ superfamily with ATPase activity required for enzyme catalysis. Moreover, they bind Mg^{2+} and ATP, and are molecular chaperones for ChlD/BchD subunits. BchD/ChlD subunits appear to have an inactive AAA^+ ATPase N-terminal domain, which is homologous to BchI/ChlI, and acts as a structural platform for BchI/ChlI (Gibson et al., 1995, Jensen et al., 1996a, Lake et al., 2004, Reid and Hunter, 2004, Stenbaek and Jensen, 2010). Functional BchH/ChlH subunits bind both PPIX and Mg-PPIX. Additionally, a porphyrin carrier protein called GUN4 is involved in PPIX substrate delivery to ChlH in oxygenic photosynthetic organisms, and also binds the product (Mg-PPIX) of Mg-chelatase (Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005). The role of this protein will be discussed in more detail in section 1.3.4.3.

Mg-chelatase enzyme complex has been extensively studied in order to understand the interactions between its three subunits, and their respective roles in Mg-PPIX synthesis (Walker and Willows, 1997). So far, our understanding of the mechanism for Mg-chelatase activity has relied upon: a) biochemical characterization of Mg-chelatase activity in whole cells of photosynthetic bacteria and in isolated chloroplasts from higher plants, b) examination of mutants deficient in Mg-chelatase activity, c) *in vitro* reconstitution of Mg-chelatase activity using purified recombinant expressed subunits, and structural studies of individual subunits and protein complexes (Fodje et al., 2001).

The more recent information on the function of Mg-chelatase subunits and interaction of subunits with each other in the complex will be described in section 1.3. Also, a comprehensive overview of Mg-chelatase with emphasis on the structure and possible mechanism for insertion will be discussed later in sections 1.4 and 1.5.

1.3 Mg-chelatase is a multifunctional enzyme

1.3.1 *In vivo* assays

Mutation analysis of the photosynthetic gene cluster from *Rhodobacter sphaeroides (R. sphaeroides) and Rhodobacter capsulatus (R. capsulatus)* indicated that all three Mg-chelatase subunits are required for Mg-PPIX biosynthesis (Gorchein et al., 1993, Bollivar et al., 1994). *In vivo* evidence in higher plants also supported this idea, as it has been previously described that *xantha-f, xantha-g* and *xantha-h* mutants of barley lacking ChllH, ChlD, and ChlI, respectively, are unable to synthesize Chl (Jensen et al., 1996a, Kannangara et al., 1997). The first Mg-chelatase activity assay was undertaken in intact cells of *R. sphaeroides* that demonstrated the biological insertion of magnesium into PPIX with the production of Mg-PPIX monomethyl ester, not Mg-PPIX (Gorchein, 1972). However, Mg-PPIX monomethyl ester treatment of the whole cells with various electron transport inhibitors or uncouplers decreased

the amount of Mg-PPIX monomethyl ester formation (Gorchein, 1973) indicating an ATP requirement for Mg-chelatase activity (Gorchein, 1972, Gorchein, 1973). The first *in vitro* Mg-chelatase activity measured from higher plants was with unpurified homogenates of etiolated wheat seedlings (Ellsworth and Lawrence, 1973, Willows et al., 1996), while the first stimulation of Mg-chelatase activity was carried out by incubating etioplasts or developing chloroplasts from cucumber cotyledons with PPIX and a mixture of longer wavelength metalloporphyrins (Zn-PPIX, Mg-PPIX and Mg-PPIX monoester) (Smith and Rebeiz, 1977). Further characterisation of Mg-chelatase was achieved with intact cucumber chloroplasts (Pardo et al., 1980, Richter and Rienits, 1980) and isolated chloroplasts from tobacco (Papenbrock et al., 1999, Papenbrock et al., 2000), barley (Jensen et al., 1996b) and pea (Walker and Weinstein, 1991a). Another study illustrated that developing chloroplasts require glutamate for Mg-chelatase activity (Castelfranco et al., 1979), which could be substituted with 10 mM ATP (Pardo et al., 1980). In addition, it has been shown that ATP hydrolysis is required for Mg-chelatase activity since non-hydrolysable analogues of ATP adenosine 5'-[β,γ-methylene] triphosphate and adenosine 5'-[β,γ-imido] triphosphate cannot support Mg-chelatase activity (Walker and Weinstein, 1991a).

1.3.2 *In vitro* assays

The first plant *in vitro* Mg-chelatase activity assay in cucumber chloroplast demonstrated preliminary kinetic analysis and generated hyperbolic curves for magnesium, ATP, and PPIX substrates which is representative of a non-regulatory effect (Richter and Rienits, 1982, Walker and Weinstein, 1991a). However, a previous study shows a sigmoidal relationship of PPIX and ATP with cucumber chloroplasts (Fuesler et al., 1981). A later study on pea chloroplast then showed a sigmoidal relationship of Mg-chelatase activity with magnesium and ATP (Guo et al., 1998). Furthermore, other studies on *Synechocystis* and *Chlorobium tepidum* (*C. tepidum*)

showed sigmoidal effects of Mg-chelatase with magnesium (Jensen et al., 1998, Reid and Hunter, 2004, Johnson and Schmidt-Dannert, 2008). These data suggested a regulatory role of the magnesium substrate.

The discovery of a two-step nature of Mg-chelatase reaction was demonstrated by biochemical characterization of Mg-chelatase activity fractions from lysed pea chloroplasts (Walker et al., 1992). The activity of Mg-chelatase from lysed pea chloroplasts and thylakoid required both a soluble (~80 %) and membranous component (Walker and Weinstein, 1991b), with a synergistic effect when mixed together, demonstrating that at least two protein components are involved (Walker and Weinstein, 1991b, Walker et al., 1992). However, in contrast to pea, only the membrane fraction was required for Mg-chelatase activity of cucumber, and the loss of activity after disruption of cucumber plastids was recovered by the addition of PPIX and ATP (Lee et al., 1992). If the pea lysate is not pre-incubated with ATP, a "lag-phase" in Mg-PPIX formation is observed over time indicative of a then unknown enzyme complex assembly (Walker et al., 1992). Preparation with ATP causes a destruction of the "lag-phase" (Walker et al., 1992). During the first stage of the Mg- chelatase reaction which is involved in complex assembly, ATP can be replaced by a slowly hydrolysable ATP analogue (adenosine 5'-[γ-thio]triphosphate), but there is an obligatory ATP requirement for the second stage of the insertion of magnesium into PPIX and formation of Mg-PPIX product (Walker and Weinstein, 1994).

The first *in vitro* assay of a bacterial Mg-chelatase was with crude BchI, BchD, and BchH proteins from *R. sphaeroides* recombinantly expressed in *E. coli* (Gibson et al., 1995), these proteins were then purified by a combination of anion exchange and gel filtration and activity demonstrated by reconstitution of all three subunits (Willows et al., 1996). The BchH protein cell lysates were red and this was attributed to PPIX bound to this overexpressed protein (Willows et al., 1996). The Mg-chelatase activity of the *R. sphaeroides* and *R. capsulatus*

requires all three proteins, magnesium and ATP, with exogenous PPIX not needed as it was bound to BchH from *Chlorobium vibrioforme* (*C. vibrioforme*) (Petersen et al., 1998), and *R. capsulatus* (Willows and Beale, 1998). The ChlH subunit from *Synechocystis* PCC6803 lose their red colour during anion exhange purification indicating a weaker association with PPIX (Jensen et al., 1996b). In all cases reconstitution of activity requires all three subunits, ATP or an ATP regenerating system along with PPIX or BchH-PPIX. Further studies on both Mg-chelatase characterization and purification from bacteria used a 6xHis-Tag on the N-terminal for protein expression due to its rapid purification by metal ion affinity chromatography (Jensen et al., 1998, Petersen et al., 1998, Gibson et al., 1999, Willows et al., 1999).

In the recent *in vitro* study hybrid Mg-chelatases containing mixtures of subunits from *Synechocystis* and *T. elongatus* enzymes were generated. Both *Synechocystis* and *T. elongatus* chelatase are closely related, but *T. elongatus* is more thermostable and showed simpler steady state kinetics compared with the *Synechocystis* enzyme. Also they react differently in response to free Mg^{2+}, where the hybrid chelatase of *T. elongatus* ChlD and *Synechocystis* ChlI and ChlH is not co-operative with Mg^{2+}, but *Synechocystis* enzyme is highly co-operative (Reid and Hunter, 2004, Davison et al., 2005). The loss of cooperativity after substituting ChlD indicates that ChlD subunit of *Synechocystis* plays an important role in the Mg^{2+} regulation of the Mg-chelatase. Thus, there might be a particular interaction that is responsible for the change in co-operativity in the *Synechocystis* subunit which is missing in *T. elongatus* subunit (Adams et al., 2014a).

1.3.3 ATPase activity site (BchI/ChlI- BchD/ChlD complex)

1.3.3.1 BchI/ChlI subunit

As a member of AAA^+ protein family, the BchI/ChlI subunit (Neuwald et al., 1999, Vale, 2000, Ogura Teru and J., 2001, Iyer et al., 2004) exhibits properties common to this family including

chaperoning (Neuwald et al., 1999) and reassembling different target molecules in an ATP-dependent manner (Iyer et al., 2004). In the case of BchI/ChlI these two functions are observed since the BchI functions as a chaperone for BchD, preventing its degradation and the ChlI/ChlD double hexamer functions as the motor complex during hydrolysis of ATP (Lundqvist et al., 2013). The addition of ATP to the AAA^+ protein family causes the formation of oligomers with large conformational changes (Gibson et al., 1999, Hansson et al., 2002). It has been shown that the ATPase activity of BchI depends on the concentration of free magnesium in *Synechocystis*, indicating that the magnesium atom could be surrounded by the I subunit (Reid et al., 2003). It was demonstrated that ATPase-deficient BchI mutants from *R. capsulatus* are still able to form oligomers if bound to ATP but ATP hydrolysis of each subunit within the BchI complex is obligatory for activity (Hansson et al., 2002). Most photosynthetic organisms only have one BchI/ChlI subunit. However, the BchI from *C. vibrioforme* has two isoforms (38 and 42 kDa) with each demonstrating a similar final Mg-chelatase activity (Peterson et al., 1998). Additionally, there are two ChlI subunits (ChlI1 and ChlI2) of *Arabidopsis thaliana (A. thaliana*) that both play a role in chelation (Rissler et al., 2002, Kobayashi et al., 2008).

The data shows that, ChlI2 has the ability to form a functional Mg-chelatase complex and participate in catalysis therefore it is able to form an active Mg-chelatase complex (Kobayashi et al., 2008, Huang and Li, 2009). However the natural mRNA levels of *chlI2* is ~4-6 fold lower than *chlI1* (Huang and Li, 2009). Since the *chlI1/chlI2* double mutant was recovered with overexpression of ChlI2 driven by the ChlI1 promoter, it is likely each isoform may take part in the Mg-chelatase complex. The differences between ChlI1 and ChlI2 include: The C-terminal of ChlI2 is larger than ChlI1 (Apchelimov et al., 2007), also ChlI2 has lower basal expression *in vivo* (Rissler et al., 2002) and has lower V_{max} and a higher K_{mATP} than ChlI1 (Kobayashi et al., 2008).

The BchI/ChlI contains sensor I/II motifs and an arginine sensor (Arg289), which is present only in AAA^+ proteins (Neuwald et al., 1999, Vale, 2000, Fodje et al., 2001). These sensor I and sensor II motifs are essential for ATP hydrolysis, whereas the arginine sensor is involved in ATP-hydrolytic conformational changes (Iyer et al., 2004). It has been demonstrated that in BchI mutants of *R. capsulatus* the substitution of lysine for arginine may lead to a complete loss of ATPase activity, which indicates the importance of this sensor (Arg289) (Hansson et al., 2002). However, there are some structural features of BchI which do not belong to the common structure of an AAA^+ protein including; a β-hairpin protruding from helix3, along with 2 other hairpins from the core of the BchI structure (Fodje et al., 2001).

Investigating seven mutants in barley *xantha-h* locus encoding ChlI showed that four are recessive and three are semidominant. Recessiveness was explained by a lack of mutant *chlI* expression (Hansson et al., 2002). Thus, in heterozygous plants there is only wild type complexes. Semidominant mutation in *bchI* led to amino acid exchange in the binding site of ATP at the interface of two BchI monomers and abolished ATPase activity, yet the mutant subunits formed hexamers in the presence of ATP. Mixed hexamers were also formed with wild type BchI. These mixtures of mutants and wild type BchI showed an intermediate ATPase activity and reduced Mg-chelatase activity. It was concluded the ATPase sites act autonomously of each other and Mg-chelatase activity requires all sites to be active (Hansson et al., 2002).

1.3.3.2 BchD/ChlD subunit

The BchD/ChlD subunit also belongs to the AAA^+ protein family but has no ATPase activity (Fodje et al., 2001). BchD forms an oligomeric structure in the absence of ATP (Hansson and Kannangara, 1997, Jensen et al., 1999, Petersen et al., 1999), and this oligomer plays an important structural role as a platform for the assembly of BchI/ChlI and BchH/ChlH subunits (Grafe et al., 1999, Axelsson et al., 2006). Mutational studies demonstrated cooperativity

between BchD subunits in the oligomeric state (Axelsson et al., 2006). The BchD and BchI subunits directly interact (Papenbrock et al., 1997, Gibson et al., 1999, Grafe et al., 1999, Jensen et al., 1999) and *in vitro* assays of Mg-chelatase activity requires excess BchI over BchD for the stability of the BchI-BchD complex. BchD aggregates in the absence of BchI *in vitro* in *R. capsulatus*, so an excess of BchI is required to maintain its solubility during refolding of BchD and formation of a functional BchI-BchD complex (Jensen et al., 1998, Willows and Beale, 1998, Gibson et al., 1999, Sawicki and Willows, 2008).The chaperone-type activity of the BchI subunit for the BchD oligomer is important for the stability of ChlD *in vivo* as the cellular concentration of ChlD is reduced in ATPase-defective *chlI* mutants from barley as well as *chlI* knockouts in *A. thaliana* (Hansson et al., 1999, Petersen et al., 1999, Lake et al., 2004, Huang and Li, 2009). In contrast, cellular levels of ChlD remained unaffected in *chlH* mutants from barley (Hansson et al., 1999, Olsson et al., 2004), *A. thaliana* (Huang and Li, 2009) and *C. reinhardtii* (Chekounova et al., 2001, Lake et al., 2004). Taken together these results indicate that the ChlH subunit does not directly affect the quantity of ChlD, but the intracellular concentration of ChlD, and hence Mg-chelatase activity, is dependent upon ChlI with a functional ATPase activity (Lake et al., 2004, Sawicki and Willows, 2008). This chaperoning role of BchI/ChlI appears to be ATP-dependent manner, as the formation of a BchI/ChlI and BchI/ChlD complex is ATP dependent *in vitro* (Sawicki and Willows, 2008).

Barley *(Hordeum vulgare*) mutants with single and multiple point mutations in Mg-chelatase genes have been characterized and analyzed (Axelsson et al., 2006). Some of the corresponding mutations have been generated in the *R. capsulatus* Mg-chelatase genes that have been expressed in *Escherichia coli*. Enzyme activity analysis of these mutant BchD proteins compared to wild type as well as mixtures of wild type and mutant subunits showed that BchD oligomers with mixed mutant and wild type subunits could form and were less active than wild type oligomers alone. Hence, cooperativity is required between BchD subunits for maximal

activity that not all subunits in a BchI-BchD complex were needed for one round of enzymatic activity (Axelsson et al., 2006). Kinetic experiments using *R. capsulatus* subunits revealed BchD behaves as the enzyme since enzymatic activity increases linearly with increasing BchD concentration and excess amounts of the other two substrates; BchI and BchH-PPIX (Sawicki and Willows, 2008).

1.3.3.3 BchI-BchD/ChlI-ChlD complex

The BchI▪BchD/ChlI▪ChlD complex is the catalytic center of Mg-chelatase and formation of the BchI6D6 complex requires preincubation of BchI/ChlI and BchD/ChlD subunits with ATP and magnesium (Walker et al., 1992, Willows et al., 1996, Jensen et al., 1998, Willows and Beale, 1998). This was initially reported in *R. sphaeroides* and *Synechocystis* with gel filtration data of complex formation showing a molecular mass of 200 kDa (Gibson et al., 1999, Jensen et al., 1999). The optimal stoichiometry of this complex also differ between the organisms, but it is always concentration dependent. ChlI▪ChlD complex formation in *Synechocystis* requires magnesium, ATP or ADP, and it is energy dependent, as non-hydrolysable ATP-analogues (adenosine 5'-[γ-thio]triphosphate) are required for preservation of a 1:1 complex (Jensen et al., 1999). In *R. sphaeroides* Mg- chelatase, a 5:1 BchI:BchD ratio is optimal at 8.5 nM BchD, though 1:1 BchI:BchD ratio is optimal at 27 nM BchD (Gibson et al., 1999, Willows, 2003). Similarly, optimal Mg-chelatase activity in *R. capsulatus* requires a 4:1 BchI:BchD ratio at 700 nM BchD, though 2.5:1 is sufficient at 1840 nM BchD (Gibson et al., 1999, Willows, 2003). This problem with ratios was solved when BchI was treated as the substrate in the reaction and BchD was treated as the enzyme (Reid and Hunter, 2004, Sawicki and Willows, 2008). Under these conditions saturation kinetics was observed forming dynamic and unstable BchI-BchD complex which was observed to be dependent on the concentration of the BchI subunit.

Consequently, *in vitro* assays according to the stoichiometry of these subunits do not represent the functional ratio of each subunit in the active complex.

Moreover, previous *in vitro* studies of *T. elangatus* have demonstrated Mg-chelatase enzyme is non-co-operative in response to free Mg^{2+}, and the ChlD subunit of *Synechocystis* plays an important role in the Mg^{2+} regulation (Adams et al., 2014a). Based on kinetic experiments in *Synechocystis*, at limiting concentrations ChlD behaves as an enzyme, while ChlI and ChlH act as saturable substrates (Jensen et al., 1998), as K_m of ChlI is 2-3 times lower than that of ChlH (Jensen et al., 1998). Thus, the interaction between ChlD and ChlI is stronger than the interaction between ChlD and ChlH (Grafe et al., 1999).

If any of BchI/ChlI, magnesium and ATP are absent or present below optimal concentrations, BchD/ChlD is degraded *in vivo* (Lake et al., 2004). A decrease in each of these components occurs during prolonged darkness in plants (Usuda, 1988, Papenbrock et al., 1999, Ishijima et al., 2003). Also, it has been shown in an *in vivo* study of tobacco plants that Mg-chelatase functionality decreases in both anti-sense and sense *chlI* transgenic tobacco plants, as formation of a ChlI-ChlD complex is essential for the activity of Mg-chelatase (Papenbrock et al., 2000).

1.3.4 Porphyrin binding complexes

1.3.4.1 BchH/ChlH (GUN5) subunit

BchH/ChlH is at the core of the Mg-chelatase complex due to binding both the substrate (PPIX) and the product (Mg-PPIX) of the enzymatic reaction (Gibson et al., 1995, Willows et al., 1996, Willows and Beale, 1998, Karger et al., 2001, Sirijovski et al., 2008). The retention of the PPIX binding to BchH or ChlH varies during purification of the protein. When BchH from *R.capsulatus* is heterologously expressed in *E. coli*, the protein naturally sequesters PPIX from the cell, and the purification of BchH can be followed by collecting this red/brown pigment which is retained throughout this process of gel filtration and ion exchange chromatography

(Willows et al., 1996, Willows and Beale, 1998). In *R. sphaeroides*, PPIX removal from BchH requires detergent and RP-HPLC separation (Hansson and Kannangara, 1997), while, incubation of BchH-PPIX from *R. capsulatus* with detergent (Tween80) and purification with nickel affinity chromatography was found to be ineffective in producing undenatured apo-BchH (BchH without PPIX bound) (Sawicki and Willows, 2010). In *Synechocystis*, PPIX is depleted from ChlH after purification by ion exchange chromatography (Jensen et al., 1998). The strength of the PPIX or deuteroporphyrin binding to BchH/ChlH is measured by titration of porphyrin and monitoring tryptophan quenching. *R. sphaeroides* BchH has a K_m of 0.15 or 0.36 μM for PPIX (Willows and Beale, 1998, Gibson et al., 1999), whereas ChlH from *Synechocystis* has a K_m^{proto} of 1.25 μM (Jensen et al., 1998). Surprisingly, *R. capsulatus* BchH has a similar K_m^{proto} to *Synechocystis* ChlH (1.23 μM) (Willows and Beale, 1998, Sawicki and Willows, 2008). Direct comparison of *R. sphaeroides* BchH and *Synechocystis* ChlH demonstrated the stronger binding of deuteroporphyrin to BchH with a dissociation constant (K_d) of 0.53 μM for BchH against 1.22 μM for ChlH (Karger et al., 2001).

The requirement of pre-activation of BchH/ChlH on Mg-chelatase activity has been studied in a number of ways. It was found that the "lag-phase" in the Mg-chelatase assay could be reduced if *Synechocystis* ChlH was pre-incubated with deuteroporphyrin in the presence of magnesium, while the ATP was not required for this pre-activation (Jensen et al., 1998). The binding of deuteroporphyrin to the *Synechocystis* ChlH appeared to be enhanced in the presence of magnesium chloride during gel filtration (Karger et al., 2001), but the K_d of ChlH binding to deuteroporphyrin was found to be 2-3 fold higher when measured by fluorescence quenching (Karger et al., 2001) which is the opposite of what should occur. Moreover, ChlH was preserved from inhibition by the cysteine-modifying molecule, NEM (N-ethyl maleimide) via preincubation of ChlH with PPIX and Mg-ATP (Jensen et al., 2000), but preincubation of ChlH with PPIX alone did not protect against NEM inactivation (Jensen et al., 2000). To account for

the discrepancy between these results, it was suggested that different conformers of ChlH may exist in solution with a slow equilibrium between the conformers depending on the porphyrin bound status and the magnesium bound status of ChlH.

1.3.4.2 ChlH (GUN5) is a multi-functional protein

Mutations in the *gun5/cch* gene of *A. thaliana* corresponding to commonly known ChlH results in increased transcription of nuclear-encoded *lhcb* (Mochizuki et al., 2001). This phenotype is a genome uncoupled phenotype and indicates a defect in plastid to nucleus signaling occurs with mutation in ChlH. (Mochizuki et al., 2001). ChlH from *A. thaliana* binds the phytohormone abscisic acid (ABA) and was reported as an ABA receptor with further work showing that mutations in ChlH affect ABA signalling and stomatal closure (Shen et al., 2006, Wu et al., 2009). However, the role of ChlH as an ABA receptor has been questioned as there is no evidence of ABA binding to the barley ChlH and barley *chlh* mutants do not appear to be defective in ABA perception or signaling (Muller and Hansson, 2009).

ChlH from *Synechocystis* has an additional role in regulating gene expression. It binds to the sigma factor, SigE, and represses gene expression and this anti-sigma factor (anti-SigE) function is both light dependent and enhanced by Mg^{2+} (Osanai et al., 2009). Recently a study of three ChlH's, from *Oryza sativa, Hordeum vulgare* and *C. reinhardtii,* expressed in *E. coli* found monomeric and multimeric forms of ChlH separable by size exclusion chromatography (Müller et al., 2014). The monomers were consistent from batch to batch and at least four times more active than multimers, whereas the multimeric form showed varied batch to batch activity, suggesting that the ChlH's were being oxidatively damaged during purification and expression (Müller et al., 2014). Once purified the monomeric form was stable and lacked PPIX suggesting that the lack of bound PPIX could be the reason for light stability (Müller et al., 2014), in

contrast the *R. capsulatus* BchH with bound PPIX aggregated and was inactive upon light exposure (Sirijovski et al., 2006).

1.3.4.3 GUN4 subunit

Genomes Uncoupled 4 (GUN4) is an additional Mg-chelatase enhancer protein which plays a major role in the magnesium chelation branch of Chl biosynthesis by binding porphyrin intermediates (PPIX and Mg-PPIX) and stimulating the activity of Mg-chelatase (Karger et al., 2001, Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005). Effective function of Mg-chelatase requires interaction of magnesium at physiological concentrations (~2 mM magnesium) as well as GUN4 with ChlH (Davison et al., 2005). It has been shown that in *A. thaliana* and *Synechocystis* this interaction increases the Mg-chelatase activity *in vitro* (Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005, Sobotka et al., 2008). A significant proportion of GUN4-PPIX complex *in vivo* is expected to be in complex with ChlH which is required for optimal Mg-chelatase activity and this GUN4-PPIX complex exists in the membrane (Adhikari et al., 2009). GUN4 has a higher affinity for both PPIX and Mg-PPIX binding compared with ChlH (Karger et al., 2001, Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005), and a 1:1 molar ratio of GUN4:ChlH was suggested sufficient for near-optimal Mg-chelatase activity during *in vitro* assays (Davison et al., 2005). GUN4 is also involved in binding and trafficking of both PPIX and deuteroporphyrin substrate to ChlH in oxygenic photosynthetic organisms (Verdecia et al., 2005). The K_m for deuteroporphyrin using GUN4 and Mg-chelatase is approximately 5-fold lower than with Mg-chelatase alone (Verdecia et al., 2005). This suggests that GUN4 binds to deuteroporphyrin and delivers it to ChlH (Verdecia et al., 2005). It is possible that GUN4 plays a role in PPIX delivery, since *Synechocystis gun4* mutant cells showed 20-30 % activity of Mg-chelatase and ferrochelatase compared with wild type (Wilde et al., 2004). GUN4 also binds Mg-PPIX monomethyl ester,

uroporphyrin III, and coproporphyrin III (Adhikari et al., 2009). Hence, GUN4 plays a critical role in porphyrin channelling to other enzymes in the Chl and heme biosynthetic pathways (Adhikari et al., 2009). It has been suggested that GUN4 plays a major role in Mg-PPIX trafficking or shielding PPIX and Mg-PPIX from collisions with molecular oxygen which might yield ROS (Larkin et al., 2003). Structural studies of GUN4 from *Synechocystis* indicates a hydrophobic surface or so called 'greasy palm" which binds the hydrophobic porphyrin substrate. It was suggested that this binding site protects the porphyrin from collisions with molecular oxygen after excitation with bright light (Verdecia et al., 2005).

In a recent study of *C. reinhardtii*, the expression of tetrapyrrole biosynthesis genes and photosynthesis-associated nuclear genes positively correlated with $O_2(a^1\Delta_g)$ levels, and GUN4 was involved in both sensing these changes and signal transmission. Therefore, two roles were proposed for GUN4 firstly as a photo protector of both PPIX and Mg-PPIX preventing $O_2(a^1\Delta_g)$ production, and secondly a sensor of excess PPIX resulting from perturbations in the tetrapyrrole biosynthesis pathway. The free excess PPIX contributing to $O_2(a^1\Delta_g)$-mediated retrograde signalling pathway (Brzezowski et al., 2014).

Additionally, association of ChlH and GUN4 *in vivo* was observed using white, red, far-red, and blue light up-regulation of *gun4* transcripts during greening of *A. thaliana* seedlings along with *chlH* (Stephenson and Terry, 2008). GUN4 is one of several *gun* mutants originally identified in *A. thaliana*, which all displayed a phenotype indicating they were impaired in plastid to nuclear signalling (Susek et al., 1993). The *gun4* gene was the first loci to be identified from this mutant screen (Susek et al., 1993, Willows and Hansson, 2003), and it has since been shown to be involved in both retrograde signalling and post-translational regulation of tetrapyrrole biosynthesis (Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005, Peter and Grimm, 2009). Both GUN4 and GUN5/CCH/ChlH are key participants in the generation of a plastid signal, although GUN4 binds to Mg-PPIX more tightly than GUN5 in *Synechocystis*

(Larkin et al., 2003). It is most likely that GUN4 acts as a regulator of Mg-PPIX synthesis, metabolic trafficking, or in porphyrin-induced changes of nuclear gene expression (Larkin et al., 2003). It was suggested that accumulation of Mg-PPIX is the direct signal to quench expression of nuclear encoded plastid-related genes (Strand et al., 2003) and this idea was used to describe the involvement of GUN4 in signalling. However, this theory has been shown to be inviable as both PPIX and Mg-PPIX concentrations never reach the levels required for this to occur (Mochizuki et al., 2008, Moulin et al., 2008, Woodson et al., 2011).

The GUN4 protein is only found in oxygenic photosynthetic organisms, with no known homolog of GUN4 in Bchl biosynthetic organisms such as *R. capsulatus* (Larkin et al., 2003, Davison et al., 2005). *gun4* from *A. thaliana* and *Synechocystis* both have a photosensitive phenotype (Larkin et al., 2003, Davison et al., 2005) and as the tetrapyrrole biosynthetic pathway involves several phototoxic intermediates, the enzymatic reactions of Chl biosynthesis must be regulated in a way to prevent the photo oxidation of the cell (Walker and Willows, 1997, Papenbrock et al., 2000). Thus GUN4 has a putative photo protective role within the chloroplast which may protect plants from reactive oxygen species (ROS) that are produced by collisions between oxygen and porphyrins under bright light (Larkin et al., 2003, Sobotka et al., 2008).

Recent determination of the K_d of the more water soluble porphyrin analogues deuteroporphryin (DPIX) and Mg-deuteroporphyrin (MgDPIX), with GUN4 from *Synechocystis* by isothermal titration caliometry (ITC) shows a stronger affinity for DPIX (0.8 μM) over MgDPIX (1.3 μM) (Chen et al., 2015a). This is the opposite of previous studies which determined K_d using tryptophan fluorescence quenching with MgDPIX (K_d of 0.3 μM/0.449 μM) and DPIX (K_d of 2.29 μM/0.865 μM) (Davison et al., 2005, Verdecia et al., 2005). ChlH has a slightly higher K_d of DPIX (1.22 μM) and MgDPIX (2.43 μM) over GUN4 using the fluorescence quenching method (Karger et al., 2001) while K_d of DPIX for BSA is 1.7 μM

(Rotenberg and Margalit, 1985). The ITC method is a direct measurement of molecular interaction through changes in enthalpy (ΔH) while fluorescence quenching relies on tryptophan residues in close proximity to the binding site therefore the tryptophan-quenching derived K_d data should be taken with caution. Discrepancies between these two methods for determining binding constants have been previously observed with K_d of ferulic acid binding to BSA where ITC showed two binding sites whilst tryptophan fluorescence showed one (Ojha et al., 2012). Weaker binding of GUN4 or ChlH to the product of the Mg-chelatase reaction (Mg-PPIX) compared to substrate PPIX seems more plausible as it would favour release of product to the next enzyme in the chlorophyll biosynthetic pathway (magnesium-protoporphyrin IX methyltransferase (ChlM)) which has a calculated K_d of 2.37 μM (Shepherd et al., 2003).

1.3.5 Biochemical interaction of Mg-chelatase subunits

The enzyme kinetic data clearly indicates that ChlH interacts with GUN4 and ChlI interacts with ChlD subsequently ChlH-GUN4-PPIX interacts with ChlI-ChlD (Sawicki and Willows, 2008). Direct interaction between Mg-chelatase subunits has been seen in tobacco, using the yeast two-hybrid system (Papenbrock et al., 1997, Jensen et al., 1998, Petersen et al., 1998). The interaction of ChlI and ChlD subunits has already been discussed in section 1.3.3.3, but in conjunction with kinetic properties it is more likely that the ChlI and ChlD subunits together as an ChlI-ChlD complex is required for the interaction with the ChlH subunit, (Papenbrock et al., 1997). The addition of BchH-PPIX/ChlH-PPIX to the BchI-BchD/ChlI-ChlD complex have been shown to increase the ATPase activity in *Synechocystis* and *R. capsulatus* (Jensen et al., 1998, Sawicki and Willows, 2008).

1.3.6 Reactive Oxygen Species and its role in retrograde signalling

Generally ROS can inactivate Mg-chelatase *in vitro* (Willows 2003), inhibit Chl biosynthesis *in vivo* (Peter and Grimm, 2009) and promote both photo bleaching and cell death (Willows, 2003, Aarti et al., 2007). More importantly, low levels of ROS alters the intracellular redox state, activation of redox-sensitive proteins, redox-sensitive parts of proteins and potentially inhibit or increase enzymatic activity. All of this indicates that ROS has a major role in cellular physiology and thus plays an important role in the life of the plants (Reinbothe et al., 2010, Bhattacharjee, 2012). Plants have shielding mechanisms for ROS and they also generate ROS for regulation of every aspect of biology from perception of environmental cues to gene expression (op den Camp, 2003, Reinbothe et al., 2010). In particular the highly reactive $O_2(a^1\Delta_g)$ is an important ROS which can lead to impaired physiological function and even cause cell death depending on the amount produced (Foyer and Noctor, 2000, Bhattacharjee, 2012).

1.4 Structural studies

1.4.1 BchI/ChlI subunit

The high resolution crystal structure of I subunit from *R. capsulatus* has been solved to 2.1 Å resolution (Figure 1.3). ATPase activity is only reported in the BchI/ChlI subunit where it has a characteristic ATP binding site (Walker A motif GX4GKSX6A and Walker B motif hhhhD(D/E)) which are responsible for ATP hydrolysis (Walker and Weinstein, 1994, Hansson and Kannangara, 1997, Peterson et al., 1998, Jensen et al., 1999, Hansson et al., 2002, Lake et al., 2004). The BchI/ChlI subunit has two domains, the N terminal domain includes both the Walker A(P-loop)(G[DH]RGTGKS) and the B nucleotide binding motif characteristic of ATP binding (Fodje et al., 2001, Sirijovski et al., 2006) and the C-terminal domain has four up-down bundled α-helices and two domains linked with a long alpha helical segment (Willows et al., 1999, Fodje et al., 2001, Willows and Hansson, 2003).

Additionally, the overall sequence based analysis of BchI/ChlI subunit, indicate it belongs to the AAA^+ protein family with the closest structural homologues to HslU, NSF-D2 and Pol-III AAA^+ proteins. The AAA^+ proteins form hexameric ring structures and use ATP hydrolysis to move parts of their characteristic domain (Neuwald et al., 1999). The formation of a hexameric structure of BchI from *R. capsulatus*, or heptameric structure of ChlI from *Synechocystis* in the presence of ATP has been revealed by EM single particle analysis (Reid et al., 2003, Willows et al., 2004). However, oligomerisation of the ChlI from *Synechocystis* is concentration-dependent, indicating that without ATP in high protein concentration (8 $mg.mL^{-1}$) oligomers form (208 kDa ChlI complex), but without ATP at lower protein concentration (0.8 and 2 $mg.mL^{-1}$) only partial oligomerisation occurs (~81 and 108 kDa sized structures) (Jensen et al., 1998).

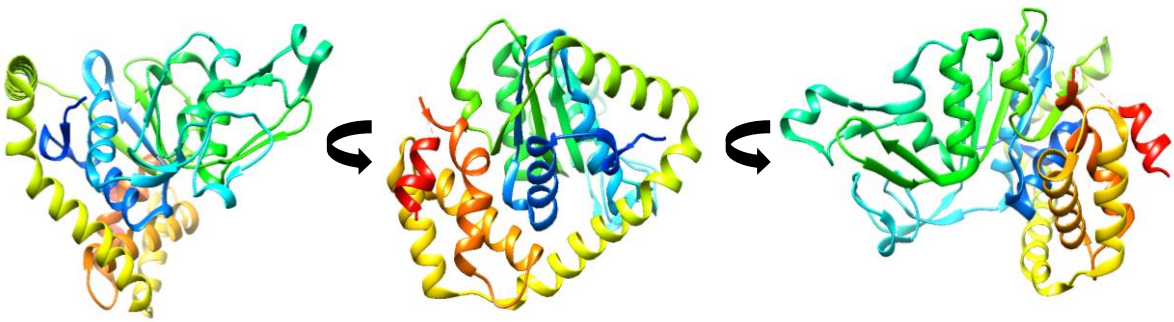


Figure 1.3 Crystal structure of BchI subunit

Different orientation of BchI crystal structure of *R. capsulatus* at 2.1 Å shown as a ribbon diagram with 90^0 rotation relatively to each other. The C-terminal four-helix bundle is connected to the N-terminal domain by a long helical region (Fodje et al., 2001).

1.4.2 BchD/ChlD subunit

There is no high resolution structure of BchD/ChlD. All information about the structure of the BchD/ChlD subunit comes from the sequence analysis of different organism, and substitution of the integrin domain with ChlI-BchD appears to form a hexameric ring in the absence of ATP (Axelsson et al., 2006). ChlD from tobacco was deconstructed to the N-terminus, C-terminus, glutamine/asparagine/proline-rich central region, and flanking sections, which are the important parts involved in protein–protein interaction and enzyme activity (Grafe et al., 1999). The glutamine/asparagine/proline-rich center along with the two flanking regions show some Mg-chelatase activity (30 %), while removal of the N-terminus from wild type ChlD retains wild-type activity and removal of the C-terminus reduces Mg-chelatase activity by approximately 50-60 %. This indicates that the C-terminus is the most essential part for the activity. However, when the N-terminus of ChlD is replaced with the N-terminus of ChlI, only ~20 % of enzyme activity remains. Therefore, the ChlD N-terminus is functionally distinct from the ChlI N-terminus despite their sequence homology (Grafe et al., 1999), and is classified as an AAA^+ protein (Fodje et al., 2001). The walker A motif for ATP/GTP binding and hydrolysis is not conserved in BchD/ChlD and maybe only partially intact depending on the species (Sirijovski et al., 2006). The C-terminal integrin-I and AAA^+ domain of the BchD subunit plays an important role in ATP binding to BchI and conformational changes of BchI to BchD (Lundqvist et al., 2010). Integrins are known to be involved in protein-protein interaction (Takagi, 2007) and potential interaction sites for the integrin I domain were found in BchI/ChlI and BchH/ChlH, with recognition sequences of RGE/D in BchI/ChlI, and LDV in BchH/ /ChlI (Fodje et al., 2001). The integrin I domain includes a metal ion coordination motif, which supports the binding of BchD/ChlD to a magnesium ion (Fodje et al., 2001). The interaction

between the BchI and BchD proteins are suggested to occur via three β-hairpin elements of BchI structure (Fodje et al., 2001). Direct interactions are reported using affinity assays between tobacco ChlD and ChlI or ChlH (Grafe et al., 1999).

1.4.3 BchI-BchD/ChlI-ChlD complex

Both BchI/ChlI and BchD/ChlD are classified as AAA^+ protein family and are able to form ring shaped hexamers (Neuwald et al., 1999, Fodje et al., 2001, Willows et al., 2004). Moreover, they associate in a 1:1 ratio complex (Elmlund et al., 2008). These two reasons aid in the hypothesis that the BchI-BchD/ChlI-ChlD complex could take the form of a double-hexameric ring (Fodje et al., 2001, Elmlund et al., 2008, Sawicki and Willows, 2008) which is a variation of type-2 AAA^+ proteins. In these proteins each monomer has two sequential AAA^+ types; one with and one without ATPase activity. The one without ATPase activity (BchD) may play a structural role in such oligomer (Vale, 2000).

The EM structure of BchI6D6 double hexameric ring of *R. capsulatus* at 7.5 Å, 14 Å and 13 Å resolution have been solved in the presence of ADP, the nonhydrolyzable ATP analogue AMPPNP, respectively (Elmlund et al., 2008, Lundqvist et al., 2010). The validation of the cryo EM structure was confirmed by the analysis of the chemical cross-linked peptides (Lundqvist et al., 2010). Incubation of BchI and BchD with a non-hydrolyzable ATP analogue AMPPNP leads to double hexameric structure which is compressed in the direction of symmetry and contracted in diameter. However, incubation with ATP results in the same contraction at the diameter, but the structure is elongated at the symmetry axis. This could be the BchH binding conformation (Lundqvist et al., 2010).

Various mechanistic and schematic models have been proposed for double hexameric ring BchI6D6 formation. Assembly of the BchD hexameric ring is the first step which occurs spontaneously in ATP-independent manner (Axelsson et al., 2006). The second step is ATP and

Mg^{2+} dependent assembly of BchI subunits in which six subunits are stepwise assembled into a hexameric ring using the BchD hexameric structure as a platform (Walker and Weinstein, 1994, Fodje et al., 2001, Reid et al., 2003, Reid and Hunter, 2004, Willows et al., 2004, Axelsson et al., 2006). Finally, the complex forms a 3-fold symmetry generated from a trimer of homodimers in each ring (Elmlund et al., 2008, Lundqvist et al., 2010). Figure 1.4 shows BchI-BchD complex structure in the presence of ADP.

In the next step of Mg-chelatase structure conformational changes, the BchI-BchD complex binds to the BchH subunit, which will be explained in detail in the functional model of Mg-chelatase activity (Figure 1.9). After the formation of Mg-PPIX, the whole enzyme complex disassembles except for the BchD hexameric ring, which then moves to the new catalytic cycle to function as a platform (Axelsson et al., 2006).

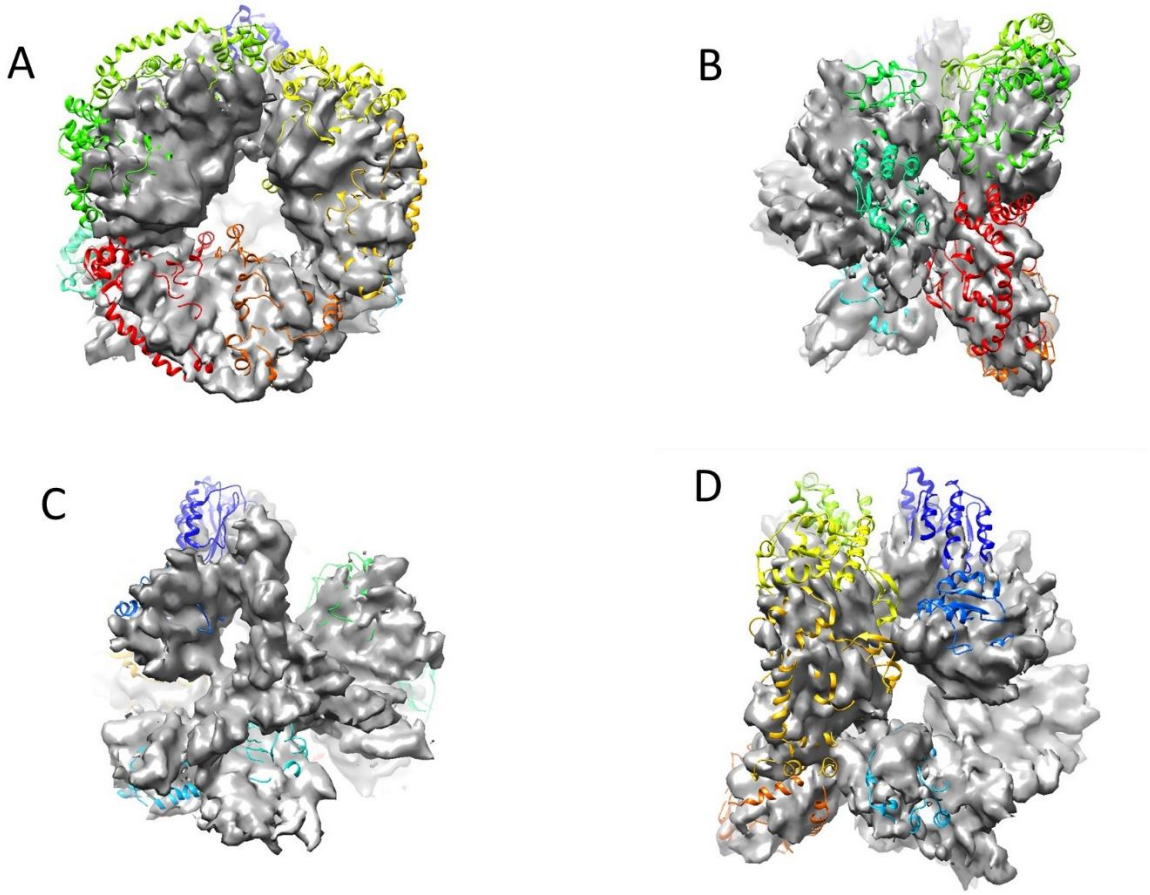


Figure 1.4 Structure of $BchI_6D_6$ complex in the presence of ADP

Reconstructed model of BchI6D6 complex in the presence of ADP at 7.5 Å. The structure was rotated 90^0 relative to each other. A)The BchI-BchD complex showing the BchI ring, B) and D) side view of the BchI-BchD complex showing the C-terminal integrin I domain and the proline-rich region of the BchD dimer, C) The BchI-BchD complex showing the BchD ring (Elmlund et al., 2008).

1.4.4 BchH/ChlH subunit

An EM structural representation of the native and PPIX-depleted *R. capsulatus* BchH revealed a structure at 25 A˚ resolution consisting of three lobe shaped domains connected at the central point. Two of the lobes have protrusions termed, thumbs and fingers, which are joined upon substrate presentation. Structural differences have been reported for apo and PPIX bound forms

of BchH, and a large conformational change is indicated upon binding to PPIX (Sirijovski et al., 2008). The structural change of BchH could be due to the interaction of BchH with the BchI and BchD subunits (Sawicki and Willows, 2008). However, the distinct residue for PPIX binding in BchH/ChlH has remained elusive. The three histidine residues of BchH/ChlH are conserved amongst plants (*A. thaliana*, barley, *Antirrhinum majus*) and bacterial species (*R. capsulatus*, *Synechocystis*). Mutant studies may assist in identifying the residues involved (Walker and Willows, 1997). Furthermore, 35–40 % of sequence identity between BchH from *R.capsulatus* and the ChlH subunits from *Synechoscystis, T. elangatus* and *A. thaliana*, and ~ 60-80 % sequence identity between ChlH from oxygenic organisms have been reported. (Qian et al., 2012).

Recently an open symmetrical cage shape structure of ChlH form *T. elangatus* has been reported by using EM data at 27 Å resolution. The apo-ChlH structure contains a small globular head joined to the whole protein by a narrow neck, and an internal cavity of ~100 nm^3 (Qian et al., 2012). The N-terminal domain of ChlH has been identified as a head domain based on labelling with a nitrilotriacetic acid-nanogold particle that attaches to the N-terminal His-tag. In addition, it was suggested that although both C-terminal and N-terminal domain are involved in PPIX binding, but the N-terminal domain contains most of the porphyrin binding residues (Sirijovski et al., 2008). Deletion of the N-terminus (566 residues) resulted in complete abolition of Mg-chelatase activity, while the PPIX binding part was unaffected (Qian et al., 2012). Additionally, it is possible that Gly-127-Phe-156 region in ChlH sequences controls the movement of head domain and also closing or opening the enclosed cavity within ChlH. Thus, mobility of head domain might be responsible for catalytic activity of Mg-chelatase enzyme, because neighbouring cavity regions with a mobile head domain provides a large enough region for PPIX insertion and Mg-PPIX product releasing (Qian et al., 2012). Small Angle X-ray Scattering (SAXS) data of apo-ChlH from *Synechocystis* revealed a structure with two domains

(small and large) at a Dmax value of 170 Å (Figure 1.5), and similarities were noticed between two apo-ChlH structures by superimposing the EM structure of the *T. elangatus* and the SAXS structure of the *Synechocystis* (Qian et al., 2012). The structural information of ChlH SAXS data appears to be related directly to previous kinetic studies of Mg-chelatase from *Synechocystis* (Jensen et al., 1998, Gibson et al., 1999, Joanne Viney et al., 2007).

It was suggested that ChlH evolved the cage shape during the evolution of oxygenic photosynthesis from anoxygenic organisms. BchH with a flexible three lobed shape structure improved to the ChlH with cage shape structure, in order to enclose the product of Mg-chelatase (Mg-PPIX) and protect it from photo oxidation (Qian et al., 2012). Both SAXS and EM data were collected and analyzed for the ChlH-deuteroporphyrin-PPIX complex from the *Synechocystis* and *T. elangatus*, but no structural differences were found between the apo-ChlH and the ChlH-D-PPIX complex (Qian et al., 2012). The crystal structure of full length ChlH protein from *Synechocystis* has been determined recently (Chen et al., 2015b), and this is shown in ribbon diagram in Figure 1.6. The overall architecture with dimensions of 133 Å × 80 Å × 75 Å is a cage shape assembly which is consistent with previous ChlH SAXS analysis (Qian et al., 2012, Adams et al., 2014b). The crystal structure shows that ChlH is composed of six domains (I–VI). Domains III–VI constituting the cage-like assembly, and domains I and II form the N-terminal ‘head’ and ‘neck’ regions (Chen et al., 2015b). The dimerization interface is located between I and V domain and the two ChlH subunits within the dimer are nearly structurally identical (Figure 1.6) (Chen et al., 2015b). This structure advances our understanding of molecular basis for substrate channelling during the magnesium-chelating process. The putative active site of ChlH protein is a hydrophobic pocket buried deeply inside the protein. This pocket is located at the interface between domains III and V. Residues surrounding the pocket are Asn467, Val468, Leu624, Asn657, Glu660 and Ile663 from domain III, and Trp985, Val1041, Phe1042, Asn1097, Tyr1102, Asn1107, Tyr1125, Lys1129,

Asp1163, Ile1167, Val1172, His1174, Tyr1175 and Asp1177 from domain V (Chen et al., 2015b These residues are conserved throughout evolution, suggesting a conserved PPIX binding mechanism for ChlH subunits of different species. The ChlH crystal structure from Synechocystis suggested that both cch and gun5 mutants of ChlH protein introduce changes in the interior of the protein and thus prevent the chelation reaction from happening in an interior pocket (Chen et al., 2015b).). Although the unliganded full length ChlH structure provides the first hint of a potential PPIX binding pocket, the confirmation of this pocket as the PPIX binding site needs to be confirmed.

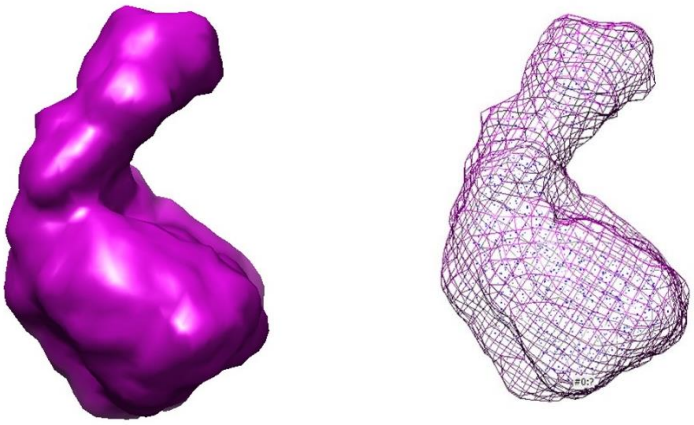


Figure 1.5 : SAXS structure of BchH

BchH solid Surface model on the left and mesh surface model with colorful balls representing excluded dummy atoms on the right. The thumb and finger protrusions fuse while binding to substrate (PPIX) (Qian et al., 2012).

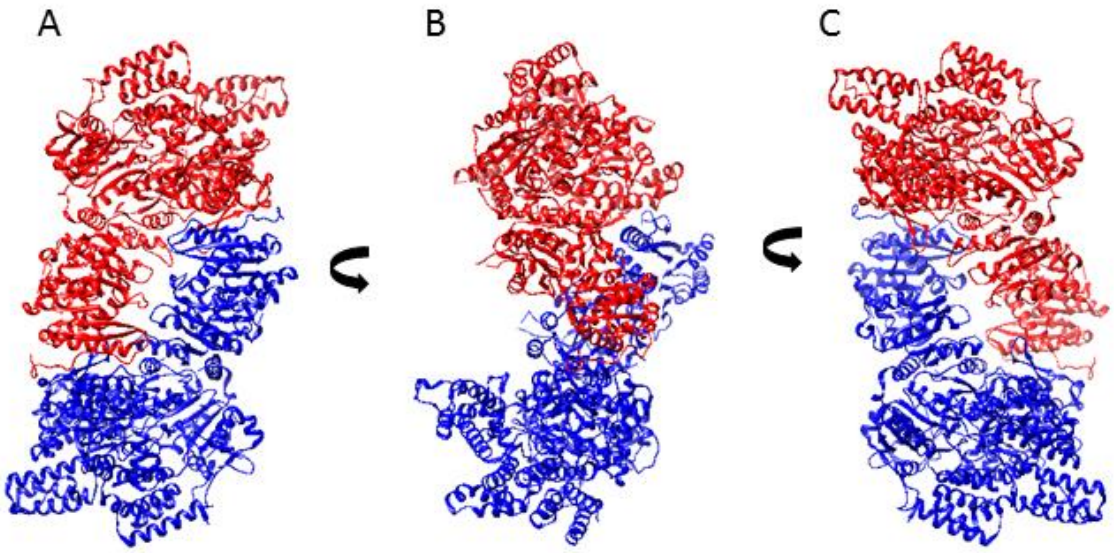


Figure 1.6 : Crystal structure of ChlH

A-C) represent different orientation of the ChlH crystal structure of *Synechocystis* at 2.5 Å (PDB entry: 4ZHJ) shown as a ribbon diagram with 90^0 rotation relatively to each other. The active site is buried deeply inside the protein interior, and the surrounding residues are conserved throughout evolution (Chen et al., 2015b). Different colors (red and blue) are representative of individual chains indicatin g the dimeric crystal structure.

