Assembly of modified and non-native proteins in *Rhodobacter sphaeroides* and *Escherichia coli*

Katie Jane Grayson

A thesis submitted for the degree of Doctor of Philosophy at the University of Sheffield

Department of Molecular Biology and Biotechnology

September 2015

EPSRC
Engineering and Physical Sciences Research Council

Summary

Photosynthesis is a highly efficient and productive process that converts solar energy into chemical energy. Much of the visible and near-infrared radiation falling on the surface of the Earth is not absorbed by photosynthesising organisms, which occupy particular spectral niches depending on the absorption of the particular pigments they synthesise. For synthetic biology applications it would be worthwhile to design and construct bacteria that could utilise a greater range of wavelengths than naturally-evolved photosynthetic bacteria. Although the incorporation of synthetic chromophores to complement native light-harvesting proteins is promising, the approach generally involves *in vitro* reassembly. In order to create tailor-made light harvesting antennas *in vivo*, we must make use of the toolbox of proteins and pigments available in nature, or create synthetic elements that are able to be created by the host organism.

To investigate the possibility of creating artificial light-harvesting antennas *in vivo*, the yellow fluorescent protein, YFP, was incorporated as a chromophore into the photosynthetic apparatus of the purple photosynthetic bacterium *Rhodobacter* (*Rba.*) *sphaeroides*. It is shown that energy absorbed by YFP can transfer to the native reaction centre and LH1 proteins, sufficient to enhance the photosynthetic growth rate in a *Rba. sphaeroides* carotenoidless mutant. The light-driven proton pump, proteorhodopsin (PR), also has potential to augment the proton motive force in *Rba. sphaeroides* and drive downstream metabolism. *Rba. sphaeroides* was engineered to express PR and its chromophore retinal.

The gene for a transmembrane synthetic peptide maquette was designed and expressed in *Rba. sphaeroides*. This work forms the basis of the bottom-up redesign of photosystem components with the aim of augmenting photosynthesis in *Rba. sphaeroides* and in the long term to create new photosynthetic complexes and membrane assemblies. In addition, the plasticity of the *E. coli* Tat pathway for the export of maquettes was investigated, as the quality control mechanism of the Tat transporter makes it a desirable system for efficient large-scale protein production to facilitate further characterisation of maquette proteins.

Acknowledgements

First I would like to thank my supervisor Neil Hunter for his continued support and advice through the years that I have worked in the lab. I would also like to thank Graham Leggett for the opportunity to work as part of the Low-Dimensional Chemistry project.

I have had the pleasure of working with and visiting several collaborators with whom I have gained many new skills and experiences. The maquette work was done in collaboration with Les Dutton, Goutham Kodali and Bohdana Discher at the University of Pennsylvania. The Tat work was done with Colin Robinson, Alex Jones and Cristina Matos at University of Kent. Thank you to everyone at Washington University in St Louis who contributed to the RC-YFP work, including Christine Kirmaer, Dewey Holten, Preston Dilbeck and Jon Yuen. I would particularly like to thank Kaitlyn Faries for the ultrafast transient absorbance data on the RC-YFP complexes. Thank you for making me feel so welcome and at home when I made my visits. Thank you to PARC who funded my visit to St Louis.

I would like to thank all the people in the Hunter lab, past and present, who have helped me with their expertise and wise advice over the course of my PhD. In particular I would like to thank Dr Pu Qian for producing the EM images and model of the LH1-RC-YFP complex, Xia Huang for the YFP lifetimes data, Dr John Olsen for help with the 77 K membrane data, Dr Dan Canniffe for help with gene design and synthesis, Dr Sarah Hollingshead for help with the *Synechocystis* work, Shuang Chi for carotenoid expertise, and Dr. Amanda Brindley and Dr Pu Qian for all their help with protein purification. Thank you to Elizabeth Martin (with magic cloning skills), Dr Dave Mothersole and Dr Michaël Cartron for teaching me all they know. Thank you to Dr Andy Hitchcock for helping me out with a new short loop version of the transmembrane maquette as I reached the last stages of thesis writing.

Thank you to all my friends in Sheffield and back at home. Thank you to Wednesday Club: Abi, Sophie, Suze, Simon, Lucy and Sarah for getting me through my PhD with weekly beers in the University Arms. Thanks for all the lunches in the park, duck watching and for always being there to answer my questions regarding thesis writing and science knowhow. Thank you also to the Janets and the Goons for cheering me on all the way. Thank you to Will Wood for keeping life in E11 lively.

Finally, I would like to say a special thank you to my family for their love, belief and support. My parents, who can now pronounce "*Rhodobacter sphaeroides*" perfectly, have always taken a strong interest in my research and bravely sat through many practice presentations and proof read sections of my thesis – thank you! Thank you to my sister, Hannah, for

providing me with time-out from my thesis by taking me to wacky art exhibits. To Alex, who also diligently sat through so many practice presentations, thank you for all your love and patience over the last 4 years – here's to the future!

Table of Contents

Chapter 1: Introduction

Chapter 2: Materials and Methods

List of Figures

List of Tables

Abbreviations

β-DDM	n-dodecyl-β-D-maltopyranoside
Å	Angstrom(s)
AFM	atomic force microscopy
ALA	5-aminolevulinic acid
amp	ampicillin
AP	amphiphilic
ATP	adenosine triphosphate
AU	absorbance units
B800	bacteriochlorophyll in LH2 with an absorbance maximum of 800 nm
B850	bacteriochlorophyll in LH2 with an absorbance maximum of 850 nm
B875	bacteriochlorophyll in LH1 with an absorbance maximum of 875 nm
BC(M)DO	beta-carotene (mono)dioxygenase
BChl(s)	bacteriochlorophyll(s)
bp	base pair(s)
X ^{o}C	X degrees Celsius
Chl(s)	chlorophyll(s)
Crt	carotenoid
cyt	cytochrome
DAD	diode array detector
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid
EM	electron microscopy
g/ µg/ ng	gram(s)/microgram(s)/nanogram(s)
GFP	green fluorescent protein
GPR	green proteorhodopsin

HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HPLC	high performance liquid chromatography
hr	hour(s)
ICM	intracytoplasmic membrane
IPTG	isopropyl β-D-1-thiogalactopyranoside
(N)IR	(near)infrared
K	Kelvin
kb	kilobase
kDa	kilodaltons
km	kanamycin
LB	Luria-Bertani medium
LDAO	lauryldimethylamine N-oxide
LH1	light-harvesting complex 1
LH2	light-harvesting complex 2
mA	milliamps
mat	mature
min	minute(s)
ml/ µl/ L	millilitre(s)/ microlitre(s)/ litre(s)
M/ mM	molar/ millimolar
mRNA	messenger RNA
nm/ µm	nanometre(s)/ micrometres(s)
OD_x	optical density (at "x" nm)
PCR	polymerase chain reaction
PGC	photosynthetic gene cluster
PR	proteorhodopsin
pre	precursor
psi	pounds per square inch
Q (UQ)	quinone (ubiquinone)
Q_B	quinone B (site)
QH_2O	distilled water purified using the Milli-Q system

Rba.	*Rhodobacter*
RBS	ribosome binding site
RNA	ribonucleic acid
Rps.	*Rhodopseudomonas*
RC	reaction centre
RC-only	reaction centre only i.e. without LH1 ring
rpm	revolutions per minute
Sec	secretory pathway
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMA	styrene maleic acid
Suc	sucrose
SYFP2 / *syfp2*	strongly-enhanced yellow fluorescent protein 2
TA	transient absorption
Tat	twin-arginine translocation pathway
TB	terrific broth
TM	transmembrane maquette
T_m	melting point
Tris	tris hydroxymethyl methylamine
UPB	upper pigmented band (immature intracytoplasmic membrane)
UV	ultraviolet
V	volts
WT	wild-type
YFP	yellow fluorescent protein

Chapter 1

Introduction

1.1 Photosynthesis

Photosynthesis is the process by which light energy is captured and stored by an organism. This stored energy is used to drive energy-requiring cellular processes. Photosynthetic organisms, including plants, algae and many species of bacteria, are the primary source of energy for the majority of life on Earth. The overall process for photosynthesis is a redox reaction and can be represented by the equation:

$$CO_2 + 2H_2A + (light) \rightarrow (C\text{-}H_2O) + 2A + H_2O \qquad \text{(van Niel, 1941)}$$

H_2A represents electron and hydrogen ion donors, such as H_2O and H_2S, which are used to reduce CO_2 to produce carbohydrate ($C\text{-}H_2O$) and A, an oxidized electron donor. Oxygenic photosynthesis occurs in plants, algae and cyanobacteria, and uses water as the electron donor with O_2 as the oxidation product. Photosynthetic bacteria, excluding cyanobacteria, perform anoxygenic photosynthesis, and they utilize a number of different compounds, including inorganic acids such as succinate and acetate yielding an oxidation product that is not molecular oxygen.

Photosynthesis can be split into two main stages: the light-dependent reactions, and the light-independent or “dark” reactions. During the light-dependent reactions photons of solar energy are harvested by pigment-protein complexes and are transduced to drive proton and electron transfer. Ultimately, a proton gradient across the membrane is formed, facilitating the synthesis of ATP. During the dark reactions CO_2 and water are fixed as organic compounds.

1.2 Photosynthetic organisms

1.2.1 Classification of photosynthetic organisms

According to classification of organisms using the phylogenetic class of their small subunit RNA, photosynthetic organisms are found within the eukarya and bacteria domains (Figure 1.1). No archael photosynthetic organisms have been found. The photosynthetic phyla are: plants (including algae), purple bacteria, green sulphur bacteria, cyanobacteria, Gram-positive bacteria, acidobacteria, and green non-sulphur bacteria. These organisms all perform chlorophyll-based photosynthesis.

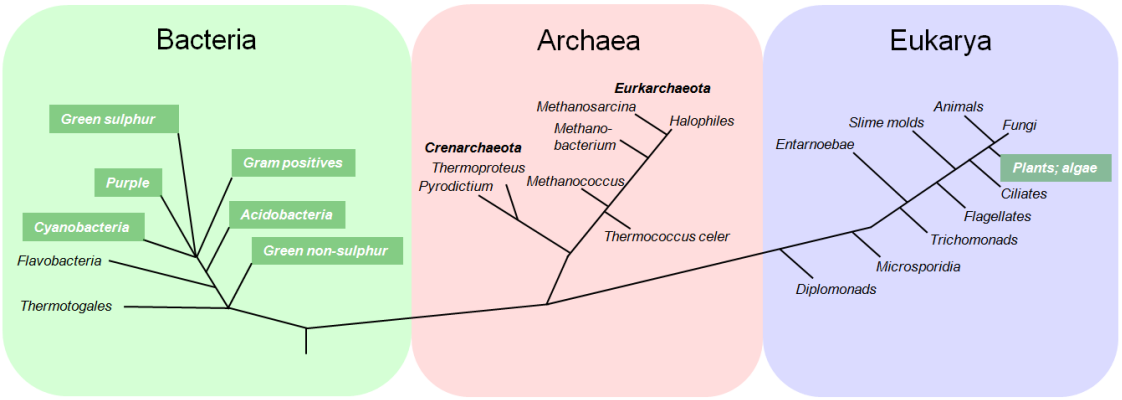


Figure 1.1 Classification of chlorophyll-based photosynthetic organisms

Phylogenetic classification is based on the small rRNA subunit. Phyla with photosynthetic organisms present are highlighted in green.

(Adapted from Mothersole, 2013)

Some organisms perform rhodopsin-based photosynthesis which is mechanistically very different from chlorophyll-based photosynthesis and involves the *cis* to *trans* isomerisation of a chromophore directly coupled to ion transport across a membrane (Lanyi, 2004). The fact that only two principal types of light-harvesting (chlorophyll-based and rhodopsin-based) are known strongly suggests that microorganisms started using solar energy for photosynthesis at the very beginning of biological evolution on Earth (Zubkov, 2009).

1.2.1.1 Purple photosynthetic bacteria

Purple photosynthetic bacteria are metabolically versatile anoxygenic phototrophs. They are capable of growing photoautotrophically, photoheterotrophically, aerobically, anaerobically and fermentatively (Imhoff, 1995). Most purple bacteria are capable of fixing molecular nitrogen. Unlike purple non-sulphur bacteria, purple sulphur bacteria are capable of using sulphur containing compounds such as H_2S as an electron donor (Brune, 1995). The Calvin cycle fixes CO_2 in all purple bacteria (Tabita, 1995).

1.2.1.2 Green bacteria

There are two types of green bacteria: green sulphur, which primarily use H_2S as an electron donor, and green non-sulphur, which do not use H_2S.

Green sulphur bacteria are obligate anoxygenic phototrophs and strict anaerobes. They are capable of nitrogen fixation and can fix CO_2 through a reverse tricarboxylic acid cycle (Sirevag, 1995). Green sulphur bacteria are found in environments with very low light levels, such as below the chemocline in stratified lakes.

Green non-sulphur bacteria are significantly more versatile than green sulphur bacteria and are capable of photoautotrophy, photoheterotrophy and aerobic growth. CO_2 is fixed using the hydroxypropionate pathway.

1.2.1.3 Heliobacteria

Photosynthetic heliobacteria are anoxygenic and strictly anaerobic. They are often found in soils and rice paddies and are capable of producing endospores. They are active nitrogen fixers but CO_2 fixation in photosynthetic heliobacteria is not well understood. Heliobacterial photocomplexes contain some similar characteristics to those found in green sulphur bacteria and cyanobacteria. The primary pigment involved in heliobacterial photosynthesis is chlorophyll *g*, which is unique to this group.

1.2.1.4 Cyanobacteria

Cyanobacteria are a large and diverse oxygenic photosynthetic group, occupying a wide range of habitats across the Earth, including extreme environments such as hypersaline bays and hot springs. They are often referred to as "blue-green algae"; this is considered by some as a misnomer as, although the photosynthesis performed by cyanobacteria is remarkably similar to eukaryotic phototrophs, they are prokaryotic. Many species of cyanobacteria can fix N_2 and some are capable of using H_2S as an electron donor (Padan, 1979). According to the endosymbiotic theory, chloroplasts found in plants and eukaryotic algae evolved from cyanobacterial ancestors.

1.2.1.5 Acidobacteria

The physiologically diverse and ubiquitous Acidobacteria have been assigned to a newly devised phylum, which includes the anoxygenic photoheterotroph *Candidatus Chloracidobacterium thermophilum*, discovered in 2007 in a hot spring in Yellowstone National Park (Bryant *et al*., 2007).

1.2.1.6 Algae and plants

One distinct phylum of the "eukarya" domain is capable of photosynthesis: the vast and diverse group of algae and plants. Although algae and plants are complex, the same basic principles of eukaryotic photosynthesis are found in oxygenic cyanobacteria.

1.2.1.7 Rhodopsin-based photosynthesis

Rhodopsin-based photosynthesis is mechanistically very different to the chlorophyll-based photosynthesis undertaken by all the above groups. Ions, either Cl^- (halorhodopsins) or H^+ (bacteriorhodopsins), are pumped across the membrane as a consequence of the action of light. For many years, it was assumed rhodopsin-based photosynthesis was only found in

extremely halophilic Archaea, however a new form of bacterial rhodopsin, known as proteorhodopsin, was recently discovered in marine proteobactera. Proteorhodopsin is discussed further in Section 1.9. (Beja *et al*., 2000; DeLong and Beja, 2010).

1.2.2 *Rhodobacter sphaeroides*

Rhodobacter (Rba.) sphaeroides is the best characterised purple non-sulphur bacterium. It is Gram-negative, rod-shaped and metabolically diverse. Its preferred growth conditions are anaerobic photoheterotrophy and aerobic chemoheterotrophy. It is also capable of dizatrophic growth and fermentation (Blankenship *et al*., 1995). *Rba. sphaeroides* is found in anoxic zones of water in deep lakes, soil, mud, sludge, sewage and waste lagoons (Siefert *et al*., 1978; Cooper *et al*., 1975).

Rba. sphaeroides grows rapidly in liquid medium in the laboratory under both anaerobic photoheterotrophic and aerobic chemotrophic conditions. Because it is not dependent on phototophy, the genes essential for anoxygenic photosynthesis can be removed, allowing for in-depth study. Its genome is small, fully sequenced and well annotated, allowing for rapid genomic manipulation. A wealth of structural information is available for the membrane protein complexes involved in photosynthesis which include the light harvesting proteins, LH2 and LH1, the reaction centre, the cytochrome bc_1 complex and ATP synthase.

1.3 Pigments in photosynthesis

1.3.1 The role of pigments in photosynthesis

All photosynthetic organisms require pigments to absorb light in order for energy to be stored. The light-harvesting complexes of *Rba. sphaeroides* primarily contain two types of pigment: bacteriochlorophyll (BChl) and carotenoids. Light-harvesting membrane proteins absorb photons; on absorbing light the molecule becomes excited. This electronic excitation energy can then be transferred to nearby BChl molecules in the peripheral complexes until it reaches the reaction centre (RC). BChls are also present in the RC where they participate in primary charge separation, in which excitation energy is transduced into electron transfer energy (Figure 1.2).

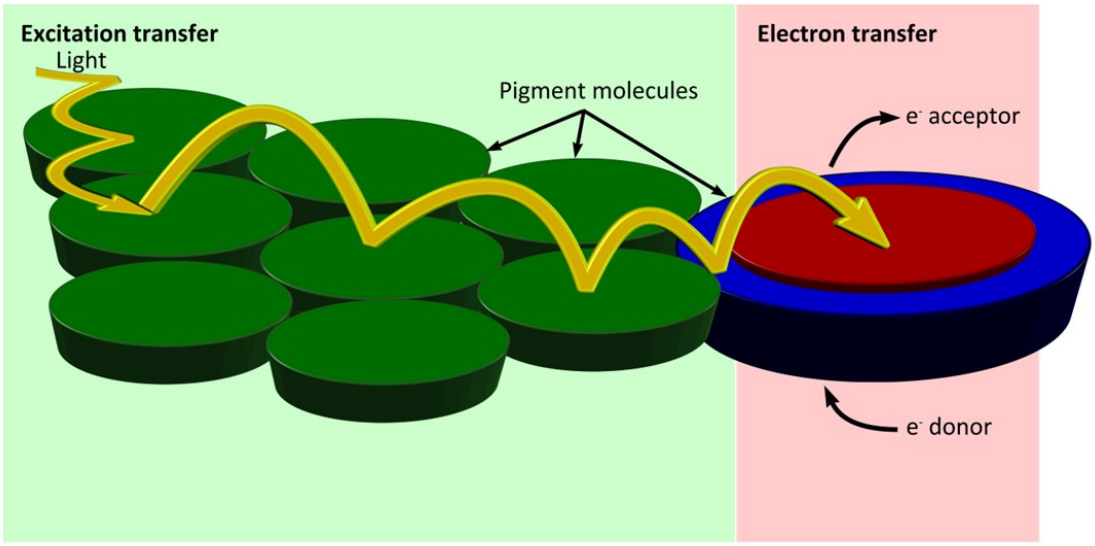


Figure 1.2 Basic concept of energy transfer from pigmented light-harvesting complexes to reaction centres in *Rba. sphaeroides*

Photons of light are absorbed by the pigment molecules in the peripheral light-harvesting LH2 complexes of *Rba. sphaeroides* (green) and create an excited state. This electronic excitation energy can be transferred from pigment to pigment until a reaction centre (red) is reached, via the surrounding LH1 complex (blue). At the reaction centre, the excitation energy is transduced to a photochemical charge separation, then ultimately as a quinol.