1.4.5 GUN4 subunit

Analysis of the known GUN4 sequence shows high sequence similarity of the C-terminal domain among all GUN4 family members (Figure 1.8). In contrast, the N-terminal is the most highly variable sequence and is only unique to some of the prokaryotic family members (Davison et al., 2005, Verdecia et al., 2005). The role of the N-terminal helical bundle within prokaryotes is still unknown, while the C-terminal domain region is suggested to be involved in PPIX binding based on site-directed mutagenesis and kinetic analysis (Davison et al., 2005). The crystal structures of the GUN4 protein from *Synechocystis* GUN4 and *T. elongatus* GUN4 have already been solved. The C-terminal domain consists of eight alpha helices in both crystal structures, indicating functional and structural conservation (Davison et al., 2005, Verdecia et

al., 2005). Figure 1.7 shows the structure of *Synechocystis* GUN4 presented in panels A–C and the structure of *T. elongatus* GUN4 presented in panels D–F. Both crystal structures are mostly similar, but the position of extended poly-peptides are differently aligned, with the position 219 and 214 Arg the most notable (Davison et al., 2005, Verdecia et al., 2005, Zhou et al., 2012). Additionally an organized, long poly-peptide region found between helix 6 and 7 on *Synechocystis* and helix 7 and 8 on *T. elongatus* is not classified as part of the helices (Davison et al., 2005, Verdecia et al., 2005). The mechanism of GUN4 activation of Mg-chelatase is different in eukaryotes compared to cyanobacteria, as the orthologous C-terminal peptide VFKTNYPSF has been identified as a phospho-peptide in *A. thaliana* (Reiland et al., 2009). All plant and algal GUN4 have YSF conserved, however, orthologous cyanobacterial GUN4 proteins are missing this C-terminal extension. Without this C-terminal peptide plant and algal GUN4 are unable to activate ChlH. It has been suggested that phosphorylation of GUN4 is required for formation of a stable chloroplastic GUN4-ChlH complex (Zhou et al., 2012). The solution structure of GUN4 from *C. reinhardtii* was determined using SAXS at 20 Å, and compared our SAXS GUN4 model with the two proposed cyanobacterial crystal structures (Davison et al., 2005, Verdecia et al., 2005). This finding will be explained in more details in chapter 3 of this thesis. Taken together it is clear that chloroplastic GUN4 has a more elongated structure, and it has an additional 9-10 amino acid C-terminal domain at the end of main molecular axis that is phosphorylated and required for Mg-chelatase activity. In addition, the N-terminal 4 helix bundle found in cyanobacterial GUN4 proteins is absent from chloroplastic GUN4 (Zhou et al., 2012). Recently, crystal structures of GUN4 with porphyrin bound (PDB codes 4xkb and 4xkc) from *Synechocystis* have been elucidated. Both are predominantly α-helical proteins comprised of a smaller N-terminal domain (five α-helices) connected to a larger C-terminal domain (eight α-helices) (Chen et al., 2015c). Finally, the crystal structure of GUN4 from *C. reinhardtii* at 3.5 Å resolution was solved, and hence the first structural details of an

eukaryotic GUN4 protein have been revealed (Tarahi Tabrizi et al., 2015). When our GUN4 crystal structure superposed with cyanobacterial GUN4 structures, the α-helical scaffolds are conserved. Whilst, the α2/α3 and α6/α7 loops previously predicted to undergo movement to accommodate porphyrin binding display markedly different conformations (Chen et al., 2015c). This diversity of fold captured by multiple crystal structures supports the notion that these loops undergo conformational rearrangement to accommodate the insertion of PPIX or Mg-PPIX into hydrophobic cleft, as observed in the recent porphyrin-bound *Synechocystis* structures (Chen et al., 2015c).

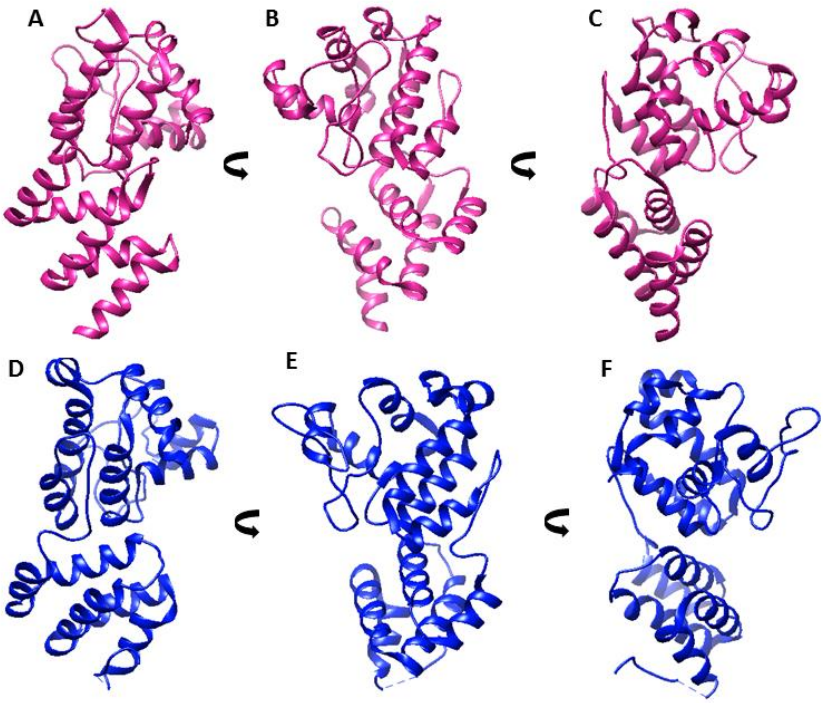


Figure 1.7: Crystal structure of GUN4 from *Synechocystis* and *T. elongatus*.
A–C) represent different orientation of the *Synechocystis* GUN4 structure coloured in pink (Davison et al., 2005). D–F) are the same orientation as A–C for the *T. elongatus* GUN4 structure coloured in blue (Verdecia et al., 2005).

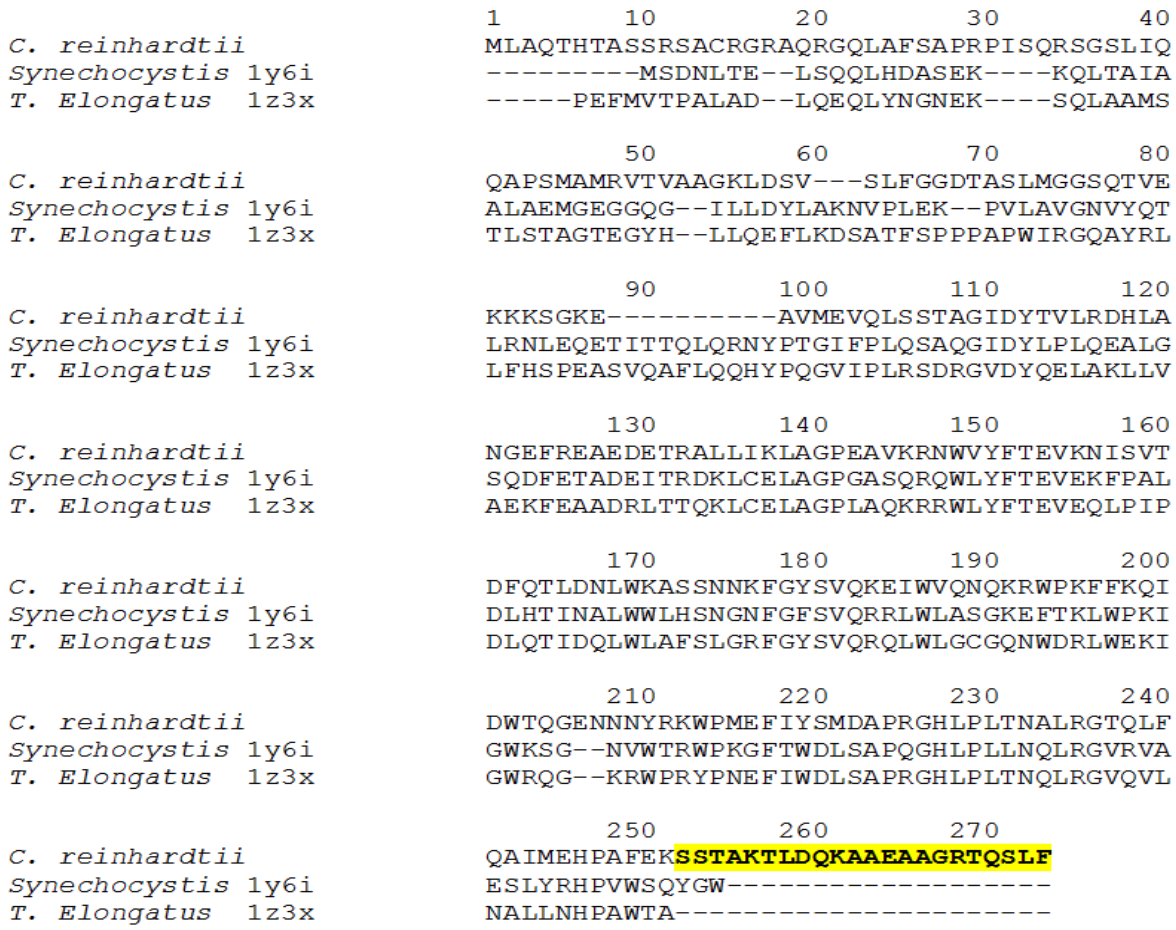
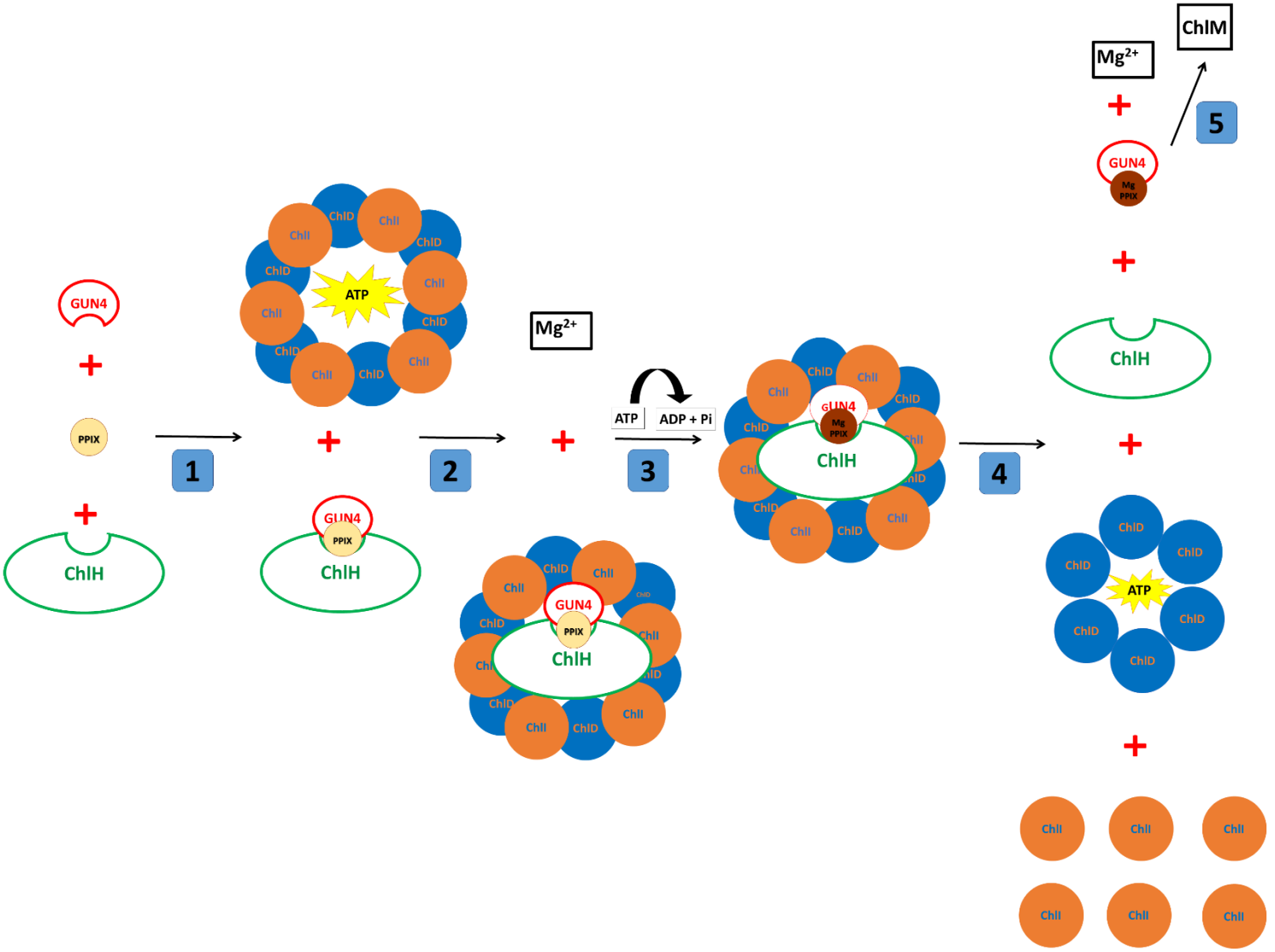


Figure 1.8 Structure-based sequence alignment of GUN4

Sequence alignment of GUN4 from *C. reinhardtii, Synechocystis,* and *T. elongatus* using PROMALS3D. The yellow labelled sequence is represented the C-terminal elongated part at end of GUN4 from *C. reinhardtii* which is absent in the other prokaryotic organisms.

1.5 Functional model of Mg-chelatase:

This model explains Mg-chelatase subunits binding, complex formation, and the Mg-chelatase enzyme assembly and disassembly.


GUN4
PPIX
ChlH
1
ChlD
ChlI
ATP
2
Mg2+
3
ATP
ADP + Pi
Mg
4
5
ChlM

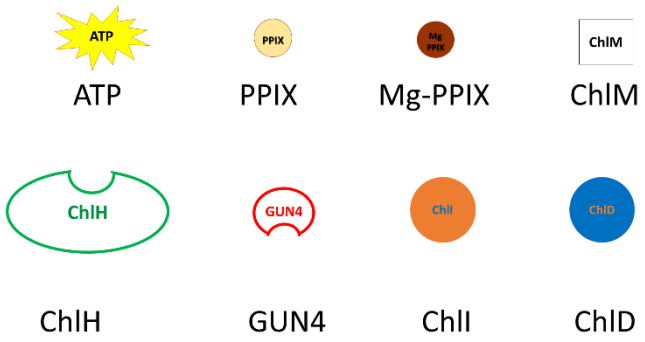


Figure 1.9 Functional model of the Mg-chelatase enzyme

1-Binding of GUN4, PPIX and ChlH, and generation of the GUN4-ChlH-PPIX complex. 2-Interaction of ChlI (ATP-dependent subunit) with ChlD (ATP-independent subunit) and formation of a ChlI-ChlD complex, then binding of this complex to the GUN4-ChlH-PPIX complex. This binding leads to the insertion of Mg^{2+} into PPIX and the assembly of the Mg-chelatase enzyme. GUN4-ChlH-PPIX complex acts as a substrate for ChlI-ChlD complex. 3-Insertion of Mg^{2+} into PPIX, and formation of Mg-PPIX. 4-Diassembly of Mg-Chelatase enzyme. The hexameric ChlD ring will remain assembled and ready to function as a platform in new catalytic cycle. 5- ChlM binding to Mg-PPIX. The enzyme directly downstream of Mg-chelatase, ChlM removes Mg-PPIX from GUN4 and/or ChlH in order to allow the chlorophyll biosynthetic pathway to proceed.

1.6 Project aims

Mg-chelatase is one of the most challenging enzyme of the chlorophyll biosynthetic pathway. Despite several decades of research many aspects of this important enzyme have yet to be elucidated.

In order to better understand the Mg-chelatase enzyme the work presented in this thesis was dedicated to both functional and structural aspects of Mg-chelatase enzyme subunits. Special attention was given to structural studies to further investigate the functions of specific subunits of Mg-chelatase enzyme. The overall aim of this work was, as such, to understand the structure and function of the GUN4 and ChlH subunit of Mg-chelatase. In order to achieve this aim, two

separate structural analysis (Small angle X-ray scattering and X-ray crystallography) were utilized to investigate the structural details. Additionally, the functional understanding of both GUN4 and ChlH subunits were carried out using two different biochemical/ biomolecular approaches. The three main aims of this project are listed below:

AIM I- Determine the structure of GUN4, and ChlH proteins using SAXS.

Structural information of GUN4 and ChlH proteins and GUN4-ChlH-PPIX complex of Mg-chelatase assists in visualizing the intricate nature of this reaction mechanism. The structural information of all mentioned proteins is discussed in chapter 3.

AIM II- Determine the crystal structure of GUN4.

The three dimensional X-ray crystal structure of GUN4 from *C. reinhardtii* is the first eukaryotic GUN4 crystal structure which gives us more structural details about hydrophobic part/PPIX binding site of GUN4 protein.

AIM III - Elucidating the mechanisms by which GUN4 is involved in plastid to nuclear signaling.

GUN4 is a chloroplast located protein and binds both PPIX and Mg-PPIX, therefore it appears to be involved in PPIX delivery or stabilization to the ChlH subunit of the Mg-chelatase (Larkin et al., 2003). GUN4 regulates the flux of intermediates into chlorophyll biosynthesis and also mediates the chloroplast to nuclear signaling mechanism (Larkin et al., 2003). The role of GUN4 and ChlH proteins in $O_2(a^1\Delta_g)$ production and chloroplast to nucleus signaling is extensively explained Chapter 5 and a functional model for this role is also provided

Chapter 2

Chapter 2. Materials and methods

2.1 Materials

Unless stated elsewhere all materials were from Sigma-Aldrich (NSW, Australia), Astra Scientific (NSW, Australia), Chem-Supply (NSW, Australia), Amresco (NSW, Australia), Gensearch (Queensland, Australia), Bio-Rad (NSW, Australia), Thermo Fisher Scientific (VIC, Australia), GE-Healthcare life science (NSW, Australia), Merck Millipore and Life Technologies (VIC, Australia).

Growth medium was from Difco Laboratories. PPIX and deuteroporphyrin were from Porphyrin Products (Logan, UT, U.S.A). Milli-Q H_2O (Millipore) was used for preparation of all media and buffers.

2.2 Cloning of *chlI1*, *chlI2 chlD*, *chlH*, and *gun4*

RNA was extracted from *C. reinhardtii* cells using a Masterpure yeast RNA purification kit (Epicenter) according to the manufacturer's instructions. mRNA was converted to cDNA using a superscript VILO cDNA synthesis kit using random primers. Details of *chlh* cloning into pET-28a was previously described in (Publication I in Appendix 2). The *chlD*, and *gun4* genes from *C. reinhardtii* were cloned from cDNA into expression vector pET28a (Merck-Novagen), while *chlI1* and *chlI2* were cloned into pGEX-6P-1 using primers and restriction sites listed in (Table 2.1). Transformation of each construct into *E. coli* strain BL21 (DE3) from Life Technologies was performed according to the manufacturer's instructions.

Table 2.1: Primers used for construction of expression plasmids.

CrchlI1-Fwd	5'-CGCGGATCCTGCAATGTGGCGACTGGAC-3'
CrchlI1-Rvs	5'-GGCCTCGAGTTACTCCATGCCGAACACCTG-3'
CrchlI2-Fwd	5'-CGCGGATCCGCCGCGAAGAAGCCGAAC-3'
CrchlI2-Rvs	5'-GGCTCGAGTTACCGACGAGGGGGCAAG-3'
CrchlD-Fwd	5'-CGCGGATCCGCCATGAAGGTGTCTGAGGAG-3'
CrchlD-Rvs	5'-GGCCTCGAGGCCACTGCACCTTGCCACCTC-3'
CrchlH-Fwd	5'-CGCGGATCCTGCAATGTGGCGACTGGAC-3'
CrchlH-Rvs	5'-GGCCTCGAGGGGAGGCCGCTTATTGGAC-3'
CrGUN4-Fwd	5'-CGCGGATCCGCGGGCAAGCTGGACTC-3'
CrGUN4-Rvs	5'-GGCTCGAGAGCCGCTGAGCTGCTGAG-3'

2.3 Protein production

Mg-chelatase subunits (ChlI1, ChlI2, ChlD, ChlH) and GUN4 from *C. reinhardtii* were heterologously expressed in *E. coli* according to the procedures described in this section.

2.3.1 Protein expression

Expression of each protein was performed separately as a fusion with an N-terminal poly-His tag (ChlD, ChlH, GUN4) or a GST-tag (ChlI1 and ChlI2) which was later cleaved with PreScission Protease (GE Healthcare). The His-tagged proteins were cloned from cDNA of *C. reinhardtii* cc124 into expression vector pET28a (Accession XM_001700843 between 141 and 4416 bp) between the *BamHI* and *EcoRI* sites (Merck-Novegen, Darmstadt, Germany) (Müller et al., 2014). The GST-tagged proteins were cloned from cDNA of *C. reinhardtii* cc124 into expression vector pGEX-6P-1(GE-Healthcare).

Glycerol stocks (20 %) were prepared from single colony transformants from solid LB medium (10 g/l Bactotryptone, 5 g/l yeast extract, 10 g/l NaCl, 20 g/l agar); with either 50 mg/l kanamycin for His-tagged proteins or 100 mg/l ampicillin for GST-tagged proteins, grown to mid-late log phase. The glycerol stock (10 µl) was spread on LB agar containing the appropriate antibiotics and grown overnight at 37 °C. Fresh LB (10 mL) was added to the plate and slowly shaken at 150 rpm at room temperature for 5 min to resuspend the colonies. The resuspended colonies were used to inoculate 1 L of LB medium (10 g/l Bactotryptone, 5 g/l yeast extract, 10 g/l NaCl) at 37 °C with either 50 mg/l kanamycin or 100 mg/l ampicillin (Müller et al., 2014) in 3 L flasks. The incubation and shaking at 180 rpm and 37 °C were continued until an OD_{600nm} of 0.5 was reached. The cells were then cooled for 30 min in an ice-water bath set at ~ 10 °C prior to induction of protein expression by the addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 0.1 mM and incubation and shaking continued for a further 15 h at 15 °C.

The cells were harvested by centrifugation (4,000 x *g*, 10 min, 4 °C) and resuspended in binding buffer (20 mM Tris-HCl pH 7.9, 0.5 M NaCl, 5 mM imidazole) for His-tagged proteins or PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 7.3) for GST-tagged proteins prior to lysis by two passes through a French press. Soluble and insoluble lysis fractions were separated by centrifugation (20,000 x *g*, 15 min, 4 °C), and stored at -80 °C in small aliquots.

2.3.2 Protein purification

His-tagged ChlH, ChlD and GUN4 were purified by immobilized metal ion affinity chromatography using a 5 ml HisTrap FF column using Ni^{2+} as the metal ion (GE Healthcare). The supernatant from lysed cells was loaded onto the column and washed with 10 volumes of wash buffer (20 mM Tris-HCl pH 7.9, 0.5 M NaCl, 60 mM imidazole) and the proteins were

eluted with elution buffer (80 mM Tris-HCl pH 7.9, 0.5 M NaCl, 1 M imidazole). The peak elution fractions were determined and selected by Bradford Reagent (BioRad). Proteins were immediately desalted into Mg-chelatase standard buffer (0.5 mM Tricine-NaOH pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, 2 mM dithiothreitol) using PD-10 columns according to the manufacturer's instructions (GE Healthcare). Proteins were concentrated to a final concentration of 2-10 mg/mL using a 10 kDa Amicon centrifugal filter device (Millipore) by centrifugation at 4,000 x *g* for 15 min at 4 °C, and stored in small aliquots (50 µl) at -80 °C.

GST-tagged ChlI1 or ChlI2 lysates in 40 ml PBS (pH 7.3) were applied to a 5 ml GST-trap column (GE Healthcare), and washed with PBS until no protein was detected via Bradford assay (Bradford, 1976). Protein was eluted with GST-elution buffer (50 mM Tris-HCl pH 8.0, and 10 mM reduced glutathione), desalted with a PD-10 into precission cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT pH 7.5), and concentrated to 5 ml.

1 M unit of PreScission Protease (GE Healthcare) was added to 100 µg GST-tagged ChlI and the digest allowed to continue for 15 h at 4 °C. The mixture was re-applied to the GST-trap column, and the run-through containing ChlI1/2 with no GST-tag collected, desalted into exchange buffer, concentrated and stored at -80 °C in small aliquots.

2.3.3 Size exclusion chromatography

ChlH and GUN4 proteins were additionally purified using size exclusion chromatography with an ÄKTA purifier 10 system (GE Healthcare). The experiments were all conducted at 4 °C. A Superose™ 6 10/300 GL column (GE Healthcare) was used for ChlH purification at a flow rate of 0.5 ml/min and 0.5 ml fractions were collected. A Superdex™ 200 HR 10/30 column (GE Healthcare) was used for GUN4 purification at a flow rate of 0.5 ml/min and 0.5 ml fractions were collected. The tubing was 0.5 mm, and multiple injections for each protein sample were conducted, with an injection volume of 500 µl. Both ChlH and GUN4 proteins were eluted into

Mg-chelatase standard buffer (50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT) and were detected using the absorbance at 280 nm and 260 nm. Peaks were confirmed to contain protein by assaying fractions using a nanodrop instrument or Bradford reagent (BioRad) (Section 2.3.5). For ChlH, the two defined peaks corresponding to an oligomer at 8 ml and monomer at 12.5 ml were pooled separately, while GUN4 only had one peak. Each protein containing peak fraction was concetrated using a 10 kDa Amicon centrifugal filter device (Millipore) by centrifugation at 4,000 x *g* for 15 min at 4 °C, to a final concentration of 2-10 mg/mL and stored in small aliquots (50 µl) at -80 °C.

2.3.4 Expression and purification of BchH protein

Expression and purification of Mg-chelatase BchH subunit from *R. capsulatus* for the $O_2(a^1\Delta_g)$ production measurement in chapter 5 of this thesis was conducted by Dr. Artur Sawicki (Sawicki and Willows, 2007). The PPIX concentration bound to the BchH was determined by separating pigment bound through addition of basic organic solvent resulting in protein precipitation and solubilised PPIX. Thus, an aliquot of purified BchH was added to an excess of 80 % (v/v) acetone and 20 % (v/v), 1 M ammonia (acetone/ammonia) and immediately vortexed. The solution was centrifuged at 18,000 x *g* for 5 min at room temperature, and the supernatant was used to measure the amount of PPIX present using a PerkinElmer Life Sciences LS 55 luminescence spectrometer by referring to a PPIX standard curve of Ex400 nm and Em630 nm (Sawicki and Willows, 2008).

2.3.5 Protein quantitation and quality analysis

Protein concentration and estimation of purity was performed using nanodrop/Bradford assays and SDS-PAGE respectively:

a) Concentration assays

A NanoDrop 1000 (BIOLAB) instrument provided a rapid estimate of protein concentration at 280 nm using a molar extinction coefficient of 41,940 $M^{-1}.cm^{-1}$ for GUN4, 152,250 $M^{-1}.cm^{-1}$ for ChlH, 31,205 $M^{-1}.cm^{-1}$ for ChlD, 12,950 $M^{-1}.cm^{-1}$ for ChlI1 and 117,140 $M^{-1}.cm^{-1}$ for ChlI2 (Gasteiger et al., 2005). For each NanoDrop reading 1 µl protein was used immediately after each protein fraction was purified from the column.

Protein concentrations were also estimated by Bradford assay using Bio-Rad protein assay reagent, according to the manufacturer's instructions using BSA to generate a standard curve (Bradford, 1976) at 595 nm. Measurements were recorded using a BMG-PHERAstar plate reader with 200 µl BSA, and 25 µl protein sample in a CELLSTAR 96 well plate (Sigma-Aldrich).

b) SDS-PAGE analysis

SDS-PAGE is used to analyse protein samples in order to judge protein purity and estimate protein size. Precision Plus Protein™ All Blue standards were used according to manufacturer's instructions as a reference. Samples were centrifuged (16,000 x *g*, 3 min), and depending on protein concentration 1-5 µL or 2-5 µg protein, 1.5 µL loading dye (4X-Bolt LDS sample buffer, Life Technologies), and water to 10 µL were mixed in an eppendorf tube, boiled for 5 min at 95 °C, and centrifuged (16000 x *g*, 3 min) again. The supernatants were then loaded onto a mini protein TGX stain-free gels (Bio-Rad). Proteins were separated at 154 V for 15 min in Tris-Glycine-SDS buffer (3.03 g/l Tris, 14.41 g/l glycine, 1 g/l SDS pH 8.9). Proteins were then left to gently shake overnight in staining solution (Coomassie Brilliant Blue G-250 (0.25 % w/v), ethanol (10 % v/v), acetic acid (10 % v/v)). The staining solution was drained, the gel was destained with water by gentle shaking for 2 hours before visualization using the Gel Doc EZ imaging system.

2.4 Porphyrin preparation

Porphyrin concentrations were determined using absorbance spectrometry with a Beckman DU640 UV-Vis spectrophotometer using a quartz cuvette of path length 1 cm in 5 % (v/v) HCl at 407.5 nm and a molar extinction coefficient of 278,000 $M^{-1}.cm^{-1}$ (Dawson et al., 1986). PPIX stock solutions were freshly prepared for each Mg-chelatase enzyme assay by dissolving a speck of solid PPIX in 1 µl of 1 M NaOH, adding 100–1000 µl water, and centrifugation at 16,000 x *g* for 5 min. The supernatant was transferred to a fresh tube a small volume added to 5 % (v/v) HCl and the concentration determined spectrophotometrically using the extinction coefficient stated above. Mg-PPIX was synthesized from PPIX according to the procedure of (Fuhrhop and Granick, 1971). Also, the Mg-PPIX preparation method was the same as PPIX with the concentration determined using the same extinction coefficient (Sawicki and Willows, 2007). Stock solutions for both PPIX and Mg-PPIX were prepared fresh daily.

2.5 Mg-chelatase enzymatic assay

Preparation of each protein solution was in assay buffer consisted of 50 mM Tricine–NaOH (pH 8.0), 15 mM $MgCl_2$, 2 mM DTT, 4 mM ATP. A ChlH-GUN4-PPIX complex was prepared by first separately mixing a four-times concentrate of ChlH-PPIX and GUN4-PPIX complexes consisting of 1000 nM of ChlH with 1000 nM PPIX, and 1000 nM of GUN4 with 1000 nM PPIX. Each tube was incubated at 22 °C for 20 min in the dark, then equal volumes of ChlH-PPIX and GUN4-PPIX were mixed and incubated at 22 °C for another 20 min in the dark to give a 2 times concentrate of ChlH-PPIX-GUN4. Preparation of the ChlI1-ChlI2-ChlD complex involved an initial preparation of a 50-fold concentrate composed of 500 nM ChlD, 500 nM ChlI2 and 3000 nM which was then diluted 25 times to get a 2 times concentrate. Final concentrations of Mg-chelatase subunits used in the assays were; 10 nM ChlD, 60 nM ChlI1, 10 nM ChlI2, 250 nM ChlH, together with 250 nM GUN4 and 500 nM PPIX. Assays were

started by adding an equal volume of ChlI1-ChlI2-ChlD to ChlH-PPIX-GUN4 and immediately measuring fluorescence in a BMG-PHERAstar plate reader. Fluorescence was measured using Ex420 nm and Em600 nm, with time intervals of 10 s for a duration of 30 min to determine the initial rate in nmol Mg-proto.min^{-1}.$nmol^{-1}$ ChlD using a Mg-proto standard curve.

2.6 Small angle X-ray scattering

2.6.1 SAXS sample preparation

Purified ChlH and GUN4 proteins were concentrated to appropriate concentrations in the range 1–10 mg/ml in Mg-chelatase standard buffer (50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT) at 22 °C. The GUN4-ChlH-PPIX complex was formed by mixing ChlH with PPIX in a stoichiometric ratio of 1:1 with 1 mM ATP, and incubated at 22 °C for 20 min. GUN4 was added to the ChlH-PPIX mixture in a stoichiometric ratio of 1:1, and incubated at 22 °C for 20 min. Proteins were diluted 3-fold in the same final buffer composition for batch mode SAXS analysis.

2.6.2 SAXS data collection

Measurements were performed at the Australian Synchrotron SAXS/WAXS beamline equipped with a Pilatus 1 M detector (170 X 170 m, effective pixel size, 172 X 172 μm). The wavelength of the X-rays was 1.0332 Å. For most measurements a sample-detector length of 1576 mm was used providing a q range of 0.006-0.4 $Å^{-1}$, where q is the magnitude of the scattering vector, which is related to the scattering angle (2θ) and the wavelength (λ) as follows: $q = 4\pi \sin(\theta) / \lambda$.

SAXS data was collected in two different methods: a) A 96 well plate using a 1.5 mm glass capillary at 12 °C under continuous flow at 2 second exposures. b) Samples were run through an inline chromatography system equipped with a Superose^{TM} 6 PC 3.2/30 SEC column (GE Healthcare) at a flow rate of 0.2 $mL.min^{-1}$, with either 2 s, 5 s or 10 s exposures. Protein

concentrations are indicated in the relevant sections. SCATTERBRAIN was used to create two-dimensional radially averaged intensity plots, normalized to sample transmission, and background subtracted (Petoukhov et al., 2007).

The raw SAXS data were normalized with the incident radiation intensity using the photodiode reading in the beam stop, which also accounts for absorption properties of the sample.

2.6.3 SAXS data analysis

The ATSAS software package (Konarev et al., 2003) was used for SAXS analysis of the buffer blank subtracted data. The radius of gyration (Rg) of the scattering complex and the scattering intensity were extrapolated to zero q to obtain I (0) using two methods. One method was Guinier analysis using the ATSAS program PRIMUS (Konarev et al., 2003), and the other method used the numerical method of Glatter (Glatter, 1977b) using the ATSAS program GNOM (Svergun, 1992). Data quality was assessed using the Guinier region with PRIMUSQT (Konarev et al., 2003). The Guinier region is expected to be linear at low q (usually less than 0.003) according to the relationship $\ln I(q) = \ln I(0) - (R_g^2/3)q^2$. Thus, plotting ln I vs q^2 gives a straight line at low q and allows calculation of Rg and I(0) from the slope and y-intercept, respectively. The lowest q point used to derive this linear fit line was also used as the lowest q point for the numerical method and Fourier transform of Glatter (Glatter, 1977a). GNOM also provided the maximum dimension (Dmax), defined by the distance (r) at which the probability distribution function p(r) approaches zero. The distribution function, P(r), was derived by the indirect Fourier transform method (Glatter, 1977a).

2.6.4 SAXS 3D shape reconstruction

DAMMIN program (Svergun, 1999) was used to build ab initio low resolution shape of randomly oriented particles in solution using their regularized scattering data provided by GNOM. DAMMIN default fast mode was used for shape determination, and the q range

between 0.034 to 0.45 $Å^{-1}$ was set for all three tested proteins (GUN4, ChlH, and GUN4-ChlH-PPIX). In each DAMMIN run, few default parameters were same for the three tested protein samples. Default parameters include R-factor: 1.5, Maximum order of harmonics: 10, Point symmetry of the particle: P1, Number of equivalent positions: 1, Minimum number of contacts: 6, Maximum number of contacts: 12.

However, Table 2.2 presents the default parameters which was calculated differently in three protein samples. Also, to validate the results of the modelling, DAMMIN was run several times and the results of separate reconstructions were compared with one another and averaged. Eight 3D shapes were finally superimposed with DAMFILT and averaged with the DAMAVER program of ATSA package (Volkov and Svergun, 2003). Modelling was carried out using Chimera (Goddard et al., 2007).

Table 2.2 Default Parameters used for SAXS shape reconstructions

Name of Protein	Final SQRT(Chi)	Number of initial dummy atoms	Number of Final dummy atoms	Number of cycles
GUN4	1.096	2819	1541	155
ChlH	1.355	6101	2562	257
GUN4-ChlH-PPIX	1.100	4266	1769	174

2.7 Crystallization methods

2.7.1 Screening

Spare matrix screening was utilized to identify conditions conducive to form crystalline material of specific soluble proteins (Jancarik and Kim, 1991). A range of commercial kits $JCSG^{+}$, JCSG I, JCSG II, JCSG III, JCSG VI (QIAGEN), MCSG–I, MCS–II, MSCG–II

(Microlytic), PEG/Iron, Salt RX, Crystal Screen, Index (Hampton research) were chosen for screening crystallization of GUN4 (Table 4.1). These commercial matrix screening kits incorporate a wide range of solution variables, salt and precipitants statistically found to be the most successful in crystallizing proteins.

Screening of the GUN4 protein for nucleation was conducted in a sitting-drop vapour diffusion format, employing a 1:1 ratio of protein:precipitant in total volumes of 1-2 µl over 50 µl reservoir using 3-well intelli-plateas (Art Robbins Instruments). Crystallization screens were set up using a robotic liquid handling dispenser (Phoenix, Art Robbins Instrument) housed at the University of New South Wales, Australia (School of Physics). Wells with buffer alone were also prepared to account for inorganic salt crystal formation. The entire plate was then sealed with adhesive seals (Crystal-Clear strip, Douglas Instrument). All crystallization plates were maintained in the dark and at 25 °C for eight weeks. All wells were examined for signs of crystal growth under microscope (100-200 fold magnification) after 24 hours and then various intervals.

2.7.2 Optimization

To produce crystals of sufficient size and quality for diffraction optimization around successful initial hit conditions utilising fresh protein samples are required. For each hit condition associated with a protein, a combination of two or more of the following optimization strategies were utilized:

a) Grid screening around initial ‘hit’ (different concentrations of precipitant, buffer, additives and pH)
b) Testing variations of protein:precipitant ratios (1:1, 1:1/2)
c) Microseeding using nucleated crystals as seeds to introduce in to new drop
d) Ligand soaking to obtain the ligand bound state in which PPIX was added to the drops

Optimization screens utilized both sitting-drop (Cryschem Plate) and hanging-drop (VDX plate, Hampton research) formats over wells containing 500 µl precipitant solution. An aliquot of 1-3 µl of this precipitant was mixed with an equal volume of protein on a silicon glass cover slip which was then inverted over each pre-greased well. Following manual set up of optimization screens, trays were stored in the dark at 25 °C for up to eight-weeks.

2.7.3 Harvesting

Crystals were harvested directly from mother liquor at (Garvan Institute of Victor Chang Cardiac Research Institute) using loops attached to a steel pin and magnetic base (Hampton Research). Harvesting was achieved by simply looping crystals from the mother liquor and then plunge-cooling them in liquid nitrogen with no cryoprotection regime. Crystals were then transported to the Australian Synchrotron for data collection.

2.7.4 Crystallographic data collection

Each of approximately 50 crystals were placed on a goniometer head, and rotated around a single axis by small progressive increments while exposed to the X-ray beam and a diffraction image was recorded during each rotation increment (Dauter, 1999), providing the information of the molecular arrangement determination in the crystal (Harp et al., 1998). The data was recorded at 100 K on MX2 (beam size 37 X 32 µm) in 0.5-l.0 ° increments. Diffraction was measured on ADSC quantum 210r and 315r detectors, respectively (ADSC, Poway, USA). Usable diffraction data, using an unattenuated beam and three second exposures was only obtained from 2 crystals. Although analysis of diffraction data starts after the data set is complete, the data collection experiment and its analysis are tightly linked. The extent to which a crystal diffracts directly determines how detailed the final reconstruction of the electron density will be (Rupp, 2010). All crystallographic data collection statistics are presented in Table 4.5 and they were performed by Dr. Stephen Harrop.

2.7.5 X-ray crystallography methods

All steps of data processing, calculation, electron density reconstruction, model building, structure refinement, validation and analysis were conducted with standard software tools. Processing starts by measuring spot positions and determining the underlying lattice that would produce such a pattern of spots ("indexing"), such that a crystallographic space group can be assigned (Lamzin and Perrakis, 2000). Once this is done, the intensities of all spots over all the images are integrated. Finally, the intensities derived from the same reflection found on different images are compared, such that the intensities on different images can be scaled ("scaling" step), allowing multiple measurements of the same spots to be averaged and assembled into a final dataset (the "merging" step), which will then be used for model refinement.

The GUN4 data were indexed using the software MOSFLM (Leslie and Powell, 2007). This software allowed was used to obtain pixel labeling, smooth scaling to avoid changes in scale factor, robust estimation of parameters as well as resolution correction for dependent radiation damage to individual reflection. MOSFLM has the advantage that the accuracy of the refined cell parameters does not depend on the accuracy of the crystal to detector distance or direct beam position, providing these are known sufficiently well to allow correct indexing of the reflections (Leslie and Powell, 2007). The program POINTLESS (Evans, 2005) was used to investigate unmerged integrated intensities and determine the lattice with the highest possible symmetry compatible with unit cell parameters. This enabled a more confident determination of the space group, symmetry and screw axes to be made (Evans, 2011). When a unique solution with the highest total probability was obtained, it was chosen as a space group solution. However, when more than one laue group was equally likely, the true space group was recognized by trialing all possible space group solutions. After determination of the most probable space group of the dataset, the program steps INTEGRATRET and CORRECT were

rerun with correct space group. The program AIMLESS was implemented for merging and scaling reflections from the different images from different datasets and outputting a final list of reflections and their average intensities (Evans and Murshudov, 2013).

2.8 Proteomic analysis

2.8.1 Trypsin in-gel digestion

Each stained gel lane was cut into 16 equal size pieces and each piece was further chopped and placed into a well of a 96-well plate. In order to destain, the gel pieces were briefly washed with 100 mM NH_4HCO_3, then twice with 200 µL of acetonitrile (50 %)/100 mM NH_4HCO_3 (50 %) for 10 min and finally dehydrated with 100 % acetonitrile. The samples were air dried and reduced with 50 µL of 10 mM DTT/NH_4HCO_3 (50 mM) at 37˚C for 1 h before alkylating in the dark with 50 µL of 50 mM iodoacetamide/ NH_4HCO_3 (50 mM) at room temperature for 1 h. They were then briefly washed with 100 mM NH_4HCO_3, 200 µL of acetonitrile (50 %)/100 mM NH_4HCO_3 (50 %) for 10 min, dehydrated with 100 % ACN and then air dried. Finally, samples were digested with 20 µL of trypsin (12.5 ng/mL in 50 mM NH_4HCO_3) for 30 min on ice and then overnight at 37 °C. Peptides resulting from trypsin digestion of proteins were extracted twice with 30 µL of acetonitrile (50 %)/formic acid (2 %), dried in a vacuum centrifuge and reconstituted to 10 µL with 2 % formic acid.

2.8.2 Mass spectrometry

Each sample was analyzed on a Q Exactive Orbitrap mass spectrometer (Thermo Scientific). Samples from each fraction were separated over a 70 minute gradients by using an Easy Nano LC 1000 (Thermo Scientific). Samples from each fraction were separated over a 70 min gradients. Ten microliters of sample was injected onto a halo C18 75 µm 2.5 cm peptide trap column and desalted with 20 µl of 0.1 % formic acid. The peptide trap was then switched on line with a halo C18 75 µm x 10cm analytical reversed phased column. Peptides were eluted

from the column using a linear solvent gradient, with steps from 1–50 % of buffer (99.9 % (v/v) ACN, 0.1 % (v/v) formic acid) for 70 min, 50-85 % of the same buffer for 2 min, hold at 85 % for 8 min with a flow rate of 300 nl/min across the gradient.

The column eluate was directed into a nanospray ionization source of the mass spectrometer. A 1.5 kV electrospray voltage was applied via a liquid junction upstream of the column. Spectra were scanned over the range 350 - 2000 amu (atomic mass unit). Automated peak recognition, dynamic exclusion, and tandem MS of the top ten most intense precursor ions at 30 % normalization collision energy were performed using Xcalibur software (Thermo scientific).

Raw files obtained from nanoLC-MS/MS were converted to mzXML files and searched against *gun4* genome sequence using the global proteome machine (GPM) software, Version 2.1.1 (http://www.thegpm.org) and the X!Tandem algorithm. Spectra were also searched against a reversed sequence database for estimation of false discovery rates (FDR). Peptide identifications with log (*e*) values less than -1 were chosen. Parent ion tolerance of 100 ppm and fragment ion mass error of 0.4 Da were used for peptide identification. Complete modification of carbamidomethylation of cysteine and potential modifications of oxidation of methionine and tryptophan were considered.

2.9 Singlet oxygen detection

2.9.1 Reconstitution of ChlH and GUN4 with PPIX

Reconstitution of ChlH or GUN4 with PPIX was achieved by mixing an equal concentration of PPIX with ChlH, pre-incubating at 22 °C for 20 min and desalting into reconstitution buffer (20 mM Tricine–NaOH, pH 8.0, 2 mM $MgCl_2$) using a 5 ml HiTrap Desalting Column (GE Healthcare). The peak elution fractions were pooled after assaying by Bradford Reagent (Biorad).

2.9.2 Histidine dependent singlet oxygen consumption measurements

The rate of light induced PPIX-generated $O_2(a^1\Delta_g)$ production was measured using histidine as the $O_2(a^1\Delta_g)$ acceptor which ultimately removes oxygen from solution and this was detected using a standard Clark-type electrode (Rank Brothers Oxygen Electrode) (DeRosa and Crutchley, 2002, Rehman et al., 2013). The concentration of consumed $O_2(a^1\Delta_g)$ was quantified following calibration of the oxygen electrode assuming oxygen-saturated water has a concentration of 240 µM (Day et al., 1985). Assays were performed in triplicate in buffer containing 5 mM Histidine, 50 mM Tricine–NaOH pH 8.0, and 2 mM $MgCl_2$ with 2.5 µM PPIX and variable concentrations of GUN4 and ChlH as stated in figure legends (Figure 5.4) in a final volume of 3 mL. The total amount of oxygen removed from solution per minute was measured with a conversion factor of $[O_2] = 2.4 \times 10^{-4}$ mol·d^{-3} which was used for calibration of the oxygen electrode using air saturated water. Each experiment was performed over a 20-60 minute period under the direct incandescent light source of 132 µmol m^{-2} s^{-1} as measured in the water jacketed chamber at 30 °C.

2.9.3 SOSG fluorescent singlet oxygen production measurements

The rate of $O_2(a^1\Delta_g)$ production was measured using $O_2(a^1\Delta_g)$ Sensor Green (SOSG) reagent; a ROS-selective fluorescent sensor (Life technologies/Molecular Probes) (Flors et al., 2006). PPIX or Mg-PPIX were used as photosensitizers to generate $O_2(a^1\Delta_g)$ at a final concentration of 10 µM unless otherwise stated in figure legends (Figure 5.2). All assays were measured under the exposure of direct incandescent light 12 cm from the plate through a glass sheet to prevent heating with a measured light intensity of 132 µmol m^{-2} s^{-1}. Assays were performed in triplicate in buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT. Concentrations of GUN4, ChlH, or BSA proteins used in the assays were 1, 2, 3, and 5 µM with 5 µM SOSG (containing a final concentration of 0.02 % (v/v) methanol in

assays) used for detection. The fluorescence spectra were measured using Ex485 nm and Em520 nm and readings were taken every 50 s for 20 min in a BMG-PHERAstar plate reader at a gain of 400 at 30 °C.

Chapter 3

Chapter 3. Structure characterization of GUN4, ChlH and GUN4-ChlH-PPIX complex by Small Angle X-ray Scattering (SAXS)

3.1 Introduction

GUN4 is the regulatory subunit of Mg-Chelatase and is required for optimal enzyme activity. Cyanobacterial GUN4 binds the chlorophyll biosynthesis intermediates (PPIX and Mg-PPIX), and stimulates Mg-chelatase activity. ChlH is the largest subunit of Mg-chelatase, and plays regulatory, catalytic and chaperoning roles in oxygenic photosynthetic organisms (Hudson et al., 1993, Gibson et al., 1996). ChlH also binds both PPIX substrate and the Mg-PPIX product. It has been suggested that ChlH chaperones the Mg-PPIX product to the active site of the ChlM (the next enzyme in chlorophyll biosynthetic pathway) in *Synechocystis* (Shepherd et al., 2005), and in tobacco (Alawady and Grimm, 2005). As discussed in section 1.3.4.3 the GUN4 protein interacts with ChlH in higher plants and cyanobacteria (Larkin et al., 2003, Wilde et al., 2004, Adhikari et al., 2011). Interaction of these two subunits heavily influences the activity of Mg-chelatase (Davison et al., 2005) by mediating PPIX binding to ChlH and releasing the Mg-PPIX from ChlH (Davison and Hunter, 2011). Previous studies have demonstrated that ChlH is the only protein that co-purified with GUN4 (Larkin et al., 2003, Sobotka et al., 2008).

Structure determination will shed more light on the structure and function of GUN4, ChlH and GUN4-ChlH-PPIX complex in the Chl biosynthesis pathway. Therefore, to better understand the mechanical basis, porphyrin binding mechanism and Mg-chelatase stimulatory activity of GUN4, we undertook SAXS analysis of the GUN4, ChlH and GUN4-ChlH-PPIX proteins from *C. reinhardtii* to determine the low resolution solution structure.

This chapter elucidates the structural analysis of both ChlH and GUN4 subunits separately and also assembly of the GUN4-ChlH-PPIX complex, which will help us to clarify the functional role of GUN4 and its interaction with the ChlH subunit. Additionally, I have compared the *C.*

reinhardtii GUN4 and ChlH proteins SAXS structures with the cyanobacterial protein crystal structures.

3.2 Protein purification

The GUN4 and ChlH proteins were purified as described in section 2.3.2 and 2.3.3. Coomassie Brilliant Blue-stained SDS–PAGE gels indicate that proteins were pure after HisTrap FF purification (Figure 3.1, lane 1 and lane 2). The total yield of protein per culture volume was 6 mg. Further purification was conducted for ChlH and GUN4 by gel filtration using Superose^{TM} 6 10/300 GL, and Superdex™ 200 HR 10/30 columns (GE- Healthcare) respectively. The gel filtration profiles from these final steps of purifications are shown in Figure 3.2 A, and B. ChlH expressed in *E. coli* produces monomers and oligomers, therefore an optimized method according to a previous study was used to achieve a large proportion of monomers (Müller et al., 2014). Hence, ChlH protein with detection of absorption at 280 nm and 405 nm had two major components, the oligomeric component eluted at 8 mL and a monomeric component eluted at ~12.5 mL (Figure 3.2.A). The GUN4 protein with detection of absorption at 280 nm eluted at ~17 mL which resulted in a mono-dispersed monomeric protein (Figure 3.2B). The high purity degree of the preparation was confirmed by the SDS-PAGE.

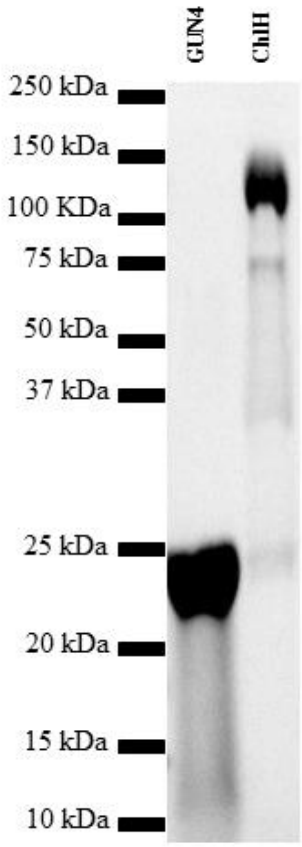


Figure 3.1 SDS-PAGE of purified proteins

SDS–polyacrylamide gel electrophoresis of the purified His-tagged proteins (2 µg of each protein was loaded).

Lane1. IMAC and Superdex 200 10/300 purified GUN4, lane 2 IMAC and Superose 6 10/300 purified ChlH monomer. The positions of molecular mass markers are indicated at the left.

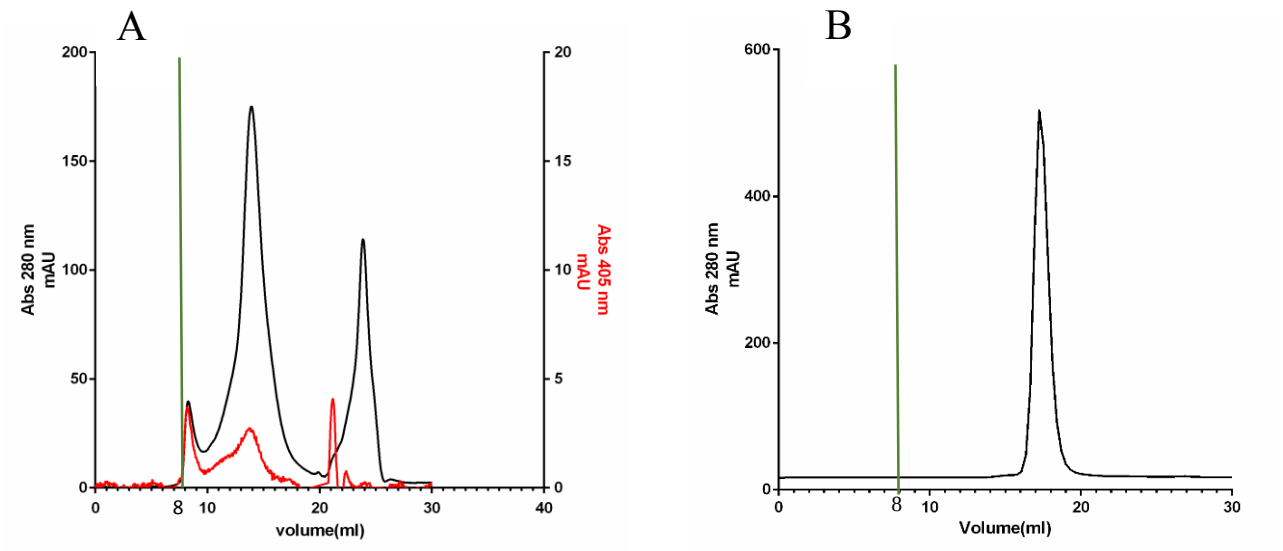


Figure 3.2 Semi preparative purification of ChlH and GUN4

Chromatographic separation was conducted in buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT.

A) Gel filtration chromatography of ChlH (black line, detection of absorption at 280 nm; red line, detection at 405nm) on Superose 6 10/300. B) Gel filtration chromatography of GUN4 on Superdex 200 10/300. The green lines in A) and B) are indicating the void volume of 8 ml for both ChlH and GUN4 calibration.

3.3 Mg-chelatase activity measurements

To ensure that the ChlH and GUN4 proteins from *C. reinhardtii* were functional, the Mg chelatase assay was used (Luo et al., 1999, Sawicki and Willows, 2008). The assay requires separate preparation of a ChlH-PPIX-GUN4 complex, and a ChlI1-ChlI2-ChlD complex, followed by their unification to start the assay. Final concentrations of 250 nM GUN4, and 250 nM of ChlH with 20 nM ChlD, 400 nM ChlI1, 20 nM ChlI2. The Mg-chelatase activity of 2.6 ± 0.1 nmol/min/nmol ChlD was determined for the monomeric ChlH. The activity for the monomer is ~5 fold higher than a previous study utilizing ChlH from *C. reinhardtii* in combination with GUN4 and remaining Mg-chelatase subunits from *Oryza sativa* (Müller et al., 2014). Oligomeric ChlH from rice and barley has variable activity compared to the

monomer, typically in the range of (0-25%) (Müller et al., 2014). The same variable results were observed using *C. reinhardtii* oligomeric ChlH.

3.4 Small angle X-ray scattering experiment

3.4.1 Significance of SAXS experiment

SAXS has become an important method to study biological macromolecules, providing the overall structure information including both architectural arrangements and conformations in the 50–10 Å resolution range in near physiological conditions (Svergun and Koch, 2002, Putnam et al., 2007).

In this study, SAXS was performed to obtain structural information for the homogeneous and monodisperse GUN4, ChlH and the GUN4-ChlH-PPIX complex in solution at 20 Å. The monomeric ChlH sample was chosen for SAXS and for preparation of the complex as it was more active in the Mg-chelatase activity assay and was found to contain monodispersed scattering particles, which are required for standard SAXS analysis. The purity of the sample preparation was confirmed using a combination of SDS-PAGE and analytical gel filtration chromatography (Figures 3.1 and 3.2).

3.4.2 Size Exclusion Chromatography SAXS (SEC-SAXS)

Scattering data was obtained from SEC-SAXS method using a superpose 6 (PC 3.2/30, 2.4ml) column (GE- Healthcare) at SAXS beamline. Final concentrations of 97.92 µM GUN4 and 5.9 µM ChlH proteins were used for SEC-SAXS analysis. High quality, aggregate free, scattering data was collected for GUN4, ChlH and GUN4-ChlH-PPIX complex proteins by running the samples through the in-line SEC column to eliminate aggregate. A table of SAXS data collection parameters for all proteins including the beamline set up is presented (Table 3.1).

Table 3.1 SAXS data collection parameters of GUN4, ChlH and GUN4-ChlH-PPIX complex

Instrument	Australian Synchrotron SAXS beamline
Beam geometry Wavelength (Å)	120 micron point source, 1.033 Å
q range ($Å^{-1}$)	0.034 to 0.45 $Å^{-1}$
Exposure time (second)	2, 5 or 10 second (sample flow at 1 µl/sec)
Concentration range loaded (mg/m1)	0.09 - 3 mg/ml
Temperature (°C)	25 °C

Gel-filtration profiles of the proteins are shown in Figures 3.3 and 3.5. All gel-filtration profiles started with 5 minutes delay between SAXS and UV-Vis at the beginning. Similarly, in Figures 3.4, 3.6 and 3.7, low q and high q profiles were of the same proteins are shown. In Figure 3.3 the GUN4 protein with detection of absorption at 280 nm eluted at ~25 min (after subtraction of 5 min delay), which is similar to the elution of GUN4 from the buffer at ~25 min presented in buffer subtraction profile in Figure 3.4. The data across the peak of GUN4 between 20 to 25 minutes was chosen for further data analysis. PRIMUS software was used for buffer subtraction (Konarev et al., 2003).

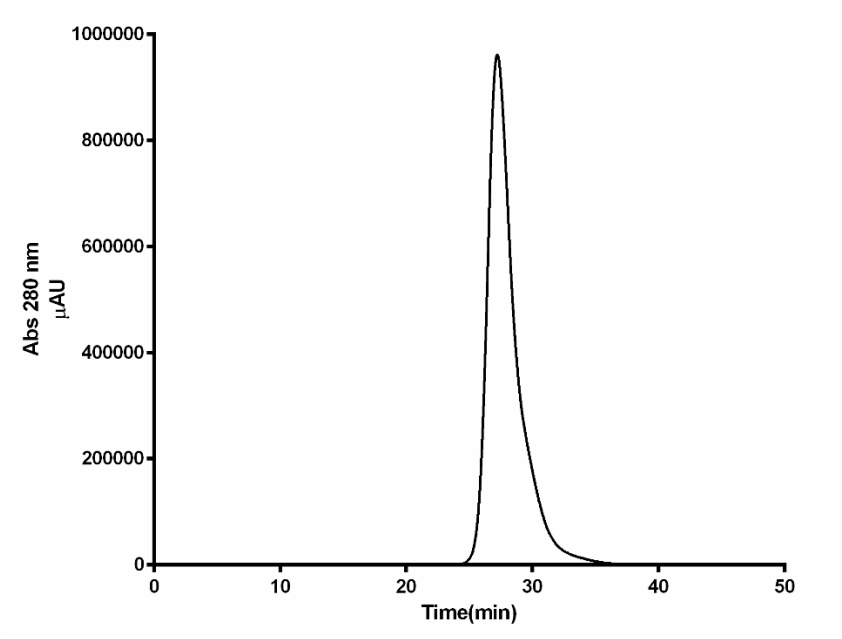


Figure. 3.3 GUN4 SAXS gel-filtration profile

Gel filtration chromatography of GUN4 in line with UV-Vis using Superose^{TM} 6 PC 3.2/30 SEC column (GE Healthcare) at absorbance 280 nm. Chromatographic separation was conducted in buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT.

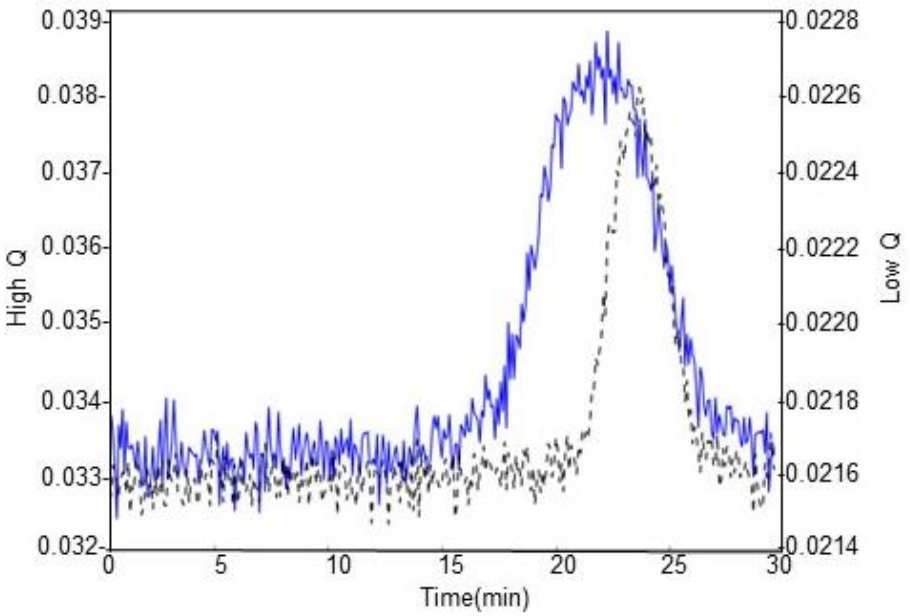


Figure. 3.4 GUN4 SAXS high Q and low Q profile

Subtraction of GUN4 protein from buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT. The blue line represents GUN4 protein, and the black line represents the buffer.

Both ChlH protein and GUN4-ChlH-PPIX complex with detection of absorption at 280 nm eluted at ~20 min (Figure 3.5). The data across the peak of ChlH and GUN4-ChlH-PPIX complex between 15 to 25 minutes was chosen for further data analysis. Buffer subtraction profile of ChlH in Figure 3.6 shows the elution of ChlH from the buffer at ~25 min, which is similar to the buffer subtraction profile of GUN-ChlH-PPIX complex in Figure 3.7. The data across the peak of the complex between 18 to 28 minutes was chosen for further data analysis.

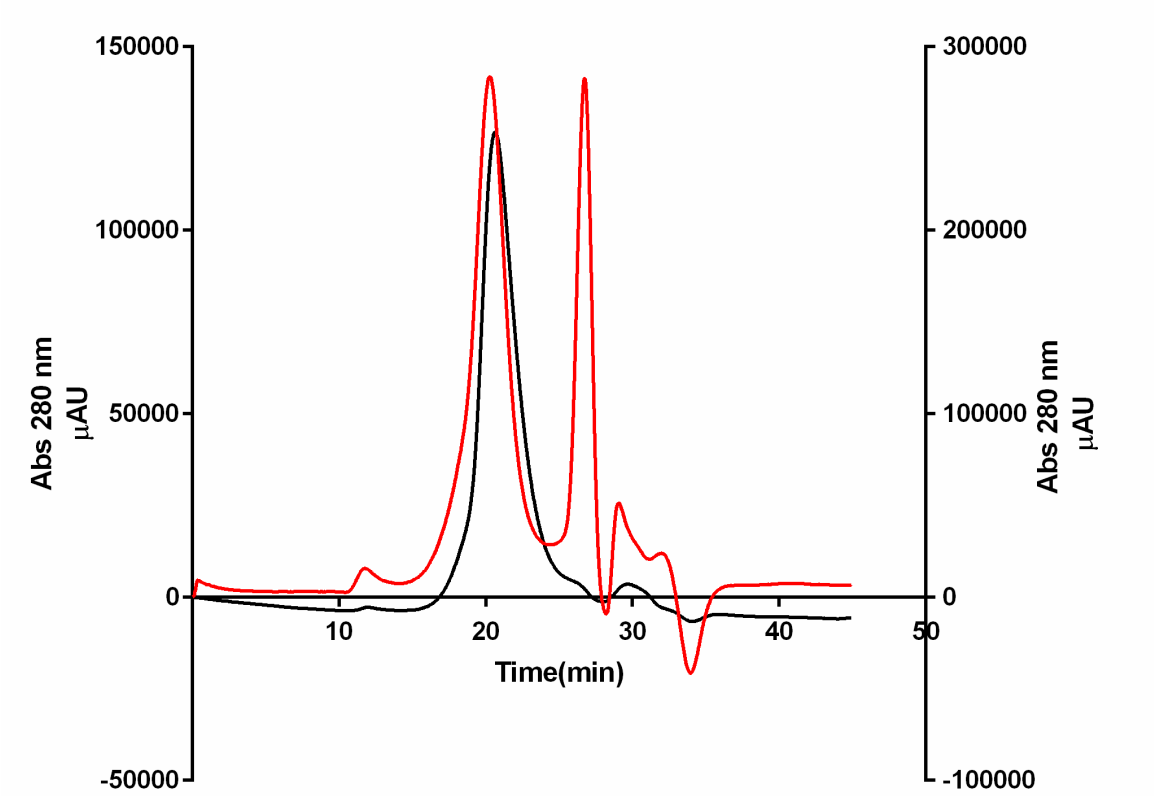


Figure. 3.5 ChlH and GUN4-ChlH-PPIX SAXS gel-filtration profile

Gel filtration chromatography of ChlH and GUN4-ChlH-PPIX complex in line with UV-Vis using Superose^{TM} 6 PC 3.2/30 SEC column (GE Healthcare). The black line represents (ChlH), and the red line represents (GUN4-ChlH-PPIX) separation at absorbance 280 nm. Chromatographic separation was conducted in buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT.

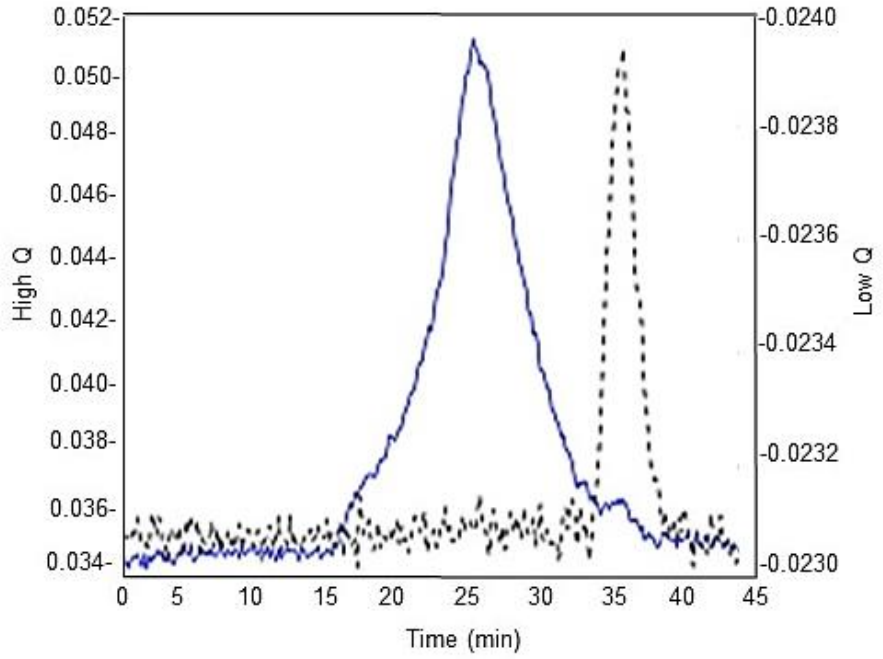


Figure. 3.6 ChlH SAXS high Q and low Q profile

Subtraction of ChlH protein from buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT. The blue line represents ChlH protein, and the black line represents the buffer.

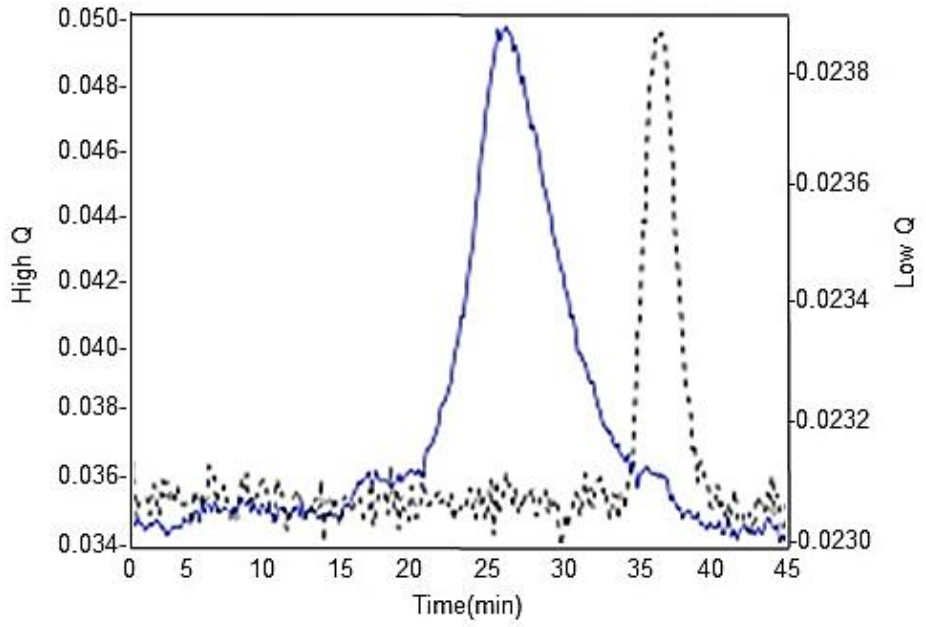


Figure. 3.7 GUN4-ChlH-PPIX SAXS high Q and low Q profile

Subtraction of GUN4-ChlH-PPIX complex from buffer consisting of 50 mM Tricine–NaOH, pH 8.0, 10 % (w/v) glycerol, 2 mM $MgCl_2$, and 2 mM DTT. The blue line represents GUN4-ChlH-PPIX complex, and the black line represents the buffer.

The data chosen across each peak of proteins was then averaged and merged using PRIMUS (Konarev et al., 2003) from ATSAS package to obtain the scattering curve. The software employed for SAXS data analysis is shown in Table 3.2.

Table 3.2 Software program employed for structure characterization of GUN4, ChlH and GUN4-ChlH-PPIX complex by SAXS

Primary data reduction	SCATTER BRAIN (Australian Synchrotron)
Data manipulation	PRIMUS
Data processing	GNOM
Ab initio analysis	DAMMIN
Validation and averaging	DAMAVER
Computation of model intensities	CRYSOL
Three-dimensional graphics representations	Chimera

3.5 SAXS structural parameters

3.5.1 Probability distribution functions

To obtain a geometrically interpretable function, the SAXS data of each protein were used to derive a probability distance distribution P(r) that represents the number of distances of length (r) present in the scattering volume (Glatter and Kratky, 1982). Also, by transforming the scattering data from reciprocal to real space it is possible to create a distance distribution plot (P(r)). The P(r) plot is positively skewed at the highest protein concentrations which is strongly indicative of elongated species (Figure 3.8B). At lower protein concentrations the P(r) is less skewed suggesting less elongated protein species (Figure 3.9B and 3.10B).

The scattering curves obtained for each protein is shown in Figures 3.8A, 3.9A, 3.10A, and Figures 3.8B, 3.9B and 3.10B present the P(r) plots. The P(r) curves of all proteins have one principal maximum peak each. The GUN4 has a principal maximum at distance 25 Å (Figure 3.8B), ChlH at distance 50 Å (Figure 3.9B) and the complex at distance 53 Å (Figure 3.10B), indicating the characteristic of globular molecules for all three proteins. The GUN4-ChlH-

PPIX complex has the highest maximum particle distance (53 Å) with no significant difference with the ChlH structure Å (Figure 3.10B).

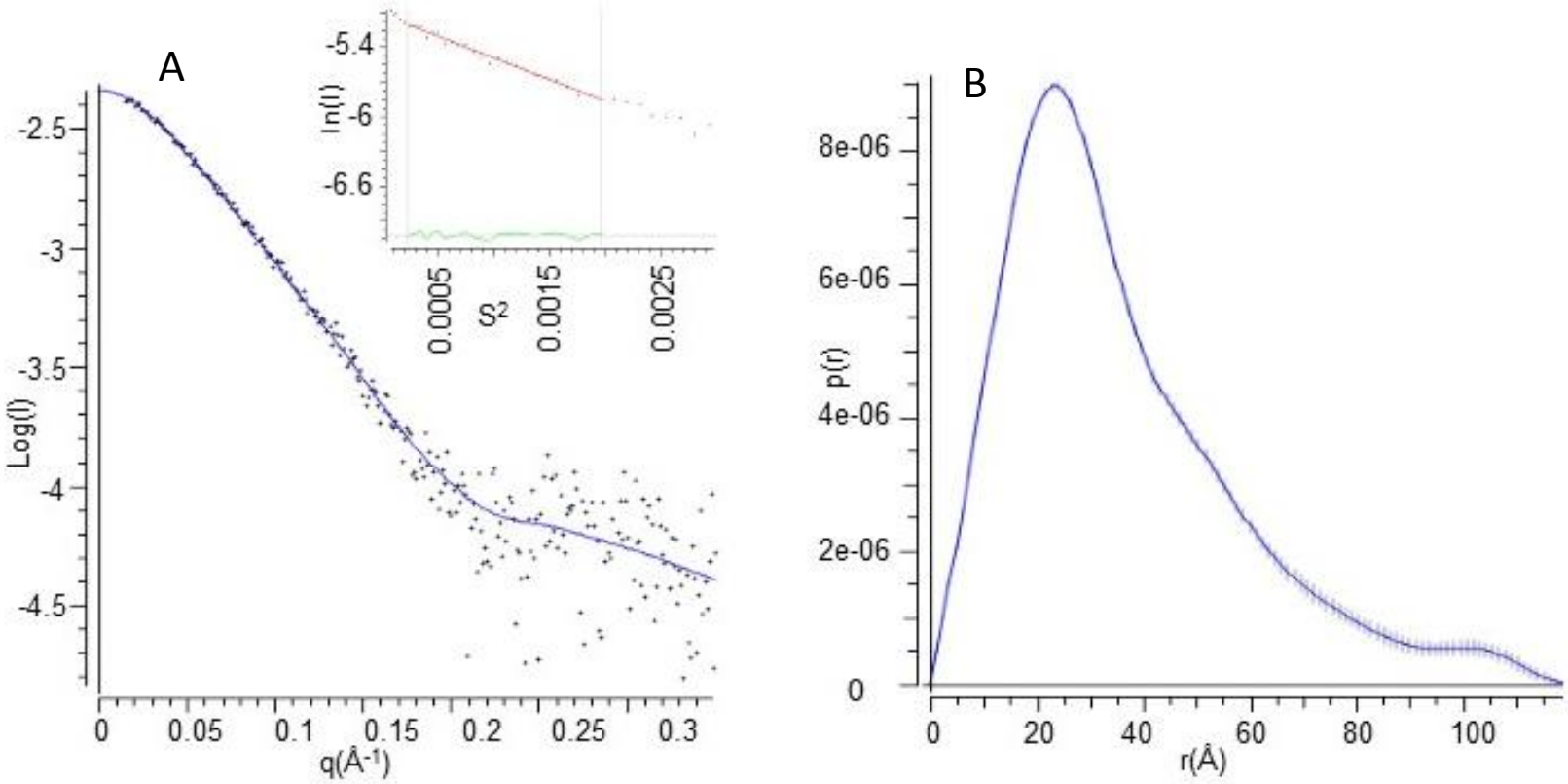


Figure 3.8 Experimental scattering curves on absolute scale for GUN4

A: GUN4 scattering curve and guinier plot with linear fit (red lines), shifted on the vertical axis for clarity, showing well determined Rg regions.

B: GUN4 P(r) curve**,** transformation of scattering curve gives radius of gyration and maximum diameter of the particle. Rg = 29.44 Å, Dmax = 120±5 Å

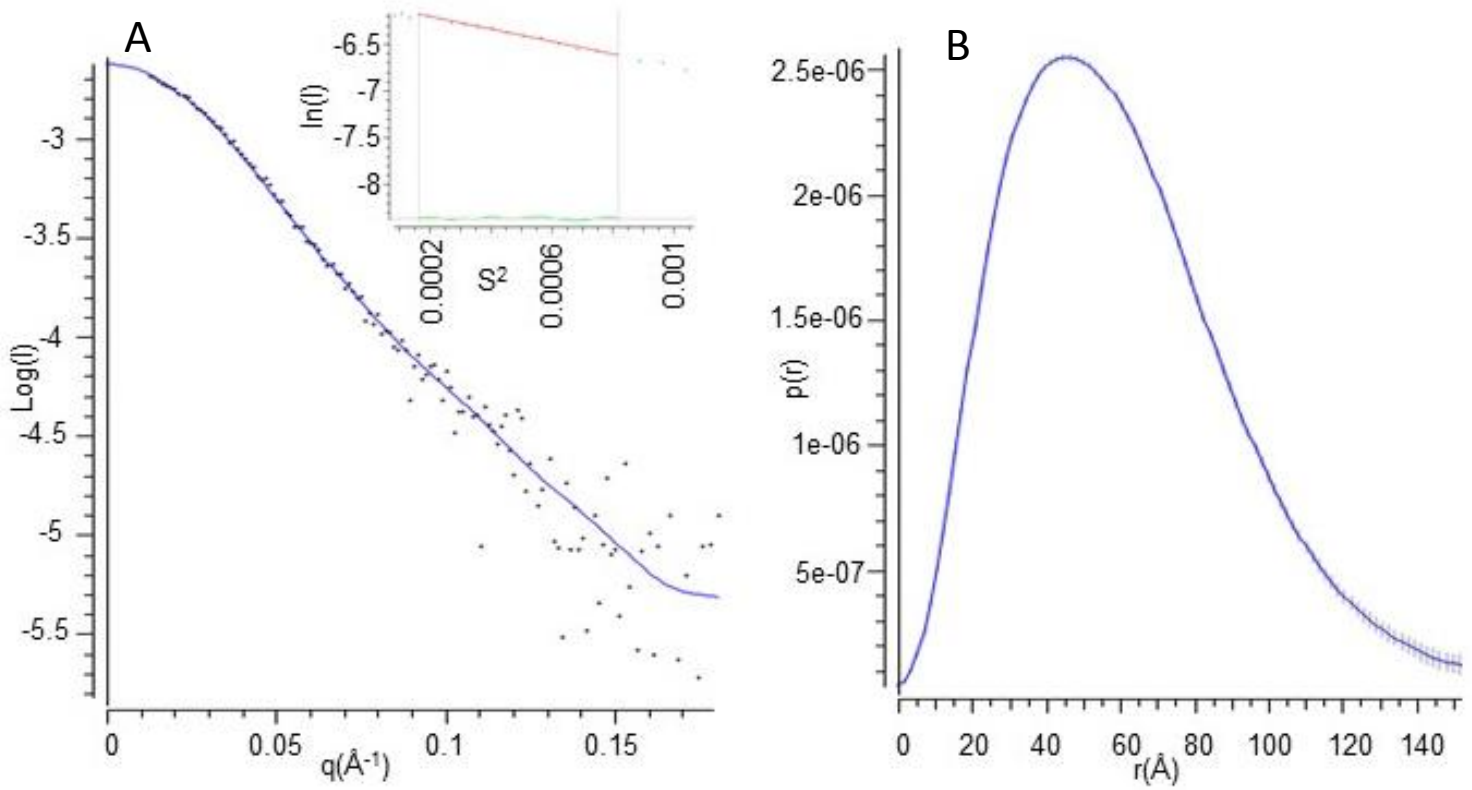


Figure 3.9 Experimental scattering curves on absolute scale for ChlH

A: ChlH scattering curve and guinier plot with linear fit (red lines), shifted on the vertical axis for clarity, showing well determined Rg regions.

B: ChlH P(r) curve, transformation of scattering curve gives radius of gyration and maximum diameter of the particle. Rg = 47.19 Å, Dmax = 155±5 Å

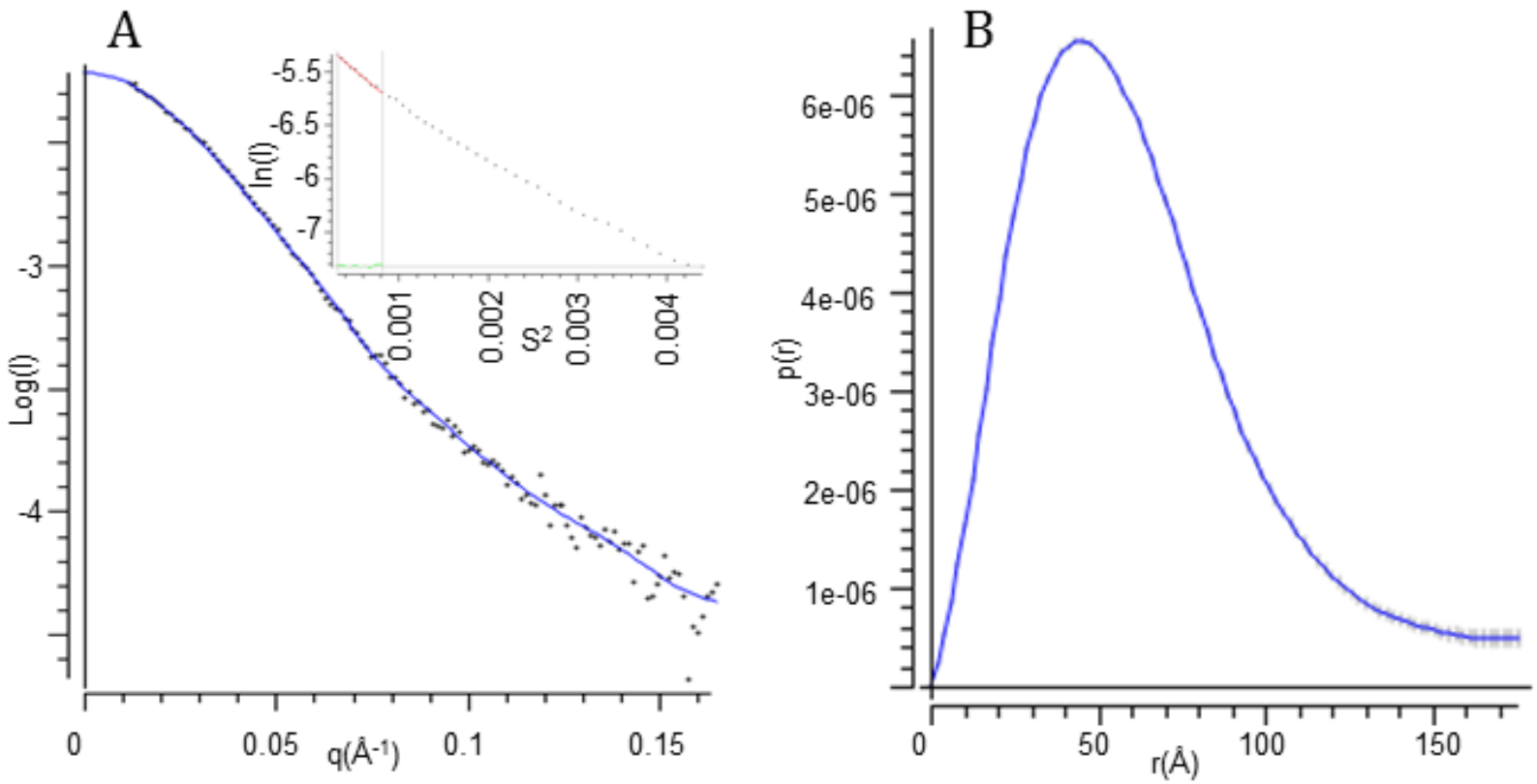


Figure 3.10 Experimental scattering curves on absolute scale for GUN4-ChlH-PPIX:

A: GUN4-ChlH-PPIX scattering curve and guinier plot with linear fit (red lines), shifted on the vertical axis for clarity, showing well determined Rg regions. **B**: GUN4-ChlH-PPIX P(r) curve, transformation of scattering curve gives radius of gyration and maximum diameter of the particle. Rg: 46.20 Å, Dmax: 175±5 Å

3.5.2 Radius of gyration

To determine whether the scattering data are from monodispersed species, radius of gyration (Rg) were determined by SAXS. Rg gives an estimation of the distribution of mass within GUN4, ChlH, and GUN4-ChlH-PPIX complex, also defines the root mean square difference of all atoms from the canters of mass. The guinier region of each data set was also plotted to check the signs of large aggregation at very low Q. In insets of Figures 3.8A, 3.9A and 3.10A guinier plots indicate that there are no protein aggregates. Guinier plots are used to derive guinier Rg values (logI vs.q^2).

3.5.3 Maximum dimension

The SAXS data provides the maximum dimension (Dmax), defined by the distance (r) at which the distribution function p(r) approaches zero or close to zero. The distribution function provides an estimation of the distribution of distances between scattering canters, giving an idea of particle size (Glatter and Kratky, 1982). Structural parameters derived from scattering data are presented for each protein in Tables 3.3, 3.4 and 3.5.

Previously, the SAXS and EM data (Qian et al., 2012) showed that ChlH forms an extended, asymmetric molecular assembly and there is no significant difference between the Rg of 47.19 Å, derived for *C. reinhardtii* ChlH at 20 Å in this study and the Rg of 46.9 Å, derived for ChlH from *Synechocystis* at 30Å calculated from SAXS data (Qian et al., 2012).

Table 3.3 SAXS structural and physical parameters of GUN4.

I(0) [from P(r)] (cm^{-1})	0.004
I(0) (from Guinier) (cm^{-1})	0.003
Rg [from P(r)] (Å)	29.44
Rg (from Guinier) (Å)	29.71
Dmax (Å)	120±5
Volume [from P(r)] ($Å^3$)	51003
Volume (from Guinier) ($Å^3$)	62998
Volume of Correlation [from P(r)] ($Å^2$)	161.211
Volume of Correlation (from Guinier) ($Å^2$)	217.62
Molecular mass [from I(0)] (Da)	9324
Molecular mass (from Guinier) (Da)	15216
Calculated monomeric molecular mass from sequence (Da)	26000

Table 3.4 SAXS structural and physical parameters of ChlH.

I(0) [from P(r)] (cm^{-1})	4.09
I(0) (from Guinier) (cm^{-1})	3.72
Rg [from P(r)] (Å)	47.19
Rg (from Guinier) (Å)	46.15
Dmax (Å)	155±5
Volume [from P(r)] ($Å^3$)	4852955
Volume (from Guinier) ($Å^3$)	3464369
Volume of Correlation [from P(r)] ($Å^2$)	930
Volume of Correlation (from Guinier) ($Å^2$)	927
Molecular mass [from I(0)] (Da)	160916
Molecular mass (from Guinier) (Da)	162330
Calculated monomeric molecular mass from sequence (Da)	154000

Table 3.5 SAXS structural and physical parameters of GUN4-ChlH-PPIX

I(0) [from P(r)] (cm^{-1})	5.20
I(0) (from Guinier) (cm^{-1})	5.60
Rg [from P(r)] (Å)	46.20
Rg (from Guinier) (Å)	47.48
Dmax (Å)	175±5
Volume [from P(r)] ($Å^3$)	457378
Volume (from Guinier) ($Å^3$)	492355
Volume of Correlation [from P(r)] ($Å^2$)	645
Volume of Correlation (from Guinier) ($Å^2$)	695
Molecular mass [from I(0)] (Da)	817000
Molecular mass (from Guinier) (Da)	926000
Calculated monomeric molecular mass from sequence (Da)	200000

3.5.4 Porod volume and volume of correlation

Some other physical parameters including scattering particle's volume (V_P), correlation length (l_c), Q and volume of correlation (V_c) could also be determined by SAXS (Glatter and Kratky, 1982). Q requires convergence of the SAXS data at high scattering vectors (q, $Å^{-1}$) in a $q^2I(q)$ versus q. V_P and l_c are undefined for flexible particles, but Rg is the only structural parameter that can be reliably derived from SAXS data on flexible systems (Rambo and Tainer, 2013).

$V_P = 2\pi^2 . I(0)/Q$ $\qquad\qquad$ $Q = \int q^2 . I(q)\, dq = c. 2\pi^2(\Delta p)^2. v$

Volume of correlation (V_c) is determined as the ratio of the particle's zero angle scattering intensity, I(0), to its total scattered intensity.

$V_c = I(0)/ \int q^{I(q)\,dq} = cV^2_p (\Delta P)^2 / cV^2_p (\Delta P)^2\, 2\pi I_c = V_P/2\pi I_c$

V_c is a newly defined SAXS invariant akin to Rg and is derived from the total scattered intensity plot and represents the ratio of the particle's volume to correlation length. This parameter will be sensitive to conformational changes and can be used to corroborate changes in Rg. Also, V_c can be combined with Rg to define Q_R ratio, for determining the molecular mass of the scattering particle.

$Q_R = (V_c{}^2 / Rg)$

The mass parameter Q_R, defined as the ratio of the square of V_c to R_g with units of $Å^3$, and is linear versus molecular mass in a log–log plot (Figure 3.11). Q_R follows a power-law (Porod's law) relationship between Q_R and molecular mass.

$I(q) = A\backslash q^4$ $\qquad$ $q^4 . I(q) = A$ $\qquad$ $q^4 . I(q) = A + c_b . q^4$

Porod's law, like guinier approximation is defined within a limited range of data. ScÅtter software was used for determining all the above scattering parameters (Rambo and Tainer, 2013), and data for this analysis are mentioned in Tables 3.3, 3.4 and 3.5.

3.5.5 Molecular mass determination by Q_R

Accurate determination of molecular mass has been one of the main difficulties in SAXS analysis. Existing methods to obtain particle mass requires an accurate particle concentration, the assumption of a compact near-spherical shape, and SAXS measurements on an absolute scale (Orthaber et al., 2000, Rambo and Tainer, 2011). However, using the ScÅtter software we determined molecular mass of proteins in solution regardless of concentration or shape assumptions (Rambo and Tainer, 2013). Rg, V_c and Q_R were calculated as described above from our SAXS profiles (Tables 3.3, 3.4 and 3.5) using ScÅtter software, and Power-law relationship between Q_R and particle mass allows direct mass determination (Rambo and Tainer, 2013):

$Mass = (Q_R / e^c)^{1/k}$ $\qquad$ $In(Q_R) = K . In(Mass) + c$

Figure 3.11 shows the parameterized power-law relationship for the set of known protein (black), mixed nucleic acid protein complexes (cyan) and RNA (red) structures from the PDB. Q_R was calculated from SAXS simulations by CRYSOL (Svergun, 1992) using highly pure protein samples resulted in the determination of k and c parameters which are specific to the class of macromolecular particle (Rambo and Tainer, 2013).

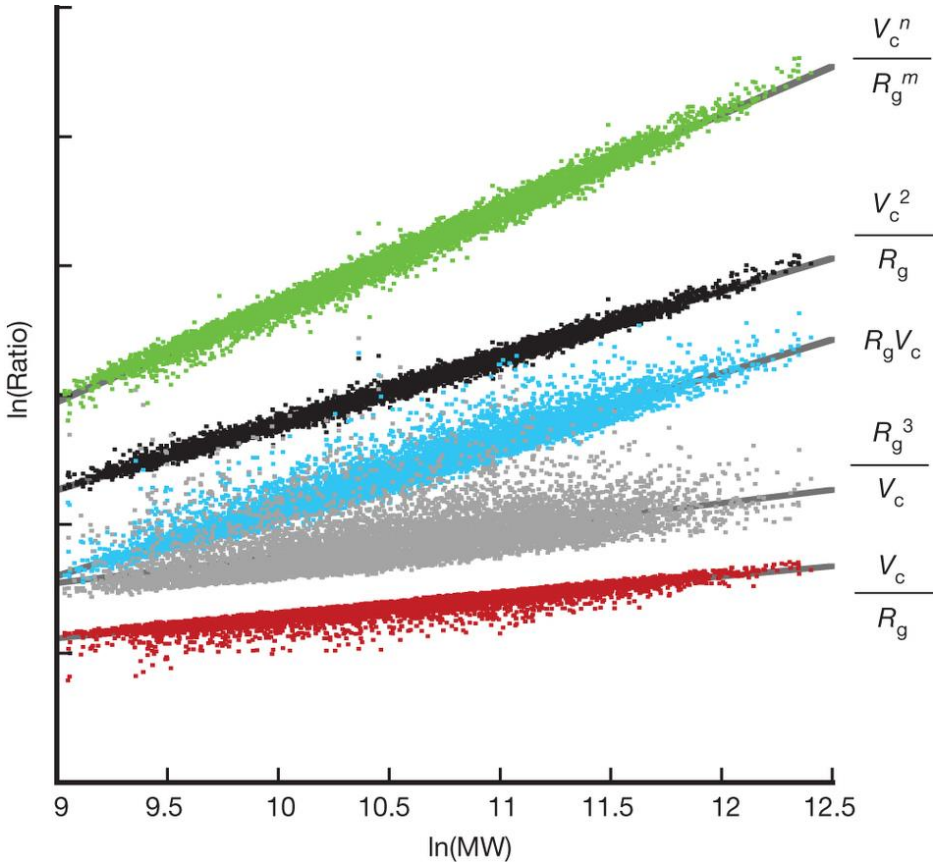


Figure 3.11 Power-law relationship between Vc, Rg and protein mass, adapted from (Rambo and Tainer, 2013) .

The linear relationship for protein observed for the ratio $V_c^{\,2}\,Rg^{\,-1}$ (black) suggests that a power-law relationship exists between the ratio and particle mass of the form ratio=c(mass)^k. The ratio, $Vc^{\,2}\,Rg^{-1}$, is defined by units of $Å^3$ with mass in Daltons, and power-law relationship (black) produces an average mass error of 4.0 ± 3.6 %. Additional ratios examined (green, cyan, grey and red) displayed asymmetric nonlinear relationships. In green, the fit included generic *m* (0.92466 ± 0.0008) and *n* (1.89260 ± 0.0005) parameters in a nonlinear surface optimization resulting in an average mass error of 4.0 ± 3.6 %. (Rambo and Tainer, 2013).

Molecular mass calculation of GUN4 from I(0) is 9324 Da and from Guinier is 15216 Da which is underestimated compared with 26000 Da calculated monomeric molecular mass from

sequence (Table 3.3). The big difference in the molecular mass calculation of GUN4 might be due to a few reasons:

a) Calculation for volume of correlation is dependent on the integral approaching a constant value at high q which is well presented and aesthetic for ChlH and GUN4-ChlH-PPIX complex in Figure 3.12 B. However, the integral does not reach a constant value at high q for GUN4 in Figure 3.12 B.
b) Many outliers were shown in molecular mass calculation of various particles (Figure 3.11). There might be few outliers in molecular mass calculation of GUN4, as it is a very small and elongated structure.
c) The scattering of GUN4 was not strong enough as it is a very small molecule, and this might affect the molecular mass calculation of GUN4.
d) C-terminal domain of GUN4 SAXS structure is elongated. This elongation is linearly effected by Rg. In $Q_R = (V_c{}^2 / Rg)$ equation for molecular mass calculation, there is a direct relationship between V_c and molecular mass, and a reverse relationship between Rg and molecular mass. Elongation in C-terminal domain of GUN4 SAXS structure increased the Rg. Thus, molecular mass was decreased as well.
e) The buffer subtraction of GUN4 from blank was very difficult, also because GUN4 is a small molecule.

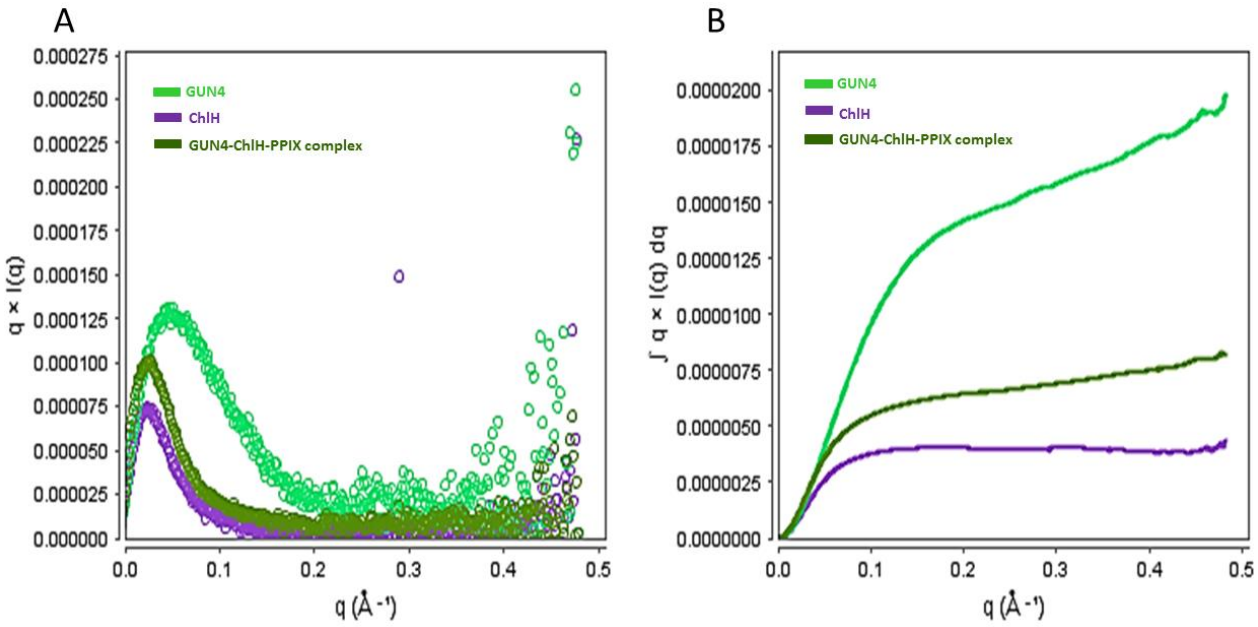


Figure 3.12 Volume of correlation for GUN4, ChlH and GUN4-ChlH-PPIX complex

A: Total scattered intensity of all three protein samples. Guinier extrapolated dataset as q . I(q).

B: Integrated area of q.I(q). The integral approaches a constant value at high q.

Integral of q.I(q) is proportional to the particle's correlation length (I_c). The ratio of (I_0) to I_c defines the volume per correlation length (V_c). Strong divergence at high q suggest poor buffer subtraction. Ratio of (V_c ^{2}/ R_g) defines an invariant proportional to mass.

3.5.6 Probability distribution P(r) comparison

No structural information has been reported for the GUN4-ChlH-PPIX complex previously. In this study we found that among all three Dmax values identified for GUN4 (120±5 Å), ChlH (155 Å), and the GUN4-ChlH-PPIX complex (175 Å), the complex has the largest Dmax. Figure 3.13 shows that there is a significant difference between the Dmax of Gun4 compared to the Dmax of ChlH and the complex, but there is no significant difference between the Dmax of ChlH and the complex. The large hollow cage shape structure of ChlH might be the reason for explanation of the particle size similarity between ChlH and the complex, as ChlH could enclose the PPIX and GUN4. This idea needs more structural and functional studies of catalytic site of Mg-chelatase enzyme.

The ChlH subunit binding to PPIX and Mg-PPIX have been described in section 1.3.4. ChlH subunit also interacts with GUN4 subunit (Larkin et al., 2003, Sobotka et al., 2008), and activates Mg-Proto-methyltransferase (Shepherd et al., 2003, Alawady et al., 2005, Shepherd et al., 2005). ChlH from *T. elongatus* is a large 140 kDa protein with 1326 amino acids. This large hollow cage shape structure was suggested to encircle the Mg-PPIX and protect it from photo oxidative damage by channelling it to the ChlM enzyme (Qian et al., 2012).

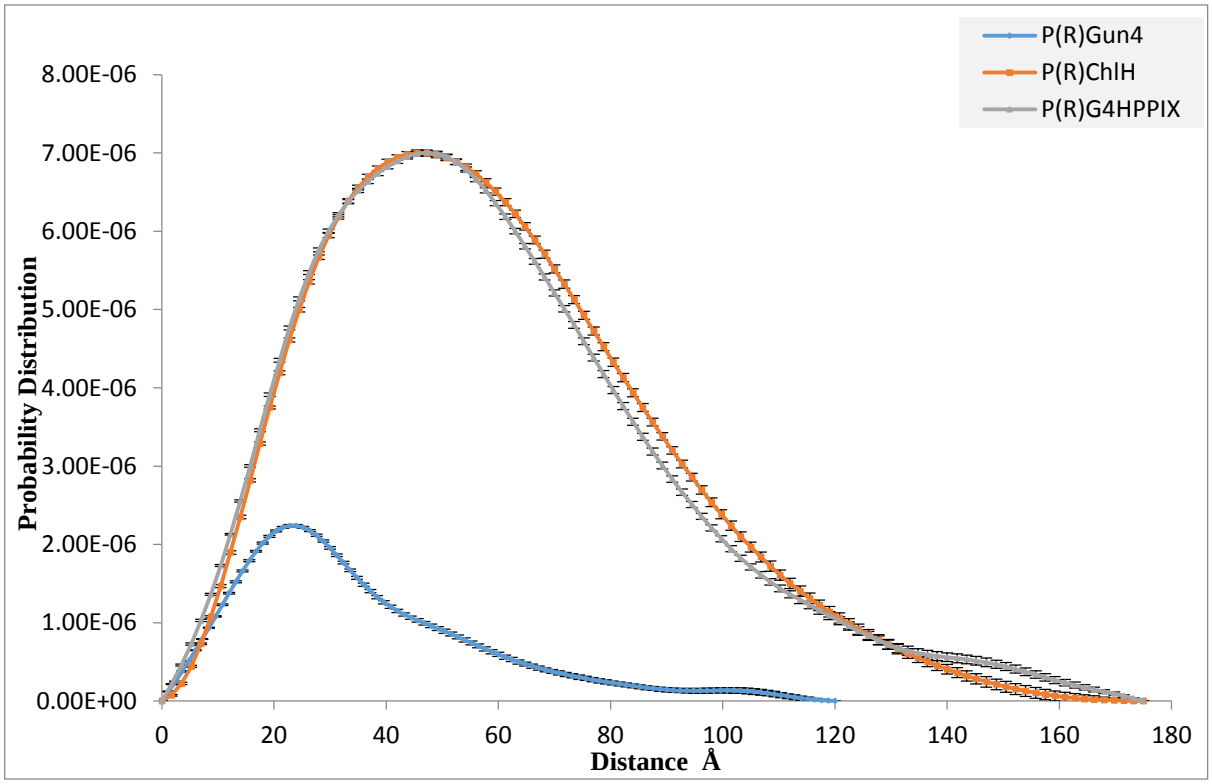


Figure 3.13 Probability distance distribution P(r) curves

Probability distance distribution P(r) curves of GUN4, ChlH, and GUN4-ChlH-PPIX complex. The area under the curve equals 1, produced with GNOM, with P(0) and P(Dmax) forced to zero.

3.6 Models of GUN4, ChlH and GUN4-ChlH-PPIX complex generated from SAXS data

3.6.1 Ab initio models generation

The scattering data were analysed and eighteen independent ab initio dummy atom models derived by DAMMIN (Svergun, 1999). Ab initio models were then superimposed and averaged to recover the three-dimensional structure from the one-dimensional scattering pattern. Figures in appendix 6, 7 and 8 show models generated from scattering pattern. The ensemble consensus structure derived by DAMFILT and averaged structure by DAMAVER (Volkov and Svergun, 2003) are shown in Figures 3.14, 3.15 and 3.16.

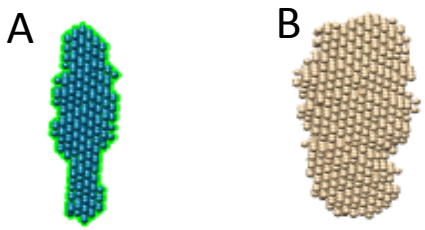


Figure 3.14 GUN4 shape determination

A: GUN4 ensemble consensus structure (DAMFILT), **B**: GUN4 averaged structure (DAMAVER)

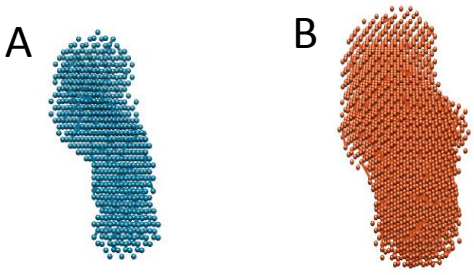


Figure 3.15 ChlH shape determination

A: ChlH ensemble consensus structure (DAMFILT), **B**: ChlH averaged structure (DAMAVER)

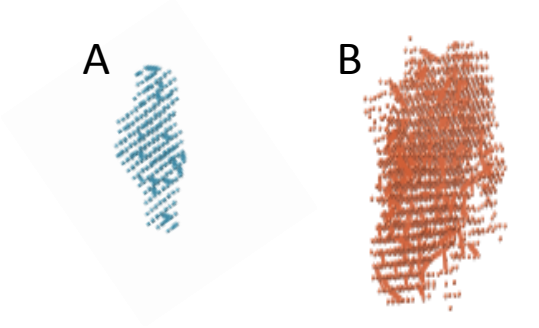


Figure 3.16 GUN4-ChlH-PPIX shape determination

A: GUN4-ChlH-PPIX complex ensemble consensus structure (DAMFILT), **B**: GUN4-ChlH-PPIX complex averaged structure (DAMAVER)

3.6.2 Comparison of SAXS models

The multiple models generated for each protein (appendix 6, 7 and 8) are visually similar and all are globular. Figure 3.17 A, B and C shows the representative ab initio dummy atom model derived by DAMMIN for GUN4, ChlH and GUN4-ChlH-PPIX complex. Surface models are displayed as mesh model and colourful bolls represent excluded dummy atoms. All structures in Figure 3.17 are presented in the same scale to facilitate the comparison.

The SAXS structure of the GUN4-ChlH-PPIX complex is very similar to the SAXS structure of the ChlH subunit suggesting that GUN4-PPIX may attach somewhere inside the cage shape structure of ChlH subunit to form a complex and this hypothesis needs more investigation on the complex formation and the complex structural analysis in future.

Moreover, the atomic model of GUN4 crystal structure with mesh surface display from *C. reinhardtii* (Tarahi Tabrizi et al., 2015) is presented in Figure 3.17 D. This model only presents the core domain of the full structure. Due to proteolysis in the long term crystallization process a part of the N-terminal domain and the last ~20 residues of C-terminal domain was removed from the full length protein (Tarahi Tabrizi et al., 2015).

Crystal structure of full length ChlH protein from *Synechocystis* has been determined very recently (Chen et al., 2015b), and this is shown in atomic model with mesh surface display in Figure 3.17 E. The overall architecture with dimensions of 133 Å × 80 Å × 75 Å is a cage shape assembly which is consistent with previous ChlH SAXS analysis (Qian et al., 2012, Adams et al., 2014b). Hence, ChlH crystal structure is in a broad structural similarity compared with our ChlH SAXS structure with a maximum diameter of 155 Å. Additionally, in Figure 3.18 the crystal structure of ChlH from *Synechocystis* (Chen et al., 2015b) fit in to our SAXS model of ChlH from *C. reinhardtii*. It is clear from Figure 3.18 that there is a high similarity in overall structural architecture.