(From Mothersole, 2013)

1.3.2 Bacteriochlorophyll *a* biosynthesis

Bacteriochlorophyll *a* (BChl *a*) is the primary pigment found in the light harvesting complexes and reaction centres of most purple bacteria, including *Rba. sphaeroides*. BChl molecules have an extensive π system of conjugated bonds, allowing them to absorb light in the visible range. BChl *a* is a planar molecule consisting of three pyrroles and one pyrroline forming a bacteriochlorin ring with a central Mg atom. The D ring of the macrocycle has a long, hydrophobic phytol tail which is important for stability and anchorage to hydrophobic proteins. It is also important for the correct assembly and function of the light-harvesting complexes (Bollivar *et al*., 1994; Addlesee and Hunter, 1999).

The biosynthesis pathway of BChl *a* in *Rba. sphaeroides* starts with the formation of 5'–amino levulinic acid (ALA) from the condensation of succinyl CoA and glycine catalysed by the enzyme HemA via the Shemin pathway (Shemin, 1956; Ohhama *et al*., 1985). Ultimately protoporphyrin IX is formed, which can then follow either the BChl or haem biosynthesis pathway depending on whether an Mg^{2+} or an Fe^{2+} ion is chelated by either magnesium chelatase or ferrochelatase respectively (Figure 1.3). Once Mg^{2+} has been chelated into

protoporphyrin IX, the molecule is transformed into BChl *a* via a series of hydration, oxidation and reduction reactions (Figure 1.4).

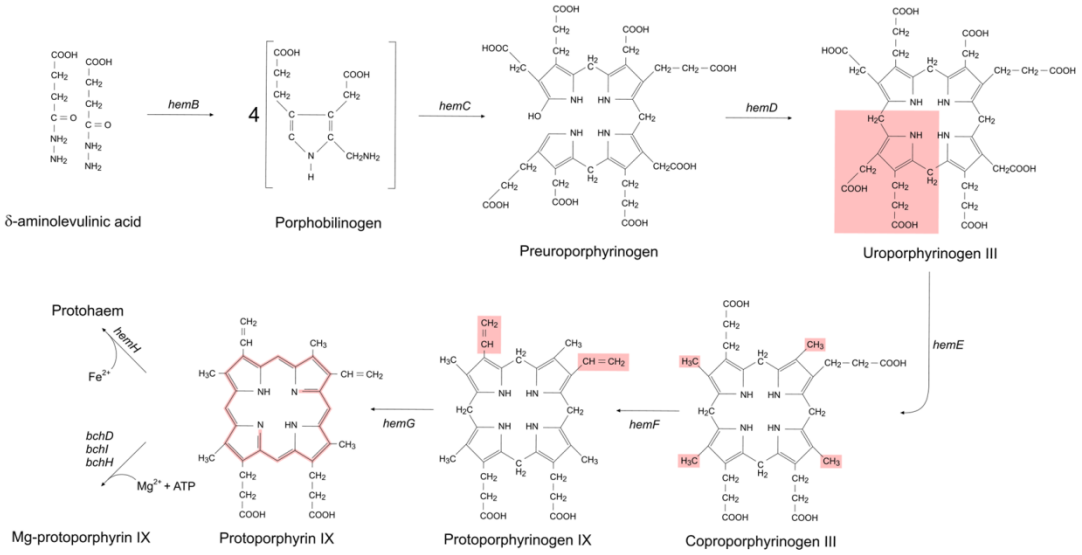


Figure 1.3 Tetrapyrrole biosynthesis from δ-aminolevulinic acid to protoporphyrin IX

Detailed schematic of the tetrapyrrole biosynthesis pathway from δ-aminolevulinic acid to magnesium protoporphyrin IX. Modifications at each step are highlighted in pink.

(From Mothersole, 2013)

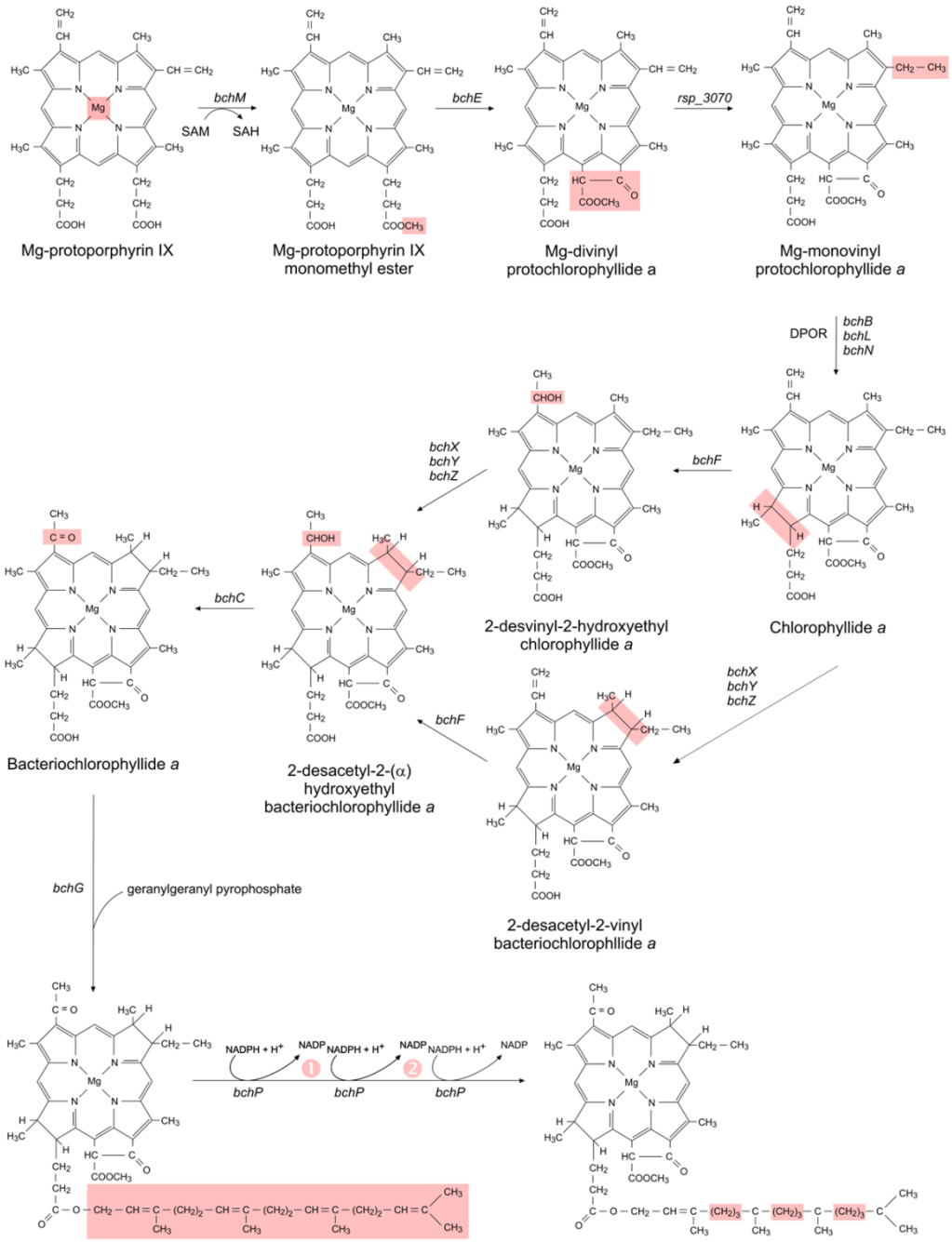


Figure 1.4 Bacteriochlorophyll *a* biosynthesis from Mg-protoporphyrin IX

Detailed schematic of the bacteriochlorophyll *a* biosynthesis pathway from magnesium protoporphyrin IX to bacteriochlorophyll synthase and geranylgeranyl reductase. Modifications at each step are highlighted in pink.

(From Mothersole, 2013)

1.3.3 Carotenoid biosynthesis

Carotenoids are hydrophobic tetraterpenoid molecules commonly associated in membrane bound pigment-protein photosynthetic complexes (Cogdell and Frank, 1987). In

photosynthetic organisms they perform three functions: light-harvesting, photoprotection and the provision of structural stability.

Carotenoids play a role in light-harvesting by absorbing light energy in the 459-570 nm range due to their conjugated π electron system. They can transfer excitation energy to neighbouring BChl *a* molecules, which do not absorb light energy from this range as efficiently (Cogdell and Frank, 1987).

In the presence of excess light, BChl can form a triplet state (Figure 1.5 A-B) capable of donating its energy to O_2 producing singlet oxygen (Figure 1.5 C), a powerful oxidant that can cause major and rapid damage to the organism. Carotenoids help prevent this photo-oxidative damage by acting as oxygen scavengers (Figure 1.5 D), directly quenching the oxygen singlet state (Foote and Denny, 1968) or by performing a trapping reaction (Figure 1.5 E). Carotenoids can also quench the BChl triplet state directly before it can interact with O_2 (Borland *et al.*, 1989); this is the most common photoprotective function of carotenoids *in vivo* (Cogdell and Frank, 1987). This photoprotective function becomes most apparent in carotenoidless mutants of *Rba. sphaeroides* such as the mutant R-26, which are extremely sensitive to the presence of oxygen, leading to photo-oxidative cell death in the presence of light (Clayton and Smith, 1980; Cogdell and Frank, 1987).

A. $BChl + hv \rightarrow {}^1BChl^*$ (Singlet excited)

B. ${}^1BChl^* + hv \rightarrow {}^3BChl^*$ (Triplet excited)

C. ${}^3BChl^* + O_2 \rightarrow BChl + {}^1\Delta gO_2$ (Singlet oxygen)

D. ${}^1\Delta gO_2 + Carotenoid \rightarrow O_2 + {}^3Carotenoid^*$

${}^3Carotenoid^* \rightarrow Carotenoid + heat$

E. ${}^3BChl^* + Carotenoid \rightarrow BChl + {}^3Carotenoid^*$

${}^3Carotenoid^* \rightarrow Carotenoid + heat$

Figure 1.5 The photoprotective function of carotenoids

In conditions of excess light, bacteriochlorophyll (BChl) can form a triplet state (A and B). This triplet state can interact with molecular oxygen to form singlet oxygen capable of causing major photo-oxidative damage. Carotenoids can quench both the singlet oxygen state (D) and the triplet BChl (E) prior to it interacting with O_2. Carotenoids return to their ground state by dissipating energy in the form of heat.

(Adapted from Cogdell and Frank, 1987)

Carotenoids are important for the stability of the photosynthetic complexes. They are essential for the formation of the *Rba. sphaeroides* LH2 light-harvesting complex; the carotenoidless mutant R-26 cannot produce fully assembled LH2 complexes (Zurdo *et al*., 1993; Lang *et al*., 1995). They also play a role in the dimerisation of the *Rba. sphaeroides* RC-LH1-PufX core complex; R-26 primarily contains monomeric core complexes (Ng *et al*., 2011).

The carotenoid biosynthesis pathway of *Rba. sphaeroides* is shown in Figure 1.6. Under different growth conditions different carotenoids are accumulated. Anaerobic growth results in the accumulation of spheroidene and hydroxyspheroidene, giving the cells a yellow-brown colour. Aerobic growth results in the accumulation of spheroidenone, giving the cells a red colour. Semi-aerobic growth results in the accumulation of hydroxyspheroidenone (Schmidt, 1978; Lang *et al*., 1995).

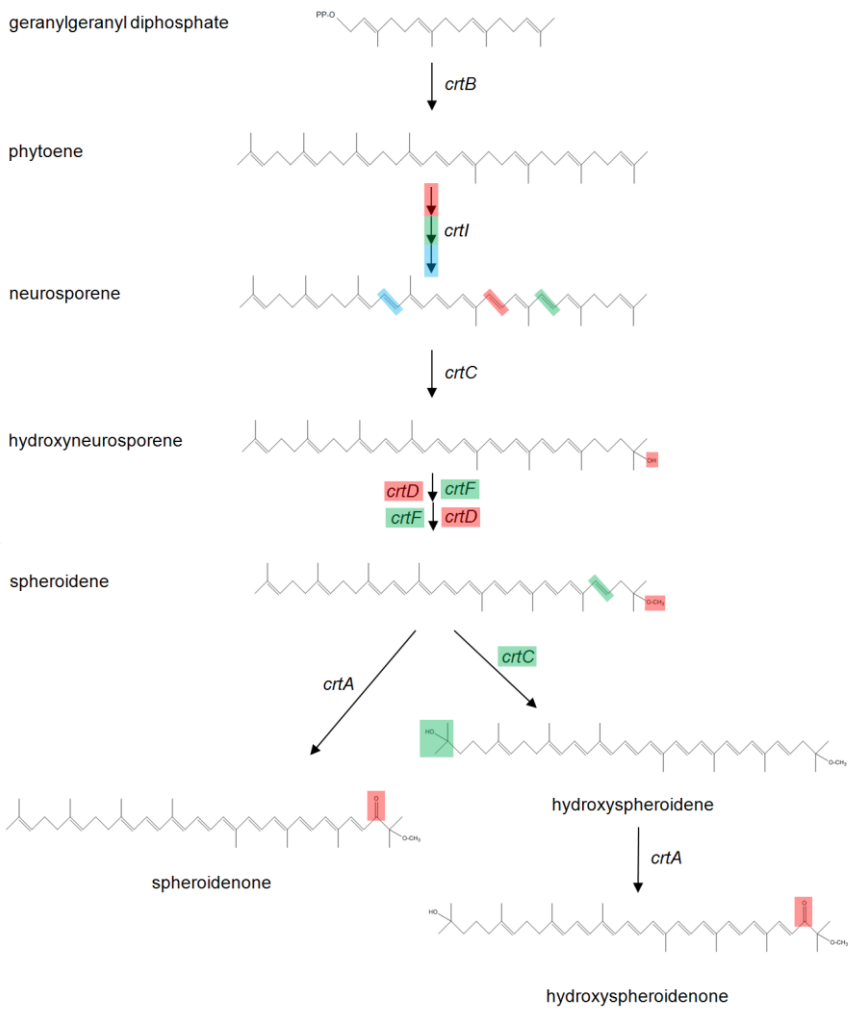


Figure 1.6 The carotenoid biosynthesis pathway of *Rhodobacter sphaeroides*

Detailed schematic of the carotenoid biosynthesis pathway in *Rba. sphaeroides* from geranylgeranyl diphosphate to spheroidenone or (hydroxy)spheroidenone depending on growth conditions.

(Modified from Mothersole, 2013)

1.4 Light harvesting in *Rhodobacter sphaeroides*

1.4.1 The photosynthetic unit of *Rhodobacter sphaeroides*

The photosynthetic unit (PSU) of purple photosynthetic bacteria is made up of a network of pigment–protein complexes that absorb solar energy as electronic excitation energy (Sener *et al*., 2007). PSUs are are located in specialised intracytoplasmic membranes (ICM) (also known as chromatophores) and are interconnected to allow excitation energy transfer to take place (Schachman *et al*., 1952).

The light-harvesting (LH) pigment-protein complexes of *Rba. sphaeroides* are composed of highly organised and efficient structural protein, BChl *a* and carotenoids. There are three different light-harvesting complexes in *Rba. sphaeroides*: the peripheral light-harvesting complexes, LH1 and LH2, and the reaction centre (RC), which is the site of energy transduction from excitation to electron transfer energy. Together, the LH1 and the RC form the "core complex" and, in wild-type *Rba. sphaeroides*, are stochiometrically fixed at a 1:1 ratio (Aagaard and Sistrom, 1972). Variation in light intensity changes the level of LH2 complexes in relation to the core complexes within the photosynthetic membrane (Aagaard and Sistrom, 1972; Adams and Hunter, 2012).

The LH2 antenna complex captures photons and transfers the resulting electronic excitation to LH1, which funnels it to the RC (Figure 1.2), where excitation energy is converted to a charge separation. Both LH complexes are made up of transmembraneous αβ-heterodimers. Energy migration is directed by the spatial organisation and energetic order of the pigments, which is determined by their binding sites within the LH complexes.

1.4.2 The peripheral light-harvesting complex LH2

LH2 complexes are found in many purple bacteria, with some exceptions including *Rhodospirillum rubrum* and *Blastochloris viridis,* the photosynthetic apparatus of which consists only of the RC-LH1 complex (Zuber and Brunisholz, 1991; Hawthornthwaite and Cogdell, 1993).

X-ray crystal structures of the LH2 complexes of *Phaeospirillum (Ph.) molischianum* and *Rhodopseudomonas (Rps.) acidophila* have identified octameric and nonameric structures respectively (McDermott *et al*., 1995; Koepke *et al*., 1996). Atomic force microscopy (AFM) of *Rhodospirillum (Rsp.) photometricum* LH2 complexes has identified nonameric, octameric and decameric oligomerisation states, leading to the suggestion that LH2 complexes may be heterogenous (Scheuring *et al*., 2004).

A 6 Å electron microscopy projection map showed that the *Rba. sphaeroides* LH2 is nonameric (Walz *et al*., 1998); this was confirmed by AFM of 2D crystals (Walz *et al*., 1998; Scheuring *et al*., 2003). As of February 2015, no high resolution 3D crystal structure of *Rba. sphaeroides* LH2 has been obtained. Therefore the structure discussed here is that of *Rhodopseudomonas acidophila,* which on the basis of its nonameric ring and the high sequence homology, is thought to be similar and has been solved to 2 Å using X-ray crystallography (Figure 1.7) (Papiz *et al*., 2003).

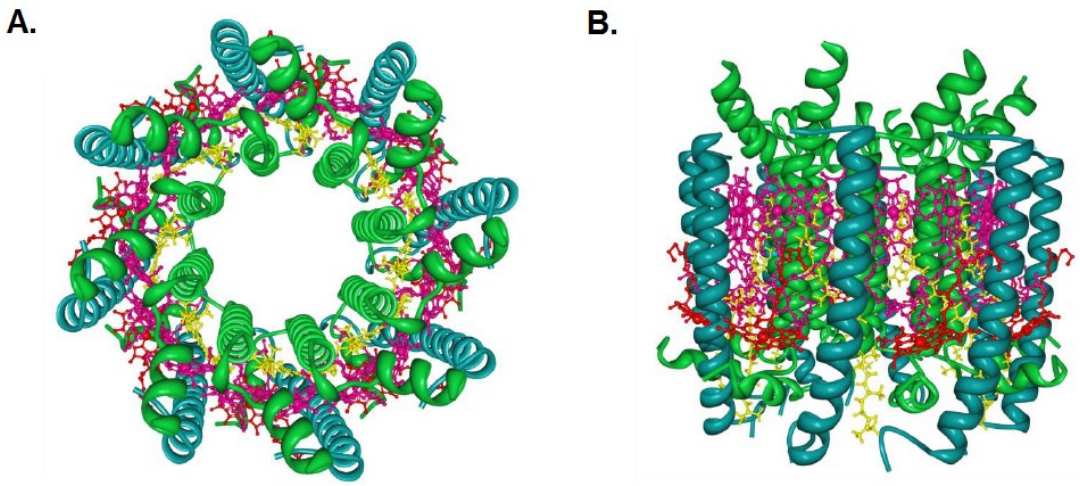


Figure 1.7 The LH2 complex of *Rhodopseudomonas acidophila*

Components have been coloured as follows: LH2 α-polypeptides in green, LH2 β polypeptides in cyan, B850 BChl in pink, B800 BChl in red and carotenoids in yellow.

- **A.** Projection view from the periplasmic side of the complex.
- **B.** Side view with the periplasmic side uppermost.

(Modified from Mothersole, 2013)

The LH2 complex consists of two concentric cylinders, each composed of nine non-covalently associated single transmembrane polypeptides; the β helical subunits form the outer ring, while the α helical subunits form the inner ring. 27 BChl *a* molecules are arranged inside the complex in two stacked rings. Nine B800 BChl *a* molecules are positioned perpendicular to the transmembrane helix axis between the outer β helices. Eighteen B850 BChl *a* molecules form an overlapping ring sandwiched between the α and β helices. Light energy is harvested and the resulting electronic excitation energy is passed from the B800 ring to the B850 ring on a timescale of 650 fs. Excitation energy transfer within the B850 ring occurs on a timescale of 100 fs due to the dense packing of the BChl *a* molecules (Jimenez *et al*., 1996). Once excitation energy is rapidly circulating around the B850 BChls, it can hop at any stage to an

adjacent ring, there is no requirement for a special interaction site with a neighbouring ring (McDermott *et al*., 1995; Koepke *et al*., 1996).

13-14 carotenoid molecules are present in the structure, in a ratio of 2:1 with BChl *a* (McDermott *et al*., 1995; Papiz *et al*., 2003). 9 of these carotenoids are closely associated with the phytyl tails of the BChl *a* molecules (Freer *et al*., 1996). In *Rba. sphaeroides* they are spheroidene or spheroidenone depending on the growth conditions (in *Rps. acidophila*, the carotenoids are rhodopin glucoside). In the absence of coloured carotenoids, the LH2 polypeptides (α and β) are synthesised but are not stably incorporated into the membrane (Lang and Hunter, 1994).

While LH2 imparts significant membrane curvature, mutants lacking LH2 produce tubular membranes consisting exclusively of core complex dimers (Siebert *et al*., 2004).

1.4.3 The major light-harvesting complex LH1

The *Rba. sphaeroides* LH1 consists of 14 αβ heterodimers forming two concentric but incomplete rings. In bacteria such as *Rsp. rubrum* and *Thermochromatium tepidum* there is a complete LH1 ring with a diameter of 116 Å, enclosing a 68 Å hole that accommodates a reaction centre (Karrasch *et al*., 1995). AFM and EM of LH1 complexes has revealed that it is capable of many different conformations, including open rings and polygonal rings (Jamieson *et al*., 2002; Bahatyrova *et al*., 2004b). AFM data of membranes from an LH1-only mutant of *Rba. sphaeroides* has led to the conclusion that cooperative association of αβ heterodimers with the RC drives LH1 ring assembly to completion (Olsen *et al*., 2014). 32 B875 BChl *a* molecules are present in an LH1 complex that overlap in a similar way to the B850 molecules in LH2. *In vivo*, LH1 is predominantly found as a component of the core complex.

1.4.4 The reaction centre

Excitation energy is converted to electron transfer energy at the reaction centre (RC). The RC from *Rba. sphaeroides* is composed of three protein subunits: L (light) and M (medium), each containing five transmembrane α-helices and related by pseudo-twofold symmetry, and the H (heavy) subunit which has a single transmembrane helix with the bulk of its mass forming a globular domain at the cytoplasmic side of the membrane (Figure 1.8). Non-covalently bound cofactors are associated with the RC. Two BChl *a* molecules form a strongly interacting dimer known as the 'special pair'. Also bound are two accessory BChl *a*, molecules two bacteriopheophytins, two quinones (ubiquinone), a non-haem Fe^{2+} ion, and a carotenoid molecule.

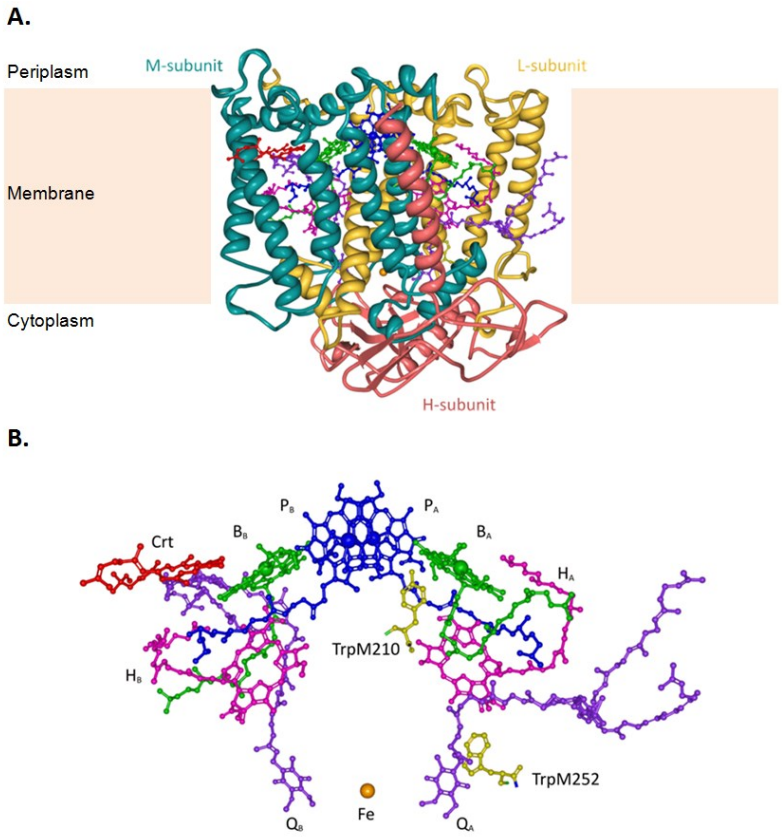


Figure 1.8 The *Rhodobacter sphaeroides* reaction centre

A. Model of the reaction centre showing the L (orange), M (cyan) and H (red) subunits. Pigment molecules are in the same colours as B.

B. The reaction centre pigment cofactors including the two molecules of BChl *a* (P_A and P_B) known as the "special pair" (blue), the two accessory molecules of BChl (B_A and B_B) (green), the two molecules of bacteriopheophytin (H_A and H_B) (pink), and the carotenoid (spheroidene) (red). Also shown are two quinones are (Q_A and Q_B) (purple), the Fe atom (orange) and two important aromatic amino acid side chains (yellow).

(From Mothersole, 2013. Modified from PDB ID 3I4D file submitted by Fujii *et al*., 2010)

High resolution X-ray crystallography structures of the RC have been obtained from *Blastochloris (Blc.) viridis* and *Rba. sphaeroides*. The *Blc. viridis* RC consists of a four-subunit complex (Deisenhofer *et al*., 1985; Deisenhofer and Michel, 1989; Deisenhofer *et al*., 1995). The *Rba. sphaeroides* structure was obtained and refined by numerous groups (Chang *et al*., 1986; Allen *et al*., 1987a; Allen *et al*., 1987b; Chang *et al*., 1991; Ermler *et al*., 1994; McAuley *et al*., 2000; Katona *et al*., 2003). At present the highest available resolution structure is 2.01 Å (Fujii *et al*., 2010). The *Blc. viridis* reaction centre has a fourth subunit, cytochrome *c,* which is replaced in *Rba. sphaeroides* by the soluble periplasmic protein cytochrome c_2.

1.4.5 The core complex

RC-LH1 complexes use harvested solar energy to power the reduction of quinone to quinol prior to the formation of the proton gradient that powers ATP synthesis. In *Rba. sphaeroides* the RC-LH1 complex forms dimeric S-shaped supercomplexes with the small polypeptide PufX; this is known as the core complex (Figure 1.9). The 3D crystal structure of the *Rba. sphaeroides* core complex has been determined to a resolution of 8 Å (Qian *et al*., 2013). Each monomer contains 14 LH1 subunits surrounding a RC in a C-shaped assembly (Qian *et al*., 2005; Qian *et al*., 2013).

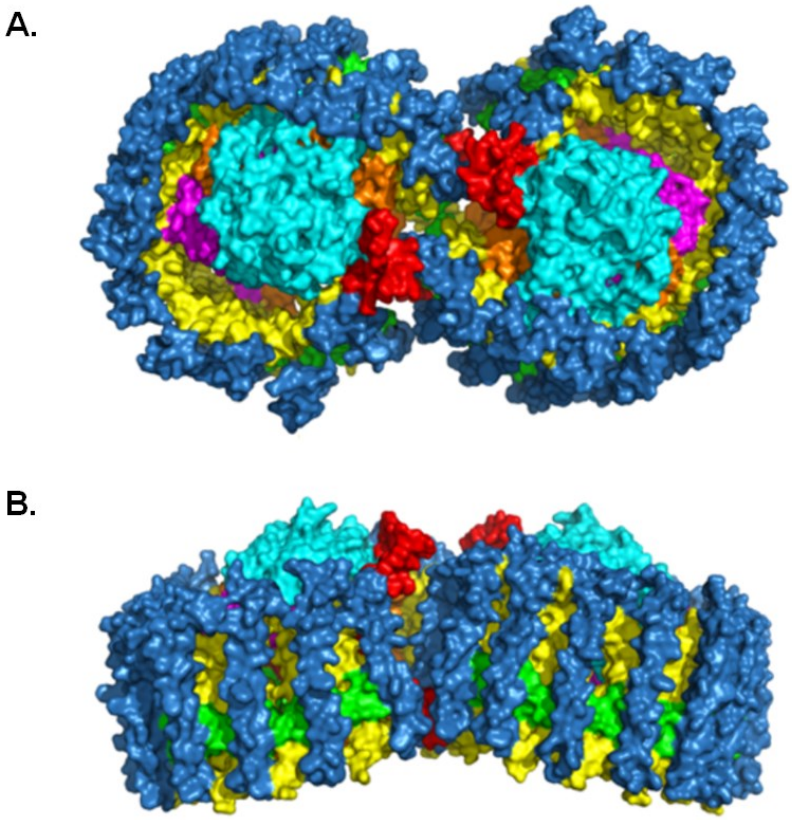


Figure 1.9 The *Rhodobacter sphaeroides* core complex dimer

LH1 β in blue, LH1 α in yellow, RC-H in cyan, RC-L in orange, RC-M in magenta, and PufX in red.

- **A.** Cytoplasmic face of the complex viewed perpendicular to the membrane.
- **B.** Complex viewed in the plane of the membrane

(Adapted with permission from Qian *et al*., 2013. Copyright American Chemical Society 2013)

The LH1 is prevented from fully encircling the RC through close contacts between PufX, an LH1 αβ subunit and the cytoplasmic domain of the RC-H subunit. This creates a channel connecting the RC Q_B site with the opening in the LH1 ring, allowing Q/QH_2 exchange with the external quinone pool (Qian *et al*., 2013).

An earlier 3D model of the *Rba. sphaeroides* RC generated to a 25 Å resolution demonstrated that the two dimer halves incline towards each other at an angle of about 146°, creating a V-shaped structure that imposes curvature on the membrane (Qian *et al*., 2008). PufX is essential to the formation of the dimeric complex. In mutants lacking PufX only monomeric RC-LH1 complexes are formed that exhibit large areas of hexagonically packed core complexes with significantly reduced membrane curvature (Francia *et al*., 1999; Qian *et al*., 2008).

Membranes of *Rba. sphaeroides* contain a mixture of monomeric and dimeric core complexes; the monomeric form predominates when grown under chemoheterotrophic conditions and the dimeric form predominates under anaerobic heterotophic growth (Ratcliffe *et al*., 2011; Crouch and Jones, 2012). Some other species of photosynthetic bacteria, such as *Rhodopseudomonas (Rsp.) palustris,* possess only monomeric core complexes (Roszak *et al*., 2003).

1.5 Transduction of excitation energy and the formation of ATP in photosynthesis

The sole purpose of the light-driven electron transfer in purple bacteria is to create a proton-motive force across the inner membrane. Excitation energy transfer begins within and between the LH2 and LH1 complexes before being passed to the reaction centre (RC). Excitation energy is passed to the "special pair" of BChls in the RC, followed by the production of charge separation and the initiation of transmembrane electron transfer. Electrons are transferred to ubiquinone, a lipid-soluble electron carrier, which picks up protons from the cytoplasm. The resulting doubly reduced compound, hydroquinone, diffuses through the cell membrane to the cytochrome bc_1 complex. Here hydroquinone transfers electrons to cytochrome c_2, a water soluble protein, whilst protons are pumped across the membrane. This generates a transmembrane proton gradient. This gradient is used by ATP synthase to produce ATP, or used directly for example to power the flagellar motor. Cytochrome c_2 shuttles electrons back to the RC reducing the oxidized special pair and the system is "reset". (Sener *et al*., 2007). This overall process is illustrated in Figure 1.10.

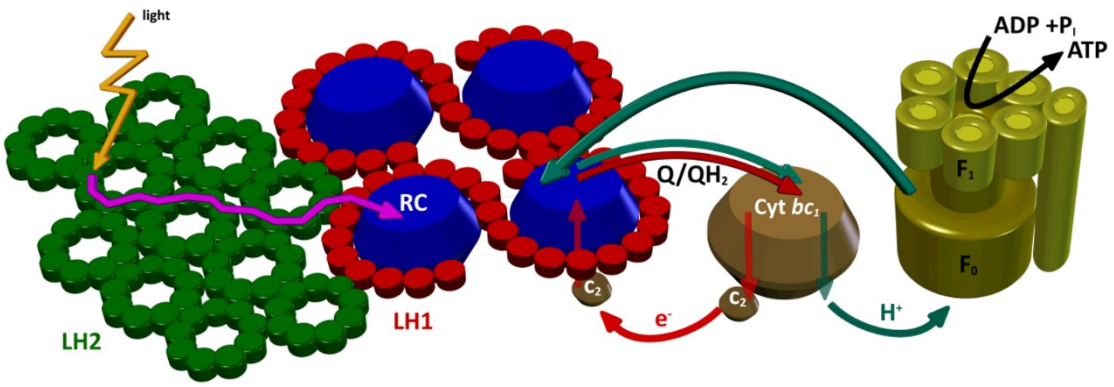


Figure 1.10 Schematic overview of electron and proton flow in the photosynthetic membranes of *Rhodobacter sphaeroides*

A schematic showing the overall process of light-harvesting and the flow of protons and electrons to transduce light energy into biochemical energy in the form of ATP by the pigment-protein complexes involved in photosynthesis in *Rba. sphaeroides*.

(From Mothersole, 2013)

1.5.1 Excitation energy transfer within LH1 and LH2

After the pigment molecules absorb photons, the energy transfer dynamics within LH1 and LH2 are ultrafast (VanGrondelle *et al*., 1994). In the case of *Rba. sphaeroides* and *Rps. acidophila,* energy transfer between B800 and B850 within LH2 takes approximately 650-800 fs at room temperature (Joo *et al*., 1996; Jimenez *et al*., 1996; Ma *et al*., 1997). Energy transfer time between neighbouring B800 molecules is approximately 0.7-1.25 ps, and between B850 molecules is approximately 110 fs, according to polarised pump-probe spectroscopy measurements (Monshouwer *et al*., 1995; Jimenez *et al*., 1996). Energy transfer between B875 molecules in *Rba. sphaeroides* LH1 is approximately 80-100 fs, according to ultrafast fluorescence depolarisation and annihilation studies (Bradforth *et al*., 1995).

Energy transfer in the light harvesting complexes also occurs between carotenoid and BChl molecules and occurs on an ultrafast timescale of a few 100 fs (Shreve *et al*., 1991; Ricci *et al*., 1996).

1.5.2 Excitation energy transfer to the reaction centre

Excitation energy is transferred between LH2 complexes, from LH2 to LH1 and finally from LH1 to the RC. Transfer between LH2 complexes and LH1 is approximately 3.3-4.6 ps (Hess *et al*., 1995a; Hess *et al*., 1995b; Nagarajan and Parson, 1997). The distance from the LH1 and the RC special pair is approximately 4.5 nm with an energy transfer time at 77 K of 35 ps. This makes it the rate limiting step in excitation trapping and reduces the probability of back

transfer from the RC to LH1 as the early events in charge separation occur within 1 ps at room temperature (Visscher *et al*., 1989; Beekman *et al*., 1994).

1.5.3 Transduction of excitation energy transfer to electron flow in the reaction centre

Upon excitation, the RC special pair (P_A P_B) promotes an electron to an excited state (P_A P_B*). This electron travels via the accessory BChl (B_A) to reduce a bacteriopheophytin (H_A), producing a charge separated state $P_AP_B^+$ H_A^-; this process takes approximately 2.8 ps (Breton *et al*., 1986). The electron is then transferred from H_A^- to the adjacent quinone, Q_A, forming a semiquinone, Q_A^-; this process takes approximately 220 ps (Holzapfel *et al*., 1990). The electron then moves within 100-200 µs to Q_B, forming the Q_B^- semiquinone, which is more stable (Gopta *et al*., 1999).

The charge separated $P_AP_B^+$ state is returned to the ground state P_AP_B through the donation of an electron from the soluble periplasmic electron carrier cytochrome c_2. This resets the cycle and the special pair is ready to receive excitation energy from LH1 again.

The flow of a second electron to Q_B, initiated by the special pair being excited by a second photon, reduces the quinone to its quinol form by the simultaneous uptake of two protons from the cytoplasm, forming Q_BH_2. Q_BH_2 then dissociates from the reaction centre and an unreduced quinone from the quinone pool in the membrane enters the Q_B site (Deisenhofer and Michel, 1991).

Electron flow in the reaction centre is reviewed in: Deisenhofer and Michel, 1991; Deisenhofer and Norris, 1993; Fleming and van Grondelle, 1994; Woodbury and Allen, 1995; Hoff and Deisenhofer, 1997; and Bixton and Jortner, 1999.

1.5.4 Completion of electron flow at the cytochrome *bc*$_1$ complex

Electron flow is completed at the cytochrome bc_1 complex, a large multisubunit integral membrane protein. The complex is also known as ubihydroquinone or cytochrome *c* oxidoreductase. There are three conserved subunits: cytochrome *b*, cytochrome c_1 and the Rieske iron-sulphur protein (ISP) (Hunte *et al*., 2008). Here quinol is oxidised and protons are pumped from the cytoplasm to the periplasm. The cycling of electrons and protons through this complex is known as the Q-cycle. The X-ray crystal structure of *Rba. sphaeroides* cytochrome bc_1 complex stabilised with an inhibitor has been solved to 2.6 Å (Esser *et al*., 2008). The complex forms an intertwined homodimer in both *Rba. sphaeroides* and *Rba. capsulatus* (Berry *et al*., 2004; Esser *et al*., 2008).

QH_2 enters the Q_0 catalytic site of the cytochrome bc_1 complex and forms a transient cytochrome *b*-QH_2-ISP complex, releasing two electrons into two different electron transfer chains; in this process two protons are released. One electron is released to the Rieske iron-sulphur cluster and the haem of cytochrome c_1 before reaching the soluble electron carrier protein cytochrome c_2. The other electron enters the Q_i site via haem b_L and haem b_H and reduces a quinone molecule to a stable semiquinone. A second molecule of QH_2 is oxidised to Q via the Q_0 site and, as a result, the semiquone in the Q_i site is reduced to QH_2, which is then released back into the membrane quinone pool. Overall, during this process, four protons are pumped into the periplasm and two cytochrome c_2 oxidases are reduced (Berry *et al*., 2004; Esser *et al*., 2006; Esser *et al*., 2008). This is shown in the equation:

$$2QH_2 + (2\text{cyt } c_2)_{ox} + Q + 2H^+_{(cytoplasm)} \rightarrow 2Q + (2\text{cyt } c_2)_{red} + 4H^+_{(periplasm)} + UQH_2$$

1.5.5 Formation of ATP by ATP synthase

The pH gradient, formed and maintained across the photosynthetic membrane by the cytochrome bc_1 complex, is used to drive the synthesis of adenosine triphosphate (ATP) from adenosine diphosphate (ADP) and inorganic phosphate by the F_1F_0 ATP synthase. ATP is universally used as the 'energy currency' in biological life and the ATP synthase is found in all known organisms, photosynthetic and non-photosynthetic.