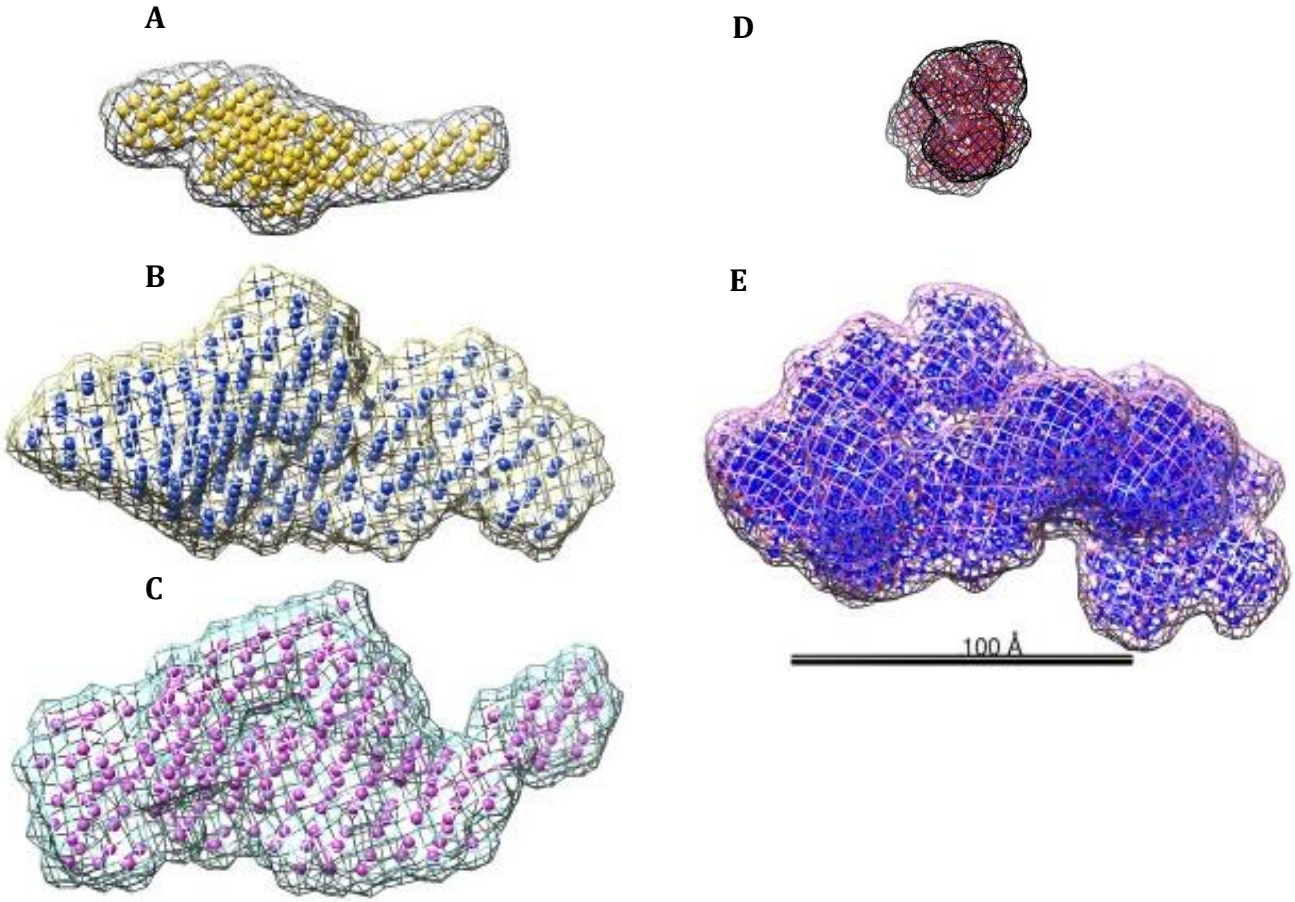


Figure 3.17 SAXS models comparison

A): SAXS model of GUN4, **B)**: SAXS model of ChlH, **C)**: SAXS model of GUN4-ChlH-PPIX Complex, **D)**: Crystal structure of GUN4 from *C. reinhardtii* (Tarahi Tabrizi et al., 2015), **E)**: Crystal structure of ChlH from *Synechocystis* (Chen et al., 2015b).
A, B, and C models are SAXS atomic models of the proteins with mesh surface presentation. Model D is the atoms model of GUN4 crystal structure with mesh surface presentation. This model only presents the core domain of the whole structure because a part of N-terminal domain and the last ~20 residues of C-terminal domain was removed due to proteolysis. Model E is the atoms model of ChlH crystal structure with mesh surface presentation. The scale bar shows the 100 Å for diameter comparison of the presented protein structures

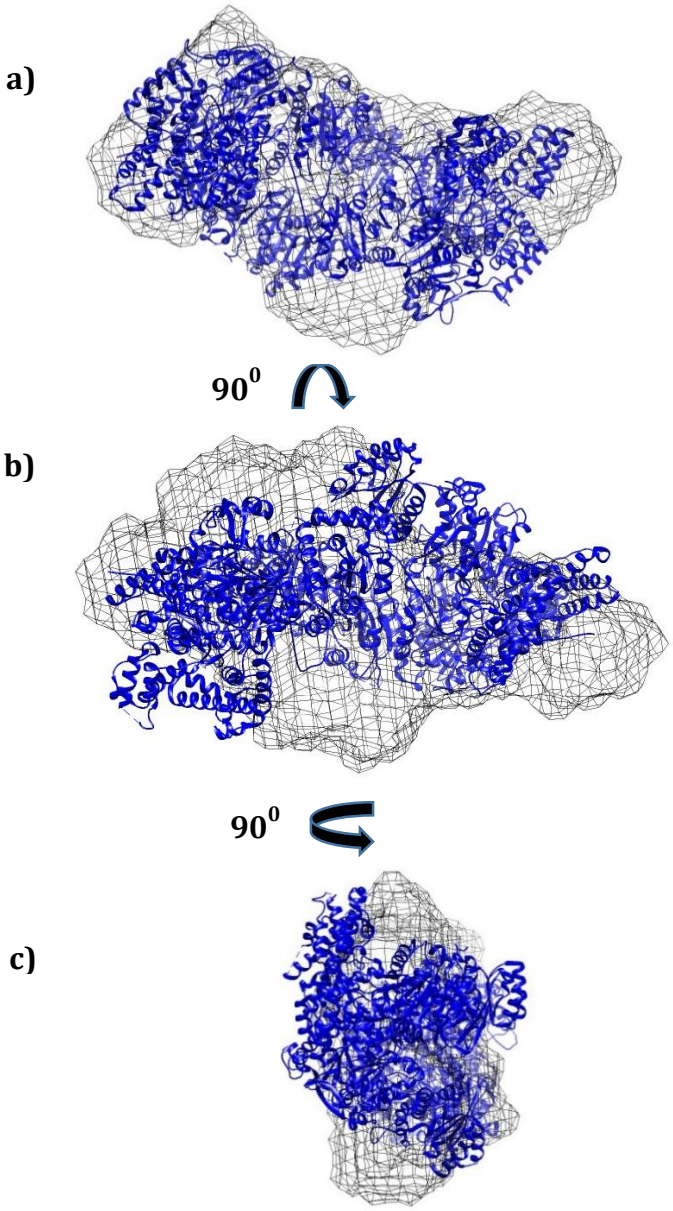


Figure 3.18 ChlH crystal structure fit of ChlH SAXS model

The blue structure represents the *Synechocystis* crystal structure of ChlH (PDB entry 4ZHJ) (Chen et al., 2015b), and the mesh model is SAXS ChlH structure from *C. reinhardtii* (data on Figure 3.9, and model on Figure 3.17 B).

3.6.3 GUN4 SAXS structure determination

It was found that GUN4 activation of Mg-chelatase requires PPIX and Mg^{2+} and GUN4 binds both PPIX and Mg-PPIX (Davison et al., 2005). On the other hand, maximum activity of Mg-

chelatase enzyme requires pre-incubation of GUN4 with ChlH (Davison et al., 2005, Adhikari et al., 2009). This has been described in details in section 1.3.4.1. The mechanism of GUN4 activation of Mg-chelatase is different in eukaryotes compared to cyanobacteria, because the orthologous C-terminal peptide VFKTNYpSF has been identified as a phospho-peptide in *A. thaliana* (Reiland et al., 2009) and all plant and algal GUN4 have the YSF as a conserved sequence. This indicates, that the orthologous cyanobacterial GUN4 proteins is missing this C-terminal extension. Additionally, plant and algal GUN4 without this C-terminal peptide is unable to activate ChlH (Figure 1.5). It has been suggested that phosphorylation of GUN4 is required for formation of a stable chloroplastic GUN4-ChlH-PPIX complex (Zhou et al., 2012). A comparison of chloroplastic GUN4 with the two proposed cyanobacterial GUN4 models created from the crystal structure of *T. elongatus* and *Synechocystis* (Davison et al., 2005, Verdecia et al., 2005) was carried out. It is clear in Figure 3.19 from the GUN4 SAXS model that chloroplastic GUN4 has an additional 9-10 amino acid C-terminal domain at the end of main molecular axis that is phosphorylated and is required for Mg-chelatase activity. Also, the N-terminal 4 helix bundle found in cyanobacterial GUN4 proteins is absent from chloroplastic GUN4 (Zhou et al., 2012).

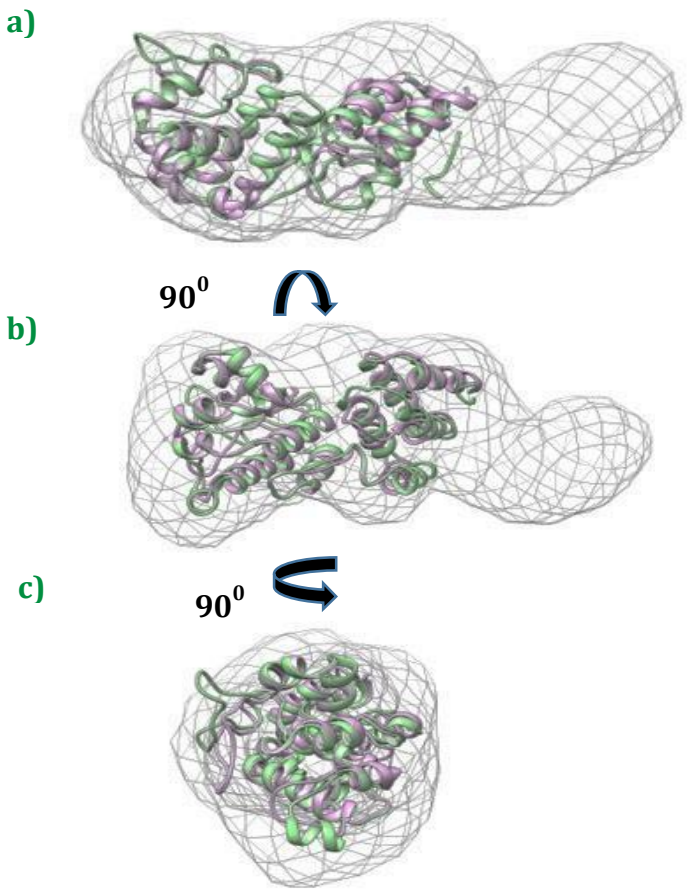


Figure 3.19 GUN4 structure fit in SAXS Model:

The green structure represents *Synechocystis* crystal structure of GUN4 (PDB entry 1y6i), and the purple structure represents *T. elongatus* crystal structure of GUN4 (PDB entry 1z3y), and the mesh model is SAXS chloroplastic GUN4 structure from *C. reinhardtii* (data on Figure 3.8 and model on Figure 3.17 A).

Chapter 4

Chapter 4. Crystallization and structure determination of GUN4 from *Chlamydomonas reinhardtii*

4.1 Introduction

4.1.1 Importance of protein crystallization

Crystallization is a commonly used technique to obtain a detailed model of the atomic structure of biological macromolecules (such as proteins, DNA, RNA, etc.), and gain insights into structure and function relationships of macromolecules in biological systems. Crystallographic studies have become of considerable importance to pharmaceutical, biotechnological and chemical industries, as well as promising tools in protein engineering, drug design and other applications in biological systems (Rhodes, 2006).

The first and the most difficult step of this method is the growth of sufficiently large crystals. A protein crystal is a three-dimensional periodic arrangement of protein molecules. Protein crystallization is not only crucial for three-dimensional protein structure identification, but is also a very interesting subject for crystal growth mechanisms studies. Protein crystals typically display relatively slow growth kinetics and have large unit cell dimensions (individual protein molecules are very large) compared to small molecule (often salts) crystal growth systems (Giacovazzo, 2002, Rhodes, 2006).

4.1.2 Complexity of protein crystallization

Protein crystallization is a very complex and difficult process due to some reasons: a) Protein molecules are composed of large, flexible molecules, and are often comprised of several subunits. b) They are chemically and physically unstable (unfolding and temperature sensitivity). Moreover, the physical and chemical properties of every protein is unique since every amino acid sequence produces a unique three-dimensional structure having distinctive

surface characteristics. Therefore, conditions applied for one protein are usually only marginally applicable to others. However, it is still impossible to predict the crystallization conditions of a protein from its other physical properties, because changes in a single experimental parameter can simultaneously influence other aspects of a crystallization experiment (Rhodes, 2006, Nemčovičová and Smatanová, 2012).

4.1.3 Principals of growing crystals

The general procedure for crystallization is to reduce the solubility of the protein and create supersaturation to pressurize the protein out of the solution while self-assembling into diffraction-quality crystals. The solubility of a protein might be reduced by either adding precipitants to the solution or removing solvent (water), or by a combination of both (Rhodes, 2006, Rupp, 2010).

4.1.4 Thermodynamics of crystal growth

Several parameters are involved in protein crystallization: concentration of both protein and precipitant, pH, and temperature are the most common parameters. Also, the parameters are varied in the search for optimal crystallization conditions (Mcpherson, 1990, Weber, 1997). In theory when the concentration of protein and precipitation combination exceeds threshold values, precipitation should occur (Rhodes, 2006).

The three stages of macromolecule crystallization are nucleation, crystal growth and cessation of growth. Nucleation is the initial formation of molecular clusters from which crystals grow. During nucleation, three dimensional associations of molecules form thermodynamically stable nuclei which provide surfaces suitable for crystal growth. In the phasing diagram inset A (Figure 4.1) the blue region shows the nucleation and growth region which requires protein and/or precipitant concentrations higher than the solubility threshold for slow precipitation. If nucleation conditions endure, many nuclei formations may occur resulting in either an

amorphous precipitate or many small crystals forming instead of a few larger ones. The phasing diagram inset B (Figure 4.1) shows the ideal crystal growth strategy, initiating with the conditions corresponding to the blue region of the phase diagram A, and after nuclei formation, move into the green region, where the growth region is (Weber, 1997, Rhodes, 2006).

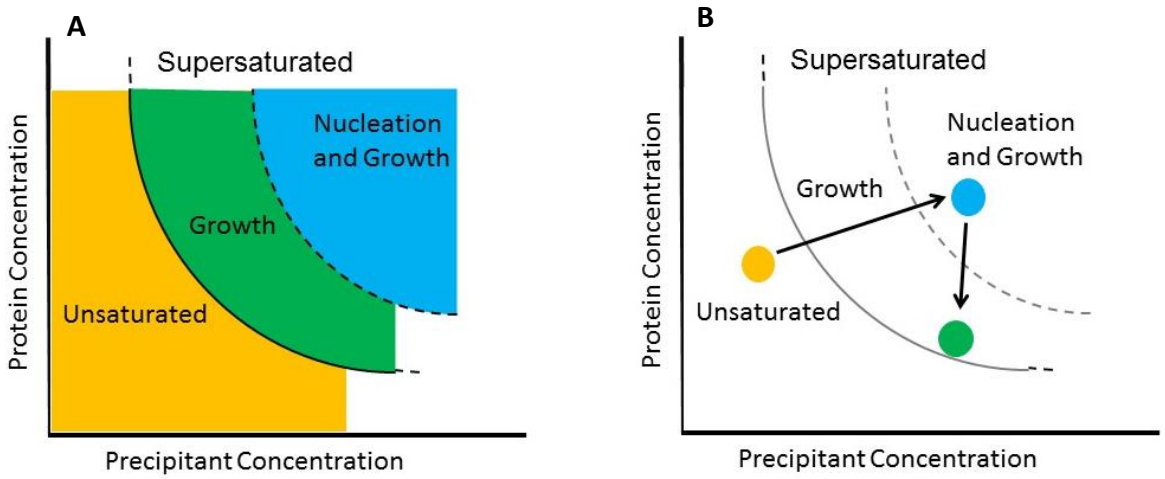


Figure 4.1 Phase diagram

The solubility curve (solid) divides phase space into regions that support the crystallization process (supersaturated solution) from those where the crystals will dissolve (unsaturated solution). The supersolubility curve (dashed) further divides the supersaturated region into higher supersaturation condition where the blue region is separated from green region.

A) Yellow region indicates concentrations of protein and precipitant where the solution is not saturated with protein. The green and blue regions indicate unstable solutions that are supersaturated with protein. Conditions in blue region support both nucleation and growth, while conditions in the green support growth only. B) An ideal strategy for growing large crystals is to let nucleation occur under conditions in the blue region, then to move to conditions in the green region until crystal growth ceases (Rhodes, 2006).

4.1.5 Crystallization methods

Screening is the first step of crystallizing a specific protein in any structure project, and it is a crucial and rate-limiting step for its structure determination. It is not possible to predict exactly

under what conditions any protein will form a crystal and thus a large number of solution variables must be thoroughly searched (Rhodes, 2006).

Vapour diffusion is a commonly used crystallization technique, in which the protein/precipitant solution is allowed to equilibrate by diffusion through air in a closed container with a larger aqueous reservoir containing the precipitant solution only. For proteins with pH-dependent solubility, supersaturation can also be achieved by diffusion of volatile acids and bases from reservoir solutions (Weber, 1997, Rhodes, 2006) Two forms of the vapour diffusion technique are the hanging-drop and sitting drop methods which have both been successfully used for protein crystal growth.

Optimization is the second step of crystallization which refines the chemical and physical parameters to produce crystals of sufficient size and quality for diffraction by utilizing the information derived from the screening experiments. In crystals produced from optimization, both environmental parameters (such as temperature) and chemical parameters (type and concentration of chemicals and pH) might be refined (Luft et al., 2007). Optimization strategies predominantly focus on decreasing the rate of vapour diffusion between the protein droplet and the reservoir precipitant solution that may slow the nucleation rate (Cudney et al., 1994).

Grid screening, is a widely used optimization strategy, in which solution pH and precipitant concentration are varied to determine optimal conditions for protein crystal growth (Cox and Weber, 1988). Some other examples of optimization techniques are screening against a suite of additives suitable for biomolecular targets (common cofactors, inhibitors, etc.), and variation of protein: precipitant ratios (1:1, 1:2, and 2:3). Also, seeding is the most powerful optimization method to separate growth and nucleation (Bergfors, 2003). In this method nucleated crystals are used as seeds and introduced into new drop equilibrated at lower levels of supersaturation.

4.2 GUN4 crystallization screening trials

Initially, GUN4 was purified in 50 mM Tricine-NaOH pH 8, 2 mM $MgCl_2$, 2 mM DTT, 10 % (w/v) glycerol. Three litters of growths culture yielded 18 mg of soluble protein for numerous crystallization trails via the purification procedure outlined in section 2.3.2 and 2.3.3.

Table 4.1 Protein production information. The N-terminal tag from the cloning vector is highlighted in yellow, while green highlighted sequence is part of the N-terminal domain of GUN4 that has been lost due to proteolysis in the crystal structure.

Source organism	*Chlamydomonas reinhardtii*
Expression vector	pET-28a
Expression host	*E. coli* BL21 (DE3) star
Complete amino-acid sequence of the construct produced	

MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDHPFTMAMRVTVAAGKLDSVSL
FGGDTASLMGGSQTVEKKKSGKEAVMEVQLSSTAGIDYTVLRDHLANGEFREAEDE
TRALLIKLAGPEAVKRNWVYFTEVKNISVTDFQTLDNLWKASSNNKFGYSVQKEIW
VQNQKRWPKFFKQIDWTQGENNNYRKWPMEFIYSMDAPRGHLPLTNALRGTQLFQ
AIMEHPAFEKSSTAKTLDQKAAEAAGRTQSLF

Given that there is a strong association of protein crystallization with purity, homogeneity and mono-dispersity of the protein sample (Benvenuti and Mangani, 2007) additional purification steps and a final gel filtration step was performed on the GUN4 protein used in subsequent crystallization trials and optimizations. Figure 4.2 shows the purity of the GUN4 sample used in all subsequent crystallization trials, with different loadings ranging from 2 to 0.125 micrograms loaded and no obvious other protein bands in the sample detected.

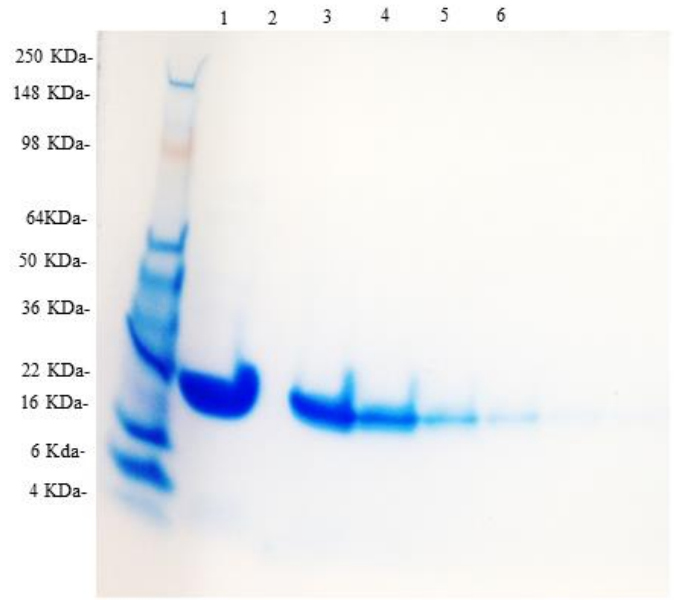


Figure 4.2 GUN4 SDS-PAGE determination

SDS–polyacrylamide gel electrophoresis of the serial dilution of purified His-tagged GUN4. In lane 1, 3, 4, 5 and 6 from left to right 2, 1, 0.5, 0.25, 0.125 µg of GUN4 protein was loaded respectively.

Sparse matrix crystal screening (employing numerous commercially supplied screens each containing 96 conditions) was used for crystallization of purified GUN4 protein at various concentrations from 3.8-35.7 mg/ml (listed in Table 4.2). Twenty one sparse matrix screens (a total of 2,016 wells) were visually inspected for crystal growth over a 48-week period and any manifestation of any crystallization or precipitation recorded. All screening methods are outlined in section 2.7.1.

Inspection of the trays /experiments 1, 2 and 3 showed mostly clear drops or, in a few instances, protein precipitation of varying intensity (light, medium and heavy), with no obvious trend relating to pH, salt type/concentration or precipitant present (listed in Table 4.2). To address clear drops, one can increase the protein concentration to increase the likelihood of achieving a supersaturated state. Hence, the protein concentration was increased from 3.8 mg/ml gradually

to 35.7 mg/ml. Additionally, among four buffer conditions (experiments 1-4), listed in Table 4.2, GUN4 protein was more stable in buffers containing 20 mM Tricine-NaOH pH 8, and 2 mM B-ME (2- β Mercaptoethanol).

Table 4.2 Determination of suitable buffer condition, and protein concentration for crystallization trials. Four different experiments with variation in GUN4 concentration, buffer condition, and crystallization screens. GUN4 concentration and buffer condition in experiment 4 was determined as the most suitable GUN4 crystallization condition.

Experiment #	**GUN4 Concentration mg/ml**	**Crystallization Screens**	**Buffer Condition**
1	3.8	MCSG I MCSG II MSCG III JCSG+	50 mM Tricine-NaOH pH 8 2mM $MgCl_2$ 2mM DTT 10 % (w/v) glycerol
2	19.3	PEG/Iron Salt RX Crystal Screen Index	50 mM Tricine-NaOH pH 8 2mM $MgCl_2$ 2mM DTT 10% (w/v) glycerol
3	34.3	PEG/Iron Salt RX Crystal Screen Index	50 mM Tricine-NaOH pH 8 2mM $MgCl_2$ 2mM DTT 10 % (w/v) glycerol
4	35.7	JSCG+ JCSG I JCSG II JCSG III JCSG VI PEG-Iron Index Crystal screen Salt RX	20 mM Tricine-NaOH pH 8 2mM B-ME(2-β Mercaptoethanol)

Crystallization was only successful in experiment 4, 2-weeks after establishing trays. In this experiment GUN4 protein was at the highest protein concentrations, and nano-crystals were observed in several conditions (listed in Table 4.3).

Table 4.3 Seventeen GUN4 protein crystallization conditions from this work

Appearance of GUN4 crystals in different time intervals

#	**Crystallization condition**	**Screen name**	**Time of appearance**
1	0.03 M Citric acid 0.07 M BIS-TRIS propane (pH 7.6) 20 % (w/v) Polyethylene glycol 3,350	PEG-Ion	2-weeks
2	0.02 M Citric acid 0.08 M BIS-TRIS propane (pH 8.8) 16 % (w/v) Polyethylene glycol 3,350	PEG-Ion	2-weeks
3	4 % (v/v) Tacsimate (pH 7.0), 12 % (w/v) Polyethylene glycol 3,350	PEG-Ion	2-weeks
4	4 % (v/v) Tacsimate (pH 8.0), 12 % (w/v) Polyethylene glycol 3,350	PEG-Ion	2-weeks
5	0.1 M Tris pH 7.0 20 % (w/v) PEG 2000 MME	JCSG I	2-weeks
6	0.2 M sodium chloride, 0.1M Na/K phosphate (pH 6.2) 40 % (v/v) PEG 400	JCSG II	2-weeks
7	0.2 M lithium sulfate, 0.1 M CHES (pH 9.5), 1.0 M sodium/potassium tartrate	JCSG IV	16-weeks
8	0.1 M Bicine (pH 9.0), 2.4 M Ammonium sulfate	JCSG IV	16-week
9	1.0 M Ammonium citrate tribasic (pH 7.0), 0.1 M BIS-TRIS propane (pH 7.0),	Salt Rx	44-weeks
10	0.7 M Sodium citrate tribasic dehydrate, 0.1 M Tris (pH 8.5)	Salt Rx	44-weeks
11	0.8 M Lithium sulfate monohydrate, 0.1 M Tris pH 8.5	Salt Rx	44-weeks
12	1.0 M Magnesium sulfate hydrate, 0.1 M Tris pH 8.5	Salt Rx	44-weeks

13	0.7 M Ammonium tartrate dibasic, 0.1 M BIS-TRIS propane (pH 7.0),	Salt Rx	44-weeks
14	0.7 M Ammonium tartrate dibasic, 0.1 M Tris(pH 8.5),	Salt Rx	44 weeks
15	0.01 M Cobalt(II) chloride hexahydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 1.0 M 1,6-Hexanediol	Crystal Screen HT	44 weeks
16	0.1 M MES monohydrate pH 6.5, 1.6 M Magnesium sulfate heptahydrate	Crystal Screen HT	44 weeks
17	0.1 M Sodium chloride, 0.1 M HEPES pH 7.5, 1.6 M Ammonium sulfate	Crystal Screen HT	44weeks

Crystals 1-6 in Table 4.3 were photographed and are shown in Figure 4.3. The size of these crystals were not satisfactory and they did not display good morphology. Thus, extensive optimization trials were required to increase the quality of crystals.

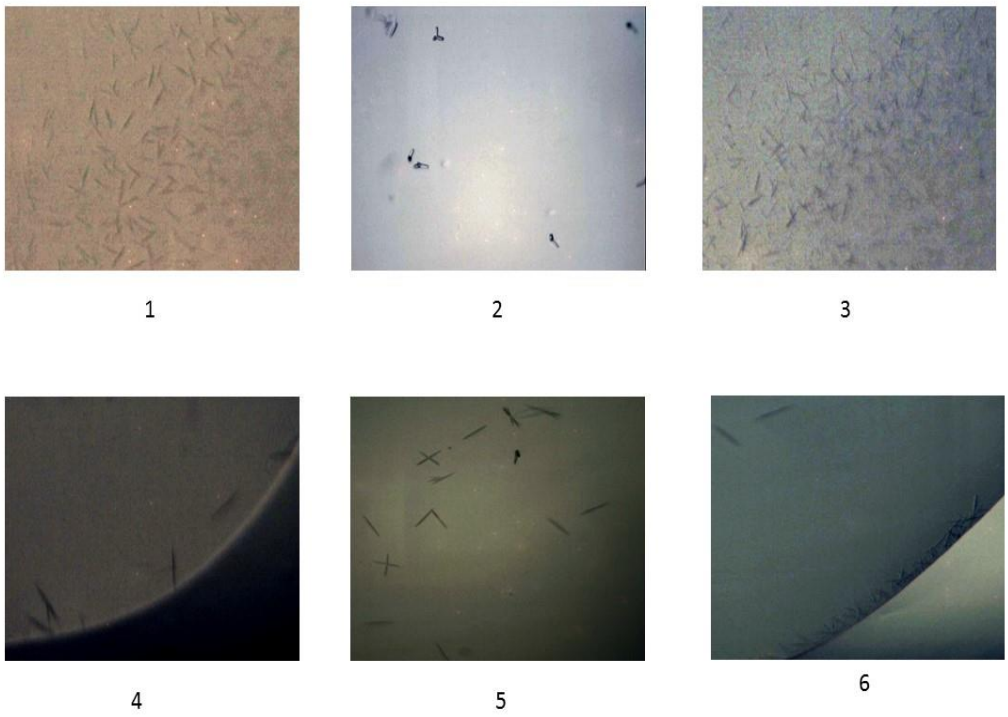


Figure 4.3 Six GUN4 nano-crystals obtained directly from spare matrix screens
Nano-crystals number 1-6 were grown using sitting drop format (1-2µl) and monitored by microscope (100-200 fold magnification), and observed after 2-weeks. Relevant growth conditions 1-6 listed in Table 4.3.

4.3 Crystal optimization

Various optimization methods are presented in section 2.7.2. The crystallization condition of crystal 2 (Table 4.3) was chosen as the best target for optimization trails as it was obtained with slightly better morphology and larger size when compared to the other five nano-crystals. A grid screening strategy (different concentration of PEG and different pH), around this 'hit' was used in conjunction with variation of protein: precipitant ratios (1:1, 1:1/2). Crystal growth following optimization was visualized under microscope (100-200 fold magnification) and recorded at various intervals, but there were no significant differences in the size and morphology of crystal formation after 8 weeks.

The grid screening may not have been successful could be because the optimal conditions for nucleation of the crystal were not ideal to support growth and/or the level of supersaturation was too high since large crystal growth requires lower supersaturation level.

Microseeding technique was then used to optimize crystal 1 (Figure 4.4). All crystals in inset 1 from Figure 4.3 were pulverized (smashed) into a crystalline slurry by glass rods and vortexed, in order to prepare seed stock (Luft and DeTitta, 1999). This seed stock was then used either in conjunction with variations of protein:precipitant:seed stock ratios (1:1:1/2) (Figure 4.3), and also streak seeding using cat whiskers (Stura and Wilson, 1990).

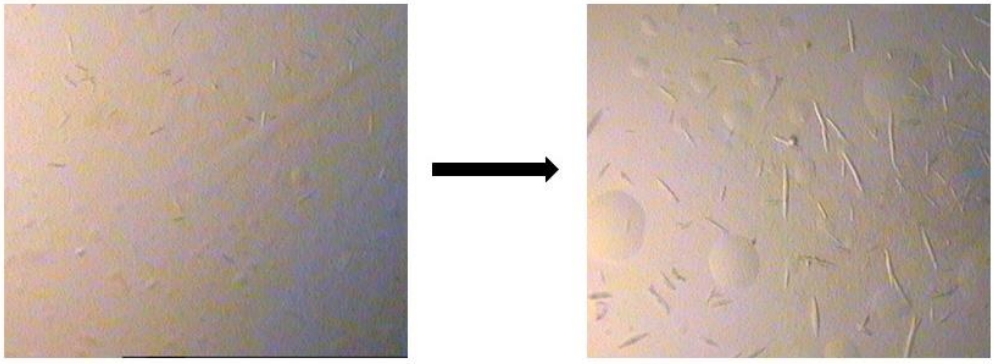


Figure 4.4 Utilizing microseeding technique for optimization

GUN4 micro crystals were grown using microseeding technique, and hanging drop format (2-3 µl). Both the size and morphology of the crystals were improved over 3-weeks.

Additionally, in order to test the possibility of protein-ligand optimization methods, ligand soaking was employed in conjugation with the microseeding method by adding PPIX to the precipitate solution and utilizing variations of protein, precipitant, seed stock ratios (1:1:1/2). Unfortunately, this method was not successful as there was no evidence of crystal formation. For these experiments inspection of trays was complicated by the requirement to use a red filter

due to the light sensitivity and photo-oxidative damage to the protein from the PPIX. From all these optimization techniques, micro seeding was found to be the most useful in terms of improving the size of the crystals (Figure 4.3).

After 16 weeks of incubation, crystals 7 and 8 (Table 4.4) were observed and photographed (Figure 4.5). Unfortunately the size of these crystals was still not sufficient for data collection.

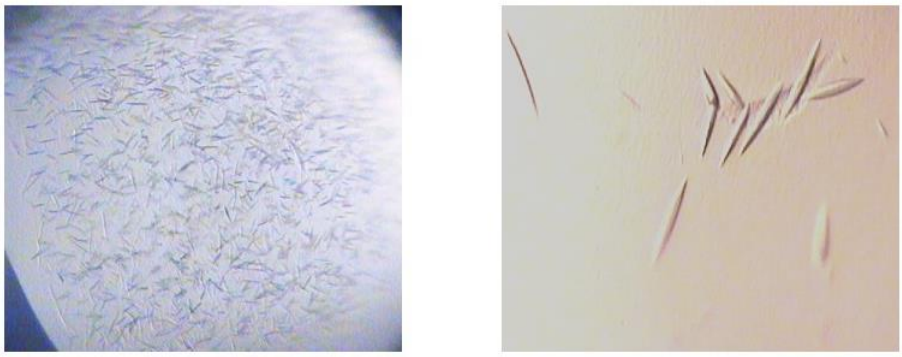


Figure 4.5 GUN4 micro crystals obtained directly from spare-matrix screens

GUN4 micro-crystals number 7-8 were grown using sitting drop format (1-2 µl) and monitored by microscope (100-200 fold magnification), and observed after 16-weeks. Relevant growth conditions 7-8 listed in Table 4.3.

4.4 Harvesting and crystallization data collection

After 44-weeks of incubation, GUN4 crystals suitable for data collection were obtained directly from spare-matrix screens. Following 1 min exposure on in-house diffractometer (Nonius FR591 rotating X-ray generator), images were collected on a Mar345 detector. Crystal 9 (Figure 4.6) showed diffraction to low (<9.0 Å) resolution. Crystallization information for crystal 9 is presented separately in Table 4.4. Harvesting and crystallographic data collection methods are outlined in sections 2.7.3 and 2.7.4.

Conditions 9-17 in Figure 4.6 provided crystals large enough to harvest, and 52 of these were plucked from mother liquor using nylon loops, then plunge-cooled into liquid nitrogen with no

additional cryoprotection regime. These crystals were transferred in a frozen state to the Australian Synchrotron (Melbourne). All crystallographic data collection statistics are presented in Table 4.5.

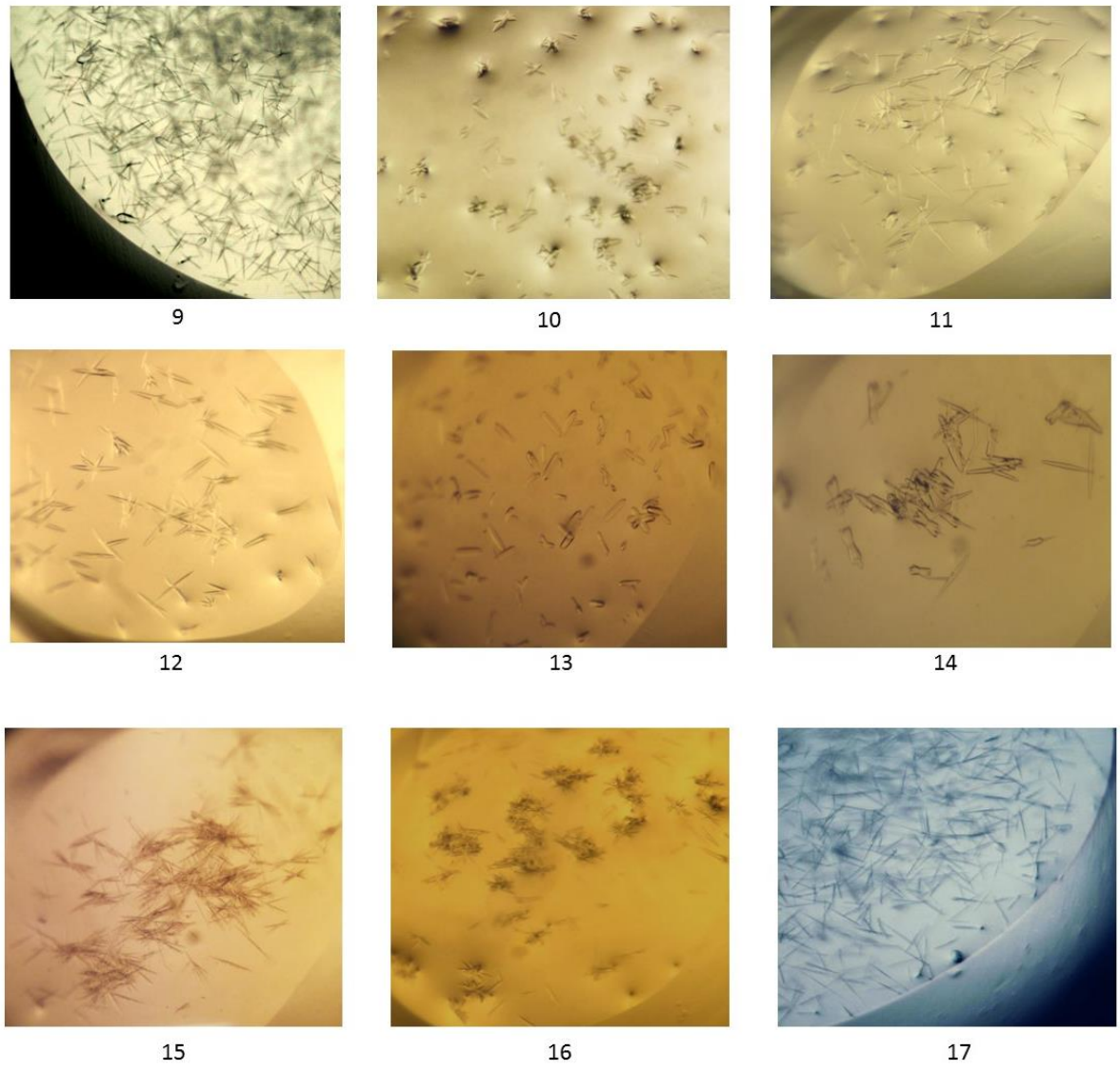


Figure 4.6 GUN4 Crystals obtained directly from spare-matrix screens

Crystals were grown using sitting-drop format (1-2 µl) and monitored by microscope (100-200 fold magnification) in various intervals and observed after 44-weeks. Growth condition of each target is indicated in Table 4.2.

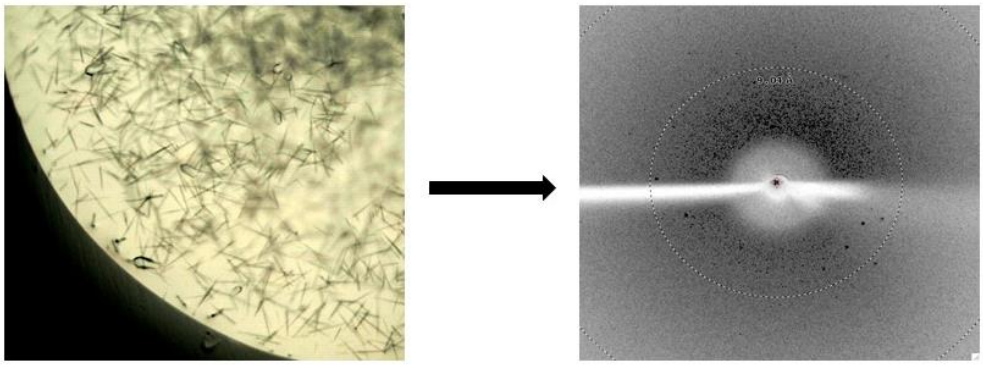


Figure 4.7 GUN4 Crystals diffraction on in-house diffractometer

GUN4 crystal directly from mother liquor diffracted to low (<9.0 Å) resolution. Diffraction quality is poor, and the resolution is low.

Table 4.4 Crystallization details of crystal 9

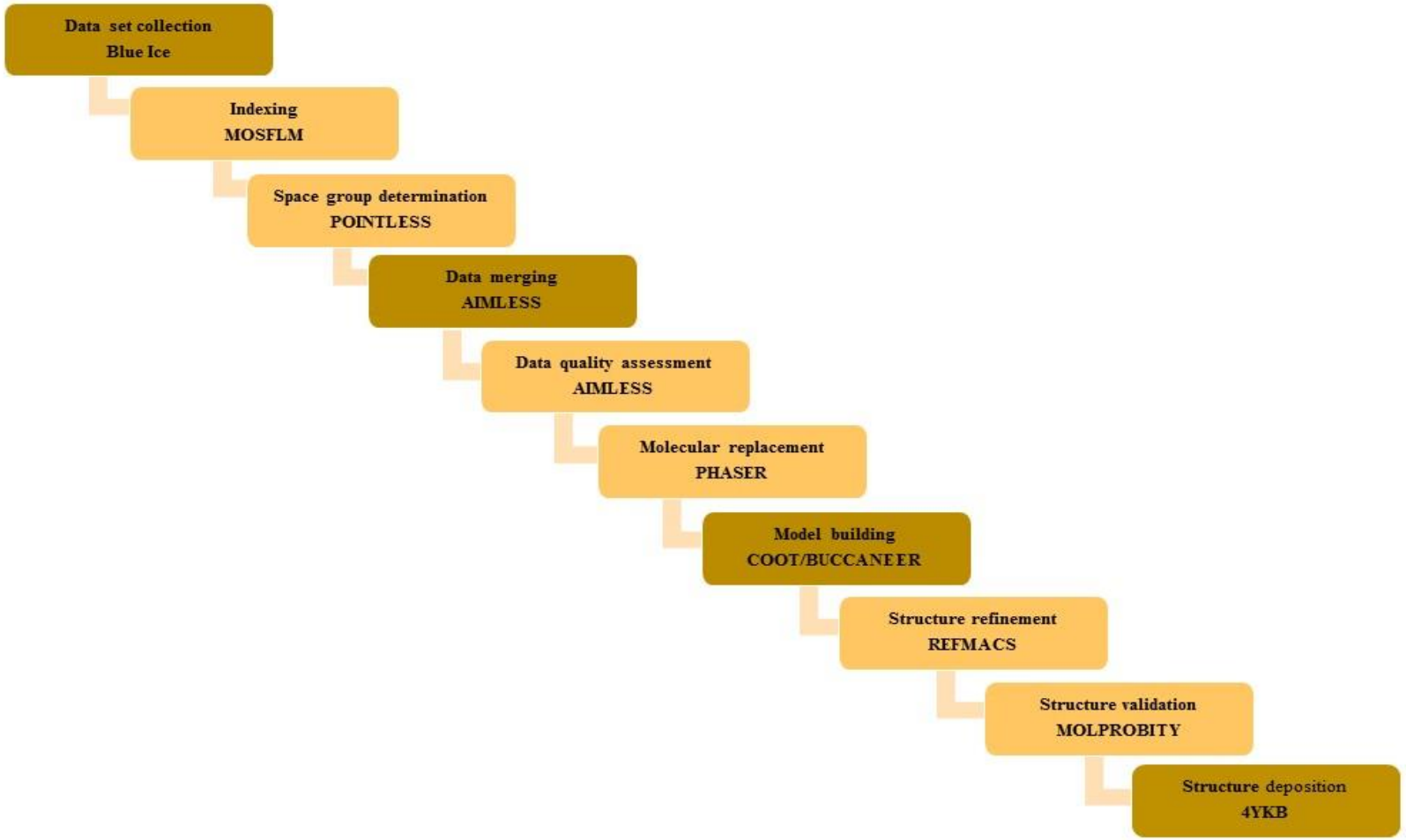
Method	Vapour diffusion, sitting drop
Plate type	24-well (Cryschem plate, Hampton Research)
Temperature (°C)	25
Protein concentration (mg ml^{-1})	35
Buffer composition of protein solution	20 mM Tricine (pH 8.0), 2mM β-mercapoethanol
Composition of reservoir solution	1.0 M ammonium citrate tribasic (pH 7.0), 0.1 M Bis Tris Propane (pH 7.0)
Volume and ratio of drop	2 µl and 1:1(protein:reservoir)
Volume of reservoir (µl)	500

Table 4.5 Diffraction data collection statistics

† As calculated by AIMLESS

Wavelength (Å)	0.9537
Space group	$P3_221$
Unit cell dimensions: a,b,c (Å); α, β, γ (°)	115.0, 115.0, 141.2; 90, 90, 120
Resolution range (Å)	3.50-44.56
Observed reflections	48483
Unique reflections †	13965
Completeness † (%)	99.1 (99.0)
Multiplicity †	3.5 (3.6)
R_{merge} †	0.21 (0.67)
Mean I/δ (I) †	5.2 (1.9)
Wilson B ($Å^2$)	78
CC(1/2)	0.964 (0.434)

Figure 4.8 Structure determination workflow



4.5 Crystallization data processing

The workflow of data processing steps and software tools are summarized in Figure 4.8 and the X-ray crystallographic methods are presented in section 2.7.5. Useful diffraction data was collected from two crystals (A: 44 frames and B: 22 frames in Figure 4.9 from crystallization condition 9 (Table 4.3). Each diffraction image was recorded over a 1° oscillation. All images were indexed and integrated to 3.5 Å with MOSFLM (Leslie and Powell, 2007). The diffraction images of two crystals are shown in (Figure 4.9).

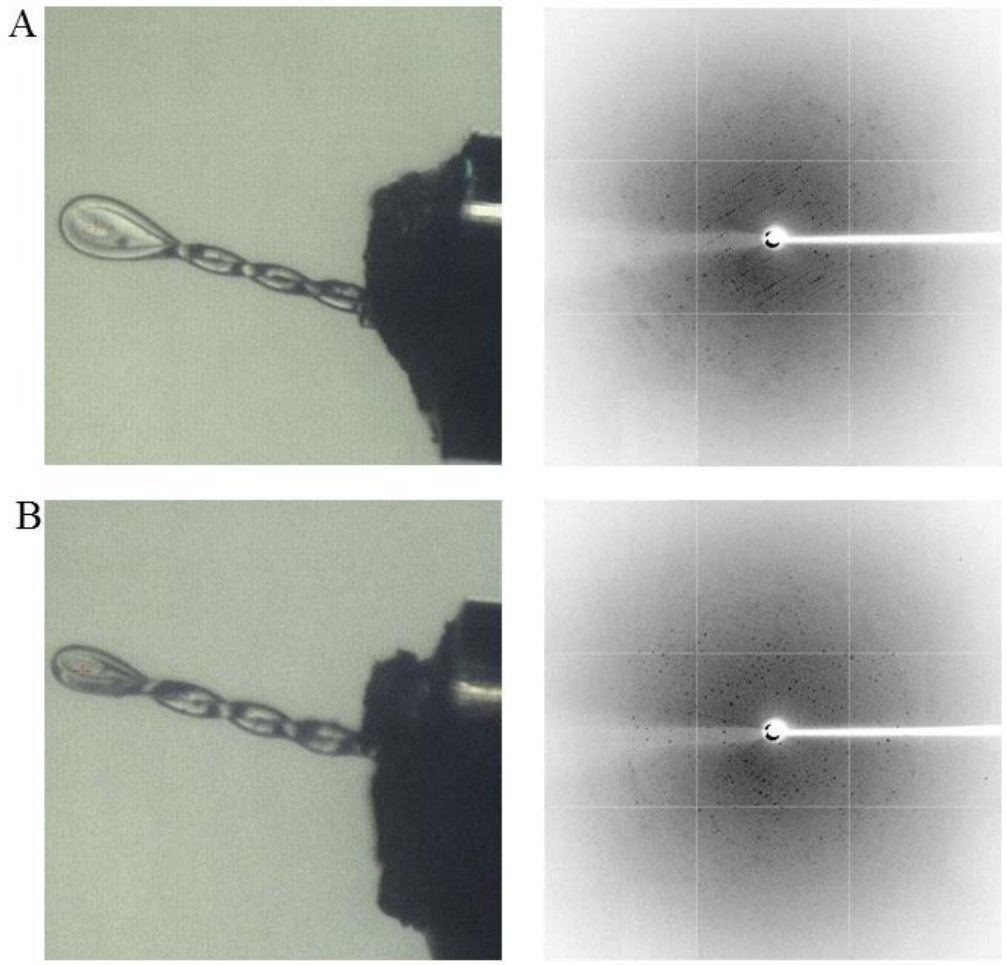


Figure 4.9 GUN4 crystal diffraction on MX2 beamline

Two GUN4 crystals directly from mother liquor diffracted to (<3.5 Å) resolution and relevant diffraction patterns.

The space group ($P3_22_1$) was confirmed with POINTLESS (Evans, 2011). Integrated reflections from these two crystals were scaled and merged with AIMLESS (Evans and Murshudov, 2013), accessed via the CCP4 software interface (Winn et al., 2011). Checks were introduced into data processing stages to ensure that both crystals were predicted to share the same space group and identical (or nearly identical) unit cell parameters. The two data sets were separately processed as outlined above, however they were scaled and merged together to increase data completeness and multiplicity. A portion of the data was assigned to be the R_{free} set to be excluded from refinement and used only for model validation.

4.6 Structure determination

4.6.1 Molecular replacement

The diffraction data from a protein crystal is a reciprocal space representation of the crystal lattice. The crystal symmetry, unit cell size and shape dictate the arrangement of the diffraction reflections during data collection experiment (Hauptman, 1997). The intensity of each diffraction reflection is very important because of proportionality to the square of the structure factor amplitude. The structure factor is a complex number containing information about the amplitude and phase of the X-ray wave. Both amplitude and phase must be known to construct the electron density map (Wilmanns and Weiss, 2005) into which the model will be built. At the diffraction experiment, it is not possible to identify the phase but only the amplitude can be identified, the so-called “the phase problem” (Hauptman, 1997). Initial phase approximation can be achieved by several ways: molecular replacement, multiple isomorphous replacement, phasing via anomalous X-ray scattering. In this study, molecular replacement (MR) was chosen for structure determination. MR is the most commonly used method of solving the phase problem. In this approach, a related 3D structure has to be known with more than 30 % of sequential identity to determine the orientation and position of the molecules within the unit

cell (Taylor, 2003). The phases from the correctly positioned search model can then be used as a starting point for structure refinement. To define the correct orientation and translation of the chosen model in the unit cell of our crystal, the calculation of three rotation and three translation parameters have to be performed (Wilmanns and Weiss, 2005).

Table 4.6 Structural homologues of GUN4

Sequence identity (%)	Source organism	PDB file
40	*Thermosynechococcus elongatus*	1Z3X
35	*Synechocystis*	1Y6I

My GUN4 structure was determined by MR with the program PHASER (McCoy et al., 2007) embedded within CCP4 suite. Among two homologues for GUN4 (Table 4.6) the search model with the highest sequence homology was taken from the *T. elongatus* WT structure (PDB entry 1Z3X). This structure is 40 % identical at the amino acid sequence level over 165 residues. The more divergent first 88 residues of this structure (the N-terminal domain) were excluded from the search model. Molecular replacement initially identified four GUN4 molecules in the asymmetric unit, consistent with the crystal comprising 48 % solvent (as calculated from the Matthews Coefficient and full-length GUN4). However, incomplete crystal packing and inspection of electron density maps suggested that additional unaccounted molecules might in fact be present. The search model was subsequently trimmed further (another 18 residues removed from the N-terminus, and removal of several loops divergent in the *T. elongatus* and *Synechocystis* structures, and side-chains truncated to the Cβ position), which allowed PHASER to fit six non-clashing molecules in the asymmetric unit (solvent content of 53 % for 6 molecules of ~150 residues). Hence, it appeared that the N-terminal domain of the molecules was missing.

4.6.2 Model building and refinement

After solution of the phase problem an initial model can be built. This initial model can be used to refine the phases. Every cycle of this process leads to the improvement of the model; therefore a new model would be applied for further structure refinement. The refinement cycle involves fitting atomic positions of the model and their corresponding B-factors (parameter that reflects the thermal motion of the atom) to the observed electron density map, usually yielding a better set of phases (Taylor, 2003). Sensible stereochemistry, hydrogen bonds and allocation of bond lengths and angles are corresponding quantities which help guide model building and refinement (Wilmanns and Weiss, 2005). The refinement process is performed until the model fails to become improved, as judged by the statistical agreement of spot intensities derived from the model with those measured in the diffraction experiment. The two statistical parameters which are monitored are the R_{work} and R_{free} factors (Table 4.7). Both factors depend on the resolution of the data. For macromolecules R-factors usually vary from 0.6 to 0.2. The R_{free} is a minor subset of the reflections which are never used during refinement, and serve only as an independent indicator of closeness of the model to the data (Brunger, 1992). Restrained B-factor refinement of GUN4 was performed with REFMAC5 (Murshudov et al., 2011), using local non-crystallographic symmetry (NCS) restraints. NCS restraints were employed as there were 6 molecules in the asymmetric unit, all with effectively the same structure. However, there is only a limited amount of data describing them, especially given the very low resolution of the data (i.e. not a very big number of reflections). These restraints effectively penalise segments of individual molecules from "wandering" too far from the average geometry of the 6 molecules, allowing refinement of more sensible models at the end of refinement. Between rounds of refinement, electron density maps and their fit to the model were examined using COOT (Emsley et al., 2010). Amino acid side chains were added if suggested by difference map electron density. Difficult segments of the model were built with sensible geometry using

BUCCANEER (Cowtan, 2006). Structure validation was performed using the MOLPROBITY web server (Chen et al., 2010). Refinement statistics are shown in Table 4.7. The coordinates for the final model have been deposited in the Protein Data Bank (PDB entry 4YKB). All structural refinement and model validation statistics are presented in Table 4.7.

Table 4.7 Structural refinement and model validation

Protein molecules/asu	6
Resolution (Å)	3.5
Atoms modeled/asu	6241
Ramachandran[#]; Favoured/Outliers (%)	92.5/ 0.71
RMSD bond lengths (Å)	0.011
RMSD bond angles (°)	1.46
R/R_{free}	0.26/ 0.31
Residues Chain A (side-chains absent)	151 (40)
Residues Chain B (side-chains absent)	147 (44)
RMSD (Å) superposed on Chain A	0.29 over 146 CA atoms
Residues Chain C (side-chains absent)	146 (39)
RMSD (Å) superposed on Chain A	0.15 over 146 CA atoms
Residues Chain D (side-chains absent)	148 (45)
RMSD (Å) superposed on Chain A	0.14 over 145 CA atoms
Residues Chain E (side-chains absent)	145 (39)
RMSD (Å) superposed on Chain A	0.15 over 145 CA atoms
Residues Chain F (side-chains absent)	141 (44)
RMSD (Å) superposed on Chain A	0.24 over 141 CA atoms
PDB entry	4YKB

4.7 Proteomic analysis of GUN4 and GUN4 crystals

The GUN4 protein (control), and two GUN4 crystal samples were analysed on LC-MS/MS. The chemical conditions of two crystals varied from each other. Crystal 9 was prepared with 1.0 M Ammonium citrate tribasic (pH 7.0), 0.1 M BIS-TRIS propane (pH 7.0), while crystal 13 was in 0.7 M Ammonium tartrate dibasic, 0.1 M BIS-TRIS propane (pH 7.0). The peptide coverage identified for GUN4 protein was used as a reference to compare with the amino acid sequences achieved using mass spectrometer. The sequence coverage for crystal 9 starts from

AA 56 to 238 and for crystals 13 is from AA 56 to 242. The last five AA in all three samples were not detected in MS analysis. The result of the MS analysis indicates the last 18 amino acids from crystal 9 and 22 amino acids from crystal 13 are truncated from the C-terminal domain of the both crystals (Figure 4.10). The details about the percentage of peptide coverage, sequence coverage, redundant peptide counts and accession are mentioned in Table 4.8.

Table 4.8 LC-MS/MS result summary of the GUN4 control, GUN4 crystal 9 and GUN4 crystal 13.

Samples	% Coverage (measured)	% Coverage (corrected)	Sequence Coverage (260 aa)	Unique Peptides	Redundant Peptide Count	Accession	Description
Crystal 9	47	56	56-238 aa	16	86	gi|159485454|	GUN4:p, tetrapyrrole-binding protein [*Chlamydomonas reinhardti*i]
Crystal 13	46	55	56-242 aa	16	91	gi|159485454|	GUN4:p, tetrapyrrole-binding protein [*Chlamydomonas reinhardtii*]
GUN4 Control	74	93	56-255 aa	39	175	gi|159485454|	GUN4:p, tetrapyrrole-binding protein [*Chlamydomonas reinhardtii*]

Figure 4.10 The last 40 amino acids of GUN4 protein (from 220-260) compared with the two GUN4 crystals. There are 22 residues missing from crystal 9 and 18 residues are missing from crystal 13.

4.8 Crystal structure of GUN4 from *Chlamydomonas reinhardtii*

Six GUN4 molecules were found in the crystallographic asymmetric unit (chains A to F). In their N-terminal domain residues (1-47) are missing due to proteolysis, and residues (48-73) are present in the crystal based on peptide mass fingerprinting of individual crystals but they are not resolved in the crystal structure (Figure 4.11). The first resolvable residue number is 81 in wild type sequence in and residue number 73 in Table 4.1, and all of them are missing the last 22-24 residues (last resolvable residues numbering 235-238, in Figure 4.11). The fold of the resolvable ~150 residues modelled are effectively the same in all chains with Root-Mean-Square-Deviation (RMSD) of 0.14-0.29 relative to chain A (Figure 4.12). Due to the low resolution of the data, between 25 and 30 % of the amino acid side chains in the various chains have not been modelled (Table 4.7). Those unmodelled are predominantly on the surface of the domain, whilst those buried in the hydrophobic core of the domain are generally well ordered. The fold is largely α-helical, as expected from the *T. elongatus* and *Synechocystis* structures (PDB entries 1Z3X and 1Y6I). The A-chain superposes with the 1Z3X structure with an RMSD

of 0.97 Å (over 143 CA positions) and with the 1Y6I structure with an RMSD of 1.1 Å (over 131 CA positions) (Figures 4.12 and 4.13).

The GUN4 molecules resolved in the crystal have undergone proteolysis, as shown by proteomic analysis of a number of crystals. The N-terminus of the resolvable domain immediately abuts a missing KKK sequence motif (analogous to the region linking the N- and C-terminal domains in related structures), whilst the C-terminal limit is adjacent to lysine 238 (the last residue modelled in chain A). The C-terminal residues missing in our structure contain sites of phosphorylation in eukaryotes, but are not sequence features found in the cyanobacterial GUN4 molecules. The extended time taken for crystal growth presumably reflects the time taken for a contaminating protease within the crystallization drop to convert sufficient molecules to the doubly truncated form such that crystal growth could occur.

4.9 Insights into PPIX binding and activity of GUN4

Proteolysis within the crystallization experiment removed 25 amino acids from the N-terminal domain based on the peptide sequence coverage from a digest of individual crystals and the last ~20 residues, leaving the bulk of the C-terminal domain intact; the domain involved in binding PPIX and Mg-PPIX. Although the structure is of relatively low resolution, the bulk of the side chains which line the "greasy palm" cleft, thought to be involved in binding the porphyrin ring, can be resolved. When superposed with cyanobacterial GUN4 structures, whilst the α-helical scaffolds are conserved, the α2/α3 and α6/α7 loops previously predicted to undergo movement to accommodate porphyrin binding, display markedly different conformations. This diversity of fold captured by multiple crystal structures supports the notion that these loops undergo conformational rearrangement to accommodate PPIX or Mg-PPIX insertion into the underlying hydrophobic cleft, as observed in the recent porphyrin-bound *Synechocystis* structures (Chen et al., 2015c).

Whilst the overall fold of the C-terminus of our *C. reinhardtii* GUN structure is highly similar to the cyanobacterial structures (Figure 4.13; *Synechocystis* structure coloured light blue and *T. elongatus* structure in dark blue), the biggest differences, and also the biggest differences between the cyanobacterial structures themselves, concern two "loop" segments; the loop linking helices 2 and 3 of the *Synechocystis* structure, and the loop linking helices 6 and 7. The α2/α3 loop is actually a short section of helix and the α6/α7 loop is a beta-turn in the porphyrin bound *Synechocystis* structures (Chen et al., 2015c), whilst the aforementioned α2/α3 and α6/α7 loops help cover the hydrophobic "greasy palm" in the unbound *Synechocystis* structure (Verdecia et al., 2005) (Figure 4.12; α2/α3 and α6/α7 loops coloured pink and orange, respectively). The *T. elongatus* structure and our *C. reinhardtii* structure are thus intermediate between the porphyrin bound and unbound *Synechocystis* structures (Figure 4.14).

The α6/α7 loop covers the greasy palm surface but projects very few residues into the cleft, namely a proline residue (conserved in all three organisms) and a leucine residue (substituted by a threonine residue in the *T. elongatus* structure and in our *C. reinhardtii* structure) (Figures 4.12 and 4.14; proline and threonine residues from our structure drawn as sticks). The porphyrin bound *Synechocystis* structures (Chen et al., 2015c) clearly show that this α6/α7 loop has a propensity for conformational dynamics to form a beta-hairpin making a more structured cleft which enables porphyrin binding.

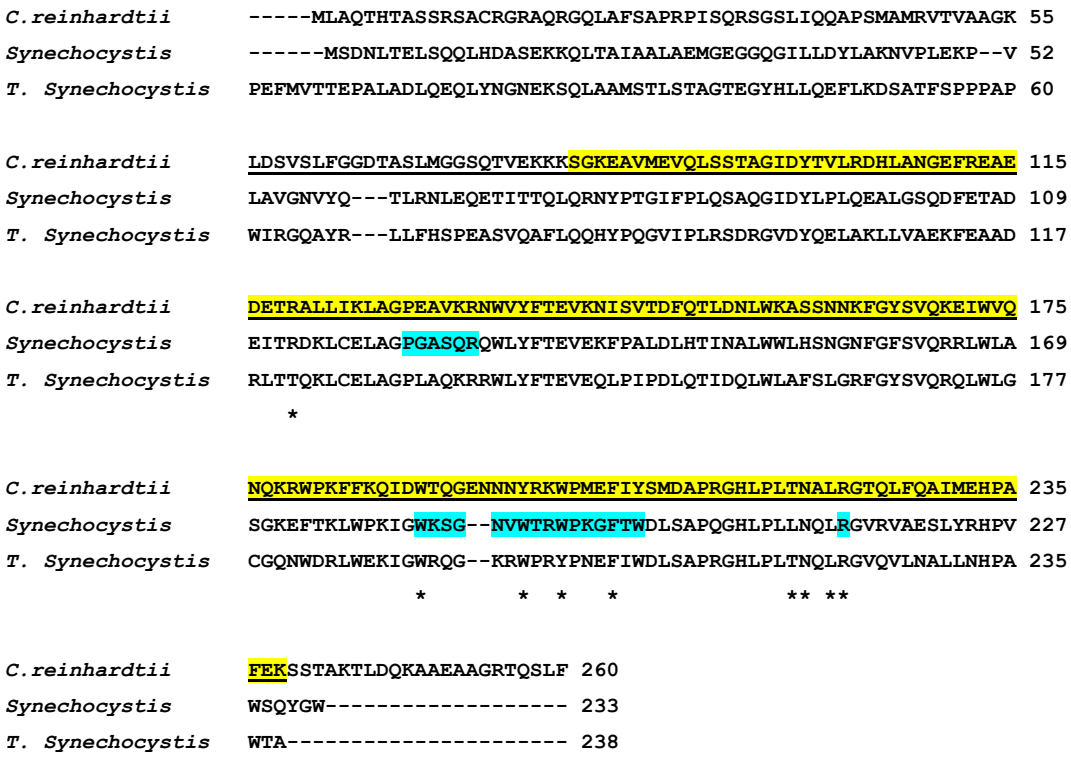


Figure 4.11 Sequence alignment of GUN4 performed with ClustalW2 (Larkin et al., 2007). Highlighted in yellow is the portion of the *C. reinhardtii* GUN4 resolvable in the crystal structure. Underlined region shows peptide coverage from proteomic analysis of crystals. Star-highlighted residues (*) are directly involved in porphyrin binding in the SyGUN4 structures and the blue highlighted residues are involved in secondary structure rearrangements (Chen et al., 2015c).

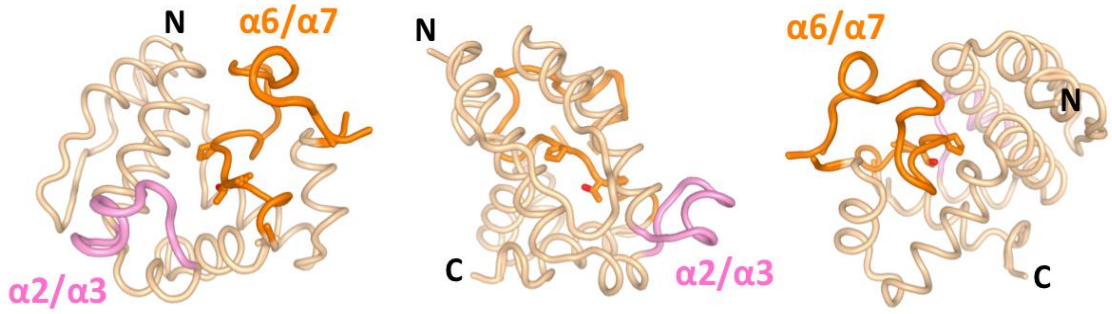


Figure 4.12 Cartoon representation of crystal structure of GUN4.

Three different perspectives are shown. The α2/α3 and α6/α7 "loops" are coloured pink and orange, respectively. The side-chains of proline 216 and threonine 218, within the α6/α7 loop, are drawn as sticks. The N and C termini are also indicated.

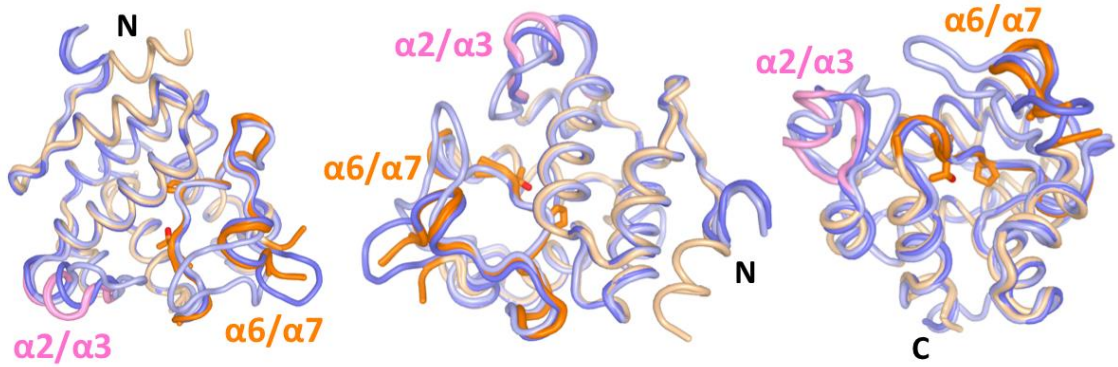


Figure 4.13 Superposition of GUN4 from *C. reinhardtii*

Three different perspectives are shown. The α2/α3 and α6/α7 "loops" are coloured pink and orange, respectively from (Figure 4.12) with cyanobacterial GUN4 crystal structures; *Synechocystis* (light blue) and *T. elongatus* (dark blue). The α2/α3 and α6/α7 loops are indicated, as are the N and C termini. The side-chains of proline 216 and threonine 218 from the *C. reinhardtii* structure are shown as sticks.

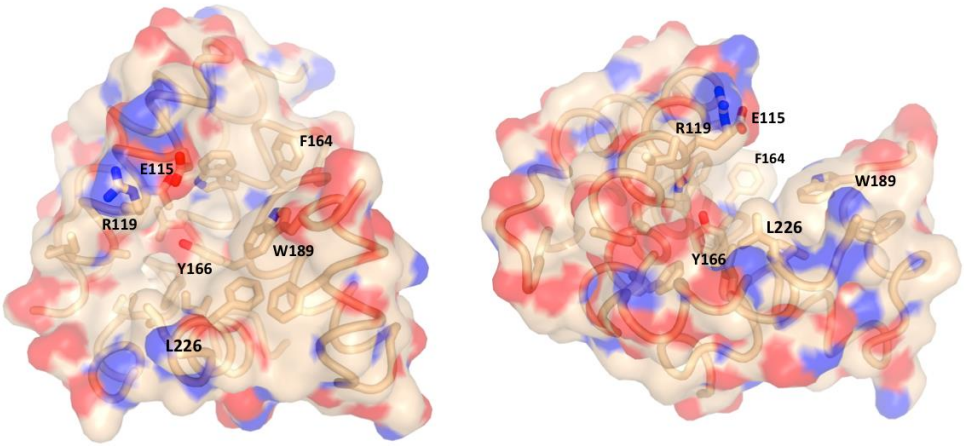


Figure 4.14 Porphyrin binding clef of GUN4 from *C. reinhardtii*

GUN4 has been trimmed of the α2/ α3 and α6/α7 loops to reveal the porphyrin-binding cleft. Both a front-on (left) and a side view (right) are presented. The side chains of numerous residues lining the channel can be viewed under the transparent surface. Tyr166 is positioned at the base and approximately in the middle of the cleft. Other residues are identified to help mark the extremities of the cleft.

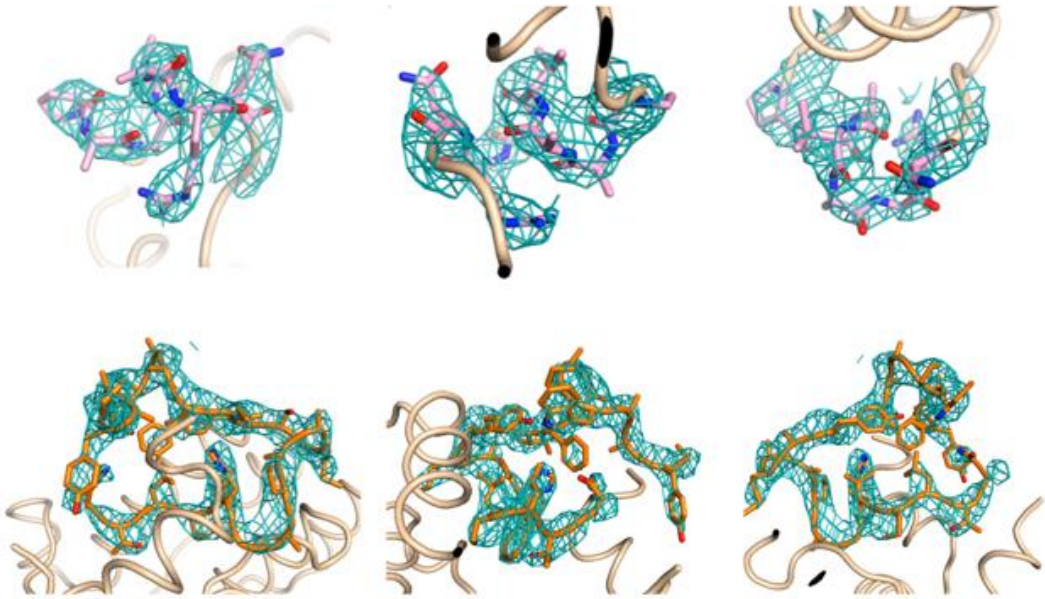


Figure 4.15 Electron density Superposition of α2/α3 and α6/α7 loops

Three different perspectives of α2/α3 loops are shown on top and coloured blue. Three different perspectives of α6/α7 loops are shown on the bottom and coloured green.

4.10 Discussion

The crystal structure of the GUN4 protein from *C. reinhardtii* was solved. Proteolysis within the crystallization experiment removed the N-terminal domain and the last ~20 residues, leaving the bulk of the C-terminal domain intact; this domain is involved in binding PPIX and Mg-PPIX. Although the structure is of relatively low resolution, the bulk of the side chains which line the 'greasy palm' cleft involved in binding the porphyrin ring can be resolved. When superposed with cyanobacterial GUN4 structures, whilst the α-helical scaffolds are conserved, the α2/α3 and α6/α7 loops previously predicted to undergo movement to accommodate porphyrin binding display markedly different conformations. This diversity of fold captured by multiple crystal structures supports the notion that these loops undergo conformational rearrangement to accommodate the insertion of PPIX or Mg-PPIX into the underlying hydrophobic cleft (Figure 4.14), as observed in the recent porphyrin-bound *Synechocystis* structures (Chen et al., 2015c)

Chapter 5

Chapter 5. GUN4 of *Chlamydomonas reinhardtii* enhances singlet oxygen production in the presence of light and protoporphyrin IX (Publication III: prepared for submission)

5.1 Abstract

The Genomes UNcoupled4 (GUN4) is a nuclear-encoded, chloroplast localized, porphyrin binding protein implicated in retrograde signalling between the chloroplast and nucleus although its exact role in this process is still unclear (Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005, Sobotka et al., 2008). Functionally, it enhances Mg-chelatase activity, the enzyme that inserts magnesium into protoporphyrin IX (PPIX) in the chlorophyll biosynthesis pathway, possibly through the removal of Mg-protoporphyrin IX (Mg-PPIX) the product of the reaction. Since GUN4 is present only in organisms that carry out oxygenic photosynthesis and because it binds PPIX and Mg-PPIX it has been suggested that it prevents production of reactive oxygen species by these photosensitizers (Larkin et al., 2003, Adhikari et al., 2011). A recent *in vivo* study shows a *chld-1/GUN4* mutant with elevated PPIX has a light-dependent up-regulation of GUN4 implicating this protein in light-dependent sensing of PPIX, with the suggestion that GUN4 reduces PPIX-generated singlet oxygen, $O_2(a^1\Delta_g)$, and subsequent oxidative damage (Brzezowski et al., 2014). In direct contrast, purified GUN4 and oxidatively damaged ChlH (also known as GUN5) increased the production of PPIX-generated singlet oxygen in the light, by a factor of five and ten, respectively, compared with PPIX alone. Additionally the functional GUN4-PPIX-ChlH complex and GUN4-Mg-PPIX generates $O_2(a^1\Delta_g)$ at a reduced rate compared with GUN4-PPIX. As O_2 $(a^1\Delta_g)$ is a potential plastid to nucleus signal, possibly through second messengers, light dependent O_2 $(a^1\Delta_g)$ generation by GUN4-PPIX is proposed to be part of a signal transduction pathway from the chloroplast to the nucleus, with GUN4-PPIX sensing the availability and flux of PPIX through the chlorophyll

biosynthetic pathway. The biochemical properties of GUN4, as a small and mobile protein, a porphyrin binding protein, a phosphor-protein and now as an $O_2(a^1\Delta_g)$ generator make GUN4 ideal for retrograde signalling.

5.2 Introduction

The consequence of organisms relying on oxygen for metabolism is the continual generation of reactive oxygen species (ROS) including superoxide ($O_2^{\cdot -}$), hydrogen peroxide (H_2O_2), hydroxyl radical ($^{\cdot}OH^-$) and singlet oxygen (O_2 ($a^1\Delta_g$)) (Triantaphylidès and Havaux, 2009, Ogilby, 2010, Shapiguzov et al., 2012). In plants the reaction center of PSII and the antenna system of the chloroplast has been reported as the major source of ROS generation (Krieger-Liszkay, 2005). A low-medium concentration of ROS triggers an acclimation response while a high level initiates cell death (Laloi and Havaux, 2015). In either case a signal is delivered from the chloroplast to the nucleus (retrograde signalling) (Shapiguzov et al., 2012, Laloi and Havaux, 2015). *In vivo* sub-lethal $O_2(a^1\Delta_g)$ levels have been generated by a Arabidopsis *flu* mutant which resulted in protochlorophyllide accumulation (Meskauskiene et al., 2001, op den Camp, 2003) and triggered signalling mechanisms which led to changes in nuclear gene expression (op den Camp, 2003). This effect was suppressed with the double mutant EXECUTER 1 and 2 thus implicating these two proteins in $O_2(a^1\Delta_g)$ dependent plastid to nucleus signalling (Wagner et al., 2004, Lee et al., 2007).

The GUN1-5 (Genome-uncoupled) proteins are also implicated in retrograde signalling since mutant *gun* plants in *Arabidopsis* still allowed transcription of photosynthesis associated nuclear-encoded genes (PhANG) for example the genes for light harvesting complex binding proteins (LHCB) (Susek et al., 1993). Exogenous feeding of Mg-PPIX resulted in reduced LHCB (Strand et al., 2003) therefore this porphyrin was suggested as a retrograde signalling molecule (Strand et al., 2003). However, intracellular concentrations of PPIX and Mg-PPIX

never reach the levels required for this to occur (Mochizuki et al., 2008, Moulin et al., 2008, Woodson et al., 2011).

The *gun4* mutants of *Arabidopsis* and *C. reinhardtii* show a pale-green phenotype indicating reduced chlorophyll biosynthesis (Larkin et al., 2003, Formighieri et al., 2012). *In vitro* it is a regulatory tetrapyrrole-binding regulatory protein involved in enhancing Mg-chelatase activity in the chlorophyll biosynthetic pathway presumably through substrate delivery and possibly release (Larkin et al., 2003, Davison et al., 2005, Verdecia et al., 2005, Sobotka et al., 2008). Mg-chelatase subunit ChlH, also known as GUN5, binds PPIX or receives PPIX from GUN4. The PPIX bound to ChlH has Mg^{2+} inserted during ATP hydrolysis by ChlI-ChlD Mg-chelatase motor complex (Zhou et al., 2012). It is proposed that all porphyrin molecules *in vivo* exist in complex with their protein partners GUN4 and/or ChlH since free porphyrins are photosensitizers with $O_2(a^1\Delta_g)$ being generated in the light under aerobic conditions (DeRosa and Crutchley, 2002, Tripathy et al., 2007). Since GUN4 appears to be exclusively found in oxygenic photosynthetic organisms, it is likely that it plays a direct or indirect role in oxygen-related stress (Larkin et al., 2003, Peter and Grimm, 2009, Formighieri et al., 2012). Additionally a Mg-chelatase deficient strain *chld-1/GUN4* that accumulates PPIX and overexpresses GUN4 implicates GUN4 in retrograde signalling through sensing and binding tetrapyrrole metabolites with the suggestion that it may prevent $O_2(a^1\Delta_g)$ production (Brzezowski et al., 2014). It is generally presumed that GUN4 plays a photo-protective role within the chloroplast, which may protect plants from ROS that are produced by collisions between O_2 and porphyrins under bright light (Larkin et al., 2003, Sobotka et al., 2008). Furthermore, it was suggested that GUN4 participates in Mg-PPIX trafficking or shields PPIX and Mg-PPIX from collisions with O_2 that might yield ROS (Larkin et al., 2003).

The crystal structure of ligand-free GUN4 suggested that its porphyrin-binding domain offered protection from collisions with molecular O_2 (Verdecia et al., 2005). However recent structural

data with PPIX bound GUN4 shows a half-open porphyrin binding pocket which is presumably important for delivery of the substrate to ChlH or ChlM or in retrograde signalling pathways (Chen et al., 2015c) and with the cleft being similar in the *C. reinhardtii* GUN4 structure (Tarahi Tabrizi et al., 2015). The porphyrin bound GUN4 has the bound porphyrin open and exposed and contrasts with a deep hydrophobic cavity which is proposed to bind PPIX in ChlH (Chen et al., 2015b).

5.3 Results

5.3.1 Protein purification

$ChlI_1$, $ChlI_2$, ChlD, ChlH, and GUN4 proteins were purified as described in the methods. The Coomassie Brilliant Blue-stained SDS–PAGE gel indicates that ChlI1, ChlI2 and ChlD were > 90 % pure after HisTrap FF purification and buffer exchange (Figure 5.1), and were used in assays without further purification. Since ChlH and GUN4 were to be used for $O_2(a^1\Delta_g)$ generation experiments, additional purification by gel filtration was performed using Superose^{TM} 6 10/300 GL, and Superdex™ 200 HR 10/30 columns respectively (Figure 3.2). ChlH expressed in *E. coli* produces monomers and oligomers, therefore an optimized method according to a previous study was used to achieve a large proportion of monomers (Müller et al., 2014). Hence, ChlH was separated into the minor oligomeric and major monomeric components eluting at 8 mL and 12.5 mL respectively (Figure 3.2A) (Müller et al., 2014). For GUN4, the previously used method was followed and resulted in a mono-dispersed protein (Figure 3.2B) which allowed its crystallization (Tarahi Tabrizi et al., 2015). GUN4 eluted as a single monomeric peak at ~17 mL on Superdex 200 which was used for all assays (Figure 3.2B).

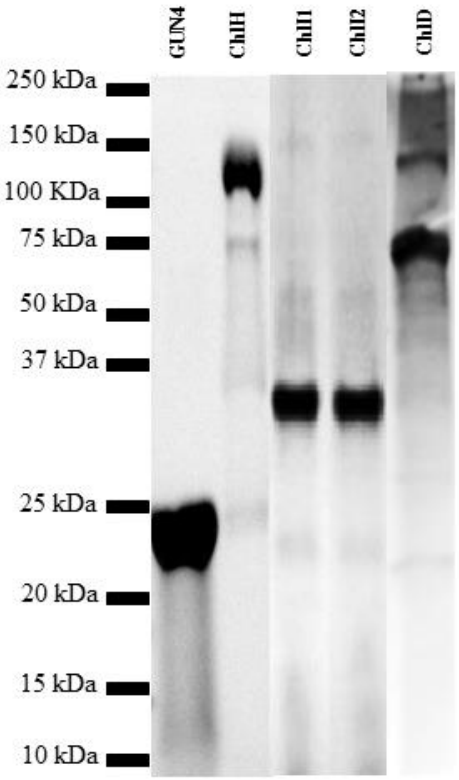


Figure 5.1 SDS-PAGE of purified proteins

Lane1 IMAC and Superdex 200 10/300 purified GUN4, lane 2 IMAC and Superose 6 10/300 purified ChlH monomer, lane 3 ChlI1, lane 4 ChlI2, lane 5 ChlD. The positions of molecular mass markers are indicated at the left.

5.3.2 Mg-chelatase activity measurements

To ensure ChlH and GUN4 proteins from *C. reinhardtii* were functional, the Mg chelatase assay was used (Luo et al., 1999, Sawicki and Willows, 2008). The assay requires the separate preparation of a ChlH-PPIX-GUN4 complex, and a ChlI1-ChlI2-ChlD complex, followed by their unification to start the assay. Final concentrations of 250 nM GUN4, and 250 nM of ChlH with 20 nM ChlD, 400 nM ChlI1, 20 nM ChlI2. The Mg-chelatase activity of 2.6 ± 0.1 nmol/min/nmol ChlD was determined for the monomeric ChlH. The activity for the monomer is 5 fold higher than a previous study utilizing ChlH from *C. reinhardtii* in combination with GUN4 and remaining Mg-chelatase subunits from *Oryza sativa* (Müller et al., 2014).

Oligomeric ChlH from rice and barley has variable activity compared to the monomer, typically very low, but in some preparations up to 25 % (Müller et al., 2014). The same variable results were observed using *C. reinhardtii* oligomeric ChlH.

5.3.3 *In vitro* singlet oxygen production detection using fluorescent sensor reagent (SOSG)

In order to measure $O_2(a^1\Delta_g)$ production of protein-porphyrin complexes in solution two methods were used, SOSG with either PPIX or Mg-PPIX used as the photosensitizer and histidine dependent $O_2(a^1\Delta_g)$ quenching with PPIX. In the first method, to confirm the increase in PPIX-generated $O_2(a^1\Delta_g)$ production by GUN4 and ChlH, the commercially available $O_2(a^1\Delta_g)$ -specific SOSG method were used in the presence of photosensitizers PPIX and Mg-PPIX. The kinetics of the SOSG fluorescence development during exposure of the assay mixture to light was evaluated as an indicator of $O_2(a^1\Delta_g)$ formation. The rate of $O_2(a^1\Delta_g)$ production for free Mg-PPIX is higher than free PPIX (Figure 5.2A), however, the rate of $O_2(a^1\Delta_g)$ production for GUN4 protein using Mg-PPIX does not show any protein concentration dependent increase (Figure 5.2B).

Conversely, there is a large ~5 fold increase in the rate of $O_2(a^1\Delta_g)$ production which is concentration dependent. As expected, addition of NaN_3 decreased the rate of $O_2(a^1\Delta_g)$ as measured with SOSG indicating that it was $O_2(a^1\Delta_g)$ that was being produced. In a separate series of experiments there was an increased rate of PPIX dependent $O_2(a^1\Delta_g)$ production for titrations with increasing concentrations of GUN4, ChlH and GUN4-ChlH but not for BSA (Figure 5.2C). GUN4 had the highest increase of $O_2(a^1\Delta_g)$ production among all tested proteins. ChlH had slightly lower $O_2(a^1\Delta_g)$ production compared with GUN4. The GUN4-ChlH-PPIX complex showed a slower rate of increased $O_2(a^1\Delta_g)$ production but was still much greater than BSA, which showed a marginal $O_2(a^1\Delta_g)$ production (Figure 5.2C).

The SOSG method is very sensitive and requires sub microgram quantities of protein the quantity of fluorescent product and fluorescent yield can be affected by SOSG binding to sensitiser and protein and these experiments required relatively long irradiation times at medium light intensity which may damage the protein during the experiment.

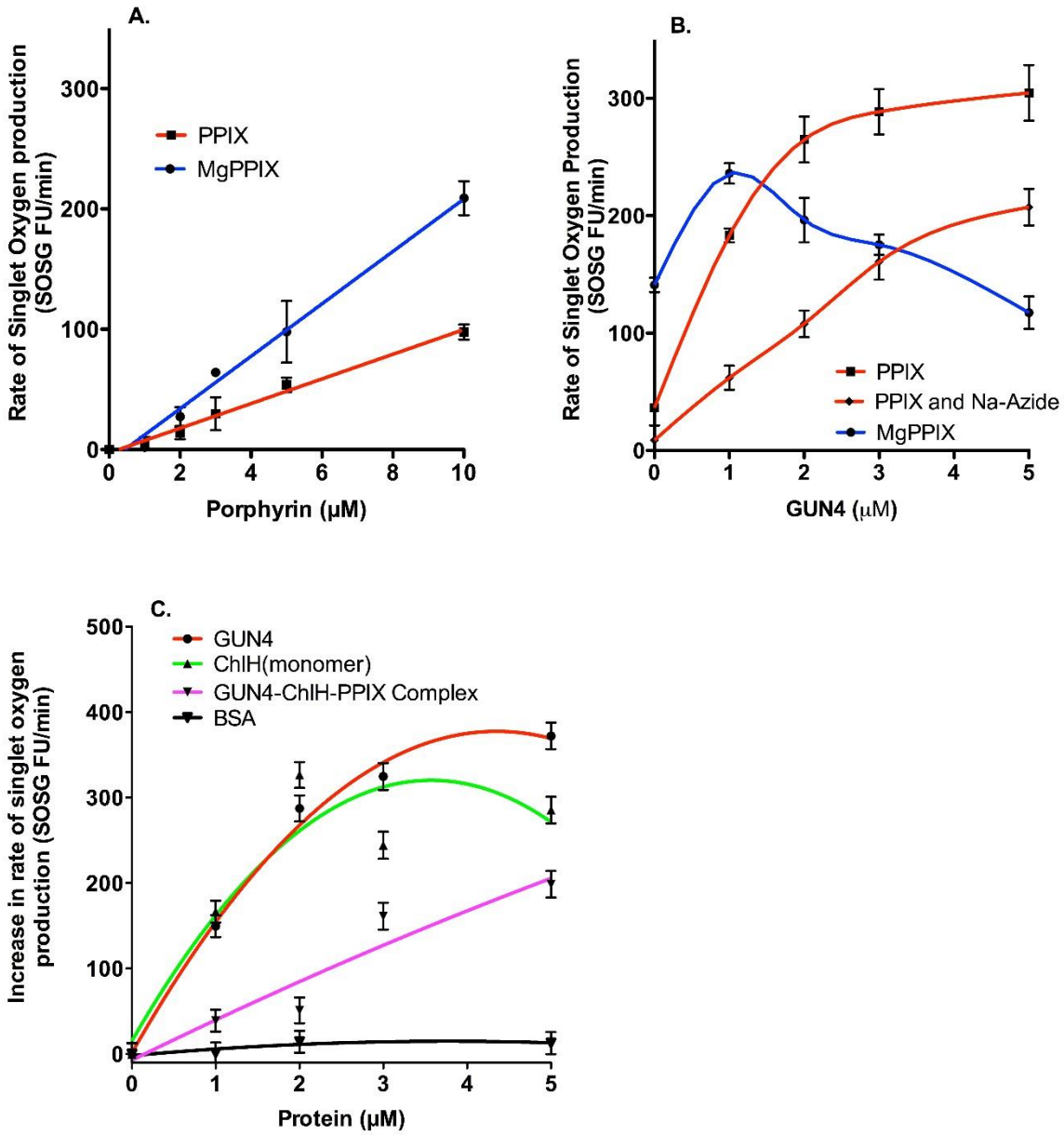


Figure 5.2. O_2 ($a^1\Delta_g$) production measured using SOSG

A. PPIX and Mg-PPIX $O_2(a^1\Delta_g)$ producers.

B. GUN4 stimulated $O_2(a^1\Delta_g)$ production from PPIX but not Mg-PPIX

Mg-PPIX or PPIX (5 μM) titration with increasing GUN4. Control assay showing quenching using 1 mM NaN_3.

C. Increase in rates of $O_2(a^1\Delta_g)$ production using GUN4-PPIX, ChlH-PPIX, GUN4-ChlH-PPIX complex, and BSA-PPIX proteins.

5.3.4 *In vitro* singlet oxygen production measurement using histidine mediated uptake

To verify the results of SOSG method an alternative method was used. In this method histidine reacts with $O_2(a^1\Delta_g)$ to produce short-lived peroxide species, which in turn destabilizes to form oxidized histidine resulting in concentration-dependent $O_2(a^1\Delta_g)$ removal from solution that can be measured using an oxygen electrode (Verlhac et al., 1984, Agon et al., 2006, Méndez-Hurtado et al., 2012, Rehman et al., 2013). Using this method the rate of $O_2(a^1\Delta_g)$ removal using PPIX ($O_2(a^1\Delta_g)$ sensitizer) was saturable with titration of PPIX concentrations (Figure 5.4) as expected under light limiting conditions. The $O_2(a^1\Delta_g)$ production rates indicate that the $O_2(a^1\Delta_g)$ production is dependent on PPIX concentration and 2.5 µM PPIX was a suitable concentration to use as it is close to concentrations found for PPIX in vivo and it is above the K_d values for porphyrin binding to the proteins tested. Using this method GUN4 protein significantly enhances the rate of $O_2(a^1\Delta_g)$ production from PPIX (Figure 5.3A) in a concentration-dependent manner and this was ~ 2.5 fold greater than BSA-PPIX. BSA was used as a control as it is known to bind PPIX (Brancaleon and Moseley, 2002) and the titration saturation observed (Figure 5.3A) is consistent with reported K_d values for PPIX binding to GUN4 and BSA. A comparison of the direct effect of PPIX bound to GUN4, and ChlH proteins in $O_2(a^1\Delta_g)$ production was tested by reconstituting ChlH and GUN4 with PPIX. The $O_2(a^1\Delta_g)$ production level of ChlH-PPIX monomer and oligomer, and GUN-PPIX proteins is increased significantly in a concentration dependent manner (Figure 5.3B). The rate of $O_2(a^1\Delta_g)$ production is enhanced in the presence of protein as found using the SOSG method. An oxidatively damaged form of ChlH-PPIX (oligomer) produces $O_2(a^1\Delta_g)$ at 2 times the rate of GUN4-PPIX, while ChlH (monomer) was half the GUN4-PPIX $O_2(a^1\Delta_g)$ production rate (Table 5.1). A comparison of these $O_2(a^1\Delta_g)$ production rates with PSII shows that these porphyrin complexes can produce $O_2(a^1\Delta_g)$ at much greater rates which has implications for signalling.

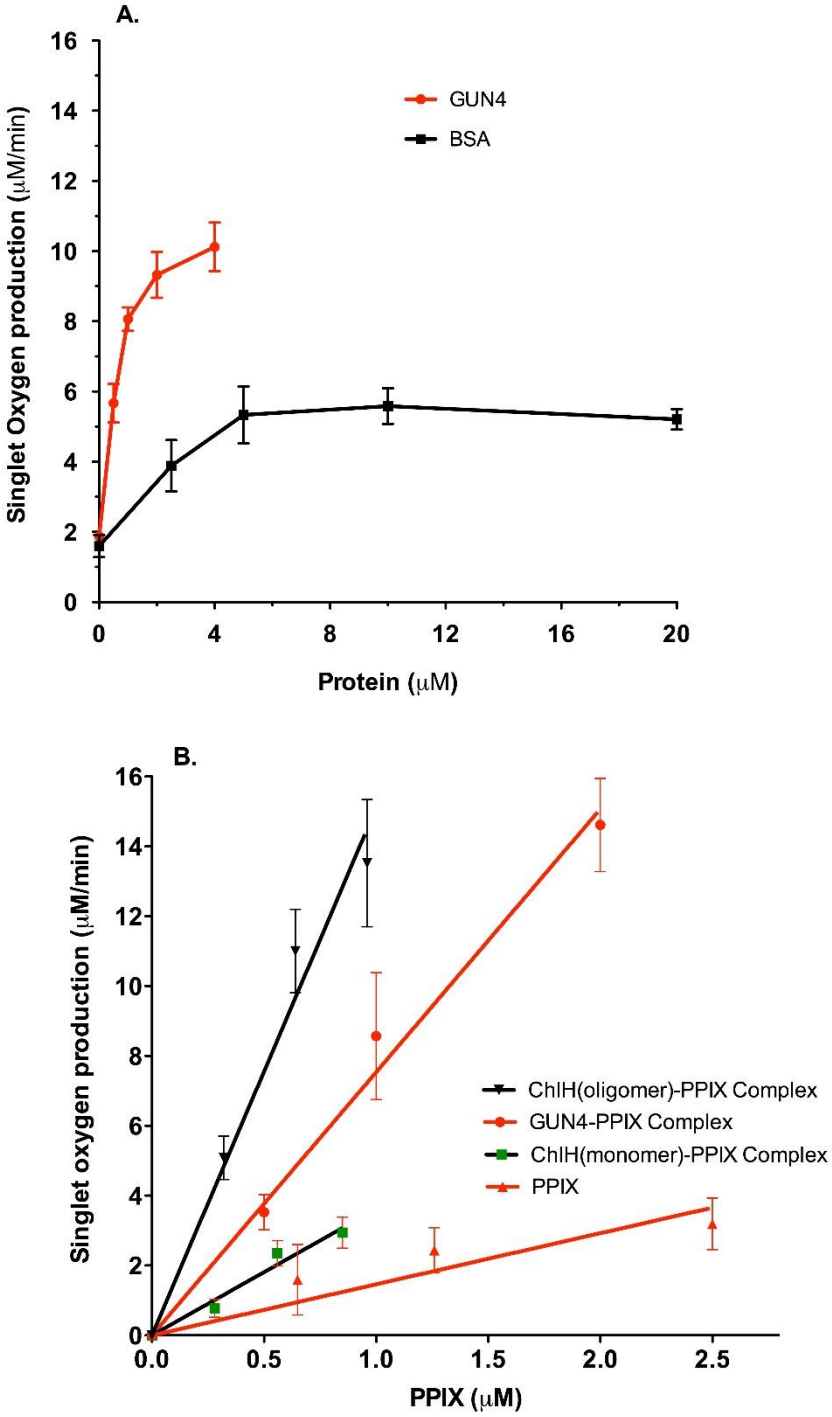


Figure 5.3 Quantification of rates of $O_2(a^1\Delta_g)$ production using histidine quenching.
A. Titration of GUN4 and BSA proteins with 2.5 μM PPIX to produce $O_2(a^1\Delta_g)$.
B. Measurement of the concentration dependence of PPIX-derived $O_2(a^1\Delta_g)$ from free PPIX, GUN4-PPIX complex, ChlH (monomer)-PPIX complex and ChlH (oligomer)-PPIX complex. The molar ratio of PPIX combined with protein was 1:1.

Table 5.1 Singlet oxygen production rate (µM singlet oxygen/min/ µM PPIX) of free PPIX and Protein complexes derived from Figure 5.3B

The data presented in this table are representative of the all slope of the lines from Figure 5.4B. PSII rate from (Telfer et al., 1994) is adjusted to light intensity of 132 umol photons /s/m^2 for comparison with our results.

*** Data calculated from supplied in this chapter and normalised for light intensity for comparison.**

Protein complexes and PPIX	µmol singlet oxygen/min/µM tetrapyrrole
ChlH(oligomer)-PPIX Complex	15.1 ± 0.8
ChlH(monomer)-PPIX Complex	3.6 ± 0.2
GUN4-PPIX Complex	7.5 ± 0.3
PPIX	1.5 ± 0.2
PS II-chlorophyll (Telfer et al., 1994)*	1.85

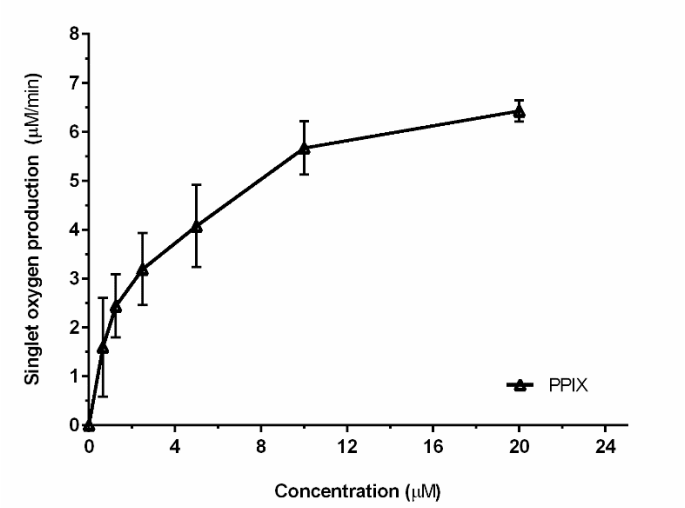


Figure 5.4 PPIX generated singlet oxygen production measurement by histidine quenching. The rates of $O_2(a^1\Delta_g)$ production using variable concentrations of PPIX.

5.3.5 Singlet oxygen production rate of BchH-PPIX complex

Details of *bchH* cloning into pET15b was outlined previously (Willows and Beale, 1998). Expression of BchH protein was performed separately as a fusion with an N-terminal poly-His tag, and monomeric BchH from *R. capsulatus* was purified as previously described (Sawicki and Willows, 2008). The $O_2(a^1\Delta_g)$ production level of BchH-PPIX complex is increased significantly in a concentration dependent way (Figure 5.5). The rate of $O_2(a^1\Delta_g)$ production is significantly higher in ChlH-PPIX (oligomer) and GUN4-PPIX compared with BchH-PPIX and ChlH (monomer).

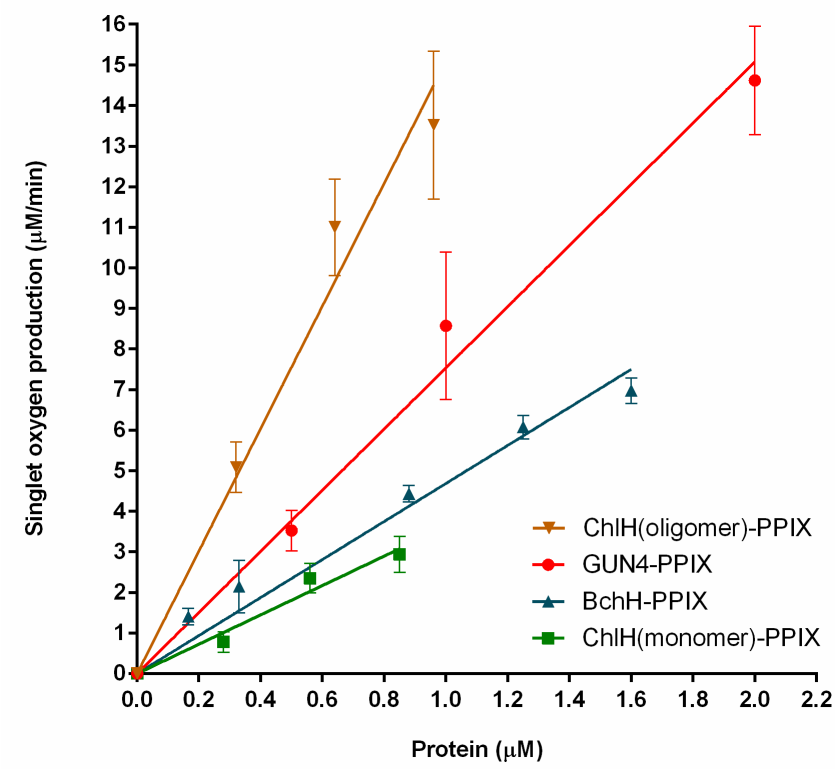


Figure 5.5 Rate of PPIX-derived singlet oxygen production of bchH-PPIX complex GUN4-PPIX complex, ChlH (monomer)-PPIX complex and ChlH (oligomer)-PPIX complex using a clark type electrode.

Rates of $O_2(a^1\Delta_g)$ production of GUN4-PPIX complex, BchH-PPIX complex, and ChlH (monomer and oligomer) complexes with PPIX. All proteins were pre-combined to 2.5 μM PPIX. The ratio of PPIX combined with protein was 1:1.

5.4 Discussion

GUN4 is involved in modulating ALA biosynthesis in response to chlorophyll biosynthesis pathway intermediates and sensing PPIX and Mg-PPIX (Peter and Grimm, 2009, Brzezowski et al., 2014), however, the mechanism is not related to any significant measurable changes in the intracellular levels of these tetrapyrroles (Mochizuki et al., 2008, Moulin et al., 2008, Woodson et al., 2011). As GUN4 accentuates PPIX-generated $O_2(a^1\Delta_g)$ rather than attenuating it as suggested (Mochizuki et al., 2008, Moulin et al., 2008, Schlicke et al., 2014) the increased production of $O_2(a^1\Delta_g)$ is an important factor that needs to be considered in the context of retrograde signalling. As such a new framework for the role of GUN4 and ChlH (GUN5) in plastid to nucleus communication were postulated. It has been shown that GUN4 and oxidatively damaged GUN5 (ChlH) in complex with PPIX are $O_2(a^1\Delta_g)$ generators. Given that GUN4 and ChlH have been localized to both membrane and soluble components within chloroplasts (Adhikari et al., 2009, Adhikari et al., 2011) and that they have a function in retrograde signalling, a new $O_2(a^1\Delta_g)$ initiated model of retrograde signalling involving GUN4 and ChlH is postulated (Figure 5.6). In this model a $O_2(a^1\Delta_g)$ signal is generated from GUN4-PPIX, and possibly oxidatively damaged ChlH-PPIX, when these molecules interact directly with a $O_2(a^1\Delta_g)$ sensing system such as the membrane associated EXECUTER 1 and 2 system. Due to the reactivity and short half-life of $O_2(a^1\Delta_g)$, the $O_2(a^1\Delta_g)$ producer must be in close proximity (10-20nm) to the $O_2(a^1\Delta_g)$ sensor (Kuimova et al., 2009, Ogilby, 2010, Bosio et al., 2013) The proximity dependence is limited by the lifetime of $O_2(a^1\Delta_g)$ which is ~1.6 µsec *in vivo*, with 3 times the lifetime corresponding to a rms radial diffusion distance of less than 20 nm (Kuimova et al., 2009, Ogilby, 2010, Bosio et al., 2013). Thus the model must take into account the relative location of the $O_2(a^1\Delta_g)$ producer to the $O_2(a^1\Delta_g)$ sensor and this must be within 20 nm. This model for $O_2(a^1\Delta_g)$ production and sensing does not require high concentrations of PPIX, but it does require the ability of GUN4-PPIX to change its location

within the chloroplasts to interact with the $O_2(a^1\Delta_g)$ sensor possibly the EXECUTOR 1 or 2 $O_2(a^1\Delta_g)$ sensing system. The nature of the signalling from EXECUTOR 1 and 2 is unknown but several candidates have been identified (Laloi and Havaux, 2015, Singh et al., 2015). This signalling would trigger differential regulation of target genes that include PPIX/chlorophyll biosynthetic enzymes, PhANG genes and other stress-related genes.

The movement of GUN4 and ChlH required by this model has been shown to occur *in organello*, as GUN4 has been shown to relocate from the soluble phase to the membrane (Adhikari et al., 2009, Adhikari et al., 2011) under various conditions. In addition the C-terminal of GUN4 is phosphorylated in plants (Reiland et al., 2009) and this phosphorylation may also influence GUN4's location within the chloroplast to sense the flux of chlorophyll intermediates. Importantly, this model also explains much of the *in vivo* data for mutant and artificially perturbed systems in which high endogenous concentrations of tetrapyrrole intermediates accumulate. The effect of tetrapyrrole accumulation would depend on the location and quantity of the tetrapyrrole and the quantity and location of tetrapyrrole binding proteins as both could potentially disrupt or complicate signalling by producing large quantities of $O_2(a^1\Delta_g)$ as well as other ROS species.

In conclusion our data assigns a role of GUN4-PPIX and oxidatively damaged ChlH-PPIX protein complexes as generators of $O_2(a^1\Delta_g)$ signals in the presence of light. It has been suggested that GUN4 is involved in sensing the flux of chlorophyll synthesis via generation of $O_2(a^1\Delta_g)$ and that this signal is transmitted to the nucleus, possibly via EXECUTOR 1 or 2, to alter nuclear gene expression. This leads to further questions to determine the impact of this signal *in vivo* in chlorophyll biosynthesis and in chloroplast biogenesis. It could be anticipated that the $O_2(a^1\Delta_g)$ signal must be either dealt with specifically through dedicated signalling mechanisms or reacted with particular antioxidant molecules to produce a secondary signalling molecule. This second possibility is particularly interesting potential secondary plastid to

nucleus signals have been identified but the source of the ROS which generates them have not been identified (Szechyńska-Hebda and Karpiński, 2013, Singh et al., 2015).

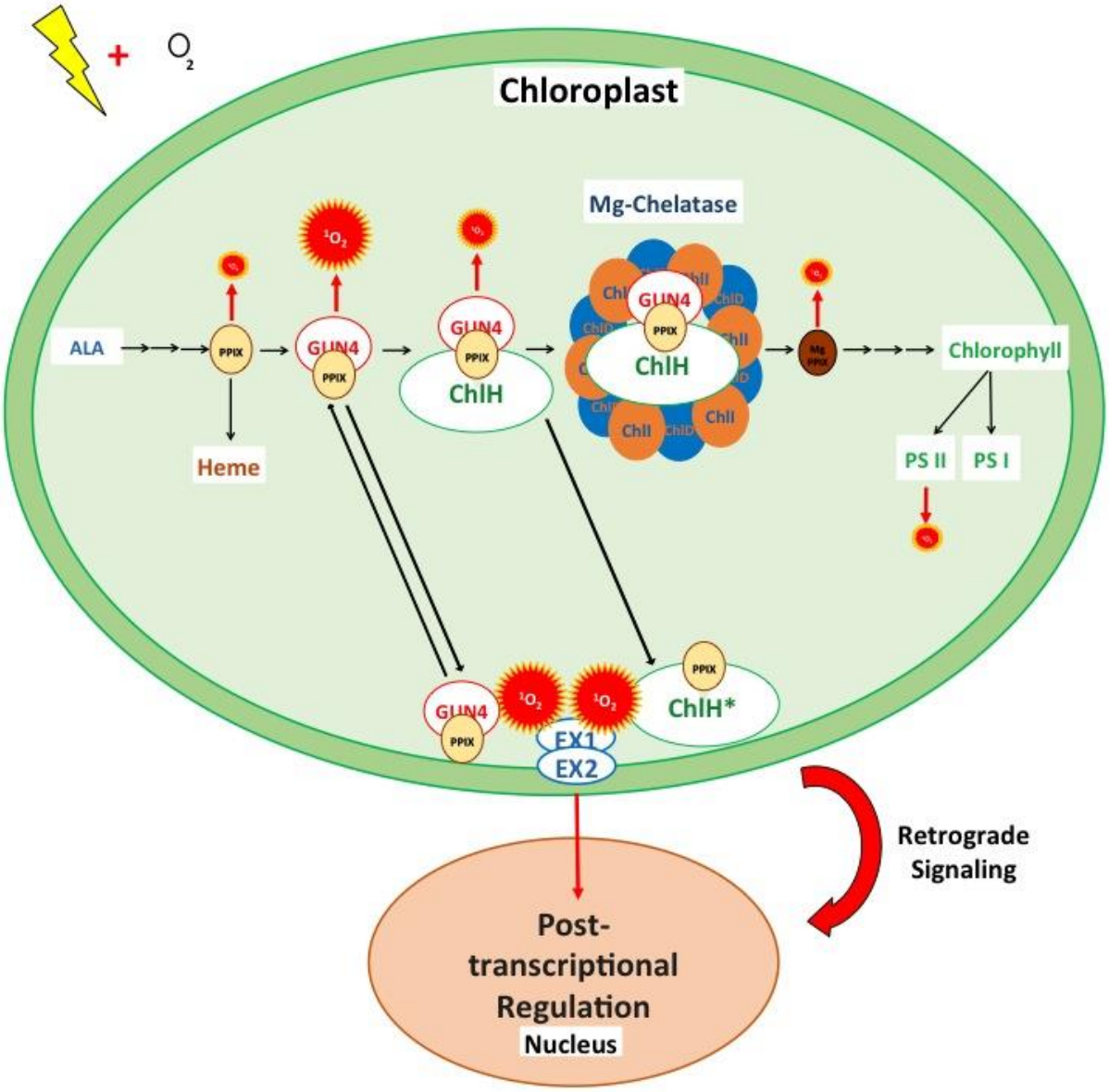


Figure 5.6 Model for function of GUN4 in $O_2(a^1\Delta_g)$ production and chloroplast to nucleus signalling.