The structure of ATP synthase from bovine heart mitochondria was solved to 2.8 Å (Abrahams *et al*., 1994). ATP synthase is a large multi-subunit protein complex, with two components: the membrane-extrinsic F_0 domain consists of the $\alpha_3\beta_3\gamma\delta\varepsilon$ subunits and it catalyses ATP synthesis using a rotary mechanism; the F_1 part of the complex forms a transmembrane ring consisting of the subunits $ab_2c_{10\text{-}15}$ that catalyses proton translocation across the membrane. Proton translocation drives the rotation of the c-ring and the attached γ subunit rotates within the $\alpha_3\beta_3$ stator. There are three active sites for substrate binding per complex and the rotation of the asymmetrical γ subunit changes their conformation allowing sequential substrate binding, catalysis and product release. Full rotation of the rotor requires one proton translocated per c subunit (10-15 protons per c-ring), which drives the formation of three ATP molecules (Feniouk and Junge, 2009).

1.6 The genetics of *Rhodobacter sphaeroides*

1.6.1 The photosynthesis gene cluster of *Rhodobacter sphaeroides*

As is the case for the majority of purple bacteria, most of the photosynthesis-related genes of *Rba. sphaeroides* are located in one region of DNA, known as the photosynthesis gene cluster

(PGC), shown in Figure 1.11. It contains at least 38 open reading frames (ORFs), as well as five genes downstream of *puhA* which may have functions related to photosynthesis (Naylor *et al*., 1999). 50% of the ORFs in this cluster are dedicated to the later stages of BChl biosynthesis, 18% for carotenoid biosynthesis, 8% for the LH and RC complexes, and 10% for other genes essential for photosynthetic growth (Naylor *et al*., 1999).

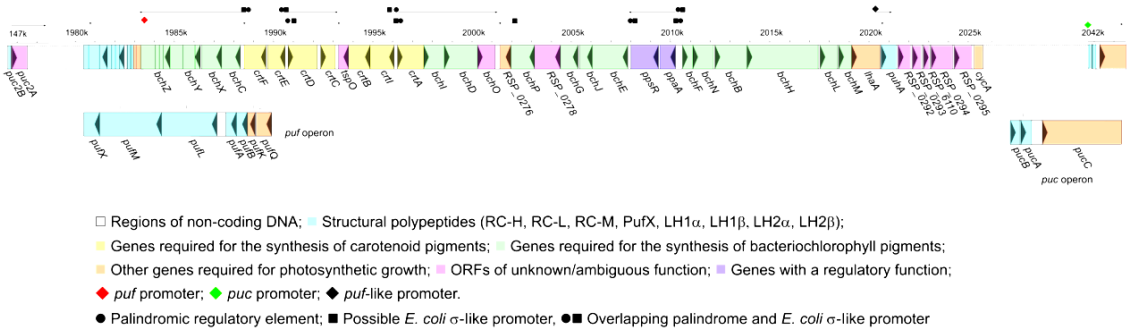


Figure 1.11 Overview of the photosynthesis gene cluster of *Rhodobacter sphaeroides*

Genes with no assigned name have been designated with their NCBI genome sequence number starting RSP_. (Chromosome 1, reference NC_007493.1). Gene lengths are approximately to scale.

(From Mothersole, 2013)

The early stages of the BChl biosynthesis pathway are shared with the haem biosynthesis pathway and are located elsewhere in the genome. The genes associated with electron transport components, such as the cytochrome bc_1 complex and ATP synthase are also not located within the cluster. Also not included in the PGC are the *puc* operons for encoding LH2 and the genes for the enzymes of the Calvin cycle (Blankenship, 2002).

The *puf* operon, which encodes the structural LH1 α and β polypeptides (*pufA* and *pufB*), the RC L and M subunits (*pufL* and *pufM*), the PufX polypeptide (*pufX),* the BChl regulation gene *pufQ* and the transcription regulator *pufK*, (Kiley *et al*., 1987; Lee *et al*., 1989; Hunter *et al*., 1991) are located in the PGC (Naylor *et al*., 1999). The reaction centre H subunit is encoded by the *puhA* ORF, close to the 5' end of the PGC (Donohue *et al*., 1986).

The genes for the LH2 β and α polypeptides (*pucB* and *pucA*) (Ashby *et al*., 1987; Kiley and Kaplan, 1987) are located about 20 kilobases (kb) downstream of the main PGC in the *pucBAC* operon. *pucC* encodes the assembly factor for LH2; interruption of *pucC* abolishes the production of LH2 complexes (Tichy *et al*., 1989). A second *pucBA* operon is found 1.36 Mb from the *pucBAC* operon, but a second *pucC* gene is not found (Zeng *et al*., 2003). 30% of LH2 complexes contain polypeptides encoded by the *puc2B* gene, but deletion of *puc1B* abolishes

the biosynthesis of LH2 entirely. The protein encoded by *puc2A* is very different to that encoded by *puc1A* and bears no resemblance beyond only 58% sequence identity in the N-terminal 48 amino acid residues. The *puc2A* gene is expressed but is not located within LH2 complexes (Zeng *et al*., 2003).

1.6.2 Control of gene expression

A 10-20 -fold molecular excess of the LH1 polypeptide transcripts over those for the RC is achieved by the synthesis and decay of the mRNA transcripts are. The mRNA transcript for the RC genes is degraded faster than that of the LH1 polypeptides (Belasco *et al*., 1985; Zhu and Kaplan, 1985). Light and oxygen intensity controls the regulation of these genes independently.

The light and oxygen conditions of the environment also control the genes for LH2 (including *puc2BA*), the genes in the BChl biosynthesis pathway, and the carotenoid biosynthesis pathway. An increase in light intensity causes a decrease in the expression of the genes for LH2, LH1 and RC. The BChl biosynthesis genes are regulated in a similar way. The carotenoid biosynthesis genes are regulated in an opposite fashion: increased light intensity results in increased expression of the carotenoid biosynthesis genes, possibly due to the photoprotective function of carotenoids under high light intensity in the presence of O_2. (Zhu Yu and Hearst, 1986; Zhu and Hearst, 1986; Kiley and Kaplan, 1987; Glaeser and Klug, 2005)

1.7 The ultrastructure of the *Rhodobacter sphaeroides* photosynthetic membrane

1.7.1 The intracytoplasmic membrane

The photosynthetic apparatus of *Rba. sphaeroides* is found in invaginations of the cytoplasmic membrane, known as the intracytoplasmic membrane (ICM). ICM is comprised of vesicle-like structures 40-60 nm in diameter, packed tightly within the cell (Vatter and Wolfe, 1958; Cohen-Bazire and Kunisawa, 1963; Peters and Cellarius, 1972). These vesicles can be isolated as photosynthetically competent 'chromatophores' following cell disruption in a French pressure cell. In some species of purple bacteria, including *Phs. molischianum* and *Blc. viridis,* the ICM are composed of stacked lamellar sheets (Drews, 1960; Miller, 1979). Certain *Rba. sphaeroides* mutants exhibit altered ICM morphology, for example LH2-deficient mutants produce tubular membranes (Siebert *et al*., 2004).

1.7.2 Maturation of intracytoplasmic membranes and precursor membranes.

Separation of the membrane faction of *Rba. sphaeroides* on continuous rate-zonal sucrose density gradients by ultracentrifugation results in the formation of two discrete bands. The slower-migrating fraction is termed the "upper pigmented band" (UPB) (Niederman *et al*., 1979); the faster moving fraction is the ICM. This difference in movement is likely to do with their membrane morphologies, particularly the spherical nature of the ICM. As determined using multiple pulse-chase radio-labelling studies with [S^{35}] methionine, the UPB develops from immature membranes and the ICM from mature membranes. Ongoing protein synthesis is necessary for the development of the ICM from the immature UPB (Niederman *et al*., 1979; Reilly and Niederman, 1986; Inamine *et al*., 1984).

Functional studies have revealed that the photosynthetic units are assembled in the following order: LH1 and RCs, activation of functional electron transfer, and finally the accumulation of LH2 (Hunter *et al*., 1979a; Hunter *et al*., 1979b; Bowyer *et al*., 1985; Koblizek *et al*., 2005) . New photosynthetic complexes are synthesised and assembled in the UPB; bacteriochlorophyll synthase is enriched here. New LH2-LH1 interactions are created as LH2 complexes are packed in between rows of core complex dimers in the UPB. A bulk LH2-only pool is formed in the ICM. (Hunter *et al*., 2005).

1.7.3 The arrangement of the photosynthetic membrane

The way in which light energy is harvested through energy transfer is related to the arrangement and stoichiometry of the light harvesting antenna and reaction centre complexes in the photosynthetic membrane. Efficient light harvesting depends on the funnelling of energy between physically close complexes (Sener *et al*., 2007). Insights into how energy is transferred so fast and efficiently throughout the membrane are gained through the wealth of knowledge of the structure of the various photosynthetic complexes and of the geography of the chromatophore. Photosynthetic membrane topology has been investigated primarily by atomic force microscopy (AFM) (Bahatyrova *et al*., 2004a). There is diversity in the arrangement of the photosynthetic membrane across different species of purple photosynthetic bacteria. Despite differences between species, the organisation of the membrane complexes does not appear to be random. There is evidence for both core clustering and the formation of peripheral light-harvesting domains. (Sturgis and Niederman, 2009).

1.8 Artificial photosynthesis

Research over several decades has attempted to recreate photosynthetic-like light harvesting processes by purely chemical means. To date, these efforts have fallen far short of the efficiency and versatility of natural systems (Harvey, 2003; Aratani and Osuka 2010; Harvey *et al*., 2011; Lindsey *et al*., 2011; Jiang *et al*., 2014). There are several main challenges to overcome. First, it is a challenge how best to make use of the diverse spectral range of incoming solar radiation. Second, singlet excited state derived from photon absorption is short lived, so reactions and energy transfer processes must be fast and efficient, also disfavouring recombination reactions. Third, it is essential to deploy large numbers of chromophores over mesoscale dimensions that funnel absorbed energy for charge separation, while retaining stringent architectural control at the atomic and sub-nanoscale level. (Jiang *et al*., 2014).

Ongoing research into the creation of biohybrid antennas seeks to fulfil these requirements through a "science of design", producing mesoscale antennas with tailorable performance specifications including spectral coverage, absorbance intensity, and efficiency of excitation delivery and transduction (Harris *et al*., 2014a; Jiang *et al*., 2014). The biohybrid approach integrates the blueprint of biological photosynthesis with the malleability imparted by synthetic chemistry.

Native-like bacterial photosynthetic antenna peptides have been used as a scaffold to attach synthetic chromophores. These biohybrid light-harvesting architectures can augment overall spectral coverage as molecules such as synthetic bacteriochlorins enhance capture of solar radiation in the near-infra red spectral region which is photon rich. (Jiang *et al*., 2014). These antennas are comprised of peptides derived from photosynthetic bacteria, or synthetic peptide analogs and bacteriochlorophyll *a*, in addition to synthetic chromophores. These biohybrid complexes are oligomers of α and β polypeptides that self assemble with 2 molecules of BChl *a* to form a dyad. Dyads then self associate to form cyclic oligomers resembling the native light-harvesting antennas, LH1 and LH2. The synthetic chromophores are covalently attached to one or both of the peptides. The use of specifically chosen synthetic chromophores enables the broadening of spectral coverage of the resulting antenna beyond that provided by the native antenna. (Jiang *et al*., 2014).

Artificial peptide maquettes (discussed in Section 1.11) have been produced with light-harvesting and reaction centre-like properties, with the potential to replace the natural components of the bacterial photosynthetic pathway (Farid *et al*., 2013; Anderson *et al*., 2014; Watkins *et al*., 2014).

Beyond recreating photosynthesis, natural photosynthetic components show promise in being used for a diverse range of technological applications. Reaction centres can be used in biohybrid devices for photovoltaics and biosensing and as nanoscale solar-powered batteries (Tan *et al*., 2013; Kamran *et al*., 2015). The *Rba. sphaeroides* reaction centre has been used as a biosensor for herbicides by using its photocurrent generation properties (Swainsbury *et al*., 2014a).

1.9 Proteorhodopsin

1.9.1 The rhodopsin family

The light-driven proton pump proteorhodopsin (PR) is a member of the rhodopsin family of proteins. Rhodopsins are found in all phyla of life (Foster *et al*., 1984; Bieszke *et al*., 1999; Beja *et al*., 2000; Finkel *et al*., 2013). The rhodopsin family has a characteristic seven-transmembrane helix structure that forms an internal binding pocket in which a chromophore, retinal, is attached. Retinal is bound covalently through a Schiff base to a lysine residue on the seventh helix. Upon photo excitation the retinal isomerises inducing a conformational change and initiating a cascade of electrochemical reactions known as the photocycle. The primary functions of rhodopsin proteins are light-driven transport of ions in microorganisms and photo-induced signal transduction in higher organisms.

The rhodopsin family can be classified into two types with similar topologies but no significant sequence homology. Type I comprises the microbial rhodopsins found in some halophilic Archaea, γ-proteobacteria, some fungi and green alga (Ruiz-Gonzalez and Marin, 2004); they function as light driven ion transporters, such as bacteriorhodopsin (H^+) and halorhodopsin (Cl^-); phototaxis receptors, sensory rhodopsin I and II; or have unknown functions, such as fungal rhodopsins (Oesterhelt and Stoeckenius, 1971; Grigorieff *et al*., 1996; Hoff *et al*., 1997; Lanyi, 1998). Type II rhodopsins comprise the visual rhodopsins found in higher eukaryotes that act as photosensory pigments found in the retinal layer in the eyes of animals and humans (Spudich *et al*., 2000; Beja *et al*., 2000). While type II rhodopsins belong to the G-protein-coupled receptor (GPCR) family, type I rhodopsins are not coupled to a G protein and share no sequence homology with GPCRs.

1.9.2 Proteorhodopsin distribution and function

PR is a light-driven proton pump, a homologue of the more widely studied bacteriorhodopsin (BR) found in some Archaea (Friedrich *et al*., 2002). The first PR to be discovered was isolated as a result of biodiversity screens from an uncultivated γ-Proteobacterium cluster “SAR 86”

from Monterey Bay, California (Beja *et al*., 2000). Since then PR has been discovered in many different organisms; about 13% of all phototrophic marine bacteria contain the PR gene (Sabehi *et al*., 2005). There are two variants of PR, green PR (GPR) and blue PR (BPR), discussed further in Section 1.9.4. PRs are not confined to a single location or taxon, and have been found in fresh water (Atamna-Ismaeel *et al*., 2008), on high mountains (Bohorquez *et al*., 2012) and in sea ice (Koh *et al*., 2010), to name but a few habitats. PR is largely absent from eukaryotes, but it has been found in the dinoflagellate, *Oxyrrhis marina* (Salmovits *et al*., 2011). Host organisms even include marine viruses (Yutin and Koonin, 2012). PR is abundant in huge numbers in the cell, occupying 20% of inner membrane surface area in *Pelagibacter ubiqu*e, a member of the SAR11 clade and one of the most abundant bacteria in the world (Giovannoni *et al*., 2005). Because PR is distributed widely throughout the oceans and at different depths it is likely that this method of photosynthesis may be a hugely important part of solar energy utilisation, energy metabolism and carbon recycling in the sea (Bamann *et al*., 2014).

There is no definitive natural biological function for PR. The large genetic diversity of PRs suggests that it could perform a range of different functions between species (Gomez-Consarnau *et al*., 2010). Since PR pumps protons across a membrane it is assumed that it is involved in the generation of proton motive force (pmf) which can be used for propelling the flagella and for ATP synthesis, which has been shown to be the case for BR (Racker and Stoeckenius, 1974).

It has been suggested that PR may be used as an alternative to respiration at times of starvation, when organic carbon is limited (DeLong and Beja, 2010). Whether cultured in a diurnal light regime or in the dark, there is no observed difference in growth rate or maximum cell yield of *Pelagibacter ubique* (Giovannoni *et al*., 2005). Light-induced survival fitness under conditions of starvation has been correlated to the transcription of PR in *Vibrio sp*. and *Dokdonia sp*. (Gomez-Consarnau *et al*., 2007; Gomez-Consarnau *et al*., 2010; Riedel *et al*., 2013). The extent to which PR can complement or substitute other means of producing energy in SAR11 and other bacterioplankton cells in the ocean remains to be quantified (Zubkov, 2009).

There is evidence to suggest that PR performs roles other than light-induced starvation survival. The PR from *Oxyrrhis marina* has a photosensoric function to detect algal prey (Hartz *et al*., 2011), although proton pumping has been demonstrated *in vitro* (Janke *et al*., 2013). In some polar bacteria the expression of PR is maintained during the dark winter months,

suggesting its use for non ATP-generating functions, such as environmental sensing or small solute transport (Nguyen *et al*, 2015).

To further investigate the function of PR various groups have expressed PR genes in *Escherichia (E.) coli*. When PR is expressed in *E. coli* along with the biosynthesis genes for retinal, there are significant light-induced increases in cellular ATP levels (Martinez *et al*., 2007). It has been shown that the pmf created by PR is sufficient to turn the flagellar motor, enabling the cells to swim (Walter *et al*., 2007). When cellular respiration is inhibited by depleting oxygen or by using the respiratory poison, azide, PR can fully replace respiration as a cellular energy source for driving flagellar rotation and for increasing cell viability in the presence of respiratory poisons and oxygen depletion (Walter *et al*., 2007). A model by the same group indicated that the maximum potential that PR is capable of generating is the same as the potential created by normal cellular respiration. This means that PR can only pump protons when pmf falls below the maximum potential, such as in times of starvation (Walter *et al*., 2007).

1.9.3 The structure and photocycle of proteorhodopsin

The PR gene encodes a 249 residue protein with a molecular weight of 27 kDa (Beja *et al*., 2000). The 3D structure of PR has been extensively studied by solid-state NMR and solution NMR (Pfleger *et al*., 2008; Hempelmann *et al*., 2011; Reckel *et al*., 2011). The 3D crystal structure of a BPR was resolved in 2013 (Ran *et al*., 2013). PR has 7 transmembrane helices (Figure 1.12), typical of the rhodopsin family.

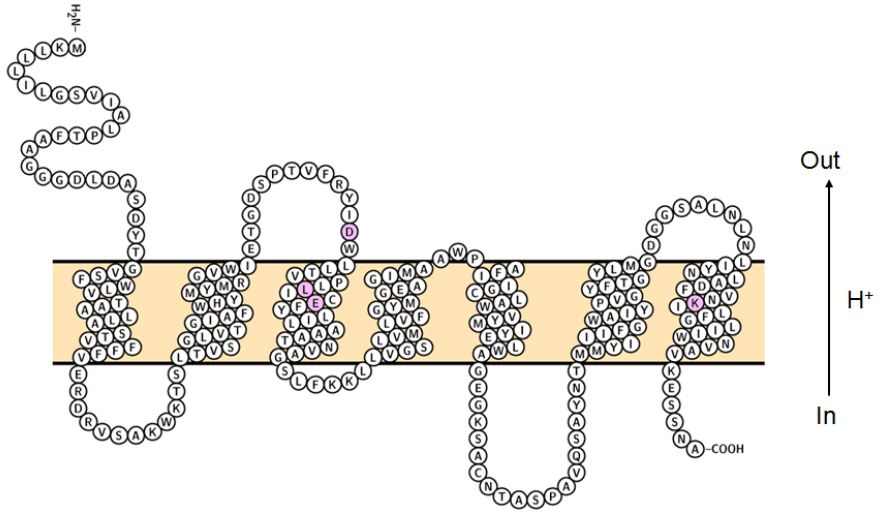


Figure 1.12 Toplogy plot of proteorhodopsin

The sequence is that of green PR (GPR, UniProt: Q9F7P4). The N terminus is towards the extracellular side and the C terminus at the cytoplasmic side. Some key residues are highlighted in pink: the proton acceptor Asp97, the blue/green spectral tuning switch Leu105, the proton donor Glu108 and the lysine Schiff base, Lys231. Figure made with Protter (http://wlab.ethz.ch/protter/start/).