Schematic diagram of the interactions between the chloroplast and nucleus. Chlorophyll biosynthetic intermediates like PPIX, Mg-PPIX act as sensitisers with molecular oxygen in the presence of light to form $O_2(a^1\Delta_g)$. It has been shown that GUN4 moves to chloroplast membrane after binding to PPIX (Adhikari et al., 2009). EXECUTER 1 and 2 proteins sense the $O_2(a^1\Delta_g)$ produced by the GUN4-PPIX which is in close proximity or bound to the Ex1 or Ex2 which sends the signal to nucleus through an unknown mechanism. The black arrows show trafficking of this GUN4-PPIX complex and movement of damaged ChlH-PPIX complex to the chloroplast membrane. The size of the red star indicates the relative production rate of $O_2(a^1\Delta_g)$ from different molecules in chloroplast and also its likely diffusion distance before decay.

Chapter 6

Chapter 6 General Discussion

6.1 Overview

The work presented in this thesis focused on combining the structural and functional studies of the Mg-chelation step in the chlorophyll biosynthetic pathway with a spotlight on the ChlH and GUN4 proteins. The SAXS structure of the GUN4 and ChlH proteins from *C. reinhardtii* and complex formation of these two along with PPIX have been discussed in chapter 3. The crystal structure of GUN4 from *C. reinhardtii* has been identified and described in chapter 4. Also, a crucial function of GUN4 and ChlH in $O_2(a^1\Delta_g)$ production with implications for chloroplast to nucleus retrograde signalling have been identified and explained in chapter 5. The key findings from this thesis are compiled into a cohesive discussion, and summarized in a general conclusion. Also, suggestions for future directions will be described in this chapter.

6.1.1 Structure of GUN4 and the PPIX binding site (ChlH-PPIX complex) of Mg-chelatase

Structural information of individual subunits and selected complexes of Mg-chelatase and GUN4 has greatly helped in visualizing the intricate nature of this reaction mechanism. The crystal structures of BchI (Fodje et al., 2001), GUN4 (Davison et al., 2005, Verdecia et al., 2005, Chen et al., 2015c), and ChlH (Chen et al., 2015b) are now known, as well as ChlI-ChlD motor complex.

As a result of this project, four protein subunits of Mg-chelatase (ChlI1, ChlI2, ChlD, ChlH) and GUN4 from *C. reinhardtii* were successfully heterologously expressed and purified in *E. coli* using a combination of methods. GUN4, ChlH and GUN4-ChlH-PPIX purified proteins were structurally observed by SAXS and GUN4 was additionally solved by X-ray crystallography and the details of the results is outline in chapter 3 and 4.

6.1.2 SAXS analysis of GUN4, ChlH and GUN4-ChlH-PPIX complex

This project aimed to determine the low resolution solution structure of GUN4, ChlH, and GUN4-ChlH-PPIX complex from *C. reinhardtii* at ~20 Å, by using SAXS. It was identified that chloroplastic GUN4 has a more elongated structure, due to an additional 9-10 amino acids at C-terminal domain at the end of the main molecular axis. This extension at the C-terminal domain is phosphorylated and required for Mg-chelatase activity, also it is involved in binding PPIX and Mg-PPIX. In addition, the N-terminal 4 helix bundle found in cyanobacterial GUN4 proteins is absent from chloroplastic GUN4 (Zhou et al., 2012). Comparison of the GUN4 crystal structure from cyanobacteria (Davison et al., 2005, Verdecia et al., 2005) with our SAXS structure of the eukaryotic GUN4 provides a model for the structure of the extra C-terminal domain of GUN4 not present in structurally determined homologues (chapter 3).

The ChlH SAXS structure is in a broad structural similarity compared with crystal structure of full length ChlH protein from *Synechocysti*s (Chen et al., 2015b). Also, structural dimensions of our ChlH SAXS structure is very similar to the previously reported cyanobacterial ChlH SAXS structure (Qian et al., 2012, Adams et al., 2014b). However, no structural information has been reported for the GUN4-ChlH-PPIX complex previously. Therefore, modelling the SAXS structure of the GUN4-ChlH-PPIX complex, provided a framework for further investigation of the structure assembly and kinetic studies of this complex. Furthermore, there is no binding affinity information for the complex formation, hence the approach in this thesis to determine the structures of these proteins represents a route towards describing the nature of complex formation. This suggest that hollow cage shape structure of ChlH might provide an open space to hold the GUN4-PPIX inside.

6.1.3 Crystal structure of GUN4

The experimental stages that have always been recognized as the greatest bottlenecks in protein structure studies are achieving sufficient level of protein concentration and protein crystallization. This thesis aimed to determine the X-ray crystal structure of GUN4 protein from *C. reinhardtii*. I was successful in achieving a high level of soluble GUN4 protein (35 $mg.ml^{-1}$) for crystallization, and crystallization of GUN4 protein required a long incubation time (308 days). This allowed us to solve and refine the GUN4 crystal structure and deposit its coordinates in the PDB (4YKB; 3.5 Å) (chapter 4). However, the structure was doubly truncated due to the long term crystallization process, with proteolysis removing the N-terminal domain (first 88 residues) and the last ~ 20 residues from C-terminal domain (paper II). Based on sequence analysis of GUN4 from *Synechocystis*, *T. elangatus*, *C. reinhardtii*, and *A. thaliana*, the N-terminal region of the cyanobacterial GUN4 protein is not highly conserved (chapter 4). The significance of the variation in the N-terminus is not clear, although the structure of all these proteins appears to be helical, which suggests that GUN4 might have multiple roles in Chl biosynthesis. However, sequence similarity of the C-terminal domain is high and suggests a conservation of structure and function. The C-terminus in both known cyanobacterial crystal structures of GUN4 contains 8 α-helices (Davison et al., 2005, Verdecia et al., 2005).

Recently, GUN4 has been crystallized with Deuteroporphyrin IX, and Mg-Deuteroporphyrin IX from *Synechocystis* (Chen et al., 2015c). In these GUN4 structures the C-terminal domain also contains eight α-helices which are connected to a smaller N-terminal domain comprising five α-helices. Using this information, the position of residues contributing to PPIX binding pocket (greasy palm) has been assigned to the *C. reinhardtii* GUN4 structure in the long loop between helices α6 and α7 (chapter 4). Structural comparison of the PPIX bound and unbound GUN4 from *Synechocystis* showed that two regions between helices α2 and α3, and the loop α6/α7 undergo significant conformational changes (Chen et al., 2015c).

Our eukaryotic GUN4 crystal structure from *C. reinhardtii* is also largely α helical, with a highly similar C-terminus domain to cyanobacterial GUN4 structures. However, the main difference between our *C. reinhardtii* GUN4 crystal structure with *Synechocystis, and T. elangatus* GUN4 structures, and also the largest differences between the cyanobacterial structures themselves, is the loop linking helices 2 and 3, and the loop linking helices 6 and 7. Among the residues in α6/α7 loop a proline residue is conserved in all three organisms but a leucine residue is substituted by a threonine residue in *T. elongatus* and in the *C. reinhardtii* structure. Thus, these findings suggest that these loops undergo conformational rearrangement to hold the insertion of PPIX and/or Mg-PPIX (paper II). Finally, the use of PPIX analogues like (platinum-PPIX) might facilitate the crystallization process towards the success of full length crystal structure of GUN4 from *C. reinhardtii*.

6.1.4 GUN4 in complex with PPIX is a singlet oxygen generator

The most surprising outcome of this thesis is identifying a novel role of $O_2(a^1\Delta_g)$ generation of GUN4 in complex with its natural porphyrin substrate, PPIX. This finding is directly opposite to current information implying that GUN4 protects cells from ROS damage. This result sheds new light on a function of GUN4 as a candidate for initiating communication between chloroplast and nucleus via $O_2(a^1\Delta_g)$. This finding will significantly affect our understanding of the role that biosynthetic intermediates play in the context of normal cellular development (chapter 5).

In plants and algae the chloroplast is the major source of ROS generation in photosynthetic oxygenic organisms. Since GUN4 only found in oxygenic photosynthetic organisms, it is likely that GUN4 plays a direct or indirect role in oxygen-related stress (Peter and Grimm, 2009, Formighieri et al., 2012). GUN4 has been suggested as a key regulator of post-translation mechanism of chlorophyll biosynthesis pathway and sensing accumulated PPIX and Mg-PPIX

levels (Peter and Grimm, 2009, Brzezowski et al., 2014) Our *in vitro* findings are the opposite of the current studies based on *in vivo* studies in which GUN4 is suggested to prevent $O_2(a^1\Delta_g)$ production (chapter 5). It has been found that GUN4 of *C. reinhardtii* increases $O_2(a^1\Delta_g)$ production in the presence of light and PPIX *in vitro*. Additionally, PPIX Binding to ChlH and GUN4 individually increased the rate of $O_2(a^1\Delta_g)$ production in the presence of light (paper III).

Identification of monomeric and oligomeric forms of ChlH protein allowed comparison of each structural form of this protein in $O_2(a^1\Delta_g)$ production. Oligomeric ChlH represents an oxidatively damaged form, corresponding to 25 % reduced Mg-chelatase activity that has been identified during overexpression in *E. coli* (Paper I). I obtained 15.2 µM (singlet oxygen/min/ µM PPIX) of ChlH oligomer-PPIX complex as the highest $O_2(a^1\Delta_g)$ rate among all the other tested protein complexes (Figure 6.1). If physiologically relevant, oligomerization of oxidatively damaged ChlH might have a functional role in the chloroplast, possibly as a signal of oxidative stress. This may be an *in vivo* $O_2(a^1\Delta_g)$ stress signal to indicate there is a sub-optimal Mg-chelatase complex or this may explain how Mg-chelatase is turned-over during periods when it is not required. However, further analysis is required to determine if the oligomeric form of ChlH is physiologically relevant or simply an artefact during overexpression in *E. coli* as originally observed (paper I), as well as, identifying the formation, and role of monomeric ChlH *in vivo*.

A high level of GUN4-PPIX *in vivo* is expected to be in complex with ChlH which is required for optimal Mg-chelatase activity and this complex exists in the membrane (Adhikari et al., 2009). The results showed that formation of GUN4-ChlH-PPIX complex generates $O_2(a^1\Delta_g)$ at a reduced rate compared with GUN4-PPIX, which suggests that PPIX binding is stronger when both GUN4 and ChlH are in complex with PPIX. Thus, PPIX is less accessible to interact with oxygen and light. It is suggested that the movement of GUN4-PPIX to the membrane might be

a reason for its interaction with the $O_2(a^1\Delta_g)$ receptors (EXECUTOR 1 or 2). The mechanism of signalling from EXECUTOR 1 and 2 to the nucleus is unknown (Uberegui et al., 2015), and $O_2(a^1\Delta_g)$ generation by GUN4-PPIX may be part of a signal transduction pathway from the chloroplast to the nucleus (chapter 5). However, this signalling would trigger differential regulation of target genes that has been anticipated to include PPIX/chlorophyll biosynthetic enzymes, PhANG genes and other stress-related genes.

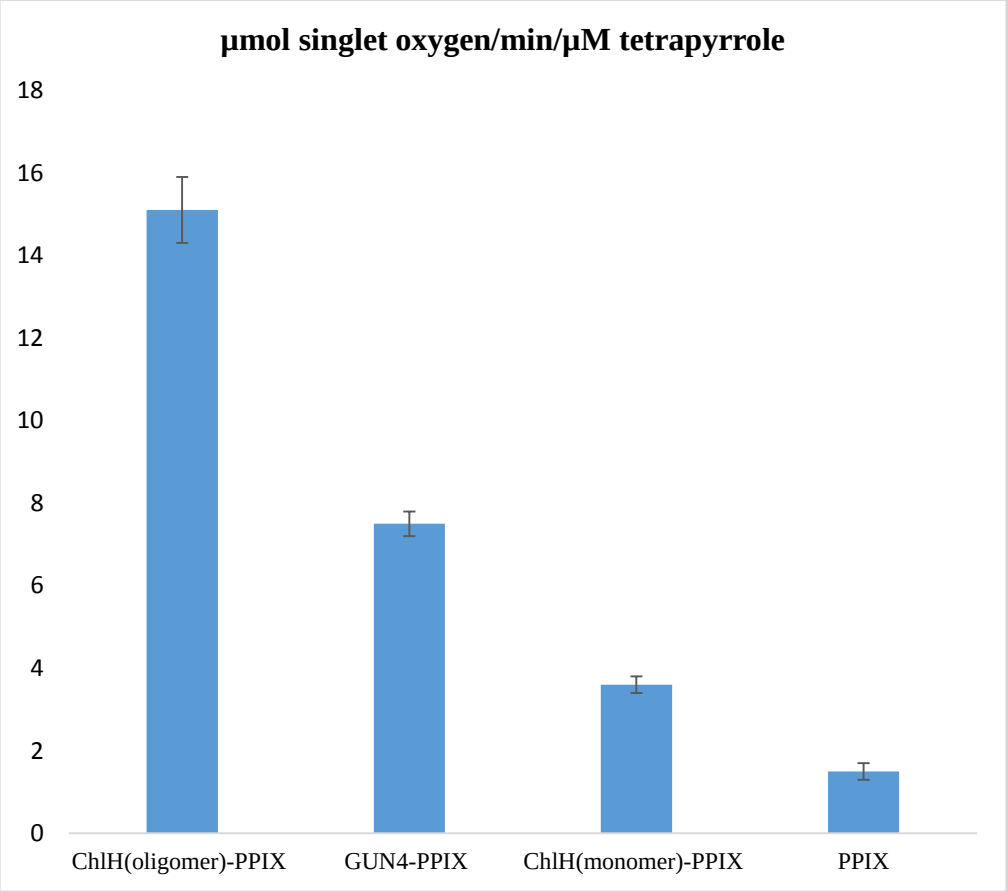


Figure 6.1 Singlet oxygen production rate (µM singlet oxygen/min/ µM PPIX) of free PPIX and protein complexes (Data derived from Table 5.1).

6.2 Conclusion

Chls and Bchls are crucial to the process of photosynthesis, and understanding of Chl and Bchl biosynthesis is essential for the regulation of assembly of the photosynthetic apparatus.

Moreover, identification of the Chl biosynthesis intermediates role in intracellular signaling, and communication between the nucleus and chloroplast are other significant points in the field of Chl biosynthesis which needs to be discussed. Numerous enzymes are involved in transforming the first committed precursor (ALA) into the final products, Chl or Bchl. My thesis focused on the Mg-chelation step of the Chl biosynthetic pathway. A key discovery was the discovery of GUN4-PPIX protein complex as a strong $O_2(a^1\Delta_g)$ generator *in vitro*. I have shown that GUN4 increases the production of PPIX-generated $O_2(a^1\Delta_g)$ in the light by a factor of five compared with PPIX alone and the $O_2(a^1\Delta_g)$ generated from GUN4 and ChlH oligomer are a strong candidates for a plastid to nucleus signal, possibly through second messengers. Additionally, determination of GUN4 crystal structure from *C. reinhardtii* has broadened our knowledge about the PPIX binding site of the protein which is in extensive agreement with previously solved cyanobacterial crystal structures. In summary this work has combined structural and functional approaches to help further define roles for ChlH of Mg-chelatase and GUN4 as $O_2(a^1\Delta_g)$ generators in the Chl biosynthetic pathway in plants and algae.

6.3 Future directions

The combination of structural and functional analysis of ChlH and GUN4 presented in this thesis provides previously unassigned functional roles of the porphyrin binding implications of these two proteins in the presence of oxygen and light. This work has opened a new outlook on plastid to nucleus signalling, the mechanism of which will require a global approach involving measurements of at least EXECUTOR 1/2 proteins, second messengers as well as GUN4/ChlH and $O_2(a^1\Delta_g)$. However initial fundamental experiments involving protein-protein interactions of ChlH and GUN4 are needed.

The interaction between ChlH-PPIX and GUN4 to form the GUN4-ChlH-PPIX complex is not well-defined. Novel protein-protein interaction studies may be required to resolve this, since

interaction of GUN4 and ChlH is very slow, and the mixture of GUN4-ChlH with PPIX has to be incubated for 5-20 min. A possible method may be to use the thermodynamic parameters of GUN4 interaction with ChlH-PPIX, and their binding affinity using ITC (Isothermal Titration Calorimetry). This method could also be used to understand the binding affinity of ChlI protein with ChlD protein and ATP, or the ChlI-ChlD complex binding affinity. It would be essential to further identify the different growth conditions for GUN4 protein *in vivo*, and the exact location of protein in plant cell. It could also be tested if GUN4 with PPIX bound can deliver PPIX to ChlH for Mg-chelatase activity. It would be advantageous to have an X-ray crystal structure of full length eukaryotic GUN4 to further identify the phosphorylation side of GUN4 in eukaryotic organisms. This could succeed in improving the quality of the GUN4 crystals from *C. reinhardtii* using analogues of PPIX (eg, platinum-PPIX).

Furthermore, small angle neutron scattering (SANS) analysis might be an interesting procedure to provide structural details of a GUN4-ChlH-PPIX complex. Studying the structural details of protein interactions will lead to a better understanding of mechanisms underlying their malfunction in important metabolic pathways. It will also be important to consider the $O_2(a^1\Delta_g)$ production rates of other intermediates in Chl biosynthesis and heme biosynthesis pathway, such as detailed mutational studies on GUN4 to identify the crucial amino acids involved in generating or quenching $O_2(a^1\Delta_g)$ by PPIX.

References

Aarti D, Tanaka R, Ito H, Tanaka A (2007) High Light Inhibits Chlorophyll Biosynthesis at the Level of 5-Aminolevulinate Synthesis During De-etiolation in Cucumber (Cucumis sativus) Cotyledons†. Photochemistry and Photobiology 83:171-176.

Adams NB, Marklew CJ, Brindley AA, Hunter CN, Reid JD (2014a) Characterization of the magnesium chelatase from *Thermosynechococcus elongatus*. The Biochemical journal 457:163-170.

Adams NB, Marklew CJ, Qian P, Brindley AA, Davison PA, Bullough PA, Hunter CN (2014b) Structural and functional consequences of removing the N-terminal domain from the magnesium chelatase ChlH subunit of *Thermosynechococcus elongatus*. The Biochemical journal 464:315-322.

Adhikari ND, Froehlich JE, Strand DD, Buck SM, Kramer DM, Larkin RM (2011) GUN4-porphyrin complexes bind the ChlH/GUN5 subunit of Mg-Chelatase and promote chlorophyll biosynthesis in *Arabidopsis*. Plant Cell 23:1449-1467.

Adhikari ND, Orler R, Chory J, Froehlich JE, Larkin RM (2009) Porphyrins promote the association of GENOMES UNCOUPLED 4 and a Mg-chelatase subunit with chloroplast membranes. The Journal of biological chemistry 284:24783-24796.

Agon VV, Bubb WA, Wright A, Hawkins CL, Davies MJ (2006) Sensitizer-mediated photooxidation of histidine residues: Evidence for the formation of reactive side-chain peroxides. Free Radical Biology and Medicine 40:698-710.

Alawady A, Reski R, Yaronskaya E, Grimm B (2005) Cloning and expression of the tobacco CHLM sequence encoding Mg protoporphyrin IX methyltransferase and its interaction with Mg chelatase. Plant Molecular Biology 57:679-691.

Alawady AE, Grimm B (2005) Tobacco Mg protoporphyrin IX methyltransferase is involved in inverse activation of Mg porphyrin and protoheme synthesis. The Plant journal 41:282-290.

Apchelimov AA, Soldatova OP, Ezhova TA, Grimm B, Shestakov SV (2007) The analysis of the ChlI 1 and ChlI 2 genes using acifluorfen-resistant mutant of *Arabidopsis thaliana*. Planta 225:935-943.

Axelsson E, Lundqvist J, Sawicki A, Nilsson S, Schroder I, Al-Karadaghi S, Willows RD, Hansson M (2006) Recessiveness and Dominance in Barley Mutants Deficient in Mg-Chelatase Subunit D, an AAA Protein Involved in Chlorophyll Biosynthesis. The Plant Cell Online 18:3606-3616.

Beale S (1999) Enzymes of chlorophyll biosynthesis. Photosynthesis research 60:43-73.

Benvenuti M, Mangani S (2007) Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. Nature protocols 2:1633-1651.

Bergfors T (2003) Seeds to crystals. Journal of Structural Biology 142:66-76.

Bhattacharjee S (2012) The Language of Reactive Oxygen Species Signaling in Plants. Journal of Botany 2012:1-22.

Bollivar DW, Suzuki JY, Beatty JT, Dobrowolski JM, Bauer CE (1994) Directed Mutational Analysis of Bacteriochlorophyll a Biosynthesis in *Rhodobacter capsulatus*. Journal of Molecular Biology 237: 622–640.

Bosio GN, Breitenbach T, Parisi J, Reigosa M, Blaikie FH, Pedersen BW, Silva EFF, Mártire DO, Ogilby PR (2013) Antioxidant β-Carotene Does Not Quench Singlet Oxygen in Mammalian Cells. Journal of the American Chemical Society 135:272-279.

Bradford MM (1976) Bradford. Analytical Biochemistry 72:248-254.

Brancaleon L, Moseley H (2002) Effects of photoproducts on the binding properties of protoporphyrin IX to proteins. Biophysical Chemistry 96:77-87.

Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355:472-475.

Brzezowski P, Schlicke H, Richter A, Dent RM, Niyogi KK, Grimm B (2014) The GUN4 protein plays a regulatory role in tetrapyrrole biosynthesis and chloroplast-to-nucleus signalling in Chlamydomonas reinhardtii. The Plant Journal 79:285-298.

Castelfranco PA, Weinstein JD, Schwarcz S, Pardo AD, Wezelman BE (1979) The Mg insertion step in chlorophyll biosynthesis. Archives of Biochemistry and Biophysics 192:592-598.

Chekounova E, Voronetskaya V, Papenbrock J, Grimm B, Beek CF (2001) Characterization of chlamydomonas mutant defective in the H subunit of Mg-Chelatase. Molecular Genetics and Genomics 566:363-373.

Chen M (2014) Chlorophyll modifications and their spectral extension in oxygenic photosynthesis. Annual review of biochemistry 83:317-340.

Chen S, Kim C, Lee JM, Lee HA, Fei Z, Wang L, Apel K (2015a) Blocking the QB -binding site of photosystem II by tenuazonic acid, a non-host-specific toxin of Alternaria alternata, activates singlet oxygen-mediated and EXECUTER-dependent signalling in *Arabidopsis*. Plant, Cell & Environment 38:1069-1080.

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta crystallographica Section D, Biological crystallography 66:12-21.

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

Chen X, Pu H, Wang X, Long W, Lin R, Liu L (2015c) Crystal structures of GUN4 in complex with porphyrins. Molecular Plants Journal 8:1125-1127.

Chew AGM, Bryant DA (2007) Chlorophyll Biosynthesis in Bacteria: The Origins of Structural and Functional Diversity. Annual Review of Microbiology 61:113-129.

Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. Acta crystallographica Section D, Biological crystallography 62:1002-1011.

Cox MJ, Weber PC (1988) An investigation of protein crystallization parameters using successive automated grid searches (SAGS). Journal of Crystal Growth 90:318-324.

Cudney B, Patel S, Weisgraber K, Newhouse Y, McPherson A (1994) Screening and Optimization Strategies for Macromolecular Crystal Growth. Acta crystallographica Section D, Biological crystallography 50:414-423.

Dauter Z (1999) Data-collection strategies Acta Cryst D Biol Crystallogr 55:1703-1717.

Davison PA, Hunter CN (2011) Abolition of magnesium chelatase activity by the gun5 mutation and reversal by Gun4. FEBS Lett 585:183-186.

Davison PA, Schubert HL, Reid JD, Iorg CD, Heroux A, Hill CP, Hunter CN (2005) Structural and Biochemical Characterization of Gun4 Suggests a Mechanism for Its Role in Chlorophyll Biosynthesis. Biochemistry 44:10.

Dawson RMC, Elliott DC, Elliott WH, Jones KM (1986) Data for biochemical research (third edition) Oxford Science Publications, OUP, Oxford ISBN 0-19-855358-7:580.

Day D, Neuburger M, Douce R (1985) Biochemical Characterization of Chlorophyll-Free Mitochondria From Pea Leaves. Functional Plant Biology 12:219-228.

DeRosa MC, Crutchley RJ (2002) Photosensitized singlet oxygen and its applications. Coordination Chemistry Reviews 233–234:351-371.

Ellsworth RK, Lawrence GD (1973) Photosynthetica 7: 73-86.

Elmlund H, Lundqvist J, Al-Karadaghi S, Hansson M, Hebert H, Lindahl M (2008) A New Cryo-EM Single-Particle Ab Initio Reconstruction Method Visualizes Secondary Structure Elements in an ATP-Fueled AAA+ Motor. Journal of Molecular Biology 375:934-947.

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta crystallographica Section D, Biological crystallography 66:486-501.

Evans P (2005) Scaling and assessment of data quality. Acta Crystallographica Section D Biological Crystallography 62:72-82.

Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta crystallographica Section D, Biological crystallography 67:282-292.

Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? Acta Crystallographica Section D 69:1204-1214.

Flors C, Fryer MJ, Waring J, Reeder B, Bechtold U, Mullineaux PM, Nonell S, Wilson MT, Baker NR (2006) Imaging the production of singlet oxygen in vivo using a new fluorescent sensor, Singlet Oxygen Sensor Green. Journal of Experimental Botany 57:1725-1734.

Fodje MN, A Hansson, Hansson M, Olsen JG, Gough S, Willows RD, Al-Karadaghi S (2001) Interplay Between an AAA Module and an Integrin IDomain May Regulate the Function of Magnesium Chelatase. Journal of Molecular Biology 311:11.

Formighieri C, Ceol M, Bonente G, Rochaix JD, Bassi R (2012) Retrograde Signaling and Photoprotection in a gun4 Mutant of *Chlamydomonas reinhardtii*. Molecular Plant 5:1242-1262.

Foyer CH, Noctor G (2000) Oxygen processing in photosynthesis:regulation and signalling. New Phytol 146:358-388.

Fuesler TP, Wright LA, Castelfranco PA (1981) Properties of Magnesium Chelatase in Greening Etioplasts: Meatl ion specifity and effect of substrate concentrations. Plant physiology 67:246-249.

Fuhrhop J, Granick S (1971) Magnesium protoporphyrin dimethyl ester and the dipotassium salt of magnesium protoporphyrin. Biochemical Preparations 13:55-58.

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server: Springer.

Giacovazzo C (2002) IUCr Texts on crystallography 7, Fundamentals of crystallography. Oxford 1992: Oxford University Press, 1-654.

Gibson LCD, Jenesen PE, Hunter CN (1999) Magnesium chelatase from *Rhodobacter* sphaeroides : initial characterization of the enzyme using purified subunits and evidence for a BchI–BchD complex. Biochemical Journal 337:243-251.

Gibson LCD, Joanne L. Marrison, Rache1 M. Leechl, Poul E. Jensen, Diane C. Bassham, Marie Cibson, Hunter CN (1996) A Putative Mg Chelatase Subunit from *Arabidopsis thaliana* cv C24. Plant Physiology 111:61-71.

Gibson LCD, Willows RD, Kannangarat CG, Wettsteint DV, Hunter CN (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. Biochemistry 92:1941-1944.

Glatter, Kratky (1982) Small angle X-ray scattering. London: Academic Press Inc(London) Ltd.

Glatter O (1977a) Data evaluation in small-angle scattering: calculation of the radial electron density distribution by means of indirect Fourier transformation. Acta Physica Austriaca 47:83-102.

Glatter O (1977b) A New Method for the Evaluation of Small-Angle Scattering Data. Journal of applied crystallography 10:415-421.
Goddard TD, Huang CC, Ferrin TE (2007) Visualizing density maps with UCSF Chimera. Journal of Structural Biology 157:281-287.
Gorchein A (1972) Magnesium Protoporphyrin Chelatase Activity in *Rhodopseudomonas spheroides*. Biochemistry Journal 127:97-106.
Gorchein A (1973) Control of Magnesium-Protoporphyrin Chelatase Activity in *Rhodopseudomonas spheroides*. Biochemistry Journal 124:833-845.
Gorchein A, Gibson LCD, Hunter CN (1993) Gene expression and control of enzymes for synthesis of magnesium protoporphyrin monomethyl ester in *Rhodobacter sphaeroides*. Biochemical Society Transactions 21:2015.
Grafe S, Saluz HP, Grimm B, Hanel F (1999) Mg-chelatase of tobacco: The role of the subunit CHL D in the chelation step of protoporphyrin IX. Biochemistry 96:1941-1946.
Guo R, Luo M, Weinstein JD (1998) Magnesium-Chelatase from Developing Pea Leaves. Plant physiology 116:605-615.
Hansson A, Kannangara CG, Westistein DV, Hasson M (1999) Molecular basis for semidominance of missense mutations in the XANTHA-H (42-kDa) subunit of magnesium chelatase. Proceedings of the National Academy of Sciences of the United States of America, The National Academy of Sciences 96:1744-1749.
Hansson A, Willows RD, Roberts TH, Hansson M (2002) Three semidominant barley mutants with single amino acid substitutions in the smallest magnesium chelatase subunit form defective AAA+ hexamers. Proceedings of the National Academy of Sciences 99:13944-13949.
Hansson M, Kannangara CG (1997) ATPases and phosphate exchange activities in magnesium chelatase subunits of *Rhodobacter sphaeroides*. Proceedings of the National Academy of Sciences 94:13351-13356.
Harp JM, Timm DE, Bunick GJ (1998) Macromolecular crystal annealing: overcoming increased mosaicity associated with cryocrystallography. Acta crystallographica Section D, Biological crystallography 54:622-628.
Hauptman H (1997) Phasing methods for protein crystallography. Current Opinion in Structural Biology 7:672-680.
Hoober JK, Kahn A, Ash D, Gough S, Kannangara CG (1988) Biosynthesis of Δ-aminolevulinate in greening barley leaves. IX. Structure of the substrate, mode of gabaculine inhibition, and the catalytic mechanism of glutamate 1-semialdehyde aminotransferase. Carlsberg Res Commun 53:11-25.
Huang YS, Li HM (2009) *Arabidopsis* CHLI2 can substitute for CHLI1. Plant physiology 150:636-645.
Hudson A, R.Carpenter, S.Doyle, E.S.Coen (1993) Olive: a key gene required for chlorophyll biosynthesis in Antirrhinum majus. The EMBO journal 12:3711 - 3719.
Ishijima S, Uchibori A, Takagi H, Maki R, Ohnishi M (2003) Light-induced increase in free Mg2+ concentration in spinach chloroplasts: Measurement of free Mg2+ by using a fluorescent probe and necessity of stromal alkalinization. Archives of Biochemistry and Biophysics 412:126-132.
Iyer LM, Leipe DD, Koonin EV, Aravind L (2004) Evolutionary history and higher order classification of AAA+ ATPases. J Struct Biol 146:11-31.
Jaffe EK (1993) Predicting the Zn(II) Ligands in Metalloproteins: Case Study, Porphobilinogen Synthase. Comments on Inorganic Chemistry 15:67-92.
Jancarik J, Kim S-H (1991) Sparse matrix sampling: a screening method for crystallization of proteins. Journal of applied crystallography 24:409-411.

Jensen PE, Gibson LC, Hunter CN (1998) Determinants of catalytic activity with the use of purified I, D and H subunits of the magnesium protoporphyrin IX chelatase from Synechocystis PCC6803. Biochemical Journal 334:335-344.

Jensen PE, Gibson LCD, Henningseni KW, Hunter CN (1996a) Expression of the chlI, chlD, and chlH Genes from the Cyanobacterium Synechocystis PCC6803 in Escherichia coli and Demonstration That the Three Cognate Proteins Are Required for Magnesium-protoporphyrin Chelatase Activity. The Journal of biological chemistry 271:16662-16667.

Jensen PE, Gibson LCD, Hunter CN (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. Biochemistry Journal 339:127-134.

Jensen PE, Reid JD, Hunter CN (2000) Modification of cysteine residues in the ChlI and ChlH subunits of magnesium chelatase results in enzyme inactivation. Biochemistry Journal 352:435-441.

Jensen PE, Willows RD, Petersen BL, Vothknecht UC, Stummann BM, Kannangara CG, von Wettstein D, Henningsen KW (1996b) Structural genes for Mg-chelatase subunits in barley: Xantha-f, -g and -h. Molecular Genetics and Genomics 250:383-394.

Joanne Viney, Paul A. Davison, C. Neil Hunter, James D. Reid (2007) Direct Measurement of Metal-Ion Chelation in the Active Site of the AAA+ ATPase Magnesium Chelatase. Biochemistry 46:12788-12794.

Johnson ET, 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. The Journal of biological chemistry 283:27776-27784.

Kannangara CG, U.C. Vothknecht, M.Hansson, Wettstein Dv (1997) Magnesium chelatase: association with ribosomes and mutant complementation studies identify barley subunit Xant ha-G as a functional counterpart of *Rhodobacter* subunit BchD. Molecular Genetics and Genomics 254:85-92.

Karger GA, Reid JD, Hunter CN (2001) Characterization of the Binding of Deuteroporphyrin IX to the Magnesium Chelatase H Subunit and Spectroscopic Properties of the Complex. Biochemistry 40:9291-9299.

Kiang NY, Siefert J, Govindjee, Blankenship RE (2007) Spectral signatures of photosynthesis. I. Review of Earth organisms. Astrobiology 7:222-251.

Kobayashi K, Mochizuki N, Yoshimura N, Motohashi K, Hisabori T, Masuda T (2008) Functional analysis of *Arabidopsis thaliana* isoforms of the Mg-chelatase CHLI subunit. Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology 7:1188-1195.

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography 36:1277-1282.

Krieger-Liszkay A (2005) Singlet oxygen production in photosynthesis. J Exp Bot 56:337-346.

Kuimova MK, Yahioglu G, Ogilby PR (2009) Singlet Oxygen in a Cell: Spatially Dependent Lifetimes and Quenching Rate Constants. Journal of the American Chemical Society 131:332-340.

Lake V, Olsson U, Willows RD, Hansson M (2004) ATPase activity of magnesium chelatase subunit I is required to maintain subunit D in vivo. European journal of biochemistry / FEBS 271:2182-2188.

Laloi C, Havaux M (2015) Key players of singlet oxygen-induced cell death in plants. Front Plant Science 6:1-9.
Lamzin VS, Perrakis A (2000) Current state of automated crystallographic data analysis. Nature Structural & Molecular Biology 7:978-981.
Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-2948.
Larkin RM, Alonso JM, Ecker JR, Chory J (2003) GUN4, a regulator of chlorophyll synthesis and intracellular signaling. Science 299:902-906.
Lee HJ, Ball MD, Parham R, Rebeiz CA (1992) Enzymic Conversion of Protoporphyrin IX to Mg-Protoporphyrin IX in a Subplastidic Membrane Fraction of Cucumber Etiochloroplasts. Plant physiology 99:1134-1140.
Lee KP, Kim C, Landgraf F, Apel K (2007) EXECUTER1- and EXECUTER2-dependent transfer of stress-related signals from the plastid to the nucleus of *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences 104:10270-10275.
Lermontova I, Kruse E, Mock H-P, Grimm B (1997) Cloning and characterization of a plastidal and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase. Proceedings of the National Academy of Sciences 94:8895-8900.
Leslie AGW, Powell HR (2007) Processing Diffraction Data with Mosflm. Evolving Methods for Macromolecular Crystallography 245:41-51.
Leslie M (2009) On the Origin of Photosynthesis. Science 323:1286-1287.
Luft J, DeTitta G (1999) A method to produce microseed stock for use in the crystallization of biological macromolecules. Acta crystallographica Section D, Biological crystallography 55:988–993.
Luft JR, Wolfley JR, Said MI, Nagel RM, Lauricella AM, Smith JL, Thayer MH, Veatch CK, Snell EH, Malkowski MG, Detitta GT (2007) Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. Protein science : a publication of the Protein Society 16:715-722.
Lundqvist J, Braumann I, Kurowska M, Muller AH, Hansson M (2013) Catalytic Turnover Triggers Exchange of Subunits of the Magnesium Chelatase AAA+ Motor Unit. Journal of Biological Chemistry 288:24012-24019.
Lundqvist J, Elmlund H, Wulff RP, Berglund L, Elmlund D, Emanuelsson C, Hebert H, Willows RD, Hansson M, Lindahl M, Al-Karadaghi S (2010) ATP-induced conformational dynamics in the AAA+ motor unit of magnesium chelatase. Structure 18:354-365.
Luo M, Weinstein JD, Walker CJ (1999) Magnesium chelatase subunit D from pea: characterization of the cDNA, heterologous expression of an enzymatically active protein and immunoassay of the native protein. Plant Molecular Biology 41:11.
McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. Journal of applied crystallography 40:658-674.
Mcpherson A (1990) Current approaches to macromolecular crystallization. European Journal of Biochemistry 189:1-23.
Méndez-Hurtado J, López R, Suárez D, Menéndez MI (2012) Theoretical Study of the Oxidation of Histidine by Singlet Oxygen. Chemistry – A European Journal 18:8437-8447.
Meskauskiene R, Nater M, Goslings D, Kessler F, op den Camp R, Apel K (2001) FLU: A negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences 98:12826-12831.

Mochizuki N, Brusslan JA, Nagatani RLA, Chory J (2001*) Arabidopsis* genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proceedings of the National Academy of Sciences of the United States of America 98:2053-2058.

Mochizuki N, Tanaka R, Tanaka A, Masuda T, Nagatani A (2008) The steady-state level of Mg-protoporphyrin IX is not a determinant of plastid-to-nucleus signaling in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America 105:15184-15189.

Mock H-P, Keetman U, Kruse E, Rank B, Grimm B (1998) Defense Responses to Tetrapyrrole-Induced Oxidative Stress in Transgenic Plants with Reduced Uroporphyrinogen Decarboxylase or Coproporphyrinogen Oxidase Activity. Plant physiology 116:107-116.

Moulin M, McCormac AC, Terry MJ, Smith AG (2008) Tetrapyrrole profiling in *Arabidopsis* seedlings reveals that retrograde plastid nuclear signaling is not due to Mg-protoporphyrin IX accumulation. Proceedings of the National Academy of Sciences of the United States of America 105:15178-15183.

Muller AH, Hansson M (2009) The barley magnesium chelatase 150-kd subunit is not an abscisic acid receptor. Plant physiology 150:157-166.

Müller AH, Sawicki A, Zhou S, Tabrizi ST, Luo M, Hansson M, Willows RD (2014) Inducing the oxidative stress response in Escherichia coli improves the quality of a recombinant protein: Magnesium chelatase ChlH. Protein Expression and Purification 101:61-67.

Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta crystallographica Section D, Biological crystallography 67:355-367.

Nemčovičová I, Smatanová IK (2012) Alternative Protein Crystallization Technique: Cross-Influence Procedure (CIP), Crystallization and Materials Science of Modern Artificial and Natural Crystals, Dr. Elena Borisenko (Ed.) InTech ISBN: 978-953-307-608-9.

Neuwald AF, Aravind L., SpougeJohn L. , V KE (1999) AAA+: A Class of Chaperone-Like ATPases Associated with the Assembly, Operation, and Disassembly of Protein Complexes. Genome Research 9:27-43.

Nisbet EG, Cann JR, Lee C, Dover V (1995) Origins of photosynthesis. Nature 373:479-480.

Ogilby PR (2010) Singlet oxygen: there is indeed something new under the sun. Chemistry Society Reviewers 39:3181-3209.

Ogura Teru, J. W (2001) AAA+ superfamily ATPases:common structure-diverse function. Genes to Cells 6:575-597.

Ojha H, Mishra K, Hassan MI, Chaudhury NK (2012) Spectroscopic and isothermal titration calorimetry studies of binding interaction of ferulic acid with bovine serum albumin. Thermochimica Acta 548:56-64.

Olsson U, Sirijovski N, Hansson M (2004) Characterization of eight barley xantha-f mutants deficient in magnesium chelatase. Plant physiology and biochemistry : PPB / Societe francaise de physiologie vegetale 42:557-564.

op den Camp RGL (2003) Rapid Induction of Distinct Stress Responses after the Release of Singlet Oxygen in *Arabidopsis*. The Plant Cell Online 15:2320-2332.

Orthaber D, Bergmann A, Glatter O (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. Journal of Applied Crystallography 33:218-225.

Osanai T, Imashimizu M, Seki A, Sato S, Tabata S, Imamura S, Asayama M, Ikeuchi M, Tanaka K (2009) ChlH, the H subunit of the Mg-chelatase, is an anti-sigma factor for SigE in

Synechocystis sp. PCC 6803. Proceedings of the National Academy of Sciences of the United States of America 106:6860-6865.
Papenbrock J, GrMe S, Kruse E, H~inel F, Grimm B (1997) Mg-chelatase of tobacco: identification of a Chl D cDNA sequence encoding a third subunit, analysis of the interaction of the three subunits with the yeast two-hybrid system, and reconstitution of the enzyme activity by co-expression of recombinant CHL D, CHL H and CHL I The Plant Journal 12:981-990.
Papenbrock J, Mock H-P, Kruse E, Grimm B (1999) Expression studies in tetrapyrrole biosynthesis: inverse maxima of magnesium chelatase and ferrochelatase activity during cyclic photoperiods. planta 208:264-273.
Papenbrock J, Mock H-P, Tanaka R, Kruse E, Grimm B (2000) Role of Magnesium Chelatase Activity in the Early Steps of the Tetrapyrrole Biosynthetic Pathway1. Plant physiology 122:1161-1169.
Pardo AD, Barbara M. Chereskin, Castelfranco PA, Franseschi VR, Wezelman BE (1980) ATP Requirement for Mg Chelatase in Developing Chloroplasts'. Plant physiology 65:956-960.
Peter E, Grimm B (2009) GUN4 is required for posttranslational control of plant tetrapyrrole biosynthesis. Molecular Plant 2:1198-1210.
Petersen BL, Jensen PE, Gibson LCD, Stummann BM, Hunter CN, Henningsen KW (1998) Reconstitution of an Active Magnesium Chelatase Enzyme Complex from the bchI, -D, and -H Gene Products of the Green Sulfur Bacterium Chlorobium vibrioforme Expressed in Escherichia coli. Journal of Bacteriology, 180,No.3:699-704.
Petersen BL, Kannangara CG, Henningsen KW (1999) Distribution of ATPase and ATP-to-ADP phosphate exchange activities in magnesium chelatase subunits of Chlorobium vibrioforme and Synechocystis PCC6803. Archives of Microbiology 146–150.
Peterson BL, G. Marianne, Miziller BM, Knud S, W. H (1998) Structure and organization of a 25 kbp region of the genome of the photosynthetic green sulfur bacterium Chlorobium uibrioforme containing Mg-chelatase encoding genes. Hereditas 129(2):131-142.
Petoukhov MV, Konarev PV, Kikhney AG, Svergun DI (2007) ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. Journal of Applied Crystallography 40:S223-S228.
Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Quarterly Reviews of Biophysics 40:191-285.
Qian P, Marklew CJ, Viney J, Davison PA, Brindley AA, Soderberg C, Al-Karadaghi S, Bullough PA, Grossmann JG, Hunter CN (2012) Structure of the cyanobacterial Magnesium Chelatase H subunit determined by single particle reconstruction and small-angle X-ray scattering. The Journal of biological chemistry 287:4946-4956.
Rambo RP, Tainer JA (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. Biopolymers 95:559-571.
Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. Nature 496:477-481.
Raymond J, Zhaxybayeva O, Gogarten JP, Gerdes SY, Blankenship RE (2002) Whole-Genome Analysis of Photosynthetic Prokaryotes. Science 298:1616-1620.
Rehman AU, Cser K, Sass L, Vass I (2013) Characterization of singlet oxygen production and its involvement in photodamage of Photosystem II in the cyanobacterium Synechocystis PCC 6803 by histidine-mediated chemical trapping. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1827:689-698.

Reid JD, Hunter CN (2004) Magnesium-dependent ATPase activity and cooperativity of magnesium chelatase from Synechocystis sp. PCC6803. The Journal of biological chemistry 279:26893-26899.

Reid JD, Siebert CA, Bullough PA, Hunter CN (2003) The ATPase Activity of the ChlI Subunit of Magnesium Chelatase and Formation of a Heptameric AAA+ Ring†. Biochemistry 42:9.

Reiland S, Messerli G, Baerenfaller K, Gerrits B, Endler A, Grossmann J, Gruissem W, Baginsky S (2009) Large-scale *Arabidopsis* phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. Plant physiology 150:889-903.

Reinbothe C, Bakkouri ME, Buhr F, Muraki N, Nomata J, Kurisu G, Fujita Y, Reinbothe S (2010) Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. Trends in Plant Science 15:614-624.

Rhodes G (2006) Crystallography Made Crystal Clear - A Guide for Users of Macromolecular Models(Third Edition) Academic Press, New York ISBN 0-12-587073-6:360.

Richter ML, Rienits KG (1980) The synthesis of magnesium and zinc protoporphyrin IX and their monomethyl esters in etioplast preparations studied by high pressure liquid chromatography. FEBS Letters 116:211-216.

Richter ML, Rienits KG (1982) The synthesis of Magnesium-Protoporphyrin IX by etiochloroplast membrane reparation Biochimica et Biophysica Acta 717:255-264.

Rissler HM, Collakova E, DellaPenna D, Whelan J, Pogson BJ (2002) Chlorophyll biosynthesis. Expression of a second chl I gene of magnesium chelatase in *Arabidopsis* supports only limited chlorophyll synthesis. Plant physiology 128:770-779.

Rotenberg M, Margalit R (1985) Deuteroporphyrin-albumin binding equilibrium. The effects of porphyrin self-aggregation studied for the human and the bovine proteins. Biochemical Journal 229:197-203.

Rupp B (2010) biomolecular Crystalography: Principles, practice, and application to structure biology New yourk, Garland Science 1-676.

Sawicki A, Willows RD (2007) S-adenosyl-L-methionine:magnesium-protoporphyrin IX O-methyltransferase from *Rhodobacter capsulatus*: mechanistic insights and stimulation with phospholipids. The Biochemical journal 406:469-478.

Sawicki A, Willows RD (2008) Kinetic Analyses of the Magnesium Chelatase Provide Insights into the Mechanism, Structure, and Formation of the Complex. Journal of Biological Chemistry 283:31294-31302.

Sawicki A, Willows RD (2010) BchJ and BchM interact in a 1 : 1 ratio with the magnesium chelatase BchH subunit of *Rhodobacter capsulatus*. The FEBS journal 277:4709-4721.

Schlicke H, Hartwig AS, Firtzlaff V, Richter AS, Glasser C, Maier K, Finkemeier I, Grimm B (2014) Induced Deactivation of Genes Encoding Chlorophyll Biosynthesis Enzymes Disentangles Tetrapyrrole-Mediated Retrograde Signaling. Molecular Plant 1211-1227.

Schon A, Krupp G, Gough S, Berry-Lowe S, Kannangara CG, Soll D (1986) The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. Nature 322:281-284.

Shapiguzov A, Vainonen JP, Wrzaczek M, Kangasjarvi J (2012) ROS-talk - how the apoplast, the chloroplast, and the nucleus get the message through. Front Plant Science 3:292.

Shen YY, Wang XF, Wu FQ, Du SY, Cao Z, Shang Y, Wang XL, Peng CC, Yu XC, Zhu SY, Fan RC, Xu YH, Zhang DP (2006) The Mg-chelatase H subunit is an abscisic acid receptor. Nature 443:823-826.

Shepherd M, James D. Reid, Hunter CN (2003) Purification and kinetic characterization of the magnesium protoporphyrin IX methyltransferase from Synechocystis PCC6803. Biochem Journal 371:351-360.
Shepherd M, McLean S, Hunter CN (2005) Kinetic basis for linking the first two enzymes of chlorophyll biosynthesis. The FEBS journal 272:4532-4539.
Singh R, Singh S, Parihar P, Singh VP, Prasad SM (2015) Retrograde signaling between plastid and nucleus: A review. Journal of Plant Physiology 181:55-66.
Sirijovski N, Lundqvist J, Rosenback M, Elmlund H, Al-Karadaghi S, Willows RD, Hansson M (2008) Substrate-binding Model of the Chlorophyll Biosynthetic Magnesium Chelatase BchH Subunit. Journal of Biological Chemistry 283:11652-11660.
Sirijovski N, Olsson U, Lundqvist J, Al-Karadaghi S, Willows Robert D, Hansson M (2006) ATPase activity associated with the magnesium chelatase H-subunit of the chlorophyll biosynthetic pathway is an artefact. Biochemical Journal 400:477.
Smith BB, Rebeiz CAA (1977) Chloroplast biogenesis: detection of Mg-protoporphyrin chelatase in vitro. Archives of Biochemistry and Biophysics 180:178-185.
Sobotka R, Duhring U, Komenda J, Peter E, Gardian Z, Tichy M, Grimm B, Wilde A (2008) Importance of the cyanobacterial Gun4 protein for chlorophyll metabolism and assembly of photosynthetic complexes. The Journal of biological chemistry 283:25794-25802.
Stenbaek A, Jensen PE (2010) Redox regulation of chlorophyll biosynthesis. Phytochemistry 71:853-859.
Stephenson PG, Terry MJ (2008) Light signalling pathways regulating the Mg-chelatase branchpoint of chlorophyll synthesis during de-etiolation in *Arabidopsis thaliana*. Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology 7:1243-1252.
Strand A, Asami T, Alonso J, Ecker JR, Chory J (2003) Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. Nature 421:79-83.
Stura E, Wilson I (1990) Analytical and production seeding techniques. Methods: Companion Methods Enzymolgy. 1:38-49.
Susek RE, Ausubel FM, Chory J (1993) Signal transduction mutants of *Arabidopsis* uncouple nuclear CAB and RBCS gene expression from chloroplast development. Cell 74:787-799.
Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. Journal of Applied Crystallography 25:495-503.
Svergun DI (1999) Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. Biophysical Journal 76:2879-2886.
Svergun DI, Koch MH (2002) Advances in structure analysis using small-angle scattering in solution. Biophysical methods 12:654-660.
Szechyńska-Hebda M, Karpiński S (2013) Light intensity-dependent retrograde signalling in higher plants. Journal of Plant Physiology 170:1501-1516.
Takagi J (2007) Structural basis for ligand recognition by integrins. Current opinion in cell biology 19:557-564.
Tamiaki H, Shibata R, Mizoguchi T (2007) The 17-Propionate Function of (Bacterio)chlorophylls: Biological Implication of Their Long Esterifying Chains in Photosynthetic Systems. Photochemistry and Photobiology 83:152-162.
Tanaka R, Tanaka A (2007) Tetrapyrrole Biosynthesis in Higher Plants. Annual Review of Plant Biology 58:321-346.

Tarahi Tabrizi S, Langley DB, Harrop SJ, Duff AP, Willows RD (2015) Structure of GUN4 from *Chlamydomonas reinhardtii*. Acta Crystallographica Section F Structural Biology Communications 71:1094-1099.

Taylor G (2003) The phase problem. Acta Crystallographica Section D 59:1881-1890.

Telfer A, Bishop SM, Phillips D, Barber J (1994) Isolated photosynthetic reaction center of photosystem II as a sensitizer for the formation of singlet oxygen. Detection and quantum yield determination using a chemical trapping technique. Journal of Biological Chemistry 269:13244-13253.

Triantaphylidès C, Havaux M (2009) Singlet oxygen in plants: production, detoxification and signaling. Trends in Plant Science 14:219-228.

Tripathy BC, Mohapatra A, Gupta I (2007) Impairment of the photosynthetic apparatus by oxidative stress induced by photosensitization reaction of protoporphyrin IX. Biochemica et Biophysica Acta 1767:860-868.

Tripathy BC, Sherameti I, Oelmüller R (2010) Siroheme: An essential component for life on earth. Plant Signaling & Behavior 5:14-20.

Uberegui E, Hall M, Lorenzo O, Schroder WP, Balsera M (2015) An *Arabidopsis* soluble chloroplast proteomic analysis reveals the participation of the Executer pathway in response to increased light conditions. Journal of Experimental Botany 66:2067-2077.

Usuda H (1988) Adenine Nucleotide Levels, the Redox State of the NADP System, and Assimilatory Force in Nonaqueously Purified Mesophyll Chloroplasts from Maize Leaves under Different Light Intensities. Plant physiology 88:1461-1468.

Vale RD (2000) AAA Proteins: Lords of the Ring. Cell Biology 150:F13–F19.

Verdecia MA, Larkin RM, Ferrer J-L, Riek R, Chory J, Noel JP (2005) Structure of the Mg-Chelatase Cofactor GUN4 Reveals a Novel Hand-Shaped Fold for Porphyrin Binding. PLoS Biology 3:0777-0789.

Verlhac J, A G, I K (1984) Water-soluble porphyrins and metalloporphyrins as photosensitizers in aerated aqueous solutions. I. Detection and determination of quantum yield of formation of singlet oxygen. New Journal of Chemistry 401–406.

Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. Journal of applied crystallography 36:860-864.

Wagner D, Przybyla D, op den Camp R, Kim C, Landgraf F, Lee KP, Würsch M, Laloi C, Nater M, Hideg E, Apel K (2004) The Genetic Basis of Singlet Oxygen–Induced Stress Responses of *Arabidopsis thaliana*. Science 306:1183-1185.

Walker C, Hupp L, Weinstein J (1992) Activation and stabilization of Mg-chelatase activity by ATP as revealed by a novel in vitro continuous assay. Plant physiology and Biochemistry Journal 30:263-269.

Walker CJ, Weinstein JD (1991a) Further Characterization of the Magnesium Chelatase in Isolated Developing Cucumber Chloroplasts1. Plant physiology 95:1189-1196.

Walker CJ, Weinstein JD (1991b) In vitro assay of the chlorophyll biosynthetic enzyme Mg-chelatase: Resolution of the activity into soluble and membranebound fractions. Biochemistry 88:5789-5793.

Walker CJ, Weinstein JD (1994) The magnesium-insertion step of chlorophyll biosynthesis is a two-stage reaction. Biochem 299:277-284.

Walker CJ, Willows RD (1997) Mechanism and regulation of Mg-chelatase. Biochemistry Journal 8:321-333.

Wang W-Y, Huang D-D, Stachon D, Gough SP, Kannangara CG (1984) Purification, Characterization, and Fractionation of the δ-Aminolevulinic Acid Synthesizing Enzymes from Light-Grown *Chlamydomonas reinhardtii* Cells. Plant physiology 74:569-575.

Weber PC (1997) Overview of Protein Crystallization Methods Methodes in enzymology, Academic press,Inc 276:13-22.

Wettstein Dv, Gough S, Kannangara CG (1995) Chlorophyll Biosynthesis. Plant Cell 7:1039-1057.

Wilde A, Mikolajczyk S, Alawady A, Lokstein H, Grimm B (2004) The gun4 gene is essential for cyanobacterial porphyrin metabolism. FEBS Lett 571:119-123.

Willows RD (2003) Biosynthesis of chlorophylls from protoporphyrin IX. Natural Product Reports 20:327-341.

Willows RD, Beale SIJ (1998) Magnesium Chelatase: Heterologous Expression of the *Rhodobacter capsulatus* BchI, -D, And -H Genes that Encode Mg Chelatase Subunits and Characterization of the Reconstituted Enzyme. Biol Chem 273:34206-34213.

Willows RD, Gibson LCD, Kanangara CG, Hunter CN, Westttstein Dv (1996) Three separate proteins constitute the magnesium chelatase of *Rhodobacter sphaeroides*. European Journal of Biochemistry 235:438-443.

Willows RD, Hansson A, Birch D, Al-Karadaghi S, Hansson M (2004) EM single particle analysis of the ATP-dependent BchI complex of magnesium chelatase: an AAA+ hexamer. Journal of Structural Biology 146:227-233.

Willows RD, Hansson M (2003) Mechanism, structure, and regulation of magnesium chelatase. . *In* KM Kadish, KM Smith, R Guilard, eds, The Porphyrin Handbook, Chlorophylls and bilins:biosynthesis, synthesis, and degradation Vol 13. Academic Press, Sydney,:1-47.

Willows RD, Mats Hansson, Samuel I. Beale, Laurbergd M, Salam Al-Karadaghi (1999) Crystallization and preliminary X-ray analysis of the *Rhodobacter capsulatus* magnesium chelatase BchI subunit. Biological Crystallography D55:689-691.

Wilmanns M, Weiss MS (2005) Molecular Crystallography. In: Encyclopedia of Condensed Matter Physics, pp 453-458 Oxford: Elsevier.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. Acta Crystallographica Section D 67:235-242.

Woodson JD, Perez-Ruiz JM, Chory J (2011) Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. Current biology : CB 21:897-903.

Wu FQ, Xin Q, Cao Z, Liu ZQ, Du SY, Mei C, Zhao CX, Wang XF, Shang Y, Jiang T, Zhang XF, Yan L, Zhao R, Cui ZN, Liu R, Sun HL, Yang XL, Su Z, Zhang DP (2009) The Magnesium-Chelatase H Subunit Binds Abscisic Acid and Functions in Abscisic Acid Signaling: New Evidence in *Arabidopsis*. Plant physiology 150:1940-1954.

Xiong J, Bauer CE (2002) Complex evolution of photosynthesis. Annu Rev Plant Biol 53:503-521.

Zeilstra-Ryalls JH, Kaplan S (2004) Oxygen intervention in the regulation of gene expression: the photosynthetic bacterial paradigm. Cellular and molecular life sciences : 61:417-436.

Zhou S, Sawicki A, Willows RD, 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.

Appendices

Appendix 1. Synopsis of publication I: Inducing the oxidative stress response in *Escherichia coli* improves the quality of a recombinant protein: Magnesium chelatase ChlH

This paper describes the ChlH protein expression and purification method from *C. reinhardtii*, *Oryza sativa,* and *Hordeum vulgare.* Size exclusion chromatography result of all three eukaryotic species showed that ChlH protein forms multimeric and oligomeric forms. The monomeric form was active in Mg-chelatase activity assay. However, the oligomeric form was found to be oxidatively damaged as it was formed because of oxidative conditions during expression and purification.

Appendix 2. Publication 1

Protein Expression and Purification 101 (2014) 61–67

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Inducing the oxidative stress response in *Escherichia coli* improves the quality of a recombinant protein: Magnesium chelatase ChlH

André H. Müller [a,b,1,3], Artur Sawicki [a,3], Shuaixiang Zhou [c], Shabnam Tarahi Tabrizi [a], Meizhong Luo [c], Mats Hansson [b,2], Robert D. Willows [a,*]

[a] Department of Chemistry and Biomolecular Sciences, Macquarie University, NSW 2109, Australia
[b] Carlsberg Laboratory, Gamle Carlsberg Vej 10, 1799 Copenhagen V, Denmark
[c] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

ARTICLE INFO

Article history:
Received 8 May 2014
and in revised form 27 May 2014
Available online 12 June 2014

Keywords:
Oxidative stress
Protoporphyrin
Chlorophyll biosynthesis
Chelatase
Magnesium

ABSTRACT

The ~150 kDa ChlH subunit of magnesium chelatase from *Oryza sativa*, *Hordeum vulgare* and *Chlamydomonas reinhardtii* have been heterologously expressed in *Escherichia coli*. The active soluble protein is found as both a multimeric and a monomeric form. The multimeric ChlH appears to be oxidatively damaged but monomer production is favoured in growth conditions that are known to cause an oxidative stress response in *E. coli*. Inducing an oxidative stress response may be of general utility to improve the quality of proteins expressed in *E. coli*. The similar responses of ChlH's from the three different species suggest that oligomerization of oxidatively damaged ChlH may have a functional role in the chloroplast, possibly as a signal of oxidative stress or damage.

Introduction

Magnesium chelatase is one of the enzymes located at the branchpoint between hemes and chlorophylls in the synthesis of tetrapyrroles. The enzyme inserts magnesium into protoporphyrin IX in an ATP hydrolysis dependent manner. The product, magnesium protoporphyrin IX, is the first committed intermediate in the synthesis of chlorophyll or bacteriochlorophyll. Magnesium chelatase consists of three subunits, BchI/ChlI (40 kDa), BchD/ChlD (70 kDa) and BchH/ChlH (150 kDa) [1] with the prefixes Bch and Chl used to distinguish the proteins from bacteriochlorophyll or chlorophyll synthesizing organisms respectively. The BchI/ChlI and BchD/ChlD subunits form a double hexameric AAA+ motor complex, $BchI_6BchD_6$, and show conformational rearrangements depending on the nucleotide present [2]. The BchH/ChlH subunit has no ATPase activity but binds the substrate and product of the reaction [3,4]. This lead to the proposal of a mechanism where ATP is bound and hydrolyzed by I_6D_6 causing a conformational change in BchH/ChlH that subsequently drives chelation of magnesium into protoporphyrin IX [5]. When studying the kinetics of the reaction the BchH/ChlH and BchI/ChlI subunits act like "substrates" in the magnesium chelatase reaction with the BchD/ChlD subunit treated as the "enzyme" [6,7].

In addition to its direct role in chlorophyll synthesis, ChlH has further been implicated in plastid to nucleus signalling as point mutants in *chlH* have a so-called *gun* (genomes *un*coupled) phenotype with disrupted control of nuclear gene expression [8]. The ChlH subunit has also been suggested to be an abscisic acid receptor [9–11]. Although its direct receptor role has been challenged [12,13], it does still appear to be involved in some way in abscisic acid signaling [13].

In chlorophyll synthesizing organisms, an accessory protein called GUN4 is required for optimal magnesium chelatase activity. GUN4 was discovered in a mutant screen of plants with defects in plastid control over nuclear gene expression [8]. Both cyanobacterial and eukaryotic GUN4 binds protoporphyrin IX and magnesium protoporphyrin IX, interacts with ChlH and stimulates the activity of magnesium chelatase [14–17].

Here we report on the heterologous expression of ChlH from three different eukaryotic organisms, each giving a monomeric and a multimeric state by size exclusion chromatography. The monomeric state gave consistent specific activity in enzymatic assays while the multimeric state varied from batch to batch. The ratio of these two states depended primarily on oxidative conditions during induction of expression.

* Corresponding author. Tel.: +61 298508146; fax: +61 298508313.
E-mail address: robert.willows@mq.edu.au (R.D. Willows).
[1] Present address: Department of Structural Biology, Stanford University School of Medicine, D143, Stanford, CA 94305, USA.
[2] Present address: Copenhagen Plant Science Center, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Copenhagen, Denmark.
[3] These authors contributed equally to the work.

http://dx.doi.org/10.1016/j.pep.2014.06.004
1046-5928/

Materials and methods

Expression clones of ChlH from Chlamydomonas reinhardtii, Hordeum vulgare and Oryza sativa

Expression constructs for ChlH used in this work were prepared from cDNA of *H. vulgare* cultivar Svalöfs Bonus, *O. sativa l. japonica Nipponbare* and *C. reinhardtii* cc124. The *O. sativa l. japonica Nipponbare* and *H. vulgare* expression plasmids, pET28a-Os-ChlH and pET15bXanF respectively, have been described previously [12,17]. The *Chlamydomonas* expression construct pET28a-Cr-ChlH was constructed by cloning the PCR product of the *chlH* from cDNA (Accession XM_001700843 between 141 and 4416 bp) between the BamHI and EcoRI sites of pET28a. The *Chlamydomonas*, *H. vulgare* and *O. sativa* proteins are differentiated by the prefixes "Cr", "Hv" and "Os" respectively.

Expression of ChlH

The expression vectors described above were used to transform *Escherichia coli* strains BL21(DE3) Star, or BL21(DE3) pLysS or BL21(DE3) Rosetta 2 (Novagen). All three strains yielded soluble expressed protein. 20% glycerol stocks were prepared from single colony transformants from solid LB medium (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, 20 g/l agar; with appropriate antibiotics: either 50 mg/l kanamycin or 100 mg/l ampicillin; and 20 mg/l chloramphenicol when using pLysS or Rosetta strains) grown to mid-late log phase. The glycerol stock (10 μL) was spread on LB agar containing the appropriate antibiotics and colonies grew overnight at 37 °C. Fresh LB (10 mL) was added to the plate and slowly shaken at 150 rpm for 5 min to resuspended the colonies. The resuspended colonies were used to inoculate 250 mL–1 L of LB media with antibiotics to OD_{600} = 0.05. After growth to OD_{600} = 0.6 at 15 °C (Os- and Cr-ChlH) or 30 °C (Hv-ChlH), the protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside to 0.1 mM and incubated for 24 h at 15 °C.

For anoxic expressions, a 30 mL culture in LB (with ampicillin and chloramphenicol) was grown from OD_{600} = 0.05 to OD_{600} = 1 at 30 °C (shaken at 200 rev. min^{-1}) and then diluted to 300 ml with LB (with 10 mM glucose and antibiotics) in a 300 ml screw cap bottle containing a magnetic stirring flea. The bottle was capped tightly with no air and stirred for 1 h at 30 °C, then induced by addition of 0.1 mM IPTG and stirred for 24 h at 18 °C. Other variations to expression conditions are as described in the results section.

Cells were harvested (10,000×*g*, 15 min, 4 °C), washed in binding buffer (50 mM Tricine-NaOH pH 8.0, 250 mM NaCl, 50 mM $MgCl_2$, 20 mM imidazole) and pelleted again (10,000×*g*, 15 min, 4 °C) to be resuspended in binding buffer for lysis by two runs through a French press. Soluble and insoluble lysis fraction were separated by centrifugation (15,000×*g*, 15 min, 4 °C).

Purification of ChlH

ChlH was purified by immobilized metal ion affinity chromatography (IMAC)[4] using a 5 mL HisTrap column using Ni^{2+} as the immobilized metal ion (GE Healthcare). The supernatant was loaded onto the column and washed with 10 volumes wash buffer (50 mM Tricine-NaOH pH 8.0, 250 mM NaCl, 50 mM $MgCl_2$, 40 mM imidazole) and the protein was eluted with elution buffer (50 mM Tricine-NaOH pH 8.0, 250 mM NaCl, 50 mM $MgCl_2$, 250 mM imidazole). The peak elution fractions were determined by Bradford Reagent (BioRad) and pooled. The yield of ChlH per liter of culture at this purification stage ranged from a low yield of 0.5 mg/L when methylviologen was added during expression to a high yield of 15 mg/L with no additives. The ChlH was either further purified and buffer exchanged directly by gel filtration chromatography as detailed below or the samples were buffer exchanged on PD-10 desalting columns (GE Healthcare) into an imidazole-free buffer (50 mM Tricine-NaOH pH 8.0, 250 mM NaCl, 50 mM $MgCl_2$, 2 mM dithiothreitol). The buffer exchanged ChlH was concentrated to between 5–10 mg/mL using a 30 K Amicon centrifugal filter device (Millipore).

Deviations from this standard procedure are detailed in the results.

Gel filtration chromatography

A SMART system (Pharmacia) or an ÄKTA system (GE Healthcare) system was used for analytical size exclusion chromatography with Superose 12 (3.2/300) column. 50 μL samples were run on these Superose columns in a buffer containing 15 mM Tris–HCl, pH 7.5, 150 mM NaCl, 2 mM dithiothreitol, 1 mM EDTA.

For larger scale purifications an ÄKTA system (GE Healthcare) with a Superdex 200 (10/300) column was used with 50 mM Tricine-NaOH, pH 8.0, 250 mM NaCl, 50 mM $MgCl_2$, 2 mM dithiothreitol. This was later modified to 50 mM Tricine-NaOH, pH 8.0, 250 mM NaCl, 2 mM dithiothreitol to allow for better control of the $MgCl_2$ concentration in assays.

Magnesium chelatase activity measurements

The assay utilized *O. sativa* ChlI, ChlD and GUN4 subunits as described previously [17]. Briefly, the assay consisted of 50 mM Tricine-NaOH, pH 8.0, 2 mM dithiothreitol, 15 mM $MgCl_2$ and 4 mM ATP with 10 nM Os-ChlD, 100 nM Os-ChlI and 500 nM Os-GUN4, 500 nM protoporphyrin IX and 500 nM ChlH. The reaction was started by mixing equal volumes of H-premix (1000 nM ChlH, 1000 nM GUN4 and 1000 nM protoporphyrin IX mixed in 50 mM Tricine-NaOH, pH 8.0, 2 mM dithiothreitol, 15 mM $MgCl_2$ and 4 mM ATP, preincubated at ambient temperature at 30 °C for 30 min prior to starting the assay) with ID premix (50 mM Tricine-NaOH, pH 8.0, 2 mM dithiothreitol, 15 mM $MgCl_2$ and 4 mM ATP with 20 nM Os-ChlD, 200 nM Os-ChlI). The fluorescence emission at 600 nm upon excitation at 420 nm was monitored in a BMG-Pherstar plate reader at 15 s intervals for 30 min. The maximum rate over a 2 min interval was determined from the linear part of the progress curves and activity units are in nmoles of magnesium protoporphyrin formed per minute per nmole of ChlD.

Results

Expression, purification and assay of ChlH proteins

The ChlH proteins were purified as described in the methods. Although the proteins appeared to be more than 90% pure as estimated by SDS–PAGE (Fig. 1) the gel filtration profiles on Superdex (Fig. 2) and Superose 12 (see an example in Fig. 3) indicated that the ChlH proteins had two major components - a multimeric component eluting at 8 mL on Superdex 200 and 0.9 mL on Superose 12 and a monomeric component eluting at ~12.5 mL on Superdex 200 and 1.25 mL on Superose 12. (see Figs. 2 and 3A) The major peak at 0.9 mL is close to the exclusion limit determined with blue dextran for Superose 12 of 0.85 mL. The smaller peak at 1.25 mL eluted at the same position as standard IgG indicating a monomeric Hv-ChlH.

[4] *Abbreviations used:* IMAC, immobilized metal ion affinity chromatography.

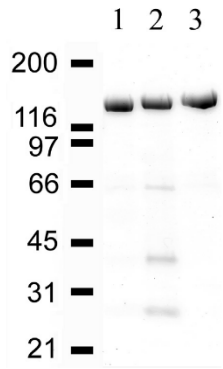


Fig. 1. SDS–PAGE of IMAC purified ChlH proteins. Position of molecular weight markers in kDa shown on the left. Lane 1 Hv-ChlH, Lane 2 Os-ChlH, Lane 3 Cr-ChlH.

Enzymatic assays of ChlH were routinely done at 0.5 μM and 1 μM of ChlH based on the ChlH protein concentration with *O. sativa* ChlI, ChlD and GUN4 subunits. The reason for choosing these ChlH protein concentrations for routine assays was to provide a meaningful comparison of ChlH batches or column fractions. If all of the ChlH in the sample was active in the assay then 0.5 μM and 1 μM would give activities close to V_{max}, based on the data in Table 1 and in Supplementary Figs. 1 and 2. Using these criteria it was found that the magnesium chelatase activity of the ChlH monomer was, batch to batch, very consistent with activities within 20% to the V_{max} values shown in Table 1. In contrast the ChlH assayed after IMAC was at most 50% of the purified monomeric ChlH and was often much lower than this (See Supplementary Fig. 2 for example). One possibility was that the multimeric ChlH's would be inactive, but somewhat surprisingly the multimeric ChlH proteins often had enzymatic activity. One ChlH multimer preparation had 25% of the activity of the monomeric ChlH subunits, although this was extremely variable and in some preparations no activity was detected for the purified multimeric ChlH.

Oligomers of ChlH

H. vulgare, *O. sativa* and *C. reinhardtii* ChlH were all produced as multimeric and monomeric soluble forms (Fig. 2), with a small amount of a dimeric form. However, in addition to the soluble monomer and multimeric forms of Os- and Cr-ChlH, insoluble misfolded Os- and Cr- ChlH was also produced. The production of this insoluble Os- and Cr- ChlH could be minimized if these *E. coli* cultures were grown and protein expression induced at 15 °C instead of growing the cultures at 30 °C and shifting to lower temperature for induction. In addition, non-reducing SDS–PAGE indicated only a small percentage of multimeric ChlH were linked via disulphide bonds (see Supplementary Fig. 3).

Sirijovski et al. [4] showed that the ChlH orthologue from *Rhodobacter capsulatus*, BchH, aggregated upon light exposure. To test if this was occurring with ChlH, separated Hv-ChlH monomers and multimers were treated in light for 2 h on ice. In contrast to the BchH, gel filtration of these light treated proteins showed both monomers and aggregates to be stable (Fig. 3B and C, respectively).

The Ni-affinity purified Hv-ChlH preparation was reddish, which indicated the presence of bound substrate protoporphyrin IX from the expression host. In the gel filtration elution profiles, the protoporphyrin IX-absorption at 405 nm was monitored in addition to protein absorbance at 280 nm. As the elution profiles showed a significant 405 nm absorbance with the multimers only, it was possible that ChlH bound protoporphyrin IX subsequently aggregated possibly in combination with light exposure. Protoporphyrin IX produces singlet oxygen upon light exposure, which might cause the aggregation of ChlH. To test this all steps of expression, lysis and purification were carried out in the dark. This purified ChlH contained almost exclusively aggregates (Fig. 3D),

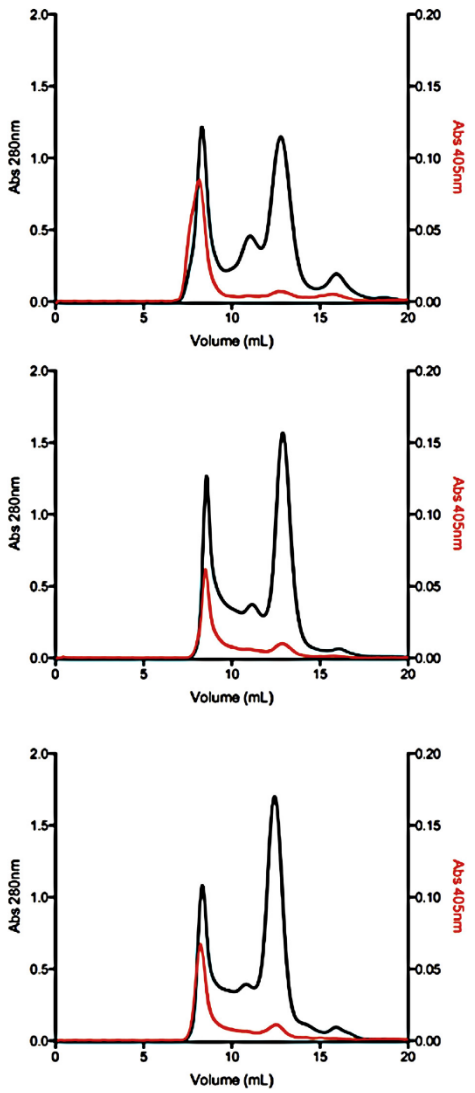


Fig. 2. Separation of IMAC affinity purified ChlH's (shown in Fig. 1) on Superdex 200 expressed under optimal final conditions with 2 μM protoporphyrin and 200 μEinsteins light exposure. Upper panel is Hv-ChlH, middle is Os-ChlH and lower is Cr-ChlH.

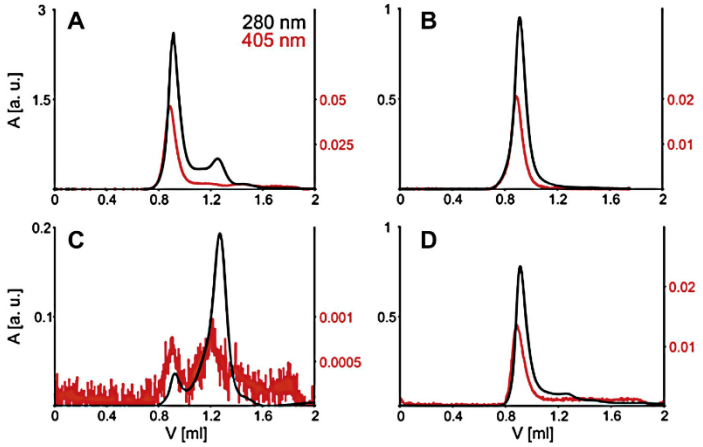


Fig. 3. Size exclusion chromatography of Hv-ChlH after Ni-affinity purification. Absorbance at 280 nm is in black, absorbance at 405 nm in red. (A) Hv-ChlH from expression culture grown under normal laboratory lighting (~45 μEinsteins). (B) Oligomeric Hv-ChlH exposed to 45 μEinsteins fluorescent light for 2 h on ice. (C). Monomeric ChlH exposed to 45 μEinsteins fluorescent light for 2 h on ice. (D) Hv-ChlH from normal expression culture incubated in complete darkness. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Comparative maximum magnesium chelatase activities of monomeric ChlH. V_{max} in nmol Mg-protoporphyrin IX/min/nmol OsChlD. (See Supplementary Figs. 1 and 2 for graphs of kinetic data).

Source of monomeric ChlH subunit in assay	$S_{0.5}$ (nM)	V_{max}	V_{max} %
Oryza sativa	300	1.0 ± 0.014	100
Chlamydomonas reinhardtii	nd[a]	0.017 ± 0.004[b]	1.7
Hordeum vulgare	150	0.17 ± 0.02	17

[a] Not determined.
[b] Due to the lower activity of Cr-ChlH in the assays a kinetic analysis was not possible. However, maximum activity was at ~0.5 μM and did not increase further at 1 μM of Cr-ChlH. The activity shown is for 0.5 μM Cr-ChlH in triplicate.

suggesting that light exposure was necessary for monomer production during expression.

Stability testing of IMAC purified Hv-ChlH, separated Hv-ChlH monomers and Hv-ChlH multimers (shown in Fig. 3) on wet ice for a week showed that there was no change in the gel filtration profile, indicating that the multimeric and monomeric states of ChlH were stable and did not interconvert once they had been purified. It was therefore something during expression or during purification in the *E. coli* extract that was causing the formation of multimers.

Shifting the proportion of monomeric and oligomeric forms of ChlH

To further test the effect of light exposure on the aggregation of ChlH, a standard expression culture and one with increased light exposure (200 μEinsteins) were grown, lysed and IMAC purified.

The light exposed expression (Fig. 4B) clearly showed a higher proportion of monomeric ChlH than the standard expression (Fig. 4A). It was noted that although the aggregates still dominated the 405 to 280 ratio of the aggregates was higher. This led us to test the addition of protoporphyrin IX (1 μM) to the expression culture at time of induction both in the presence and absence of light. The ratio of monomeric to multimeric ChlH in these protoporphyrin IX fed cultures increased but the combination of light (200 μEinsteins) and protoporphyrin IX resulted in the highest yield of monomers (Fig. 5A).

As protoporphyrin IX in the presence of oxygen produces singlet oxygen, methylviologen and methylene blue, which also produce ROS were also added to the expressing cultures to see the effect of ROS during induction. Methylviologen (Paraquat) produces superoxide anion radicals in the light and methylene blue also produces superoxide anions. In the presence of methylviologen or methylene blue the ratio of monomeric Hv-ChlH to multimeric Hv-ChlH increased (Fig. 5B and C), although the overall protein yield was lower possibly due to toxicity of these compounds or

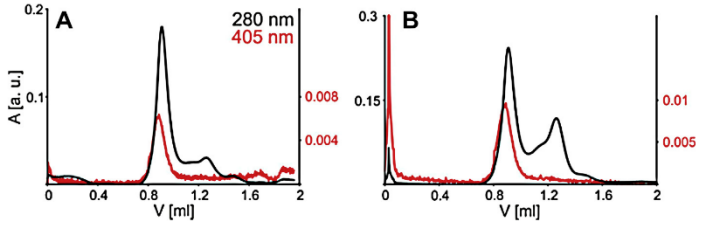


Fig. 4. Gel filtration of Hv-ChlH (Ni-affinity purified) expressed under different light conditions. Absorbance at 280 nm is in black, absorbance at 405 nm in red. (A) Hv-ChlH from standard expression culture. (B) Hv-ChlH from expression culture under increased light exposure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

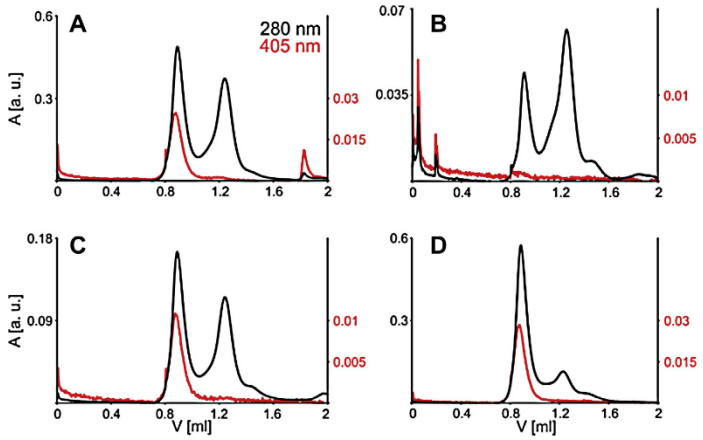


Fig. 5. Gel filtration of Hv-ChlH (Ni-affinity purified) expressed under different conditions. (A) Protoporphyrin IX 1 μM added at induction with 200 μEinsteins light exposure. (B) Methylviologen added to 1 mM at induction with 45 μEinsteins light. (C) Methylene blue to 12.5 μM added at induction with 45 μEinsteins light. (D) δ-Aminolaevulinic acid to 200 μM added at induction with 45 μEinsteins light.

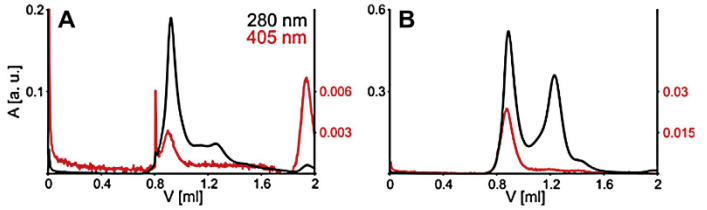


Fig. 6. Gel filtration of Hv-ChlH (Ni-affinity purified) expressed under anoxic conditions. (A) Hv-ChlH purified in normal buffers. (B) Hv-ChlH purified in the absence of dissolved oxygen (25 mg/l sodium dithionite in buffers during lysis and purification).

the global effects of ROS. In contrast, adding the protoporphyrin IX precursor δ-aminolaevulinic acid (ALA) to the cultures upon induction had very little effect on the ratio of monomer to multimer (compare Figs. 3A and 5D).

It thus appeared that the presence of ROS during expression in *E. coli* improved the ratio of monomeric to multimeric Hv-ChlH. Excess ROS cause an oxidative stress response in *E. coli* which results in increased production of enzymes such as SOD and catalase so the bacteria can cope with the increase in ROS species [18]. It was thus possible that it was this ROS stress response allowing the *E. coli* to better cope with the ROS being produced which increased the proportion of monomeric Hv-ChlH. To test this we expressed Hv-ChlH under strict anaerobic conditions and purified this anaerobically expressed Hv-ChlH in the presence and absence of oxygen. (Fig. 6).

Oligomeric forms dominated if Hv-ChlH was purified from anaerobic cultures in normal oxygen containing buffers (Fig. 6A). However, if oxygen was removed from the purification buffers by addition of sodium dithionite (25 mg/l), the ratio of monomers and aggregates was comparable to ChlH produced in the presence of protoporphyrin IX in light (compare Figs. 6B and 5A).

Discussion

The concomitant existence of two soluble forms of ChlH in overexpression cultures and the complex connection of their ratio to the incubation conditions was surprising. Dark grown cultures contained primarily oligomers and had low specific activity, while light grown ones had varying ratios of monomers and oligomers with generally higher specific activity.

The monomeric form of ChlH was readily separated from the oligomeric form of ChlH. The batch to batch activity of ChlH monomer was very consistent, with magnesium chelatase activities consistently 80–100% of the Vmax values shown in Table 1 when assayed at 0.5 and 1.0 μM ChlH. In contrast the specific magnesium chelatase activity of the oligomeric form of ChlH's varied seemingly randomly from trace activity to 25% of the V_{max} activities shown in Table 1. This suggests that the oligomer was damaged or modified compared to the monomer.

The ratio of oligomeric to monomeric forms was influenced by the presence of ROS. Light and the photodynamic compounds protoporphyrin IX, methylene blue and methylviologen cause singlet oxygen and/or superoxide anion radicals formation and more monomers as products of the overexpression. If the presence of these ROS were preventing the multimer formation, then growing *E. coli* under anaerobic conditions should yield multimer exclusively. However, that was not the case, as a substantial amount of monomer resulted as long as the dissolved oxygen concentration was minimized not only in the actual expression, but also in cell lysis and purification. Thus, both oxidative stress and absence of oxygen resulted in monomers of ChlH.

In order to explain these findings, the focus needs to be widened from ChlH and its oligomerization to the entire expression system. The presence of ROS causes *E. coli* to induce protective

mechanisms against oxidative stress [19–21]. The response produces enzymes like superoxide-dismutase and catalase that prevent accumulation of ROS. A possible scenario to explain the findings for ChlH expression emerges in which this cellular stress response also serves to protect the overexpression product ChlH. ChlH as such appears to be highly sensitive to oxidation by ROS and if ROS detoxifying mechanisms are not induced, the ChlH is oxidized and forms multimers. The high susceptibility of ChlH to oxidation during expression and purification compared to the high stability after purification indicate the oxidative sensitivity is both during translation and during the purification phase.

Monomers eluting from gel filtration have very low 405 absorbance indicating no protoporphyrin bound. This contrasts with BchH from *R. capsulatus* [4] which does have protoporphyrin bound and aggregates on light exposure. The increased susceptibility of BchH to light might be due to the presence of both the bound protoporphyrin and a FeS cluster with both being possible photosensitisers [22]. It is thus possible that the lack of protoporphyrin bound to the ChlH monomers is a factor in their light stability after purification.

The sensitivity of these three ChlH's to oxidation and subsequent multimerization raises the question: Does multimeric ChlH form *in vivo*? *In vivo* biochemical studies of ChlH have mainly been in *Arabidopsis thaliana*. Given the sequence similarity between the disparate ChlH proteins here and the *Arabidopsis* ChlH (At-ChlH) it is very likely that the At-ChlH oligomerizes in certain conditions in a fashion comparable to that described here with Os- Cr- and Hv-ChlH's. At-ChlH co-sediments with membranes [16] which may be interpreted as membrane association or could be sedimentation of multimer.

So if multimeric ChlH does form *in vivo* how does it form and what does it do? The protein GUN4 interacts with and mediates protoporphyrin IX binding to ChlH [23], and magnesium protoporphyrin IX release from ChlH [14]. It has also been suggested that GUN4 may sequester tetrapyrrole intermediates [24]. Thus a potential involvement of GUN4 is to prevent the oligomerization of ChlH. It could do this by sequestering the protoporphyrin IX from ChlH which would otherwise cause photodynamic oxidation and subsequent oligomerization of the ChlH. Thus, the multimeric form of ChlH would only form *in vivo* in conditions where protoporphyrin or magnesium protoporphyrin remains bound to ChlH causing oxidation and multimerisation of ChlH. This scenario could occur when the ratio of GUN4 relative to ChlH is low or when the concentration of protoporphyrin is high enough such that both ChlH and GUN4 have porphyrin bound. The presence of oxidized multimeric ChlH then signals a need for regulation of tetrapyrrole synthesis and other processes within the chloroplast.

ChlH from barley and Chlamydomonas was found to be active with ChlI, ChlD and GUN4 from rice. Interestingly, the $S_{0.5}$ value determined was lower than that for Os-ChlH although the Hill coefficient was significantly lower. The availability of this fully recombinant system has great advantages over the semi-recombinant system used before in which the constant quality of plastid lysate is of crucial importance [12]. A fully recombinant system allows more variation of assays and greater reproducibility, mix and match experiments involving subunits from different organisms are possible and limitations of the homologous magnesium chelatase of *R. capsulatus*, which has no Gun4 and poor solubility of BchD [5,6], can be overcome in this hybrid system.

Conclusions

Soluble multimers of ChlH are formed from monomers of ChlH due to reactive oxygen species. Protective mechanisms induced during an oxidative stress response in *E. coli* minimize this damage. The oligomeric states of ChlH could explain some of the characteristics of ChlH reported.

Acknowledgments

This work was supported by the Danish Research Council, Denmark, The Carlsberg Foundation, Denmark, The Research School for Biotechnology (FOBI), Denmark, the National Natural Science Foundation of China (Grant No. 30971748), China, and by a Macquarie University SNS grant, Australia.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.pep.2014.06.004.

References

[1] R.D. Willows, M. Hansson, Mechanism, structure and regulation of magnesium chelatase, in: K.M. Kadish, K. Smith, R. Guilard (Eds.), The Porphyrin Handbook II, Academic Press, 2003, pp. 1–48.
[2] J. Lundqvist, H. Elmlund, R.P. Wulff, L. Berglund, D. Elmlund, C. Emanuelsson, H. Hebert, R.D. Willows, M. Hansson, M. Lindahl, S. Al-Karadaghi, ATP-induced conformational dynamics in the AAA+ motor unit of magnesium chelatase, Structure 18 (2010) 354–365.
[3] N. Sirijovski, J. Lundqvist, M. Rosenback, H. Elmlund, S. al-Karadaghi, R.D. Willows, M. Hansson, Substrate-binding model of the chlorophyll biosynthetic magnesium chelatase BchH subunit, J. Biol. Chem. 283 (2008) 11652–11660.
[4] N. Sirijovski, U. Olsson, J. Lundqvist, S. al-Karadaghi, R. Willows, M. Hansson, ATPase activity associated with the magnesium chelatase H-subunit of the chlorophyll biosynthetic pathway is an artifact, Biochem. J. 400 (2006) 477–484.
[5] E. Axelsson, J. Lundqvist, A. Sawicki, S. Nilsson, I. Schroder, S. al-Karadaghi, R. Willows, M. Hansson, Recessiveness and dominance in barley mutants deficient in mg-chelatase subunit D, an AAA protein involved in chlorophyll biosynthesis, Plant Cell 18 (2006) 3606–3616.
[6] A. Sawicki, R.D. Willows, Kinetic analyses of the magnesium chelatase provide insights into the mechanism, structure, and formation of the complex, J. Biol. Chem. 283 (2008) 31294–31302.
[7] J. Viney, P. Davison, C. Hunter, J. Reid, Direct measurement of metal-ion chelation in the active site of the AAA+ ATPase magnesium chelatase, Biochemistry 46 (2007) 12788–12794.
[8] N. Mochizuki, J. Brusslan, R. Larkin, A. Nagatani, J. Chory, Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction, Proc. Natl. Acad. Sci. USA 98 (2001) 2053–2058.
[9] Y. Shen, X. Wang, F. Wu, S. Du, Z. Cao, Y. Shang, X. Wang, C. Peng, X. Yu, S. Zhu, R. Fan, Y. Xu, D. Zhang, The Mg-chelatase H subunit is an abscisic acid receptor, Nature 443 (2006) 823–826.
[10] F.-Q. Wu, Q. Xin, Z. Cao, Z.-Q. Liu, S.-Y. Du, C. Mei, C.-X. Zhao, X.-F. Wang, Y. Shang, T. Jiang, X.-F. Zhang, L. Yan, R. Zhao, Z.-N. Cui, R. Liu, H.-L. Sun, X.-L. Yang, Z. Su, D.-P. Zhang, The magnesium-chelatase H subunit binds abscisic acid and functions in abscisic acid signaling: new evidence in Arabidopsis, Plant Physiol. 150 (2009) 1940–1954.
[11] Y. Shang, L. Yan, Z.Q. Liu, Z. Cao, C. Mei, Q. Xin, F.Q. Wu, X.F. Wang, S.Y. Du, T. Jiang, X.F. Zhang, R. Zhao, H.L. Sun, R. Liu, Y.T. Yu, D.P. Zhang, The Mg-chelatase H subunit of Arabidopsis antagonizes a group of WRKY transcription repressors to relieve ABA-responsive genes of inhibition, Plant Cell 22 (2010) 1909–1935.
[12] A.H. Müller, M. Hansson, The barley magnesium chelatase 150-kd subunit is not an abscisic acid receptor, Plant Physiol. 150 (2009) 157–166.
[13] T. Tsuzuki, K. Takahashi, S. Inoue, Y. Okigaki, M. Tomiyama, M.A. Hossain, K. Shimazaki, Y. Murata, T. Kinoshita, Mg-chelatase H subunit affects ABA signaling in stomatal guard cells, but is not an ABA receptor in *Arabidopsis thaliana*, J. Plant Res. 124 (2011) 527–538.
[14] P.A. Davison, C.N. Hunter, Abolition of magnesium chelatase activity by the gun5 mutation and reversal by Gun4, FEBS Lett. 585 (2011) 183–186.
[15] R.M. Larkin, J.M. Alonso, J.R. Ecker, J. Chory, GUN4, a regulator of chlorophyll synthesis and intracellular signaling, Science 299 (2003) 902–906.
[16] N.D. Adhikari, J.E. Froehlich, D.D. Strand, S.M. Buck, D.M. Kramer, R.M. Larkin, GUN4-porphyrin complexes bind the ChlH/GUN5 subunit of Mg-chelatase and promote chlorophyll biosynthesis in arabidopsis, Plant Cell 23 (2011) 1449–1467.
[17] S. Zhou, A. Sawicki, R.D. Willows, M. Luo, 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 (2012) 205–210.
[18] H.M. Hassan, I. Fridovich, Intracellular production of superoxide radical and of hydrogen peroxide by redox active compounds, Arch. Biochem. Biophys. 196 (1979) 385–395.

[19] E. Cabiscol, J. Tamarit, J. Ros, Oxidative stress in bacteria and protein damage by reactive oxygen species, Int. Microbiol. 3 (2000) 3–8.
[20] L.F. Agnez-Lima, P.D. Mascio, B. Demple, C.F.M. Menck, Singlet molecular oxygen triggers the soxRS regulon of *Escherichia coli*, Biol. Chem. 382 (2005) 1071–1075.
[21] S.Y. Kim, E.J. Kim, J.-W. Park, Control of singlet oxygen-induced oxidative damage in *Escherichia coli*, J. Biochem. Mol. Biol. 35 (2002) 353–357.
[22] N. Sirijovski, F. Mamedov, U. Olsson, S. Styring, M. Hansson, *Rhodobacter capsulatus* magnesium chelatase subunit BchH contains an oxygen sensitive iron-sulfur cluster, Arch Microbiol. 188 (2007) 599–608.
[23] P.A. Davison, H.L. Schubert, J.D. Reid, C.D. Iorg, A. Heroux, C.P. Hill, C.N. Hunter, Structural and biochemical characterization of Gun4 suggests a mechanism for its role in chlorophyll biosynthesis, Biochemistry 44 (2005) 7603–7612.
[24] E. Peter, B. Grimm, GUN4 is required for posttranslational control of plant tetrapyrrole biosynthesis, Mol. Plant 2 (2009) 1198–1210.

Appendix 3. Synopsis of publication II: Structure of GUN4 from *Chlamydomonas reinhardtii*

This publication is all about the crystal structure of GUN4 protein from *C. reinhardtii*, which has been deposited to Protein Data Bank (PDB). Also, the PDB validation report is presented in appendix 9 of this thesis.

Appendix 4. Publication II

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structure of GUN4 from *Chlamydomonas reinhardtii*

Shabnam Tarahi Tabrizi,[a] David B. Langley,[b] Stephen J. Harrop,[c] Anthony P. Duff[d] and Robert D. Willows[a]*

[a]Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW 2109, Australia, [b]Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW 2010, Australia, [c]Australian Synchrotron, 800 Blackburn, Clayton, Melbourne, VIC 3168, Australia, and [d]ANSTO, Lucas Heights, NSW 2234, Australia. *Correspondence e-mail: robert.willows@mq.edu.au

Received 28 May 2015
Accepted 26 June 2015

Edited by T. C. Terwilliger, Los Alamos National Laboratory, USA

Keywords: genomes uncoupled; chlorophyll biosynthesis; plastid signalling; magnesium chelatase; protoporphyrin.

PDB reference: GUN4, 4ykb

Supporting information: this article has supporting information at journals.iucr.org/f

The genomes uncoupled 4 (GUN4) protein stimulates chlorophyll biosynthesis by increasing the activity of Mg-chelatase, the enzyme that inserts magnesium into protoporphyrin IX (PPIX) in the chlorophyll biosynthesis pathway. One of the roles of GUN4 is in binding PPIX and Mg-PPIX. In eukaryotes, GUN4 also participates in plastid-to-nucleus signalling, although the mechanism for this is unclear. Here, the first crystal structure of a eukaryotic GUN4, from *Chlamydomonas reinhardtii*, is presented. The structure is in broad agreement with those of previously solved cyanobacterial structures. Most interestingly, conformational divergence is restricted to several loops which cover the porphyrin-binding cleft. The conformational dynamics suggested by this ensemble of structures lend support to the understanding of how GUN4 binds PPIX or Mg-PPIX.

1. Introduction

In plants and algae, tetrapyrrole biosynthesis occurs in the chloroplast, and the genes for light-harvesting proteins are encoded in the nuclear genome (Formighieri *et al.*, 2012). Retrograde signalling pathways between these two organelles were discovered through the *Arabidopsis thaliana* genomes uncoupled (GUN) mutants (Susek *et al.*, 1993; Willows, 2003). Five *gun* mutants have been reported (*gun*1–*gun*5), among which four (*gun*2–*gun*5) are on the same retrograde signalling pathway and encode proteins that function in tetrapyrrole metabolism. *gun*2 codes for haem oxygenase, *gun*3 for phytochromobilin synthase and *gun*5 for the H subunit of Mg-chelatase. Mg-chelatase is the enzyme complex which inserts the Mg^{2+} ion into protoporphyrin IX (PPIX), generating Mg-PPIX, in an ATP-dependent manner, and which constitutes the precursor for chlorophyll biosynthesis. Mg-chelatase consists of three protein subunits: ChlI/BchI (38–42 kDa), ChlD/BchD (60–74 kDa) and BchH/ChlH (140–150 kDa; Gorchein *et al.*, 1993; Jensen *et al.*, 1996; Papenbrock *et al.*, 1997; Wang *et al.*, 1974). These subunits are conserved from bacteria to higher plants and are responsible for different functions. The ChlI/BchI subunits are members of the AAA^+ superfamily with ATPase activity required for catalysis. They bind Mg^{2+} and ATP and also function as molecular chaperones for the ChlD/BchD subunits. The ChlD/BchD subunits appear to have an inactive AAA^+ ATPase. The ChlH/BchH subunits are the largest within the complex and form the core of the magnesium chelatase enzyme, and they bind both PPIX and Mg-PPIX.

The *gun*4 gene was the first locus to be identified in the screen for genome uncoupled mutants (Susek *et al.*, 1993;

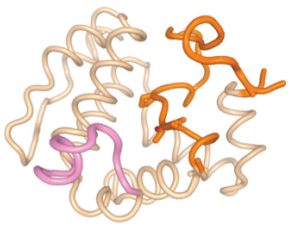


Table 1
Protein-production information.

Source organism	*C. reinhardtii*
Expression vector	pET-28a
Expression host	*E. coli* BL21(DE3)
Complete amino-acid sequence of the construct produced	MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDHP-FTMAMRVTVAAGKLDSVSLFGGDTASLMGGSQ-TVEKKKSGKEAVMEVQLSSTAGIDYTVLRDHL-ANGEFREAEDETRALLIKLAGPEAVKRNWVYF-TEVKNISVTDFQTLDNLWKASSNNKFGYSVQK-EIWVQNQKRWPKFFKQIDWTQGENNNYRKWPM-EFIYSMDAPRGHLPLTNALRGTQLFQAIMEHP-AFEKSSTAKTLDQKAAEAAGRTQSLF

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	24-well (Cryschem plate, Hampton Research)
Temperature (°C)	25
Protein concentration (mg ml^{-1})	35
Buffer composition of protein solution	20 m*M* tricine pH 8.0, 2 m*M* β-mercaptoethanol
Composition of reservoir solution	1.0 *M* ammonium citrate tribasic pH 7.0, 0.1 *M* bis-tris propane pH 7.0
Volume and ratio of drop	2 µl; 1:1 protein:reservoir
Volume of reservoir (µl)	500

Willows & Hansson, 2003), and it has since been shown to be involved in both retrograde signalling and post-translational regulation of tetrapyrrole biosynthesis (Larkin *et al.*, 2003; Peter & Grimm, 2009). The *gun*4 gene product (GUN4) has also been shown to be required for optimal Mg-chelatase activity through its interaction with the ChlH subunit (Davison *et al.*, 2005; Verdecia *et al.*, 2005). In doing so, it acts as a regulatory subunit and facilitates delivery of the PPIX substrate to ChlH, and also binds the product (Mg-PPIX; Chen *et al.*, 2015; Davison *et al.*, 2005; Larkin *et al.*, 2003; Verdecia *et al.*, 2005). Two crystal structures of GUN4 proteins have been elucidated from the cyanobacteria *Synechocystis* (PDB entries 1y6i, 4xkb and 4xkc; Verdecia *et al.*, 2005; Chen *et al.*, 2015) and *Thermosynechococcus elongatus* (PDB entries 1z3x and 1z3y; Davison *et al.*, 2005). The structures with PDB codes 4xkb and 4xkc are with porphyrin bound. Both are predominantly α-helical proteins comprised of a smaller N-terminal domain (five α-helices) connected to a larger C-terminal domain (eight α-helices). Structural conservation is stronger in the C-terminal domains, consistent with the high sequence similarity of the C-terminal domains among GUN4 family members across a broad spectrum of species. The N-terminal domain is more variable and is missing in some prokaryotic family members (Davison *et al.*, 2005; Verdecia *et al.*, 2005). Here, we report the crystal structure of GUN4 from *Chlamydomonas reinhardtii* at 3.5 Å resolution and hence reveal the first structural details of an eukaryotic GUN4 protein.

2. Methods and materials

2.1. Protein expression and purification

The GUN4 gene from *C. reinhardtii* was cloned from cDNA into expression vector pET-28a (Merck–Novagen) and was expressed in *Escherichia coli* strain BL21(DE3) as a fusion with an N-terminal poly-His tag (Table 1). The IPTG-induced cells were grown at 37°C in LB medium. His_6-tagged GUN4 was purified by immobilized Ni^{2+}-affinity chromatography using a 5 ml HisTrap FF column (GE Healthcare). The supernatant from cleared lysate was loaded onto the column and washed with ten volumes of wash buffer (20 m*M* Tris–HCl pH 7.9, 0.5 *M* NaCl, 60 m*M* imidazole) and the immobilized proteins were eluted in elution buffer (80 m*M* Tris–HCl pH 7.9, 0.5 *M* NaCl, 1 *M* imidazole). The peak protein fractions were determined using the Bradford reagent (Bio-Rad). Proteins in pooled fractions were immediately desalted using a PD-10 column (GE Healthcare) to a final buffer consisting of 20 m*M* tricine–NaOH pH 8.0, 2 m*M* β-mercaptoethanol. The GUN4 samples were concentrated to 20–35 mg ml^{-1} using a 10K Amicon centrifugal filter device (Millipore) prior to crystallization experiments.

2.2. Crystallization and data collection

The GUN4 protein was screened for crystallization using a robotic liquid-handling dispenser (Phoenix, Art Robbins Instruments) at the University of New South Wales, Sydney, Australia. Crystals were grown using the vapour-diffusion technique and a sitting-drop format, in which equal volumes of protein and well solutions were combined in total drop volumes of 1–2 µl. Crystallization was monitored by a microscope at various time intervals. Numerous crystallization hits (17) were observed from these sparse-matrix screens. The small (100 µm in the longest dimension) rice-grain-like hexagonal crystals used in X-ray diffraction experiments appeared only after a prolonged period of room-temperature incubation, 44 weeks after the trays were established. The crystallization condition was solution B5 from the SaltRx HT screen (Hampton Research), which consisted of 1.0 *M* ammonium citrate tribasic pH 7.0, 0.1 *M* bis-tris propane pH 7.0. The protein for this sample was originally supplied at 35 mg ml^{-1}. Harvesting was achieved by simply looping crystals from the mother liquor and then plunge-cooling them in liquid nitrogen with no cryoprotection regime. Crystals were transported to the Australian Synchrotron for data collection. Diffraction data were collected at 100 K on beamline MX2. Dozens of crystals were investigated; however, useable diffraction was only obtained from several crystals using an unattenuated beam and 3 s exposures. Crystallization information is summarized in Table 2.

2.3. Structure determination and refinement

Diffraction data from two crystals (44 frames and 22 frames, respectively, each recorded over 1° oscillations with 3 s exposures to an unattenuated beam) were indexed and integrated to 3.5 Å resolution with *MOSFLM* (Leslie & Powell, 2007). The space group ($P3_221$) was confirmed with *POINTLESS* (Evans, 2011). Integrated reflections from these two crystals were scaled and merged with *AIMLESS* (Evans &

Table 3
Diffraction data-collection statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	0.9537
Space group	$P3_221$
Unit-cell parameters (Å, °)	$a = b = 115.0$, $c = 141.2$, $\alpha = \beta = 90$, $\gamma = 120$
Resolution range (Å)	44.56–3.50 (3.83–3.50)
Observed reflections	48483 (11641)
Unique reflections	13965 (3267)
Completeness† (%)	99.1 (99.0)
Multiplicity†	3.5 (3.6)
R_{merge}†	0.21 (0.67)
$R_{p.i.m.}$†	0.19 (0.58)
Mean $I/\sigma(I)$†	5.2 (1.9)
Wilson B factor (Å^2)	78

† As calculated by *AIMLESS*.

Table 4
Structure refinement and model validation.

Protein molecules in asymmetric unit	6
Atoms modelled in asymmetric unit	6241
Ramachandran plot‡	
Favoured (%)	92.5
Outliers (%)	0.71
R.m.s.d., bond lengths (Å)	0.011
R.m.s.d., bond angles (°)	1.46
R_{work}/R_{free}	0.26/0.31
Residues, chain *A* (side chains absent)	151 (40)
Residues, chain *B* (side chains absent)	147 (44)
R.m.s.d., chain *B* superposed on chain *A* (Å)	0.29 over 146 C^{α} atoms
Residues, chain *C* (side chains absent)	146 (39)
R.m.s.d., chain *C* superposed on chain *A* (Å)	0.15 over 146 C^{α} atoms
Residues, chain *D* (side chains absent)	148 (45)
R.m.s.d., chain *D* superposed on chain *A* (Å)	0.14 over 145 C^{α} atoms
Residues, chain *E* (side chains absent)	145 (39)
R.m.s.d., chain *E* superposed on chain *A* (Å)	0.15 over 145 C^{α} atoms
Residues, chain *F* (side chains absent)	141 (44)
R.m.s.d., chain *F* superposed on chain *A* (Å)	0.24 over 141 C^{α} atoms
Average B factor (Å^2)	64.4
PDB entry	4ykb

† As calculated by *MolProbity*.

Murshudov, 2013), accessed *via* the *CCP*4 software interface (Winn *et al.*, 2011). The data from multiple crystals were combined to maximize the data completeness and multiplicity. Data-collection and processing statistics are summarized in Table 3.