The amino acid residues forming the binding pocket in archael rhodopsins are conserved in PR. PR is a bacterial homologue of BR, although there are several pronounced differences including spectral tuning, the presence of a histidine residue in transmembrane helix B, and the absence of the BR-typical proton release group including two glutamic acids (Figure 1.13) (Bamann *et al*., 2014).

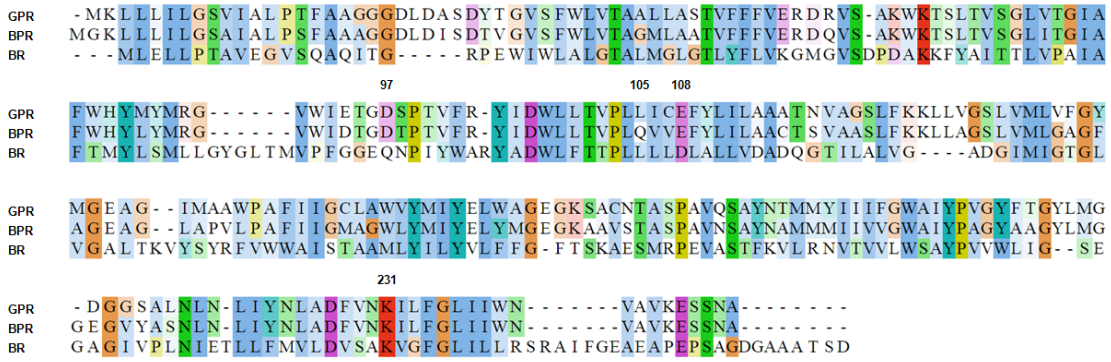


Figure 1.13 Sequence alignment for different proteorhodopsin variants and bacteriorhodopsin

ClustalWS alignment using Jalview. Uniprot sequence identifiers: Green-light absorbing proteorhodopsin (GPR) Q9F7P4, Blue-light absorbing proteorhodopsin (BPR) Q9AFF7 and bacteriorhodopsin (BR) P02945. Key residues from GPR are numbered: the proton acceptor D97, the spectral tuning switch L105, the proton donor E108 and the lysine Schiff base K231.

The chromophore of PR is all-*trans* retinal which is bound to K231 to form a protonated Schiff base. Retinal has a large variation in its absorption spectrum depending on interactions with the apoprotein (Birge, 1990). In methanol, the protonated retinal Schiff base has an absorption maximum of 440 nm, but this shifts to a longer wavelength in a protein microenvironment (Yan *et al*., 1995). Upon light excitation, retinal undergoes ultrafast isomerisation from all-*trans* to 13-*cis* (Figure 1.14), triggering a series of protein conformational changes and several proton transfer reactions. (Spudich *et al*., 2000).

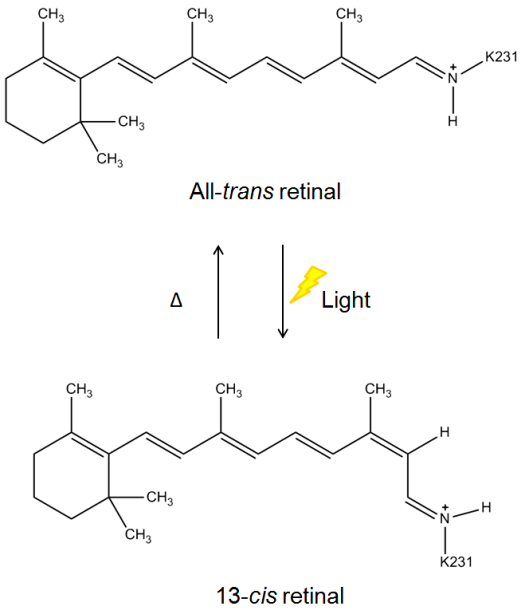


Figure 1.14 Retinal isomerisation in proteorhodopsin

In GPR, retinal is bound to Lysine 231 to form a protonated Schiff base. Retinal undergoes an isomerisation from *trans* to *cis* on light excitation. This induces conformational changes in the protein and leads to proton transport across the protein.

The pumping rhodopsins have a characteristic fast photocycle with a typical time of <30 ms compared to sensory rhodopsins which are slow cycling pigments with a halftime of >300 ms (Spudich *et al*., 2000; Spudich and Jung, 2005). The PR photocycle resembles that of BR in which protonation reactions at the Schiff base lead to proton transfer across the membrane (Beja *et al*., 2000; Dioumaev *et al*., 2002). The PR photocycle is characterised by different intermediates termed K, the product of photoisomerization, M1 and M2, the deprotonated Schiff base, N, the re-protonated Schiff base and PR'(O), the late intermediate (Figure 1.15) (Friedrich *et al*., 2002; Varo *et al*., 2003). The functionally important residues in PR include the proton acceptor D97, the Schiff base counter ions R94 and D227, and the internal proton donor E108, which reprotonates the Schiff base during M to N transition (highlighted in Figures 1.12 and 1.13) (Dioumaev *et al*., 2002; Dioumaev *et al*., 2003).

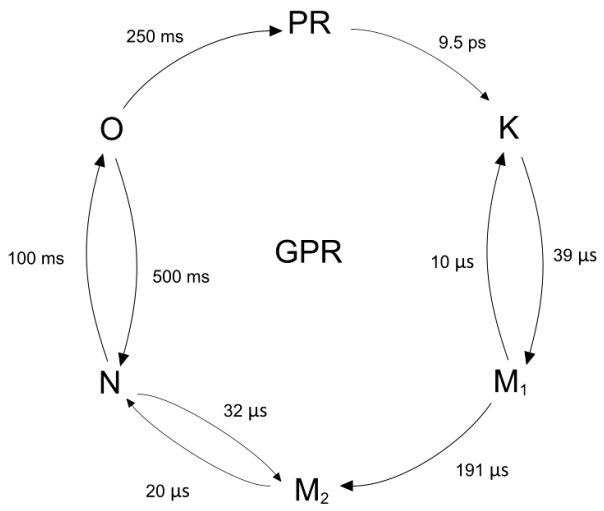


Figure 1.15 The green proteorhodopsin photocycle

Values calculated from laser-flash induced transient absorption changes at pH 9. K to M1 transition is entropy driven.

(Adapted from Mehler *et al.*, 2013)

It has been suggested that the direction of proton transfer through PR varies with pH. D97 has an unusually high pKa of around 7.5; this is stabilised by H75 through the formation of a pH dependent H-bond (Hempelmann *et al*., 2011). H75 is a highly conserved residue not found in BR and keeps the proton acceptor deprotonated during the photocycle (Bergo *et al*., 2009; Hempelmann *et al*., 2011; Balashov *et al*., 2012). The high pKa of D97 leads to a composition of initial states that are different in their protonation at physiological pH. There are functional differences in the proton pumping ability and the photocycles of the alkaline form (D97-) of GPR and acidic form (D97-H) (Bamann *et al*., 2014). At alkaline pH, protons are recruited from the cytoplasmic side of PR and are released towards to extracellular side (Beja *et al*., 2000). At acidic pH, the direction of proton transfer is inverted (Moltke and Heyn, 1995). This inward current is highly dependent on pH and voltage; inward proton pumping has been observed at pH 7.5 under hyperpolarizing potentials (Friedrich *et al*., 2002).

1.9.4 Spectral tuning in proteorhodopsin

There are two main types of PR, green PR (GPR) with absorbance maximum of 525 nm and blue PR (BPR) with absorbance maximum of 490 nm (Man *et al*., 2003). This spectral tuning allows the optimisation of light absorption at different depths in the sea. GPRs are isolated from surface waters, such as Monterey Bay; BPRs are found at greater depths where blue

light can still penetrate but other wavelengths have been filtered out, such as the central North Pacific (Beja *et al*., 2001).

GPRs and BPRs share >78% sequence identity, but their absorption peaks differ by 40 nm. The spectral properties of the chromophore are tuned by the amino acids in its direct vicinity. The green to blue adaption is characterised by a single amino acid substitution in the place of residue 105, from leucine to glutamine (Man *et al*., 2003). This single change causes a highly localised distortion of the chromophore structure and recent research has shown that the end of the retinal molecule is the "hot spot" for colour tuning (Mao *et al*., 2014).

The photocycle of BPR is ten times slower than that of GPR. The slower photocycle reflects the involvement of efficient light detection and faster photocycle indicates efficient pumping action. This has led to the suggestion that BPR functions as a sensory rhodopsin rather than a proton pump. (Wang *et al*., 2003).

The absorbance properties of PR have been altered through the mutation. An unexpected 20 nm red shift of GPR has been observed through a single mutation in the EF loop, A178R, along with an increase in the pKa of D97 (Yoshitsugu *et al*., 2008). This mutation is at a site far away from the retinal and is highly position specific, leading to speculation that there is a long range "interaction channel" between the EF loop and the retinal (Yoshitsugu *et al*., 2009). The native EF loop plays a role in proton uptake from the cytoplasmic side of GPR, and could be used in the transmission of signals across the membrane for additional functions of GPR, such as sensing or signalling. Directed evolution has also been used to tune the absorption maximum of *Gloeobacter violaceus* rhodopsin, causing increased levels of fluorescence in the far red (McIsaac *et al*., 2014).

In addition to directed evolution, retinal analogs have been used to shift the absorbance wavelength of PR (Jensen *et al*., 2013). Although they are unable to be synthesised in nature, this work demonstrates the great tuneability of PR.

1.9.5 Ultrastructure of PR in the membrane

In lipid bilayers, cryo-electron microscopy on 2D crystalline preparations, GPR forms doughnut-shaped complexes with a diameter of about 40 Å (Shastri *et al*., 2007; Klyszejko *et al*., 2008). AFM revealed that these complexes are formed by hexamers and pentamers (Klyszejko *et al*., 2008). It has been speculated that the radial arrangement of PR in these complexes has an advantage for light harvesting in the sea and, compared to monomers, results in a better quantum yield per pumped proton (Bamann *et al*., 2014).

Rhodopsins have been known to build up small antenna complexes by forming a stable entity with an additional chromophore factor. Xanthorhodopsin from *Salinibacter ruber* enlarges its spectral sensitivity by binding an additional carotenoid, salinixanthin (Balashov *et al*., 2005). It has been suggested that the ability to bind salinixanthin relies on a single glycine residue, which also exists in PR, suggesting that PR may also be able to interact with a carotenoid antenna (Imasheva *et al*., 2009). The PR from *Gloeobacter violaceus* can be reconstituted with a light harvesting carotenoid antenna (Imasheva *et al*., 2009).

1.10 Yellow fluorescent protein

1.10.1 The discovery of fluorescent proteins

Fluorescent proteins (FPs) are a fundamental tool in biological research enabling the visualisation of previously invisible processes such as cell development and protein localisation. The most famous FP is the original green fluorescent protein (GFP) discovered in the early 1960s from the *Aequorea victoria* jellyfish. In this organism, GFP transmutes blue chemiluminescence from another photoprotein into green fluorescence (Shimomura *et al*., 1962). The first written report of such bioluminescence was from Pliny the Elder in the first century AD who observed the bright glow of certain jellyfish in the bay of Naples and who cultivated slime from these organisms to make various articles luminescent (Cubitt *et al*., 1999). The Nobel Prize for Chemistry in 2008 was awarded to Martin Chalfie, Osamu Shimomura and Roger Y. Tsien for their discovery and development of GFP.

1.10.2 The function of fluorescent proteins

Fluorescent GFP-like proteins function as secondary emitters within bioluminescent systems. In *Aequorea*, GFP partners with the blue chemiluminescent protein aequorin to control the colour of emission from a broad blue emission to a sharp green peak (Tsien *et al*., 1998; Shagin *et al*., 2004). This has been hypothesised to be an adaptation to suit the visual systems of potential observers, and serves to scare potential predators away (Partridge and Cummings, 1999).

GFP homologues have been cloned from the Anthozoa species, which is not bioluminescent, indicating a function of FPs beyond bioluminescence (Matz *et al*., 1999). These proteins exhibit colour diversity that includes yellow and red fluorescent proteins (Labas *et al*., 2002). The coral Montastraea have cyan coloured GFP-like proteins with a separate evolutionary history, which suggests that the cyan colouration serves a specific function (Kelmanson and Matz, 2003). It is, as yet, unclear what role these FPs play.

In 1944, Kawaguti hypothesised that FPs might be photoprotective (Kawaguti, 1944). This theory has support in the finding of photoprotection in endosymbiotic algae. In these algae, the proposed mechanism of action is the dissipation of excess energy at wavelengths of low photosynthetic activity, and by the reflection of visible and UV light by FP chromophores (Salih *et al*., 2000). This theory remains controversial as it was found that the efficacy of the intrinsic photoprotection mechanisms of algae is far better than that expected of FPs (Gorbunov *et al*., 2001; Mazel *et al*., 2003). If the photoprotection hypothesis is true, the reason for the observed diversity of FP colours in Anthozoa remains a mystery. It is possible that FPs have functions other than photoprotection in non-bioluminescent organisms. Other possible functions for FPs include photosynthesis aid in which wavelength-transformation and back-scattering by FPs provide light enhancement for photosynthesis (Kawaguti, 1969, Schlichter *et al*., 1986; Salih *et al*., 2000), photoreception functions and the generation of colour effects aimed at the outside observer.

Bilaterian animals have a GFP-like domain, G2FP. G2FP is clearly a homologous fold to GFP but there is less than 10% amino acid sequence identity (Hopf *et al*., 2001). Proteins with the G2FP fold are not coloured or fluorescent and have roles as protein binding modules during development (Willem *et al*., 2002; Tunggal *et al*., 2003).

1.10.3. The structure of fluorescent proteins

The structure of FPs consists of a rigid 11-β-sheet barrel surrounding a central α helix (Figure 1.16) (Ormo *et al*., 1996). A few amino acids located near the centre of the β barrel form the principal chromophore. Many of the interior amino acids in FPs are charged or polar, which is unusual for a soluble protein. These residues bind water molecules, which locks them into rigid conformations within the protein. An imidazole ring with extended conjugation is formed as the result of a reaction occurring between key amino acids. The fluorescence of the protein is highly dependent on the environment surrounding the chromophore (Tsien, 1998; Follenius-Wund *et al*., 2003).

Examination of over 1000 naturally occurring variations of FP found only four absolutely conserved residues (Remington, 2006). Mutation of G67 obliterates chromophore formation as it is essential for cyclisation of the chromophore through nucleophilic attack. Y66 is also involved in chromophore formation and it has been hypothesised that it is responsible for stero-electronic steering of the maturation reaction, helping to avoid undesirable side reactions (Remington, 2006). R96 and E222 are essential to the maturation process and are catalytic residues positioned near to the chromophore. As most other residues are not

conserved, FPs can accommodate a high degree of modification to alter the physical properties of the protein (Shaner *et al*., 2007; Day and Davidson, 2009).

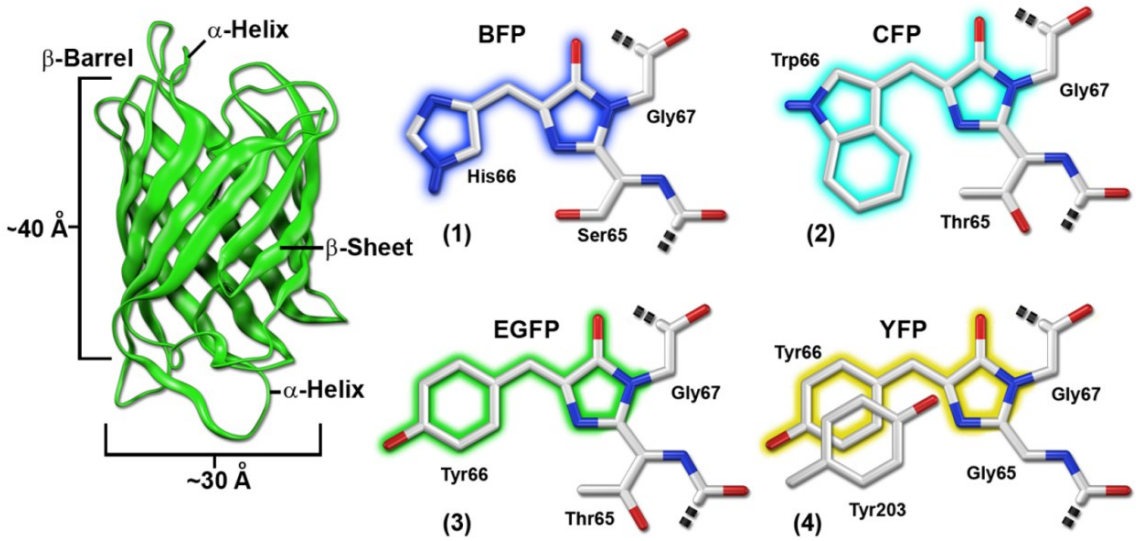
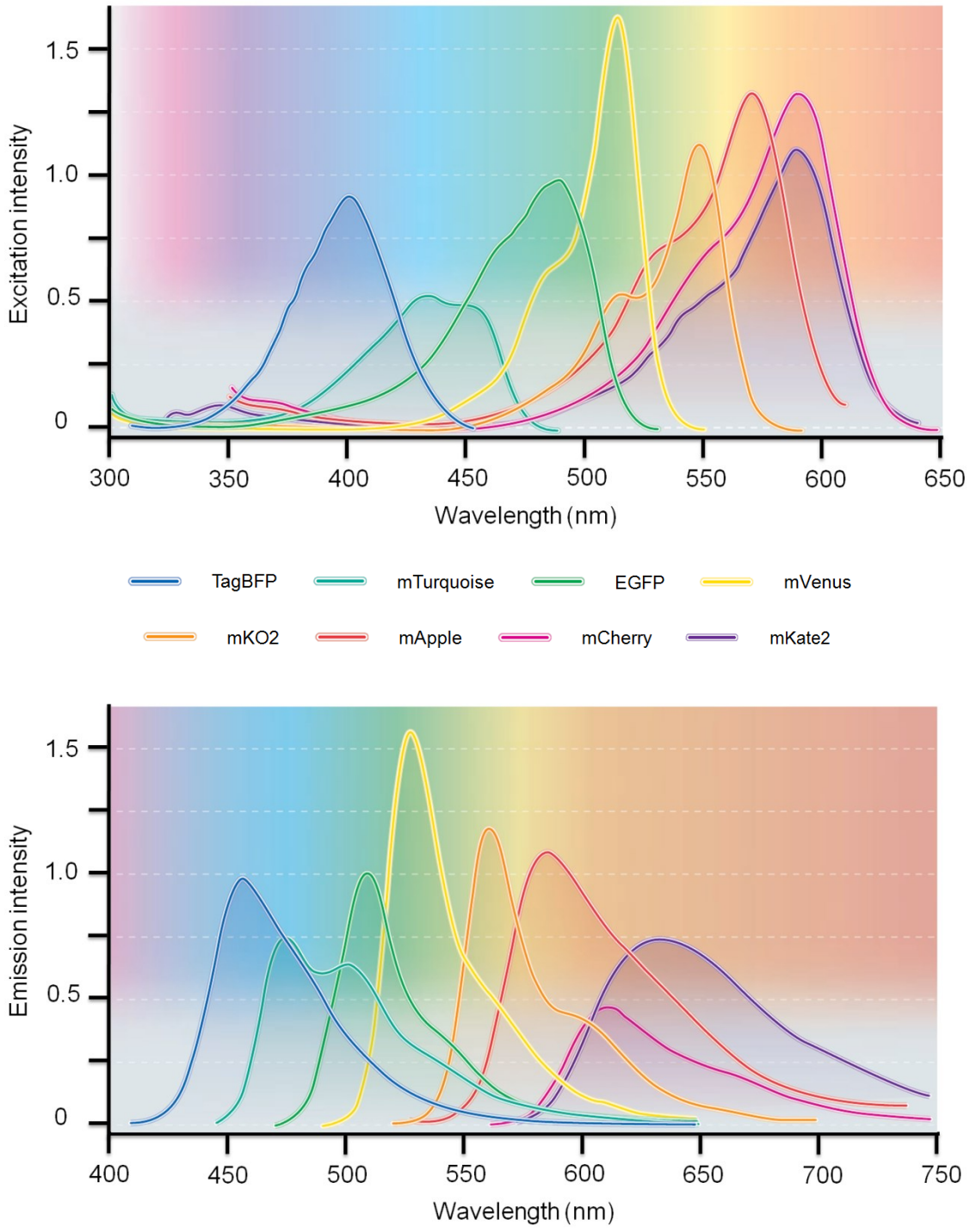


Figure 1.16 Fluorescent protein β-barrel architecture and approximate dimensions and the chromophore structures of common *Aequorea* FP derivatives

(1) BFP, (2) CFP, (3) EGFP, (4) YFP. Portions of the chromophores that are conjugated and give rise to fluorescence are shaded with colours corresponding to the emission spectral profile. The tryptophan residue (Trp66) in (2) is illustrated in the cis conformation as occurs for cerulean derivatives (Malo *et al*., 2007) rather than the *trans* isomer that is common to CFP and related variants.