The structure was determined by molecular replacement with *Phaser* (McCoy *et al.*, 2007). The search model was taken from the *T. elongatus* wild-type structure (PDB entry 1z3x), which is 40% identical at the amino-acid sequence level over 165 residues. The more divergent first 88 residues of this structure (the N-terminal domain) were excluded from the search model. Molecular replacement initially identified four GUN4 molecules in the asymmetric unit, consistent with the crystal comprising 48% solvent (as calculated from the Matthews coefficient and full-length GUN4). However, incomplete crystal packing and inspection of electron-density maps suggested that additional unaccounted-for molecules might in fact be present. The search model was subsequently trimmed further (a further 18 residues were removed from the N-terminus, with the removal of several loops that were divergent in the *T. elongatus* and *Synechocystis* structures and with side chains truncated to the C^{β} position), which allowed *Phaser* to fit six non-clashing molecules in the asymmetric unit (solvent content of 53% for six molecules of ~150 residues). Hence, it appeared that the N-terminal domain of the molecule was missing.

Restrained *B*-factor refinement was performed with *REFMAC*5 (Murshudov *et al.*, 2011), using local noncrystallographic symmetry (NCS) restraints. Between rounds of refinement, electron-density maps and their fit to the model were examined using *Coot* (Emsley *et al.*, 2010). Amino-acid side chains were added if suggested by difference map

```
C.reinhardtii  -----MLAQTHTASSRSACRGRAQRGQLAFSAPRPISQRSGSLIQQAPSMAMRVTVAAGK 55
Synechocystis  ------MSDNLTELSQQLHDASEKKQLTAIAALAEMGEGGQGILLDYLAKNVPLEKP--V 52
T.elongatus    PEFMVTTEPALADLQEQLYNGNEKSQLAAMSTLSTAGTEGYHLLQEFLKDSATFSPPPAP 60
                                          α1                          α2
C.reinhardtii  LDSVSLFGGDTASLMGGSQTVEKKKSGKEAVMEVQLSSTAGIDYTVLRDHLANGEFREAE 115
Synechocystis  LAVGNVYQ---TLRNLEQETITTQLQRNYPTGIFPLQSAQGIDYLPLQEALGSQDFETAD 109
T.elongatus    WIRGQAYR---LLFHSPEASVQAFLQQHYPQGVIPLRSDRGVDYQELAKLLVAEKFEAAD 117
                        α2/α3       α3          α4                 α5
C.reinhardtii  DETRALLIKLAGPEAVKRNWVYFTEVKNISVTDFQTLDNLWKASSNNKFGYSVQKEIWVQ 175
Synechocystis  EITRDKLCELAGPGASQRQWLYFTEVEKFPALDLHTINALWWLHSNGNFGFSVQRRLWLA 169
T.elongatus    RLTTQKLCELAGPLAQKRRWLYFTEVEQLPIPDLQTIDQLWLAFSLGRFGYSVQRQLWLG 177
                  *
                    α6                                          α7
C.reinhardtii  NQKRWPKFFKQIDWTQGENNNYRKWPMEFIYSMDAPRGHLPLTNALRGTQLFQAIMEHPA 235
Synechocystis  SGKEFTKLWPKIGWKSG--NVWTRWPKGFTWDLSAPQGHLPLLNQLRGVRVAESLYRHPV 227
T.elongatus    CGQNWDRLWEKIGWRQG--KRWPRYPNEFIWDLSAPRGHLPLTNQLRGVQVLNALLNHPA 235
                          *       *  *   *           ** **

C.reinhardtii  FEKSSTAKTLDQKAAEAAGRTQSLF 260
Synechocystis  WSQYGW------------------- 233
T.elongatus    WTA---------------------- 238
```

Figure 1
Sequence alignment of GUN4 performed with *ClustalW*2. Highlighted in yellow is the portion of *C. reinhardtii* GUN4 that is resolvable in the crystal structure. The underlined region shows peptide coverage from MSMS proteomic analysis of the crystals. Residues highlighted by stars are directly involved in porphyrin binding in the *Synechocystis* GUN4 structures and the blue highlighted residues are involved in secondary-structure rearrangements (Chen *et al.*, 2015). α-Helix sections are shown in bold, with the labelling above the sequences the same as used in the text and the other figures.

electron density. Difficult segments of the model were built with sensible geometry using *Buccaneer* (Cowtan, 2006). Structure validation was performed using the *MolProbity* web server (Chen *et al.*, 2010). Refinement statistics are shown in Table 4. The coordinates for the final model have been deposited in the Protein Data Bank (PDB entry 4ykb).

3. Result and discussion

3.1. Overall structure

Six GUN4 molecules were found in the crystallographic asymmetric unit (chains *A*–*F*). All of them are lacking the N-terminal domain (the first resolvable residues are 80–83), and all of them are lacking the last 22–25 residues (the last resolvable residues are 235–238) (see Fig. 1). The fold of the resolvable ~150 residues modelled is effectively the same in all chains (r.m.s.d. of 0.14–0.29 relative to chain *A*; see Fig. 2). Owing to the low resolution of the data, between 25 and 30% of the amino-acid side chains in the various chains have not been modelled (Table 4). Those unmodelled are predominantly on the surface of the domain, whilst those buried in the hydrophobic core of the domain are generally well ordered.

The fold is largely α-helical, as expected from the *T. elongatus* and *Synechocystis* structures (PDB entries 1z3x and 1y6i, respectively). The *A* chain superposes with the 1z3x structure with an r.m.s.d. of 0.97 Å (over 143 C^{α} positions) and with the 1y6i structure with an r.m.s.d. of 1.1 Å (over 131 C^{α} positions) (Fig. 2).

The GUN4 molecules resolved in the crystal have quite clearly undergone proteolysis, as confirmed by MSMS proteomic analysis of tryptic digests of a number of crystals. The N-terminus of the resolvable domain immediately abuts a (missing) KKK sequence motif (analogous to the region linking the N- and C-terminal domains in related structures), whilst the C-terminal limit is adjacent to Lys238 (the last residue modelled in chain *A*). The C-terminal residues missing in our structure contain sites of phosphorylation in eukaryotes, but are not sequence features that are found in the cyanobacterial GUN4 molecules. The extended time taken for crystal growth presumably reflects the time taken for a contaminating protease within the crystallization drop to convert sufficient molecules to the doubly truncated form such that crystal growth could occur.

3.2. Comparison with cyanobacterial structures

Whilst the overall fold of the C-terminus of our *C. reinhardtii* GUN structure is highly similar to the cyanobacterial structures (Fig. 2; the *Synechocystis* structure is coloured light

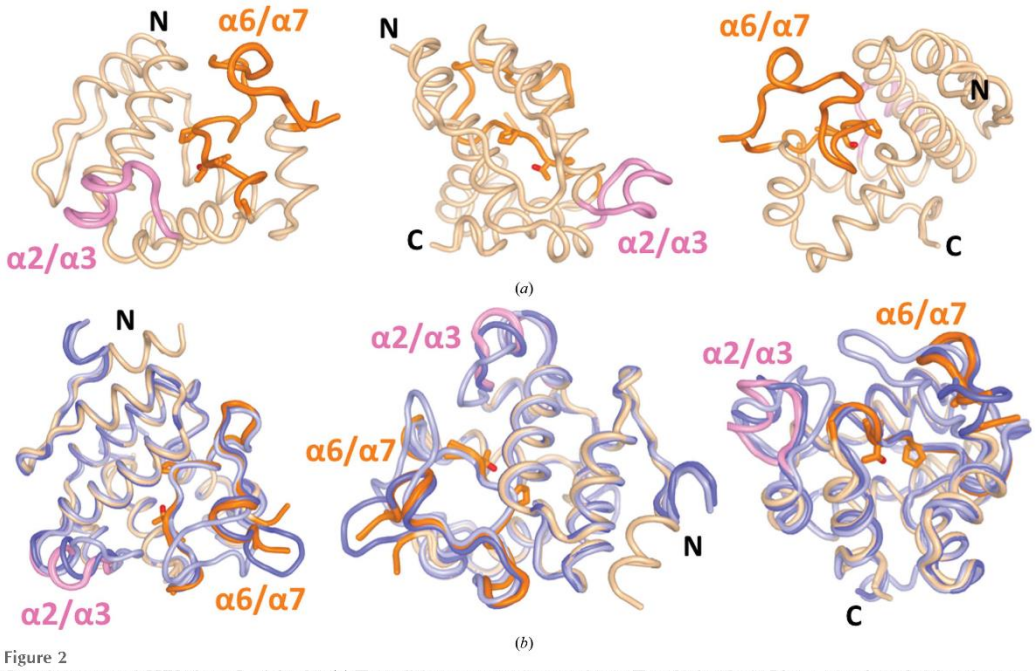


Figure 2
Crystal structure of GUN4 from *C. reinhardtii*. (*a*) Three different perspectives are shown. The α2/α3 and α6/α7 loops are coloured pink and orange, respectively. The side chains of Pro216 and Thr218 within the α6/α7 loop are drawn as sticks. The N- and C-termini are also indicated. (*b*) Superposition of GUN4 from *C. reinhardtii* [tan, orange and pink; as in (*a*)] with cyanobacterial GUN4 structures from *Synechocystis* (light blue) and *T. elongatus* (dark blue)

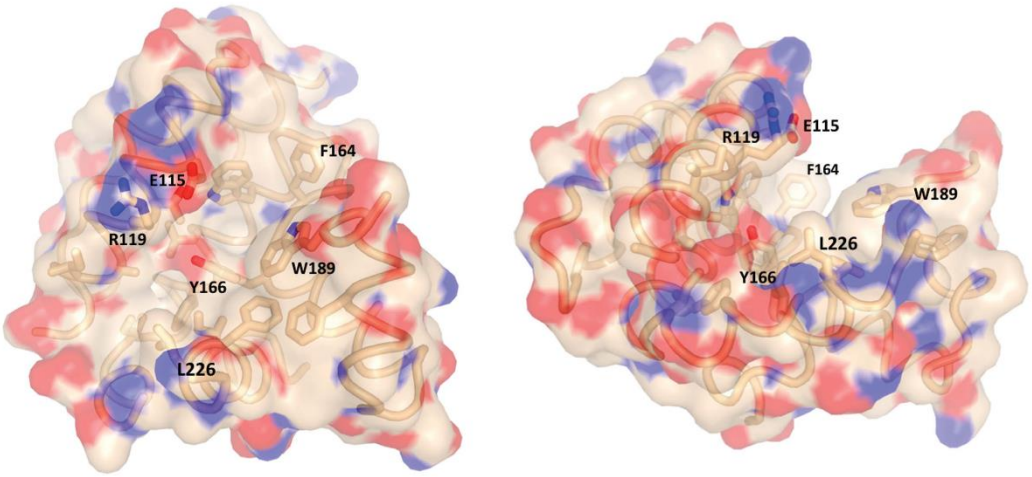


Figure 3
Porphyrin-binding cleft of GUN4 from *C. reinhardtii*. GUN4 has been trimmed of the α2/α3 and α6/α7 loops to reveal the porphyrin-binding cleft. Both a front-on (left) and a side view (right) are presented. The side chains of numerous residues lining the channel can be viewed under the transparent surface. Tyr166 is positioned at the base and approximately in the middle of the cleft. Other residues are identified to help mark the extremities of the cleft.

blue and the *T. elongatus* structure dark blue), the largest differences, and also the largest differences between the cyanobacterial structures themselves, concern two 'loop' segments: the loop linking helices 2 and 3 of the *Synechocystis* structure and the loop linking helices 6 and 7. The α2/α3 loop is actually a short section of helix and the α6/α7 loop is a β-turn in the porphyrin-bound *Synechocystis* structures (Chen *et al.*, 2015), whilst these loops help to cover the hydrophobic 'greasy palm' in the unbound *Synechocystis* structure (Verdecia *et al.*, 2005) (Fig. 2; α2/α3 and α6/α7 loops). The *T. elongatus* structure and our *C. reinhardtii* structure are thus intermediate between the porphyrin-bound and unbound *Synechocystis* structures (Fig. 2).

The α6/α7 loop covers the greasy-palm surface but projects very few residues into the cleft, namely a proline residue (conserved in all three organisms) and a leucine residue (substituted by a threonine residue in the *T. elongatus* structure and in our *C. reinhardtii* structure) (Fig. 2; proline and threonine residues from our structure are drawn as sticks). The porphyrin-bound *Synechocystis* structures (Chen *et al.*, 2015) clearly show that this α6/α7 loop has a propensity for conformational dynamics to form a β-hairpin, making a more structured cleft which enables porphyrin binding.

4. Conclusion

We have solved the crystal structure of the GUN4 protein from *C. reinhardtii*. Proteolysis within the crystallization experiment removed the N-terminal domain and the last ∼20 residues, leaving the bulk of the C-terminal domain intact; this domain is involved in binding PPIX and Mg-PPIX. Although the structure is of relatively low resolution, the bulk of the side chains which line the 'greasy palm' cleft involved in binding the porphyrin ring can be resolved. When superposed with cyanobacterial GUN4 structures, whilst the α-helical scaffolds are conserved, the α2/α3 and α6/α7 loops previously predicted to undergo movement to accommodate porphyrin binding display markedly different conformations. This diversity of fold captured by multiple crystal structures supports the notion that these loops undergo conformational rearrangement to accommodate the insertion of PPIX or Mg-PPIX into the underlying hydrophobic cleft (Fig. 3), as observed in the recent porphyrin-bound *Synechocystis* structures (Chen *et al.*, 2015).

References

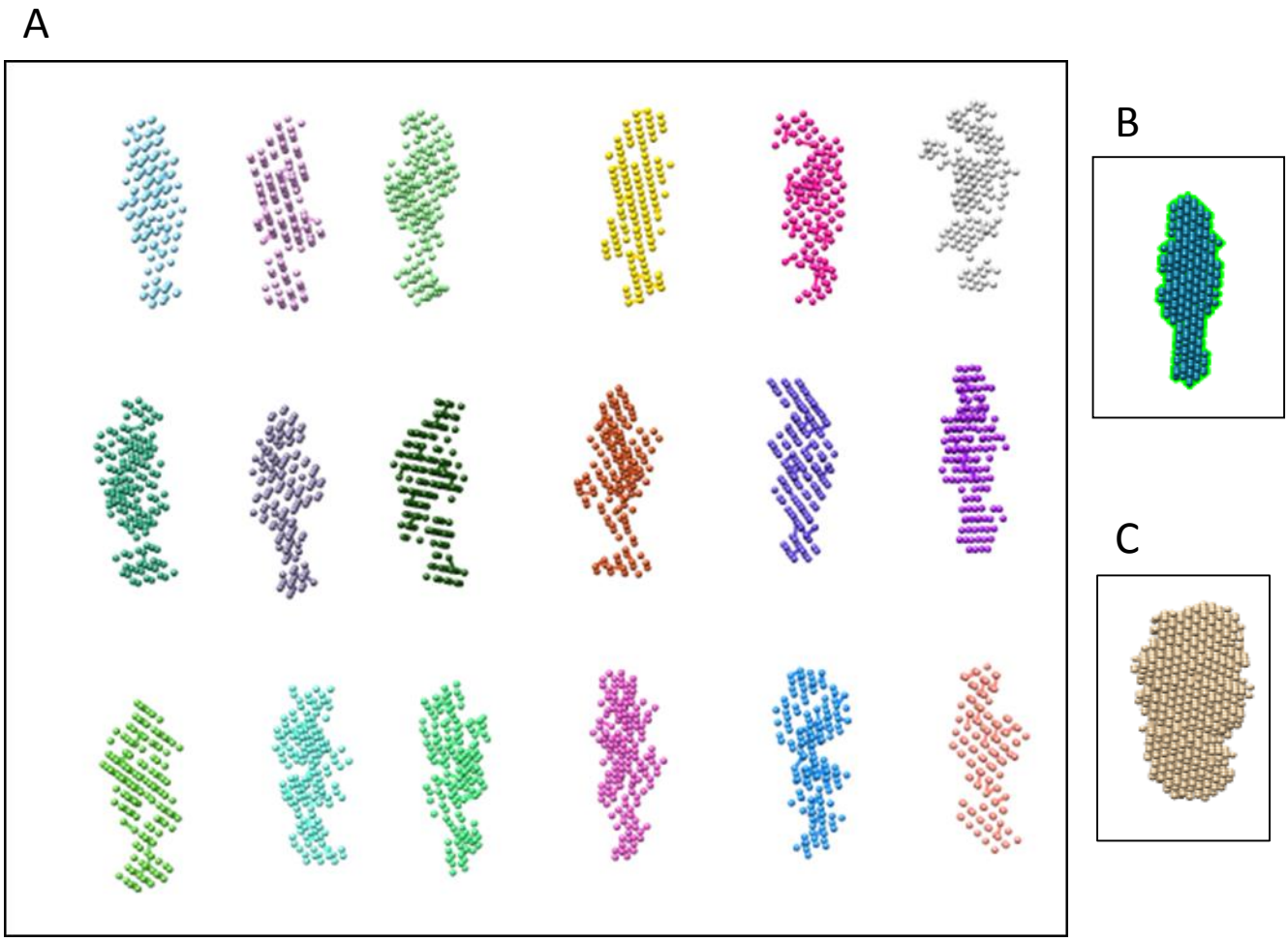
Chen, X., Pu, H., Wang, X., Long, W., Lin, R. & Liu, L. (2015). *Mol. Plant.*, doi:10.1016/j.molp.2015.04.013.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Davison, P. A., Schubert, H. L., Reid, J. D., Iorg, C. D., Heroux, A., Hill, C. P. & Hunter, C. N. (2005). *Biochemistry*, **44**, 7603–7612.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Formighieri, C., Ceol, M., Bonente, G., Rochaix, J.-D. & Bassi, R. (2012). *Mol. Plant.* **5**, 1242–1262.
Gorchein, A., Gibson, L. C. D. & Hunter, C. N. (1993). *Biochem. Soc. Trans.* **21**, 2015.
Jensen, P. E., Gibson, L. C. D., Henningseni, K. W. & Hunter, C. N. (1996). *Biol. Chem.* **271**, 16662–16667.

Larkin, R. M., Alonso, J. M., Ecker, J. R. & Chory, J. (2003). *Science*, **299**, 902–906.
Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Papenbrock, J. S., Gräfe, S., Kruse, F., Hänel, F. & Grimm, B. (1997). *Plant J.* **12**, 981–990.
Peter, E. & Grimm, B. (2009). *Mol. Plant.* **2**, 1198–1210.
Susek, R. E., Ausubel, F. M. & Chory, J. (1993). *Cell*, **74**, 787–799.
Verdecia, M. A., Larkin, R. M., Ferrer, J.-L., Riek, R., Chory, J. & Noel, J. P. (2005). *PLoS Biol.* **3**, e151.
Wang, W.-Y., Wang, J. B., Boynton, J. E. & Gillham, N. W. (1974). *J. Cell Biol.* **63**, 806–823.
Willows, R. D. (2003). *Nat. Prod. Rep.* **20**, 327–341.
Willows, R. D. & Hansson, M. (2003). *The Porphyrin Handbook: Chlorophylls and Bilins: Biosynthesis, Synthesis and Degradation*, edited by K. Kadish, K. M. Smith & R. Guilard, pp. 1–48. San Diego: Academic Press.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Appendix 5. Synopsis of publication III (Prepared for submission)

This research article investigated the $^1O_2$ production of GUN4 and ChlH proteins in *vitro.* The generated $^1O_2$ was detected using histidine-mediated chemical trapping and SOSG fluorescent methods. The experiments are designed to measure the rate of $^1O_2$ production for GUN4, ChlH, and GUN4-ChlH proteins in complex with PPIX or without PPIX. This chapter is presented as a publication research article and reports that GUN4-PPIX increases the singlet oxygen production in *vitro*, and the role of this protein in chloroplast to nucleus retrograde signalling. We compared the singlet oxygen production rate. We also generated a model for function of GUN4 in $^1O_2$ production and chloroplast to nucleus signaling. We aimed to identify GUN involvement in $^1O_2$ production is chloroplast biosynthesis pathway.

Appendix 6. GUN4 SAXS shape determination



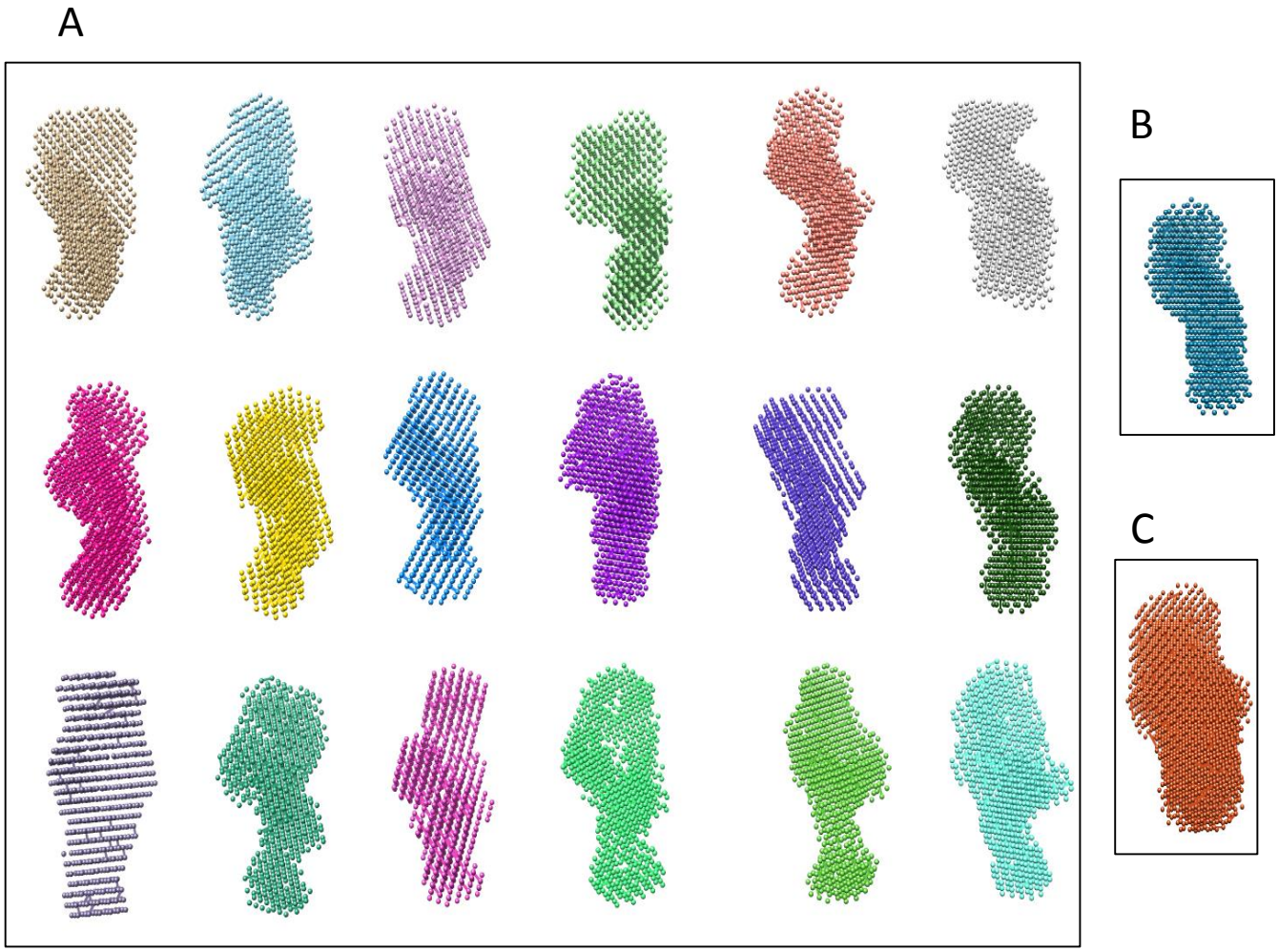
Appendix 6. GUN4 SAXS shape determination

A: Independent ab *initio* reconstructions (DAMMIN)

B: Ensemble consensus structure (DAMFILT)

C: Averaged structure (DAMAVER

Appendix 7. ChlH SAXS shape determination



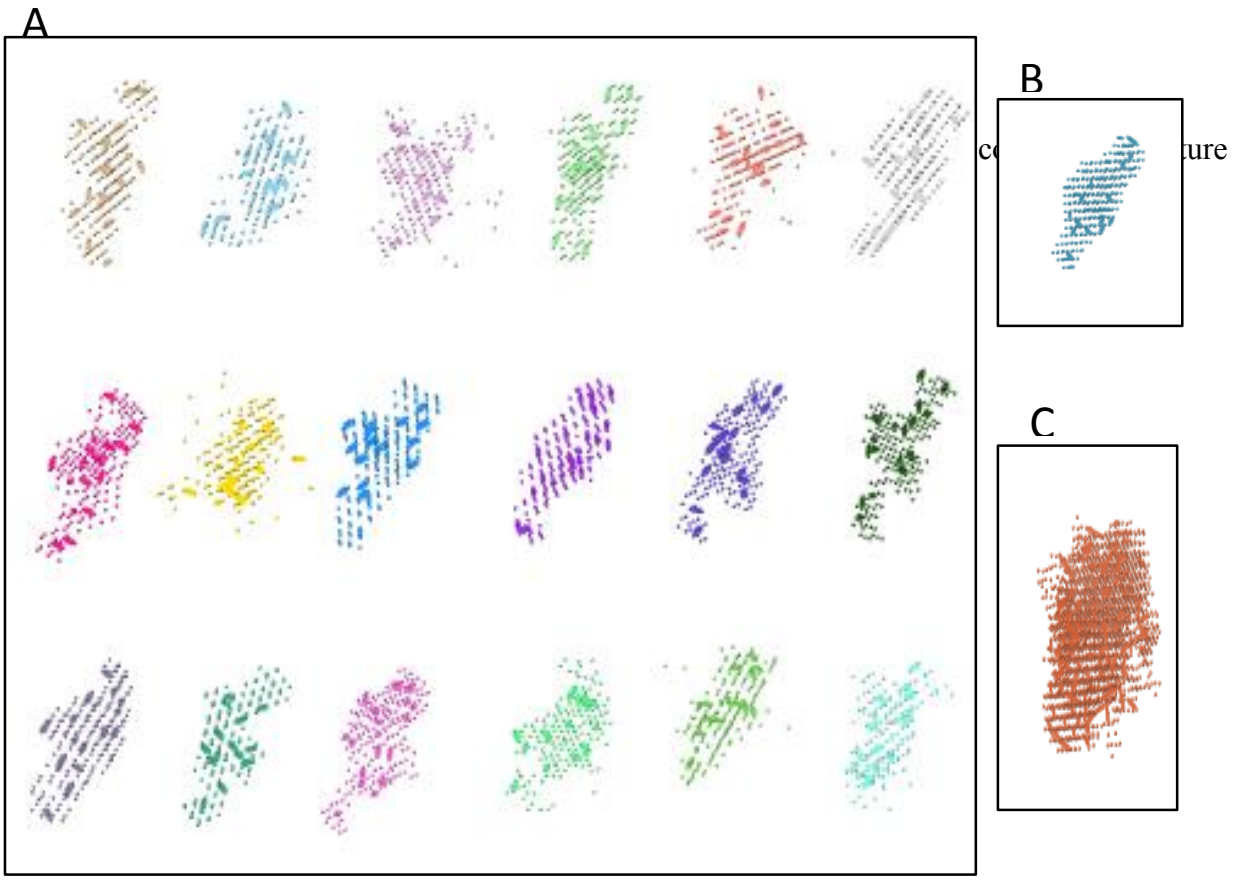
Appendix 7. ChlH SAXS shape determination

A: Independent ab *initio* reconstructions (DAMMIN)

B: Ensemble consensus structure (DAMFILT)

C: Averaged structure (DAMAVER)

Appendix 8. GUN4-ChlH-PPIX SAXS shape determination



Appendix 8. GUN4-ChlH-PPIX SAXS shape determination

A: Independent ab *initio* reconstructions (DAMMIN)

B: Ensemble consensus structure (DAMFILT)

C: Averaged structure (DAMAVER)

Appendix 9. GUN4 crystal structure validation report

Full wwPDB X-ray Structure Validation Report (i)

Mar 4, 2015 – 05:56 PM EST

PDB ID	:	4YKB
Title	:	Structure of GUN4 from Chlamydomonas reinhardtii
Authors	:	Tabriz, S.T.; Langley, D.B.; Willows, R.D.; Duff, A.P.; Harrop, S.J.
Deposited on	:	2015-03-04
Resolution	:	3.50 Å(reported)

DISCLAIMER
This is a preliminary version of the new style of wwPDB validation report.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24195
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24195

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6241 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Tetrapyrrole-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	151	Total	C	N	O	S	0	0	0
			1089	704	181	201	3			
1	B	147	Total	C	N	O	S	0	0	0
			1030	665	174	188	3			
1	C	146	Total	C	N	O	S	0	0	0
			1044	673	178	190	3			
1	D	148	Total	C	N	O	S	0	0	0
			1034	669	171	192	2			
1	E	145	Total	C	N	O	S	0	0	0
			1054	684	175	193	2			
1	F	141	Total	C	N	O	S	0	0	0
			990	642	163	183	2			

There are 126 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	MET	-	initiating methionine	UNP A8I5N5
A	33	GLY	-	expression tag	UNP A8I5N5
A	34	SER	-	expression tag	UNP A8I5N5
A	35	SER	-	expression tag	UNP A8I5N5
A	36	HIS	-	expression tag	UNP A8I5N5
A	37	HIS	-	expression tag	UNP A8I5N5
A	38	HIS	-	expression tag	UNP A8I5N5
A	39	HIS	-	expression tag	UNP A8I5N5
A	40	HIS	-	expression tag	UNP A8I5N5
A	41	HIS	-	expression tag	UNP A8I5N5
A	42	SER	-	expression tag	UNP A8I5N5
A	43	SER	-	expression tag	UNP A8I5N5
A	44	GLY	-	expression tag	UNP A8I5N5
A	45	LEU	-	expression tag	UNP A8I5N5
A	46	VAL	-	expression tag	UNP A8I5N5
A	47	PRO	-	expression tag	UNP A8I5N5
A	48	ARG	-	expression tag	UNP A8I5N5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	GLY	-	expression tag	UNP A8I5N5
A	50	SER	-	expression tag	UNP A8I5N5
A	51	HIS	-	expression tag	UNP A8I5N5
A	52	MET	-	expression tag	UNP A8I5N5
B	32	MET	-	initiating methionine	UNP A8I5N5
B	33	GLY	-	expression tag	UNP A8I5N5
B	34	SER	-	expression tag	UNP A8I5N5
B	35	SER	-	expression tag	UNP A8I5N5
B	36	HIS	-	expression tag	UNP A8I5N5
B	37	HIS	-	expression tag	UNP A8I5N5
B	38	HIS	-	expression tag	UNP A8I5N5
B	39	HIS	-	expression tag	UNP A8I5N5
B	40	HIS	-	expression tag	UNP A8I5N5
B	41	HIS	-	expression tag	UNP A8I5N5
B	42	SER	-	expression tag	UNP A8I5N5
B	43	SER	-	expression tag	UNP A8I5N5
B	44	GLY	-	expression tag	UNP A8I5N5
B	45	LEU	-	expression tag	UNP A8I5N5
B	46	VAL	-	expression tag	UNP A8I5N5
B	47	PRO	-	expression tag	UNP A8I5N5
B	48	ARG	-	expression tag	UNP A8I5N5
B	49	GLY	-	expression tag	UNP A8I5N5
B	50	SER	-	expression tag	UNP A8I5N5
B	51	HIS	-	expression tag	UNP A8I5N5
B	52	MET	-	expression tag	UNP A8I5N5
C	32	MET	-	initiating methionine	UNP A8I5N5
C	33	GLY	-	expression tag	UNP A8I5N5
C	34	SER	-	expression tag	UNP A8I5N5
C	35	SER	-	expression tag	UNP A8I5N5
C	36	HIS	-	expression tag	UNP A8I5N5
C	37	HIS	-	expression tag	UNP A8I5N5
C	38	HIS	-	expression tag	UNP A8I5N5
C	39	HIS	-	expression tag	UNP A8I5N5
C	40	HIS	-	expression tag	UNP A8I5N5
C	41	HIS	-	expression tag	UNP A8I5N5
C	42	SER	-	expression tag	UNP A8I5N5
C	43	SER	-	expression tag	UNP A8I5N5
C	44	GLY	-	expression tag	UNP A8I5N5
C	45	LEU	-	expression tag	UNP A8I5N5
C	46	VAL	-	expression tag	UNP A8I5N5
C	47	PRO	-	expression tag	UNP A8I5N5
C	48	ARG	-	expression tag	UNP A8I5N5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	49	GLY	-	expression tag	UNP A8I5N5
C	50	SER	-	expression tag	UNP A8I5N5
C	51	HIS	-	expression tag	UNP A8I5N5
C	52	MET	-	expression tag	UNP A8I5N5
D	32	MET	-	initiating methionine	UNP A8I5N5
D	33	GLY	-	expression tag	UNP A8I5N5
D	34	SER	-	expression tag	UNP A8I5N5
D	35	SER	-	expression tag	UNP A8I5N5
D	36	HIS	-	expression tag	UNP A8I5N5
D	37	HIS	-	expression tag	UNP A8I5N5
D	38	HIS	-	expression tag	UNP A8I5N5
D	39	HIS	-	expression tag	UNP A8I5N5
D	40	HIS	-	expression tag	UNP A8I5N5
D	41	HIS	-	expression tag	UNP A8I5N5
D	42	SER	-	expression tag	UNP A8I5N5
D	43	SER	-	expression tag	UNP A8I5N5
D	44	GLY	-	expression tag	UNP A8I5N5
D	45	LEU	-	expression tag	UNP A8I5N5
D	46	VAL	-	expression tag	UNP A8I5N5
D	47	PRO	-	expression tag	UNP A8I5N5
D	48	ARG	-	expression tag	UNP A8I5N5
D	49	GLY	-	expression tag	UNP A8I5N5
D	50	SER	-	expression tag	UNP A8I5N5
D	51	HIS	-	expression tag	UNP A8I5N5
D	52	MET	-	expression tag	UNP A8I5N5
E	32	MET	-	initiating methionine	UNP A8I5N5
E	33	GLY	-	expression tag	UNP A8I5N5
E	34	SER	-	expression tag	UNP A8I5N5
E	35	SER	-	expression tag	UNP A8I5N5
E	36	HIS	-	expression tag	UNP A8I5N5
E	37	HIS	-	expression tag	UNP A8I5N5
E	38	HIS	-	expression tag	UNP A8I5N5
E	39	HIS	-	expression tag	UNP A8I5N5
E	40	HIS	-	expression tag	UNP A8I5N5
E	41	HIS	-	expression tag	UNP A8I5N5
E	42	SER	-	expression tag	UNP A8I5N5
E	43	SER	-	expression tag	UNP A8I5N5
E	44	GLY	-	expression tag	UNP A8I5N5
E	45	LEU	-	expression tag	UNP A8I5N5
E	46	VAL	-	expression tag	UNP A8I5N5
E	47	PRO	-	expression tag	UNP A8I5N5
E	48	ARG	-	expression tag	UNP A8I5N5

Continued on next page...

Continued from previous page...

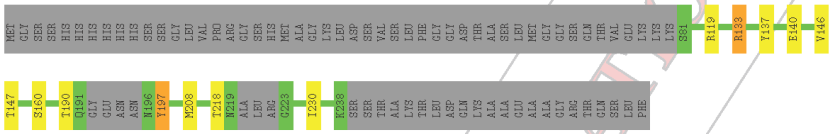
Chain	Residue	Modelled	Actual	Comment	Reference
E	49	GLY	-	expression tag	UNP A8I5N5
E	50	SER	-	expression tag	UNP A8I5N5
E	51	HIS	-	expression tag	UNP A8I5N5
E	52	MET	-	expression tag	UNP A8I5N5
F	32	MET	-	initiating methionine	UNP A8I5N5
F	33	GLY	-	expression tag	UNP A8I5N5
F	34	SER	-	expression tag	UNP A8I5N5
F	35	SER	-	expression tag	UNP A8I5N5
F	36	HIS	-	expression tag	UNP A8I5N5
F	37	HIS	-	expression tag	UNP A8I5N5
F	38	HIS	-	expression tag	UNP A8I5N5
F	39	HIS	-	expression tag	UNP A8I5N5
F	40	HIS	-	expression tag	UNP A8I5N5
F	41	HIS	-	expression tag	UNP A8I5N5
F	42	SER	-	expression tag	UNP A8I5N5
F	43	SER	-	expression tag	UNP A8I5N5
F	44	GLY	-	expression tag	UNP A8I5N5
F	45	LEU	-	expression tag	UNP A8I5N5
F	46	VAL	-	expression tag	UNP A8I5N5
F	47	PRO	-	expression tag	UNP A8I5N5
F	48	ARG	-	expression tag	UNP A8I5N5
F	49	GLY	-	expression tag	UNP A8I5N5
F	50	SER	-	expression tag	UNP A8I5N5
F	51	HIS	-	expression tag	UNP A8I5N5
F	52	MET	-	expression tag	UNP A8I5N5

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

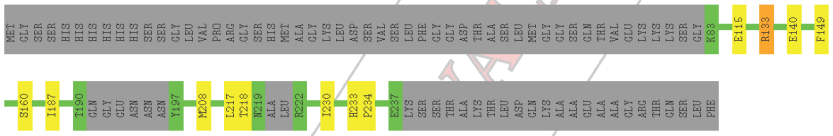
- Molecule 1: Tetrapyrrole-binding protein

Chain A:



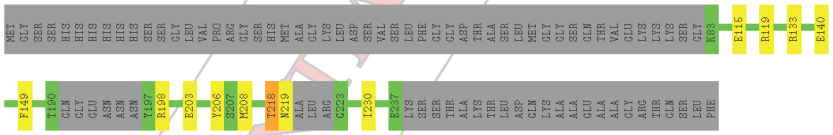
- Molecule 1: Tetrapyrrole-binding protein

Chain B:



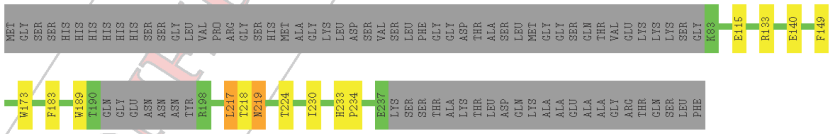
- Molecule 1: Tetrapyrrole-binding protein

Chain C:



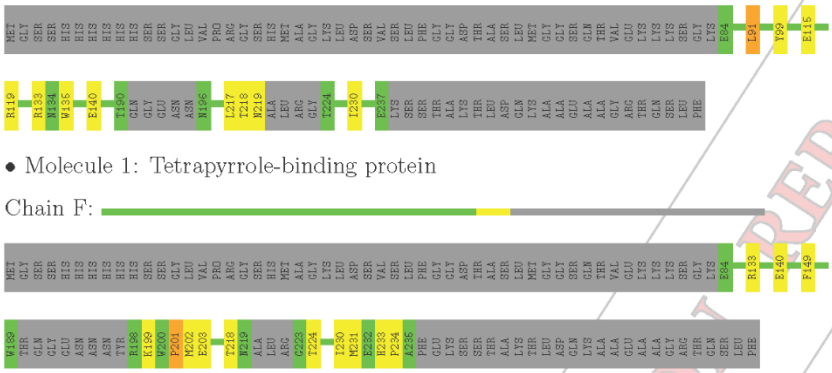
- Molecule 1: Tetrapyrrole-binding protein

Chain D:



- Molecule 1: Tetrapyrrole-binding protein

Chain E:



• Molecule 1: Tetrapyrrole-binding protein

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	115.00Å 115.00Å 141.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.50 – 3.50 42.55 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.50-3.50) 99.2 (42.55-3.50)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.256 , 0.313 0.255 , 0.309	Depositor DCC
R_{free} test set	753 reflections (5.71%)	DCC
Wilson B-factor (Å^2)	79.9	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 74.9	EDS
Estimated twinning fraction	0.049 for -h,-k,l	Xtriage
L-test for twinning	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Outliers	1 of 13941 reflections (0.007%)	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6241	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.14% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.56	0/1118	0.76	2/1534 (0.1%)
1	B	0.54	0/1055	0.79	3/1453 (0.2%)
1	C	0.54	0/1069	0.77	1/1468 (0.1%)
1	D	0.52	0/1062	0.70	1/1467 (0.1%)
1	E	0.56	0/1084	0.77	2/1492 (0.1%)
1	F	0.56	0/1018	0.76	2/1405 (0.1%)
All	All	0.54	0/6406	0.76	11/8819 (0.1%)

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	187	ILE	CG1-CB-CG2	-9.54	90.42	111.40
1	E	91	LEU	CA-CB-CG	6.87	131.10	115.30
1	A	119	ARG	NE-CZ-NH1	6.74	123.67	120.30
1	C	208	MET	CG-SD-CE	6.66	110.86	100.20
1	E	119	ARG	NE-CZ-NH1	6.17	123.38	120.30
1	F	231	MET	CG-SD-CE	6.00	109.80	100.20
1	B	187	ILE	CB-CG1-CD1	5.49	129.27	113.90
1	F	231	MET	CB-CG-SD	5.36	128.48	112.40
1	B	133	ARG	NE-CZ-NH1	5.15	122.87	120.30
1	D	115	GLU	OE1-CD-OE2	-5.10	117.18	123.30
1	A	133	ARG	NE-CZ-NH2	-5.03	117.79	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit,

and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1089	0	923	15	0
1	B	1030	0	875	9	0
1	C	1044	0	900	7	0
1	D	1034	0	855	8	0
1	E	1054	0	887	12	0
1	F	990	0	808	6	0
All	All	6241	0	5248	48	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

All (48) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:115:GLU:OE2	1:C:119:ARG:NH1	1.79	1.14
1:D:189:TRP:CZ2	1:D:217:LEU:HD12	2.35	0.61
1:B:115:GLU:OE2	1:B:217:LEU:HG	2.02	0.59
1:A:137:TYR:CZ	1:E:135:TRP:CH2	2.93	0.57
1:B:218:THR:HG21	1:B:230:ILE:HD12	1.86	0.57
1:D:218:THR:O	1:D:219:ASN:HB2	2.03	0.56
1:C:218:THR:HG21	1:C:230:ILE:HD12	1.87	0.56
1:E:218:THR:HG21	1:E:230:ILE:HD12	1.87	0.56
1:A:218:THR:HG21	1:A:230:ILE:HD12	1.88	0.56
1:D:218:THR:O	1:D:219:ASN:CB	2.53	0.56
1:F:218:THR:HG21	1:F:230:ILE:HD12	1.88	0.55
1:A:133:ARG:HH22	1:A:140:GLU:CD	2.11	0.54
1:D:218:THR:HG21	1:D:230:ILE:HD12	1.90	0.54
1:A:147:THR:HG22	1:C:206:TYR:H	1.71	0.54
1:C:133:ARG:HH22	1:C:140:GLU:CD	2.11	0.53
1:D:133:ARG:HH22	1:D:140:GLU:CD	2.12	0.53
1:F:133:ARG:HH22	1:F:140:GLU:CD	2.11	0.53
1:A:137:TYR:CZ	1:E:135:TRP:HH2	2.28	0.51
1:E:133:ARG:HH22	1:E:140:GLU:CD	2.12	0.51
1:B:133:ARG:HH22	1:B:140:GLU:CD	2.12	0.51
1:F:133:ARG:NH2	1:F:140:GLU:OE1	2.45	0.50
1:E:218:THR:O	1:E:219:ASN:CB	2.60	0.49
1:E:133:ARG:NH2	1:E:140:GLU:OE1	2.46	0.49
1:B:115:GLU:OE2	1:B:217:LEU:N	2.42	0.49
1:C:133:ARG:NH2	1:C:140:GLU:OE1	2.46	0.49
1:A:190:THR:HG22	1:A:197:TYR:CD2	2.47	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:137:TYR:CZ	1:E:135:TRP:CZ2	3.01	0.49
1:B:133:ARG:NH2	1:B:140:GLU:OE1	2.46	0.49
1:A:137:TYR:CE1	1:E:135:TRP:CZ2	3.01	0.49
1:D:133:ARG:NH2	1:D:140:GLU:OE1	2.46	0.48
1:A:133:ARG:NH2	1:A:140:GLU:OE1	2.46	0.48
1:A:137:TYR:CE1	1:E:135:TRP:HZ2	2.33	0.47
1:A:190:THR:HG22	1:A:197:TYR:CE2	2.50	0.47
1:F:201:PRO:O	1:F:202:MET:CB	2.62	0.45
1:B:115:GLU:OE2	1:B:217:LEU:CB	2.66	0.44
1:A:146:VAL:HG23	1:C:203:GLU:HA	2.00	0.43
1:A:137:TYR:CE2	1:E:135:TRP:HH2	2.38	0.42
1:F:199:LYS:O	1:F:203:GLU:CB	2.67	0.42
1:D:173:TRP:HA	1:D:183:PHE:CD1	2.54	0.42
1:E:91:LEU:HD22	1:E:99:TYR:HB2	2.03	0.41
1:B:115:GLU:OE2	1:B:217:LEU:CG	2.67	0.41
1:B:160:SER:HA	1:B:208:MET:SD	2.61	0.41
1:D:233:HIS:CG	1:D:234:PRO:HD2	2.55	0.41
1:A:160:SER:HA	1:A:208:MET:SD	2.61	0.41
1:A:147:THR:CG2	1:C:206:TYR:H	2.34	0.41
1:E:115:GLU:OE2	1:E:217:LEU:N	2.46	0.40
1:B:233:HIS:CG	1:B:234:PRO:HD2	2.56	0.40
1:F:233:HIS:CG	1:F:234:PRO:HD2	2.57	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	145/229 (63%)	132 (91%)	13 (9%)	0	100	100
1	B	141/229 (62%)	133 (94%)	8 (6%)	0	100	100
1	C	140/229 (61%)	130 (93%)	8 (6%)	2 (1%)	16	70
1	D	144/229 (63%)	132 (92%)	10 (7%)	2 (1%)	16	70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	139/229 (61%)	130 (94%)	9 (6%)	0	100	100
1	F	135/229 (59%)	124 (92%)	9 (7%)	2 (2%)	15	69
All	All	844/1374 (61%)	781 (92%)	57 (7%)	6 (1%)	30	83

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	198	ARG
1	D	219	ASN
1	D	224	THR
1	C	218	THR
1	F	224	THR
1	F	201	PRO

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	93/193 (48%)	92 (99%)	1 (1%)	84	96
1	B	86/193 (45%)	85 (99%)	1 (1%)	82	96
1	C	90/193 (47%)	88 (98%)	2 (2%)	64	92
1	D	85/193 (44%)	83 (98%)	2 (2%)	61	92
1	E	90/193 (47%)	90 (100%)	0	100	100
1	F	80/193 (42%)	79 (99%)	1 (1%)	80	96
All	All	524/1158 (45%)	517 (99%)	7 (1%)	80	96

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	197	TYR
1	B	149	PHE
1	C	149	PHE
1	C	219	ASN
1	D	149	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	217	LEU
1	F	149	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	151/229 (65%)	-0.16	0	100	100	40, 61, 85, 97	0
1	B	147/229 (64%)	-0.17	0	100	100	44, 64, 82, 93	0
1	C	146/229 (63%)	-0.21	0	100	100	35, 61, 78, 94	0
1	D	148/229 (64%)	-0.12	0	100	100	36, 61, 81, 102	0
1	E	145/229 (63%)	-0.18	0	100	100	39, 62, 79, 108	0
1	F	141/229 (61%)	0.01	0	100	100	40, 66, 94, 104	0
All	All	878/1374 (63%)	-0.14	0	100	100	35, 63, 84, 108	0

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.

Appendix 10. Biosafety ethics approval

MACQUARIE UNIVERSITY

Biosafety Committee

REQUEST FOR AMENDMENT FORM

This form is to be completed and forwarded to the Biosafety Secretary for **ALL** amendments/modifications including extensions to approved Biosafety protocols. This includes the addition or removal of personnel.

This form must be completed by the Chief Investigator (if you are an Honours, postgraduate or HDR student, your supervisor must be nominated as the Chief Investigator) and emailed to the Biosafety Secretary at Biosafety@mq.edu.au

(July 2010 Version)

Hand written forms will not be accepted.

1. **Chief Investigator/Supervisor:**

   Faculty: Science

   Department: CBMS

   Email: robert.willows@mq.edu.au

(Note: If the project is to be undertaken by an Honours/postgraduate/HDR student, the supervisor will be considered the Chief Investigator. The student may be named as a co-investigator.)

2. **Project Title:**

Expression of proteins involved in essential metabolic processes from photosynthetic and non photosynthetic organisms

3. **Biosafety Committee Approval Reference No:**

5201001087

4. **Name/s of Student/s (if applicable):**

Andre Hallen

5. **Names of Co-Investigators/Associate Supervisors/Research Assistants:**

Angela Moncrieff, Artur Sawicki, Zane Duxbury, Tony Jerkovic

6. **Does the proposed amendment involve dealings with GMOs?**

   X Yes ☐ **No**

If you have not already received approval to conduct dealings with GMOs then you must submit a new application to the Committee for review. Please download an application from
http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/biosafety_research_ethics/forms

7. **Reason for amendment/s:**
Please provide details of the changes you propose and explain why these are necessary (in your response please outline what laboratory procedures are in place, how samples will be transported, handled and disposed. Please also indicate what measures are in place to minimize any risks that might be incurred as a result of this proposed amendment).

New student working in the lab will be working on the project, Shabnam Tarrahi Tabrizi. Andrew Scafaro is working on this project. Andre Hallen is no longer working in my lab on this project.

8. **Adding/removing personnel:** **X** Adding X Removing
If **Adding**, provide the following (If more than one, please copy this page).
If **Removing,** please complete name and title sections only.

Name:	Shabnam Tarrahi Tabrizi
Title:	Ms
Personnel type:	Student
Staff / Student no.: (Mandatory)	42661706
Qualifications:	Bachelor degree and Masters degree
Positions held: (if student, specify Faculty, Department, degree and course in which enrolled)	Science CBMS PhD
What specific skills will the new personnel bring to the project?	General laboratory skills
Full mailing address:	104 Agincourt Road Marsfield NSW 2122 AUSTRALIA
Tel No. (W):	X8219
Tel No: (H):	
Mobile No:	0432036004
Fax number:	

Name:	Andrew Scafaro
Title:	Dr
Personnel type:	Postdoctoral associate
Staff / Student no.: (Mandatory)	20058332
Qualifications:	PhD
Positions held: (if student, specify Faculty, Department, degree and course in which enrolled)	Research associate
What specific skills will the new personnel bring to the project?	Molecular Biology and protein expression.
Full mailing address:	Macquarie University Department of CBMS North Ryde 2109
Tel No. (W):	9850 8219
Tel No: (H):	
Mobile No:	
Fax number:	

Name:	Andre Hallen
Title:	Mr
Personnel type:	Student
Staff / Student no.: (Mandatory)	
Qualifications:	
Positions held: (if student, specify Faculty, Department, degree and course in which enrolled)	
What specific skills will the new personnel bring to the project?	
Full mailing address:	
Tel No. (W):	
Tel No: (H):	
Mobile No:	
Fax number:	

9. **Expected date of implementation of amendments to research:**

Date: **[**10/7/12**] Immediately.**

10. **Will the amendment/s impact on the research funding arrangements?** ☐ Yes X No ☐ N/A
If YES, please provide details.

11. **Any other information** ☐ Yes ☐ No ☐ N/A
If YES, please provide details and attach these.

CHECK LIST

Compliance with the guidelines set out by the OGTR X Yes ☐ No

Compliance with legislative requirements: X Yes ☐ No

For further information see:

http://www.ogtr.gov.au/

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/biosafety_research_ethics

IMPORTANT NOTICE: ELECTRONIC SUBMISSION OF THIS FORM IS EQUIVALENT TO THE SIGNATURE OF THE CHIEF INVESTIGATOR.

From: **Biosafety @mq.edu.au**

Biosafety@mq.edu.au

BS

Subject: Biosafety Application (Ref: 5201300777 - Exempt Dealing) - Final Approval
Date: 19 November 2013 at 10:30 AM
To: A/Prof Robert Willows robert.willows@mq.edu.au
Cc: Ms Angela Moncrieff angela.moncrieff@mq.edu.au, Dr Artur Sawicki artur.sawicki@mq.edu.au, Ms Shabnam Tarahi Tabrizi shabnam.tarahi-tabrizi@students.mq.edu.au

Dear A/Prof Willows:

Re: "Expression of proteins involved in essential metabolic processes from photosynthetic and non photosynthetic organisms" (Ref: 5201300777)

Thank you for your recent correspondence. Your responses have been reviewed by the Institutional Biosafety Committee (IBC) and approval of the above Exempt Dealing has been granted, effective 19 November 2013.

Approval has been granted subject to your compliance with the Office of the Gene Technology Regulator's standard conditions for exempt work listed below:

1. The project must be conducted in accordance with the OGTR Guidance Notes for the Containment of Exempt Dealings (http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/ExemptDealGuide Sept11_2-htm)

2. The Guidance Notes are only applicable to exempt dealings conducted under the Gene Technology Act 2000. They do not provide guidance for laboratory safety, good laboratory practice or work health and safety issues

For these purposes, refer to AS/NZS 2243.3:2010.

3. You must inform the Institutional Biosafety Committee if you complete or abandon the exempt dealings with GMOs.

Please note the following standard requirements of approval:

1. Approval will be for a period of 5 years subject to the provision of annual reports. If, at the end of this period the project has been completed, abandoned, discontinued or not commenced for any reason, you are required to submit a Final Report. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. Please contact the Committee Secretary at biosafety@mq.edu.au in order to obtain a

report.

A Progress/Final Report for this project will be due on: 19 November 2014

3. If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of Final Approval to an external organisation as evidence that you have Final Approval, please do not hesitate to contact the Committee Secretary at biosafety@mq.edu.au or by phone 9850 4063.

Please retain a copy of this email as this is your formal notification of final Biosafety approval.

Yours Sincerely

A/Prof Subramanyam Vemulpad
Chair, Macquarie University Institutional Biosafety Committee