(From Shaner *et al*., 2007. Reproduced with permission from Journal of Cell Science.)

Blue (BFP), cyan (CFP), yellow (YFP), orange and red FP variants have been discovered or created, these spectral shifts are generally attributed to differences in covalent structure and extend of the π-orbital conjugation of the chromophore (Figures 1.16 and 1.17). FPs can be split into two major classes based on their maturation chemistry. Cyan to green emitting proteins have chromophores identical to GFP, but yellow to red FPs undergo an additional oxidation reaction, forming an acylimine within the polypeptide backbone. This additional oxidation step causes the conjugated portion of the chromophore to extend over the polypeptide backbone. The chromophore is physically larger, which leads to a longer emitted wavelength. (Remington, 2006).



1.17 Excitation and emission spectra of various fluorescent proteins

Fluorescent proteins have been discovered or designed to cover almost the entire spectrum of visible light. This figure shows the excitation and emission spectra of some of the brightest FPs.

(Adapted with permission from Journal of Cell Science. From Kremers *et al.*, 2011)

1.10.4. The uses of fluorescent proteins in research

FPs have been a staple of biological research ever since the *Aequorea victoria* wild-type GFP was used to highlight sensory neurons in the nematode (Chalfie *et al.*, 1994). Purified GFP is highly stable and remains fluorescent at up to 65 °C, pH 11, 1% sodium dodecyl sulphate

(SDS) or 6M guanidinium chloride; it resists most proteases for many hours (Bokman and Ward, 1981). New and improved versions have been created that are brighter, with a broader spectral range, enhanced photostability, reduced oligomerisation, acid resistance and faster maturation rates. While all FPs are large, around 25kDa in size, they have the advantage of being able to be "built in" using transgenic approaches and can be fused to their target protein in a 1:1 ratio (Patterson *et al*., 1997).

Fluorescence resonance energy transfer (FRET) is a useful technique for monitoring protein-protein interactions. It is a general, non-destructive, predictable and quantitative, spectroscopic method that is used to monitor the dynamic association of macromolecular partners in living cells (Tsien *et al*., 1993). The technique allows imaging in single cells with high temporal and spatial resolution. One protein is fused with a donor fluorophore and the other with an acceptor fluorophore. The donor emission spectrum must overlap significantly with that of the acceptor, while the overlaps between the two absorption spectra and the two emission spectra should each be minimized (Cubitt *et al*., 1999). The most common FRET pair is eYFP and eCFP.

In a recent study many alternative oligomeric forms of GFP have been engineered in which the connection of GFP molecules is driven by specific disulphide bond formation or metal ion addition (Leibly *et al*., 2015). This has potential for use in synthetic biology for example by attaching together metabolically coupled enzymes or as a scaffold to bring together pairs of proteins in close proximity.

1.10.5. The development of yellow fluorescent protein

Aequorea victoria GFP has been extensively modified to improve its application as a biological probe. Slow maturation has been an obstacle to the use of GFP for visualisation, especially when expressed at 37°C and/or targeted to certain organelles. Folding of FPs in *E. coli* is inefficient and fluorescence brightness benefits from mutations that facilitate protein folding (Chang *et al*., 2005). A commonly used GFP is eGFP which has a stabilised anionic chromophore and enhanced folding (Tsien, 1998; Shaner *et al*., 2005). It has been advantageous to produce different coloured FPs to allow for the identification of two or more different proteins within a system. Protein-protein interaction studies involving fluorescence resonance energy transfer (FRET) require dual colour imaging. A broad range of FPs spanning nearly the entire visible spectrum has been created. Yellow FPs (YFPs) are among the brightest and most versatile of the engineered FPs (Shaner *et al*., 2007).

Wild type GFP was modified to produce YFP and the enhanced folding and stability variant eYFP, which have a phenolate anion with a stacked π electron system (Tsien, 1998; Cubitt *et al*., 1999). The single mutation, T203Y, is required for this shift in absorbance and was introduced through rational design on the basis of the crystal structure of GFP S65T (Table 1.1) (Wachter *et al*., 1998). Compared to GFP, these early YFP variants were less sensitive to acid and, uniquely, could be quenched by the chloride ion, Cl^-. YFP has often been used as an acceptor for FRET in combination with CFP.

YFP Variant	Mutations
YFP	T203Y
eYFP(Q96K)	S65G V68L Q69K S72A T203Y
Venus	F46L F64L S65G V68L S72A M153T V163A S175G T203Y
SYFP2	F46L F64L S65G S72A M153T V163A S175G T203Y A206K

Table 1.1 Overview of mutations in YFP variants

Annotation based on wtGFP amino acid sequence (GenBank Accession number M62653).

(Adapted from Kremers *et al*., 2006)

An improved version of eYFP called "Venus" was produced that further decreased the sensitivity to pH and Cl^-, improving FRET efficiency by enabling more reliable FRET signals (Table 1.1) (Nagai *et al*., 2002). Maturation of the eYFP chromophore is inefficient at temperatures above 28 °C; this is not ideal for use in for example, *E. coli*, which preferentially grows at 37 °C. The mutations F64L and F46L in Venus dramatically improve this (Cormack *et al*., 1996). Other mutations include F64L, M153T, V163A and S175G which facilitate good folding (Crameri *et al*., 1996). V163A and M153T also increase protein solubility. S65G increases fluorescence intensity (Cormack *et al*., 1996). F46L accelerates the oxidation of the chromophore at 37 °C, which is the rate-limiting step of maturation (Nagai *et al*., 2002).

The variant SYFP2 was developed for use in FRET and is almost 2 fold brighter in bacteria than Venus (Table 1.1). First, the mutation A206K was introduced to create a monomeric form of the protein, termed mVenus. Monomeric FPs are useful when studying protein-protein interactions it is important that fluorescent proteins do not interact themselves (Kremers *et al*., 2006). Mutation V68L was included in Venus but eliminated in SYFP2 as, although it accelerated protein folding, fluorescence development in bacteria was delayed (Kremers *et al*., 2006). Table 1.1 summarises the mutations found in the various YFP variants described here.

1.11 Artificial peptide maquettes

1.11.1 Synthetic proteins

Synthetic proteins are man-made molecules that mimic the function and structure of natural proteins. Synthetic protein scaffolds can incorporate parts from natural proteins or be constructed through *de novo* design. The predictive engineering of novel proteins can further our understanding of how natural proteins are built and work, and potentially provide improved designs (Burton *et al*., 2013). Man-made proteins have an important role to play in the synthetic biology goal of constructing functional parts and devices for incorporating into biological or artificial organisms or systems (Channon *et al*., 2008). To be useful in synthetic biology, functional artificial proteins must interact productively with natural proteins and substrates, fully assemble *in vivo* and bind the intended cofactors in order to perform the desired function (Anderson *et al*., 2014). The challenge is that it is not easy to predict how changes in protein chemistry might affect the protein's shape, stability and function (Currin *et al*., 2015). Catalytic, electron transfer, and substrate binding components have been created so far using natural protein scaffolds and are reviewed in Watkins *et al*., 2014. This work focuses on synthetic peptides known as maquettes.

The complex and obscure interactions found in natural proteins have arisen through millennia of blind natural selection. This results in the accumulation of complexity and fragility, including individual amino acids becoming irreversibly dependent on each other, an extension of what is known as Muller's ratchet (Figure 1.18) (Muller, 1964; Dutton and Moser, 2011). Because of this complexity, it is often impossible to identify the functional roles of any one amino acid, or to determine the identities of what amino acids support any one function. As a result, importing natural sequences or motifs into man-made proteins does not always ensure successful import of the selected function. Pioneered by Les Dutton, completely synthetic self-assembling peptide structures known as "maquettes" can be designed to eliminate this unwanted evolutionary baggage.

A maquette is a protein or peptide that emulates the function of natural cofactor binding proteins whilst minimising undesirable complexity and increase engineering freedom. Each amino acid in the maquette sequence has been selected for a particular reason; this facilitates each iterative round of design and determination of the overall success of the design process (Lichtenstein, 2010). The original maquettes aimed to mimic and understand natural oxidoreductases, but it is becoming clear that they may have uses and functions that go beyond this.

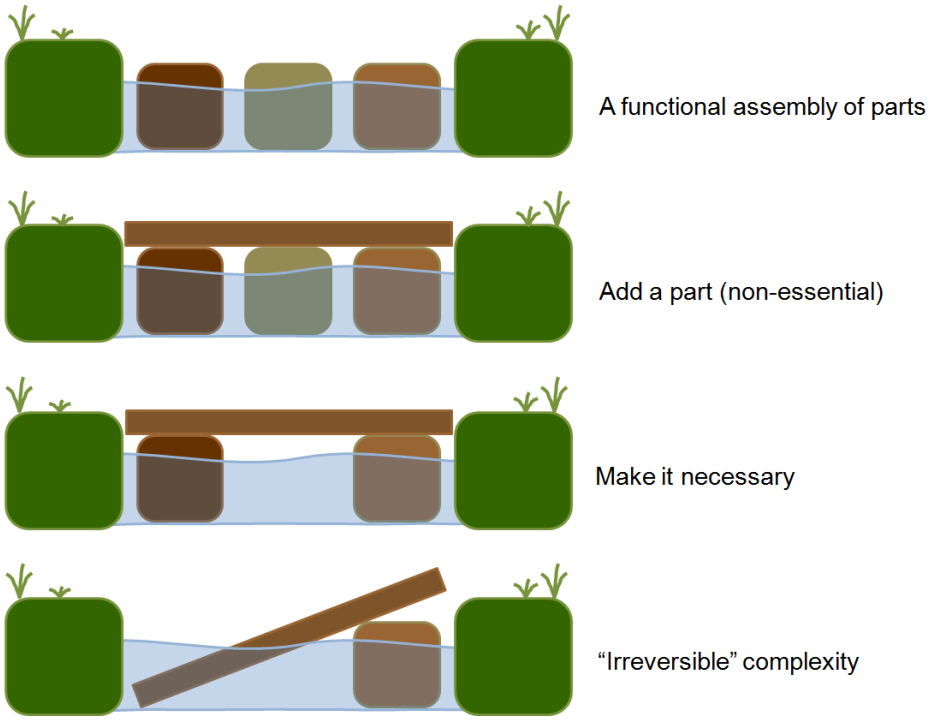


Figure 1.18 Muller’s ratchet

Muller’s ratchet is the process by which genomes acquire deleterious mutations in an irreversible manner. The principles behind this theory can be used to describe the irreversible complexity found in natural proteins. This figure illustrates how Muller’s ratchet can work.

(Adapted from Dutton and Moser, 2011)

1.11.2 Maquette design

The origin of maquettes lies in folding studies of repeated amino acid heptads that form 4 α-helical bundles free of intended function and so simple that the functionalities of each individual amino acid are few and understood (Figure 1.19) (Regan and DeGrado, 1988). A heptad of amino acids with high helical forming propensities, such as alanine, will form two turns of an alpha helix (Figure 1.19 B); a series of heptad repeats will allow alpha helices of differing lengths to be built. Helices are linked through loop regions containing residues with low helix-forming propensities, such as glycine and serine (Figure 1.19 D). A soluble protein has a hydrophilic exterior and hydrophobic interior, with protein folding being driven by packing of the hydrophobic interior; this is achieved in maquettes by carefully patterning the amino acids within a helix (Figure 1.19 B).

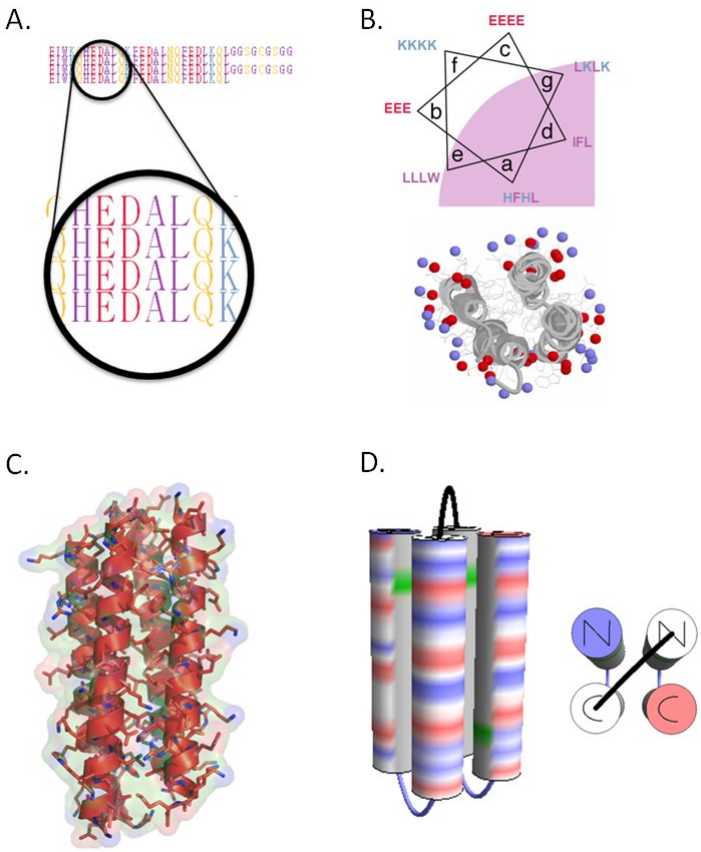


Figure 1.19 Maquette design

- **A.** The primary sequence is simple and minimalist. It is informed by reliable knowledge of natural proteins but there is no mimicry. The sequence is assembled from scratch, are reversibly iterative, flexible and robust. All amino acids are reconciled with roles.
- **B.** Certain amino acids have helical forming propensities. Residues are arranged in such a way that hydrophobic side chains are on the interior and the majority of charged side chains are towards the exterior.
- **C.** The designs allow tertiary structure NMR and X ray crystallography.
- **D.** Schematic representation of a maquette. The helices are linked together with loop regions.

(Reproduced with permission from P.L. Dutton)

The first de novo four-helix bundle proteins were designed by DeGrado in the late 1980s (Regan and DeGrado, 1988), and haem binding protein maquettes were designed by Dutton to emulate the arrangement and biophysical properties of the haem cofactors in natural proteins.

To confer a function on a maquette, design principles have been incorporated from the natural oxidoreductase family of enzymes. Oxidoreductases are electron transfer enzymes which cover a large and diverse range of functions including the light harvesting and transducing capabilities of photosynthetic complexes (Barber, 2009). Rather than mimicking

sequence and structure of any one natural enzyme, bioengineering principles have been taken to create customisable and economical man-made maquettes.

The first maquette was based on the b_H and b_L bis-histidine haem binding sites in the respiratory bc_1 complex (Robertson *et al*., 1994). This was termed H10H24, named for the sites of haem-ligating histidines, and adopted a four-helix bindle structure consisting of a noncovalent homodimer of cystine-crosslinked monomer helices. The sequence was based on binary patterning of hydrophobic and hydrophilic residues around a heptad repeat. The haem-ligating residues are located along the hydrophobic helix interface. Early maquettes based around this design were capable of differentiating the redox midpoints of the bound haems, however they did not have a singular structure. A mutant protein created by iterative design, H10H24-L6I-L13F, increased the steric bulk of the interhelix volume and NMR and crystal structures of the apoprotein could be determined (Skalicky *et al*., 1999; Huang *et al*., 2003). These structures revealed that the non-covalent monomers aligned in *anti*-topology.

The maquette HP1 was designed to increase the likelihood of obtaining a solution structure of a haem-bound maquette (Figure 1.20 A), but lost the midpoint potential differentiation between the two anti-cooperatively binding haem sites, although haem binding affinity was maintained (Huang *et al*., 2004). HP7 has nearly identical properties to HP1 (Figure 1.20 B), but allows full NMR backbone assignments, and has *syn*-orientated helix-loop helix monomers created by a covalent "candelabra" geometry via disulphide bond formation (Koder *et al*., 2006; Koder *et al*., 2009). HP7 has a ferrous affinity for O_2 over CO_2 (Koder *et al*., 2009).

These early maquettes were composed of a single synthesized peptide that forms disulphide-linked loops and self assembles into a dimeric 4-helix bundle (Rothlisberger *et al*., 2008). Recent maquette designs involve a single chain maquette assembly and remove the symmetry induced constraints of earlier designs (Figure 1.20 C) (Anderson *et al*., 2014). They have the advantage over the early homodimer maquettes in that they can incorporate a single site mutation or covalent modification, they can also be expressed in *E. coli* (Anderson *et al*., 2014).

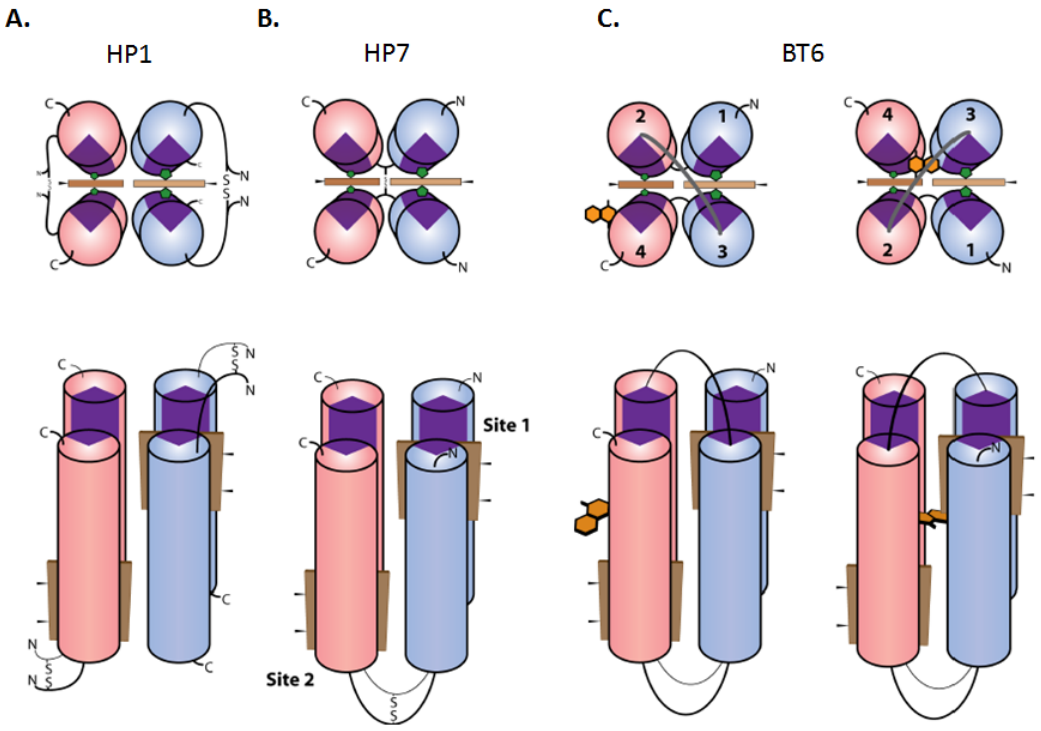


Figure 1.20 Protein topologies of multichain haem-binding maquettes

Schematic representation of the maquettes HP1 (A), HP7 (B), BT6 (C). Helices are coloured according to whether they have their N-termini (blue) or C-termini (pink) pointed towards the viewer (top) or up (bottom).
The significant change in topology between HP1 and HP7 was loop rearrangement; haem binding site and orientation of helices did not significantly vary. Between HP7 and BT6, the major topology change was the construction of a cross-bundle loop between helices 2 and 3, and the elimination of the disulphide bridge between the loops of HP7. BT6 is shown bound with hypothetical napthoquinone residue.

(From Lichtenstein, 2010)

1.11.3 Cofactor incorporation

Function is conferred on a maquette through the incorporation of cofactor molecules, commonly a haem (Figure 1.21). Haem can be ligated through two histidine residues on the interior faces of neighbouring helices. A diagonally ligated haem confers more stability than haem ligated between two helices located directly next to each other. Multiple haems can be bound within a monomeric scaffold, with haem *b* redox potentials spanning hundreds of millivolts (Shifman *et al*., 2000; Farid *et al*., 2013).

Early maquettes bound haem *b* non-covalently. A helix-loop-helix version of H10H24 was converted from a *b*-type binding maquette to a *c*-type by using the conserved *c*-type binding motif, CX_1X_2CH (Ishida *et al*., 2004). Despite having a completely unnatural protein sequence,

a synthetic *c*-type cytochrome can be fully assembled in *E. coli*. A periplasmic export tag and a constitutively expressed *c*-type maturation operon was necessary for the covalent binding of haem B through the haem vinyl groups and protein cysteine residues (Anderson *et al*., 2014). This *c*-type maquette was the first example of *in vivo* incorporation of a redox cofactor in a site specific manner. Introducing a second non-heme C binding site that self assembles with haem B in *vitro* equips the protein with an intraprotein electron transfer chain. To introduce a light activated function, the Fe of the original haem *c* can be replaced with Zn, creating a Zn-porphyrin photocentre.

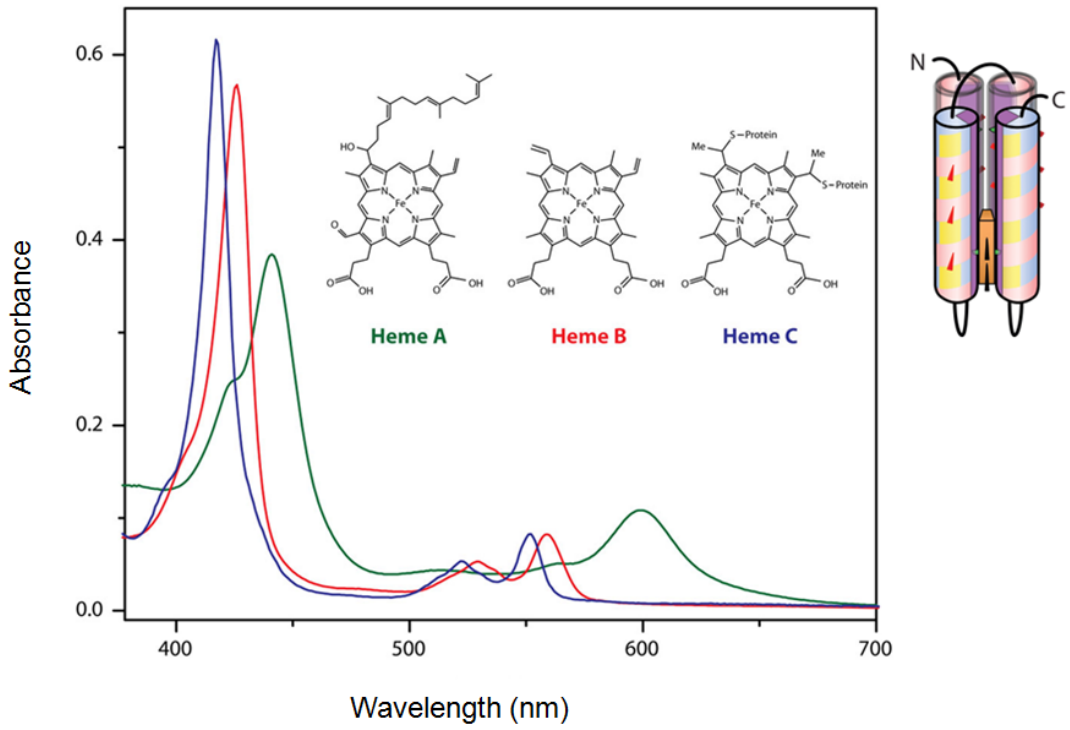


Figure 1.21 Absorbance spectra of maquettes binding different forms of haem

Left: schematic representation of a maquette with haem bound (orange).

(Reproduced with permission from P.L. Dutton.)

Maquettes are remarkably flexible and have been able to accommodate non haem cofactors such as flavin and quinone (Lichtenstein *et al*., 2012; Farid *et al*., 2013). Zinc porphyrins can be non-covalently bound, as well as chlorins and chlorophyllides, and zinc haem C (Noy *et al*., 2005; Farid *et al*., 2013; Anderson *et al*., 2014). Carotenoids can enter the hydrophobic core of an 8-helix bundle, but they are not ligated to any of the side chains. Iron sulphur clusters have been supported by consensus binding motifs in the loop regions (Gibney *et al*., 1996).

Maquettes can also incorporate metal ions, quinones, synthetic chlorophylls and other elements (Lichtenstein *et al*., 2012).

1.11.4 Maquette functions

Maquette chassis have been proven to be remarkably diverse in terms of functional elements, cofactors and assembly (Figure 1.22). Simple functions conferred on the first maquettes using these principles include proton coupling (Shifman *et al*., 1998), ligand exchange (Anderson *et al*., 2014), electrochemical charge coupling (Grosset *et al*., 2001), and the incorporation of light- and redox-active cofactors for oxidation and reduction (Sharp *et al*., 1998). While these early maquettes were homodimers, the latest maquettes are 4 α-helical monomers which exploit sequence asymmetry to confer diverse and versatile oxidoreductase functions which operate with activities comparable to their natural counterparts.

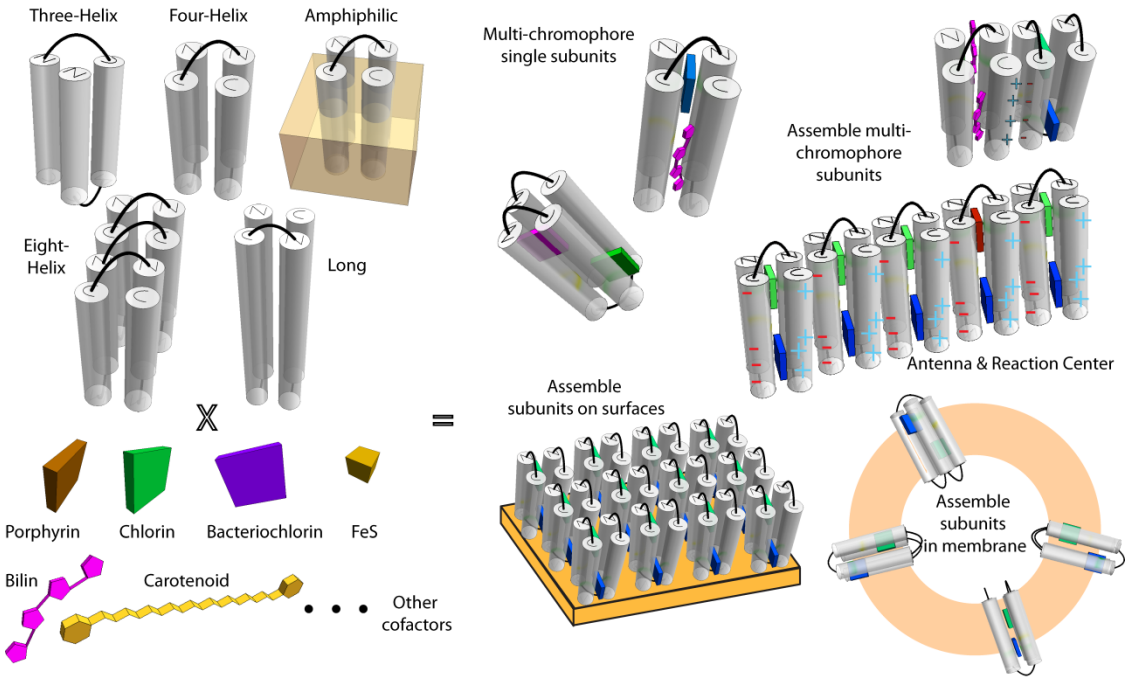


Figure 1.22 The diversity of maquettes

(Reproduced with permission from P.L. Dutton)

Amphiphilic haem binding maquettes have been designed for membrane insertion and to perform transmembrane electron transfer (Noy *et al*., 2005; Discher *et al*., 2005). Amphiphilic maquettes expand the potential of maquettes towards emulating natural membrane

proteins, for example manmade terminal oxidase proteins that could create a proton motive force within a living organism.

1.11.5 Expression of maquettes *in vivo*

Until recently, maquettes have been synthesized and assembled *in vitro*, or expressed as apoproteins in *E. coli* requiring *in vitro* addition of the selected cofactor (Robertson *et al*., 1994; Farid *et al*., 2013). Currently, there is no maquette expressed in *E. coli* that 100% binds haem B, despite nanomolar binding affinities *in vitro*, possibly due to the unfolding of the maquette in the crowded *E. coli* cytoplasm, or insufficient stimulation of the haem B biosynthetic pathway (Farid *et al*., 2013; Watkins *et al*., 2014). However, the *c*-type maquette (described in Section 1.11.3) can fully assemble *in vivo* provided the maquette is transported to the periplasmic space and is co-expressed with the type I *c*-type cytochrome maturation (Ccm) machinery (Anderson *et al*., 2014).

1.12 The twin-arginine translocation pathway

1.12.1 Protein translocation in cells

Biological membranes are tightly sealed and yet must transport molecules from ions to proteins across them without compromising membrane integrity. The best understood mechanism of protein transport across the membrane is via the general secretory (Sec) export pathway. Sec-type export is responsible for the import of newly synthesised secretory proteins into the endoplasmic reticulum, the bacterial plasma membrane, and the inner and outer membrane allowing protein import into the mitochondria and chloroplasts.

While many Sec pathways in mitochondria and chloroplasts are dependent on ATP hydrolysis some proteins, such as some subunits of the thylakoid oxygen evolving complex, require no nucleotide triphosphates for their transport; instead, transport is fully dependent on the pH gradient across the membrane via the twin-arginine translocation (Tat) pathway (Cline *et al*., 1992). Proteins transported via both the Sec and Tat pathways possess a similar N-terminal peptide signal sequence, but Tat proteins are characterised by an essential twin-arginine motif.

In contrast to Sec translocation which exports unfolded proteins, the Tat pathway exists primarily to transport fully- or largely-folded proteins across the plasma and thylakoid membranes of most free-living bacteria. While proteins being exported by the Sec pathway are kept unfolded as they are threaded through a pore, Tat-exported proteins can be transported even when internal crosslinking has been included in the protein to prevent

unfolding (Clark and Theg, 1997). There is evidence to suggest that correct protein folding is a prerequisite for Tat transport (DeLisa *et al*., 2003).

1.12.2 Occurrence of the Tat pathway

The Tat pathway is found in plants, Archaea and most bacteria including *Rba. sphaeroides*. The Tat pathway is important for many processes such as energy metabolism, biofilm formation and bacterial pathogenesis, as well as many more (Berks, 1996; Palmer *et al*., 2005). The exact mechanism of translocation is poorly understood, in part due to the lack of high resolution structural information on the components of the complex. The Tat pathway has been the most widely studied in *E. coli*.

1.12.3 Tat substrates

It is thought that the Tat pathway evolved for the translocation of complex redox proteins (Berks, 1996). Many known substrates of the Tat pathway are periplasmic proteins that bind a range of redox cofactors, such as FeS or NiFe, that can only be inserted in the cytoplasm and require substantial or complete folding of the protein prior to export (Palmer *et al*., 2005). This eliminates the need for cofactor export to the periplasm, and the subsequent catalysis of cofactor insertion in the periplasm (Berks, 1996). Cytochrome *c* will not be exported by the Tat pathway if haem has not been inserted during folding in the cytoplasm (Sanders *et al*., 2001). Thus, the Tat pathway is used when proteins must be folded in the cytoplasm (Halbig *et al*., 1999), require export in conjunction with associated subunits via a "hitchhiker" mechanism (Rodrigue *et al*., 1999), or when the protein folds too fast for the Sec system (Palmer and Berks, 2012).

1.12.4 Properties of the Tat signal peptide.

Proteins translocated by both the Sec and the Tat pathways have N-terminal signal sequence consisting of three major regions: a charged amino-terminal region, a hydrophobic region and, on the carboxy-terminal end of the sequence, a signal peptide cleavage site. In most cases the signal sequence is cleaved from the precursor protein during or immediately after translocation, liberating the mature protein. In the case of the *E. coli* TorA signal peptide (Figure 1.23), four amino acid residues from the sequence are conserved in the mature protein after translocation.

The key feature of the Tat signal sequence is a twin arginine motif, with the majority of instances in *E. coli* occurring: S/T-R-R-*x*-F-L-K, where *x* is any polar amino acid. The twin-arginine motif is essential for Tat translocation in the chloroplast (Chaddock *et al*., 1995). In bacteria, Tat translocation does not always require the presence of the consensus RR

dipeptide in the conserved sequence motif, a basic side chain such as R or K is required at the first position, and the second position can accept Q or N as well as basic amino acids (DeLisa *et al*., 2002). Mutation of RR to KK prevents Tat export of many proteins, and is therefore often used as a control when exporting foreign proteins (Buchanan *et al*., 2001). Other mutations that hinder Tat export include the substitution of the R at the second position with K, resulting in the protein to be translocated getting stuck in the membrane and the Tat signal peptide unable to be cleaved (Ren *et al*., 2013).

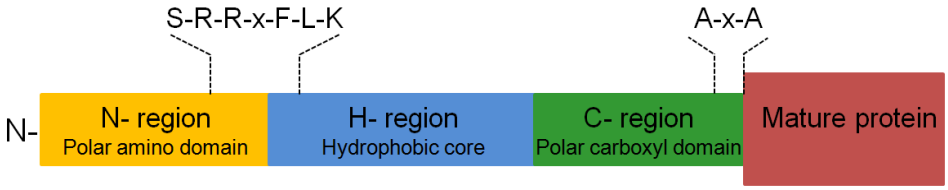


Figure 1.23 The Tat signal peptide

The Tat signal peptide is located at the N-terminus of the substrate protein and is composed of a polar amino domain (N- region), a hydrophobic core (H- region) and a polar carboxyl domain (C- region). TAT signal peptides are, on average, less hydrophobic than Sec signal peptides and are longer (38 compared to 24 amino acids). Tat signal peptides have a characteristic twin-arginine motif in the N-region and an A-x-A motif in the C- terminal region for removal of the signal peptide by signal peptidase.

(Adapted from Patel, 2014)

The twin-arginine motif alone is not sufficient to prevent mistargetting to the Sec pathway. Tat signal peptides are less hydrophobic than Sec signal peptides (Sanders *et al*., 2001). The C-region of a Sec signal peptide does not usually contain basic residues, but Tat signal peptides often contain these which hinder engagement with Sec machinery (Bogsch *et al*., 1997).

1.12.5 The Tat subunits

In *E. coli*, the minimum set of components for Tat translocase assembly is TatA, TatB and TatC. These proteins are all integral membrane proteins and reside in the cytoplasmic membrane as a Tat(A)BC substrate binding complex with a separate TatA complex (Patel *et al*., 2014).

TatA, a 9.6 kDa protein, consists of a short periplasmic N-teminal region and a transmembrane helix linked to an amphipathic helix which lies along the surface of the membrane on the cytoplasmic side. The C-terminal region is highly unstructured (White *et al*., 2010; Lange *et al*., 2007). TatB, 18.5 kDa, shares 20% sequence similarity with TatA and has a

very similar predicted secondary structure, however TatA and TatB carry out different functions within the Tat translocase (Sargent *et al*., 1999).

TatC has a molecular mass of 28.9 kDa and has six transmembrane helices (Punginelli *et al*., 2007). TatC has a major role in protein translocation playing a role in substrate recognition and binding and recruitment of other Tat components (Schreiber *et al*., 2006).

1.12.6 The mechanism of Tat protein export

The exact mechanism of how the Tat system transports folded globular proteins without compromising membrane integrity remains a mystery. The sequences of the Tat transport proteins bear no resemblance to other protein translocating subunits (Robinson and Bolhuis, 2001). The general consensus is that the substrate binds initially to TatBC, independent of other Tat components (Mori and Cline, 2002). The TatA complex then associates with the TatBC complex in the presence of a pH gradient across the membrane (Cline and Mori, 2001).

Other than cross linking studies, the active translocon has never been captured so our understanding of the translocation process is still vague. There are currently two models for the mechanism of translocation: the translocation pore, or trap door, model and the membrane destabilisation model (Figure 1.24) (Hauer *et al*., 2013).

The trapdoor mechanism is based on the theory that TatA can flip its amphipathic helix from its resting position along the cytoplasmic face of the membrane to within the membrane (Figure 1.24 A). TatA can form complexes of variable size in *E. coli* and is highly abundant relative to TatB and TatC, suggesting a role as a pore (De Leeuw *et al*., 2001; Oates *et al*., 2005). A 3D model constructed by single particle electron microscopy showed that TatA forms a cupped pore-like structure with a range of complex sizes (Oates *et al*., 2005; Gohlke *et al*., 2005; Beck *et al*., 2013). TatA uses complementary charge interactions to form a hairpin fold between the amphipathic helix and the C-terminus (Walther *et al*., 2013). This would provide an internal hydrophilic coating to the pore.

Further evidence to support the trapdoor model is the finding that the Tat substrate itself carries a motif that is capable of interacting with TatA (Maurer *et al*., 2010). The binding conformation is what would be expected if TatA surrounds the substrate as if it were a pore. It has been suggested that the interaction of TatA with the substrate leads to a seeding of monomeric TatA to form a pore (Froebel *et al*., 2012).

If the pore theory is correct, the pore must be able to become very big. The Tat-substrate, TMAO reductase, is (at its smallest) 5 nm in diameter when folded and is too large to be fully enclosed within the membrane, protruding above and below the membrane (Robinson and

Bolhuis, 2001). The membrane shows no loss of permeability when operating at maximal Tat-export capacity; how this is achieved is a mystery (Teter and Theg, 1998). To maintain membrane integrity, the pore must be able to adjust to the size of the substrate, perhaps varying numbers of subunits are required to create the channel, effectively coating the protein.

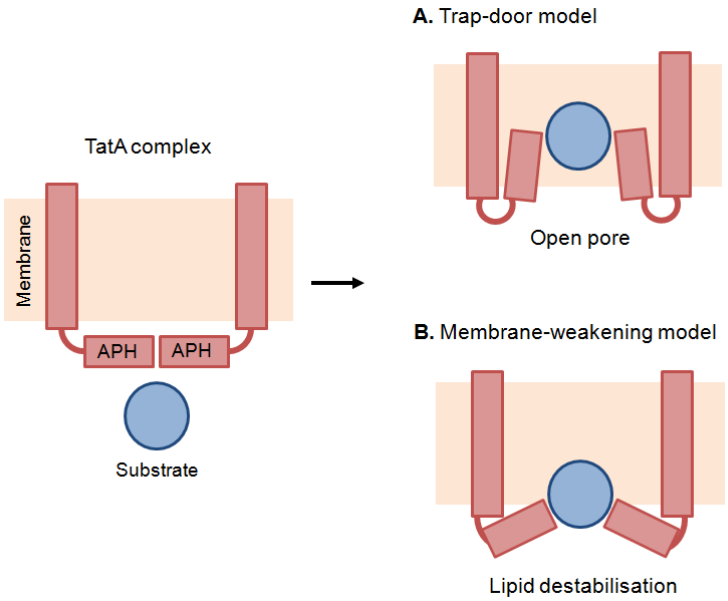


Figure 1.24 Models for the mechanism of Tat translocation

A substrate featuring a twin-arginine signal peptide is recognised by the TatBC complex in the membrane. This figure shows two proposed models for the subsequent translocation event involving the TatA complex.

- **A.** The trap-door model. Prior to substrate recognition, the amphipathic helix (APH) lies along the cytoplasmic face of the membrane. Upon substrate recognition the APH flips into the bilayer. This provides a regulated pore for the translocation of the substrate in the presence of a membrane potential.
- **B.** The membrane-weakening model. The APH of TatA aligns parallel to the cytoplasmic membrane. Recognition of a substrate induces a topological change in the APH so that it partially perturbs the lipid bilayer. This leads to the destabilisation of the lipid bilayer. Subsequently, this permits the translocation of the substrate in a less regulated manner.

(Adapted from Patel *et al.*, 2014)

Bruser and Sanders have hypothesized that TatA does not form a pore but aggregates in an unordered manner, resulting in Tat complexes large enough to destabilise the membrane (Figure 1.24 B) (Bruser and Sanders, 2003). Tha4 is a TatA orthologue found in the membrane of the chloroplast; it has a relatively low abundance compared to the excessive level found in *E. coli*. This leads to the conclusion that Tha4 is less capable of binding a substrate and rapidly forming a membrane-inserting structure around it. Recent studies have indicated that the topology of TatA is not as flexible as has been previously thought, and is only partially

incorporated into the membrane. This would imply a “lipid disrupting” property. (Walther *et al*., 2010).

Structural characterization of the TatAd complex from *Bacillus subtilis* suggests that the pore may not be able to accommodate substrate (Beck *et al*., 2013). This suggests that the TatA assembly is not a pore but instead functions to destabilize the membrane or to stabilize the translocon after it has been recruited to the substrate-bound complex (Froebel *et al*., 2011).

The flaw of the membrane destablisation model is that it compromises the specificity of substrate transport. There, therefore, must be a stabilisation system to counteract the destablisation and to prevent the leakage of ions across the membrane. It has been proposed that the phage shock protein, PspA in *E. coli* and it chloroplast homolog, VIPPI mediates this (DeLisa *et al*., 2004; Lo and Theg, 2012). There is evidence of phylogenetic conservation and proof of the ability of TatA to interact with PspA that makes this hypothesis a possibility (Kudva *et al*., 2013).

1.12.7 Translocation of non-native Tat substrates

The abilities and mechanism of the Tat system have attracted much interest as the translocase can transport fully folded globular proteins. Many studies have involved exploring the potential of non-native Tat substrates for translocation via the Tat pathway; some of these are reviewed in Robinson and Bolhuis, 2001.

Wild type GFP cannot fold correctly in the periplasm and so cannot be exported by the Sec system in an active form (Feilmeier *et al*., 2000). Fully active GFP has been observed in the periplasm when a Tat-specific target signal was fused to it (Thomas *et al*., 2001). Human immunoglobulin heavy chain variable domains (VH) are promising scaffolds for antigen binging, however it is an unstable and aggregation-prone protein. The Tat system has been used for protein folding quality control step to aid directed evolution of this protein (Kim *et al*., 2014). Tat signal peptides have also been used to export polymer-based drug delivery vehicles (Nori *et al*., 2003).

1.13 Aims of this work

Existing approaches to create artificial photosynthesis, such as biohybrid architectures, have great potential *in vitro*. However, replicating this *in vivo* is a challenge, in part due to the inability of the bacterium to produce the designer chromophores. To create tailor-made light harvesting antennas *in vivo*, we must use the existing toolbox of proteins and pigments available in nature, or create synthetic elements that the host organism is capable of

creating. The aim of the work in this thesis is to use a variety of synthetic biology approaches to begin to expand the range of possibilities for bacterial photosynthesis. These include:

1. Expanding the range of light absorbers for photosynthesis, by attaching an extra chromophore (YFP) to a photosystem component, in this case the reaction centre.
2. The assembly of proteorhodopsin in *Rba. sphaeroides* intracytoplasmic membranes with the eventual aim of augmenting the proton motive force that drives downstream metabolism.
3. The expression of genes encoding maquettes in *Rba. sphaeroides* to form the basis for a bottom-up redesign of photosystem components.
4. An investigation of the plasticity of bacterial protein export pathways for synthetic biology purposes, initially challenging the Tat export pathway of *E. coli* to export maquettes for the first time.

Chapter 2

Materials and Methods

2.1 Standard buffers, reagents and media

Unless otherwise stated, all buffers and culture media were prepared as in (Sambrook *et al*., 1989). All media and solutions were prepared using distilled water purified using the Milli-Q system from Millipore. Growth media were made following the manufacturer's instructions, using distilled water and sterilised by autoclaving for at least 20 minutes at 15 psi. All other solutions were sterilised by filtration through 0.2 µm filters. Heat labile solutions, such as vitamins and antibiotics, were added to culture medium only after they had cooled to 50 °C or below.

2.2 Nucleic acid manipulation

2.2.1 Small-scale preparation of plasmid DNA (mini-prep)

Small quantities of plasmid DNA were prepared using a Sigma mini-plasmid purification kit, according to the manufacturer's instructions. Transformed *E. coli* cultures were grown in sterile 25 ml universal tubes containing 5 ml of LB medium with the appropriate antibiotic. Cultures were grown overnight at 37 °C and shaken at 300 rpm. DNA was eluted from the columns using 50 µl warm QH_2O and stored at -20 °C.

2.2.2 Polymerase chain reaction (PCR)

DNA amplification by polymerase chain reaction (PCR) was performed using either ACCUZYME™ or MyTaq™ (Bioline). Primers were produced by Sigma-Aldrich and resuspended to 100 µl with QH_2O. All primers used in this work are shown in Table 2.3.

ACCUZYME™ reactions were performed in a total volume of 50 µl containing a final concentration of 200 nM of each primer, 25 µl of 2 x reaction mix (2.5 units ACCUZYME™, 2 mM dNTPs, 4 mM $MgCl_2$), 1 µl DMSO, and 10 ng of template DNA. ACCUZYME™ was used for amplification of DNA destined for vectors or sequencing due to its low error rate of 1.6×10^{-6} errors/base (Lundberg *et al*., 1991).

MyTaq™ reactions were performed in a total volume of 20 µl containing a final concentration of 200 nM of each primer, 4 µl 5x MyTaq reaction buffer, 1 µl MyTaq DNA polymerase, and

10 ng of template DNA. MyTaq was developed for colony screening reactions and has tolerance for a wide range of common PCR inhibitors.

Reactions were carried out using conditions appropriate to the T_m of the primers and the length of the fragment to be amplified. Primers were denatured for 3 min at 95 °C followed by 30 cycles of amplification. For ACCUZYME™ reactions, the denaturation step was 96 °C for 30 seconds, the annealing step was 58 - 62 °C for 30 seconds, and the extension time was 0.5 kb per minute at 72°C. There was a final extension for 10 minute at 72 °C. For MyTaq reactions, the denaturation step was 95 °C for 15 seconds, the annealing step was 57 - 58 °C for 15 seconds, and the extension time was 1 kb per 15 seconds at 72°C. There was a final extension for 1 minute at 72 °C.

Reactions were subsequently cleaned up via gel purification or using a Sigma-Aldrich GenElute™ PCR Clean-Up Kit.

2.2.3 Restriction enzyme digestions

Restriction enzymes were purchased from Promega and New England Biolabs. Digests were performed according the suppliers' directions, in a total volume of 20-25 µl. The samples were incubated at 37 °C for 1-2 hours. Following digestion plasmid DNA, or DNA produced from the digestion of a plasmid, was gel purified. Linear DNA produced by PCR was cleaned up using a Sigma-Aldrich GenElute™ PCR Clean-Up Kit.

2.2.4 Agarose gel electrophoresis of DNA

Restriction digests and PCR products were analysed by electrophoresis using a 1 % agarose gel containing 0.5 µg ml^{-1} ethidium bromide in TAE running buffer (40 mM Tris-acetate, 1 mM EDTA) running buffer (Sambrook *et al.*, 1989). Samples were mixed with 5X DNA loading buffer (Bioline) prior to loading onto the gel. Typically, 10-550 ng of DNA was loaded per lane. 5 µl of DNA marker Hyperladder 1 kb (Bioline) was run alongside the samples in order to estimate the sizes of DNA fragments. The samples were run at 80 V for 1 hour. DNA was visualised under 254 nm UV light.

2.2.5 Recovery of DNA from agarose gels

DNA was excised from agarose gels using the GenEluteTM Gel Extraction Kit (Sigma-Aldrich) as per the manufacturer's instructions. DNA was eluted from the columns using 50 µl warm QH_2O.

2.2.6 Ligation of DNA fragments

Typically, 10-50 ng of vector DNA and varying molar equivalents of the desired insert fragment were ligated in a final volume of 10-20 µl. The reaction included 10x T4 ligation buffer and 1 unit T4 DNA ligase (New England Biolabs). Ligations were left at room temperature overnight before transformation into chemically competent *E. coli* JM109 cells.

2.2.7 QuikChange mutagenesis

Point mutations, for the replacement of amino acids or the deletion or insertion of single or multiple adjacent amino acids, were introduced to plasmid DNA using the QuikChange II Site Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions. Primers containing the required mutation complementary to opposite strands of the vector were designed using the QuikChange Primer Design tool (http://www.genomics.agilent.com/primerDesignProgram.jsp).

2.3 DNA sequencing

Plasmid DNA was prepared by mini-prep (Section 2.2.1) or by PCR amplification (Section 2.2.2), purified by gel extraction and were diluted to a concentration of 100 ng ul^{-1} in a 30 µl volume. Samples were sent to GATC Biotech for sequencing. Results were analysed by CodonCode Aligner v3.7.1.

2.4 Preparation of *Rhodobacter sphaeroides* genomic DNA

1 ml cells were spun down at 2, 100 x *g* for 25 minutes and resuspended in 500 µl TE buffer (10 mM Tris pH 8, 1 mM EDTA). 50 µl SDS was added and the sample was heated at 70 °C for 15 minutes in order to denature protein. 50 µl phenol:chloroform:isoamyl alcohol (25:24:1) was added and the sample was vortexed; the sample should go white. The sample was centrifuged at 7,400 x *g* for 10 minutes. The top layer containing the DNA was taken without disturbing the interface. 1 volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to 1 volume sample. The sample was centrifuged at 7,400 x *g* for 10 minutes and the top layer containing the DNA taken. DNA was precipitated using 25 µl 3 M sodium acetate and 500 µl cold 100% ethanol. The sample was vortexed and the DNA appeared as a stringy white substance. In order to maximise yield of DNA, the sample was incubated at -20 °C for 1 hour. To pellet the DNA, the sample was centrifuged at 12,500 x *g* for 10 minutes. The pellet was washed with 200 µl cold 70% ethanol and centrifuged at 12,500 x *g* for 10 minutes. The DNA pellet was dried by a flame until it went clear, and was then resuspended in 50 µl QH_2O. 10 µl DNA solution (diluted up to 1 ml) was put in a UV cuvette and absorbance readings taken at

260 and 280 nm. For the concentration of DNA in µg, multiply A260 by 5. The ratio of A260:A280 should be greater than 1.4. The DNA was diluted to 100 ng per µl and aliquotted into 50 µl samples.

2.5 *Escherichia coli* strains and plasmids

Three *E. coli* strains were primarily used for this work: S17, JM109 and BL21(DE3). Chemically competent S17-1 cells (Simon *et al*., 1983) were used for plasmid transfer to *Rba. sphaeroides*. JM109 chemical competent cells were obtained from Invitrogen and stored as 25 µl aliquots at -80°C.

All *E. coli* plasmids used in this work are listed in Table 2.1. Unless otherwise stated, *E. coli* strains were grown in Luria-Bertani (LB) medium (Sambrook *et al*., 1989) at 37 °C and agitated at 300 rpm. Where required, antibiotics were added at the following concentrations: 25 µg ml^{-1} kanamycin, 200 µg ml^{-1} ampicillin.

2.6 Production of chemically competent *Escherichia coli* cells

1 ml of an overnight starter was used to inoculate a 500 ml flask containing 50 ml LB. The cells were grown to an OD_{600} of 0.4-0.6 then cooled on ice. The cells were centrifuged at 2,100 x *g* for 10 minutes and resuspended in 20 ml cold 0.1 M $MgCl_2$. The cells were centrifuged again and the pellet was resuspended in 20 ml cold 0.1 M $CaCl_2$. After further centrifugation, the cells were resuspended in 1 ml 0.1 M $CaCl_2$ 10% glycerol and aliquotted. The cells were flash frozen and stored at – 80 °C.

2.7 Chemical transformation of competent *Escherichia coli* cells

An aliquot of chemically competent *E. coli* cells was thawed on ice. 10-15 ng of plasmid DNA in QH_2O (or 2 µl ligation mixture) was added to 20 µl of cells. The mixture was incubated on ice for 20 min before a heat shock at 42 °C for 40 seconds. The mixture was returned to ice for 2 minutes. 750 µl LB medium was added and the cells were incubated at 37 °C for an hour. Cells were pelleted by centrifugation, resuspended in 50 µl LB medium and plated onto LB agar containing the appropriate antibiotic. Plates were incubated overnight at 37 °C to allow colonies to grow.

2.8 Plasmid induction in *Escherichia coli*

Overnight culture was used to inoculate the required volume of LB and induced with the appropriate inducer at an OD_{600} of 0.6 – 0.8. Cells were induced for 1-48 hours at 15 – 37 °C depending on the required conditions, and shaken at 300 rpm unless otherwise specified.

2.9 Preparation of *Escherichia coli* whole cell fractions

Whole cell fractions were prepared by harvesting 1 ml of culture and resuspending in 100 µl SDS PAGE sample buffer, Samples were incubated at 50 °C for 10 min and analysed by SDS PAGE.

2.10 Fractionation of *Escherichia coli* cells

Cells were separated into periplasmic, cytoplasmic and membrane fractions using a procedure based on the EDTA/lysozyme/osmotic shock method (Randall and Hardy, 1986). Cells were harvested and resuspended in 1 ml chilled buffer 1 (100 mM Tris-acetate pH 8.2, 0.5 M sucrose, 5 mM EDTA) followed by the addition of 40 µl of 2 mg ml^{-1} lysozyme. 500 µl of chilled QH_2O was added before incubation on ice for 5 min followed by the addition of 20 µl $MgSO_4$. The spheroplasts were pelleted by centrifugation at 16,600 x *g* in a cooled microcentrifuge and the supernatant was collected as the periplasmic fraction. Spheroplasts were washed in 1 ml chilled buffer 2 (50 mM Tris-acetate pH 8.2, 0.25 mM sucrose, 10 mM $MgSo_4$) and pelleted by centrifugation at 4 °C. The supernatant was discarded and the spheroplasts were resuspended in 1 ml chilled buffer 3 (50 mM Tris-acetate pH 8.2, 2.5 mM EDTA). Spheroplasts were lysed by sonication and membranes were separated from the cytoplasmic fraction by centrifugation at 336,140 x *g* (avg) for 30 min at 4 °C. The membranes formed a pellet and the supernatant was collected as the cytoplasmic fraction. Membranes were then resuspended in 500 µl of detergent-containing buffer (20 mM Tris-HCl pH 8, 10% v/v glycerol, 50 mM NaCl and 1 % SDS).

2.11 Membrane preparation from *Escherichia coli*

Pelleted cells were resuspended in 20 mM Tris, 5 mM EDTA, pH 8. The cells were French pressed 3 times at 18,000 psi. The cells were kept on ice throughout. Unbroken cells were removed by centrifugation at 6,794 x *g* (avg) at 4 °C for 20 minutes. The supernatant was loaded onto a discontinuous sucrose gradient of 55 – 30 % (w/w) sucrose with 4.5 ml steps of 5%. The gradients were centrifuged at 30,000 rpm (110,527 x *g* (avg)) for 18 hours in a Beckman SW 32 Ti rotor. The bands were harvested. The upper band is the inner membrane

and the lower band is the outer membrane. Membranes were harvested from the gradients, pooled and pelleted by centrifugation at 125,000 x g for 2 hours. Membranes were resuspended in binding buffer (20 mM Tris, 0.5 mM EDTA, 500 mM NaCl and 5 mM imidazole, pH 7.4) and solubilised in 0.1 % triton X and 0.5 % LDAO for 16 hours.

2.12 *Rhodobacter sphaeroides* strains and growth

2.12.1 *Rhodobacter sphaeroides* strains

All *Rba. sphaeroides* strains used and created for this work are listed in Table 2.2. *Rhodobacter sphaeroides* or *Rba. sphaeroides* refers to wild type *Rhodobacter sphaeroides* strain 2.4.1. Strains were grown in M22 medium (See Appendix) (Hunter and Turner, 1988) and supplemented with 10,000 X vitamins (0.08 M nicotinic acid, 0.01 M thiamine, 7.3 mM 4-aminobenzoic acid, 0.4 mM d-biotin), to a final concentration of 1 X, at 24°C. 0.1% casamino acids was used to supplement liquid cultures (See Appendix). Where relevant, kanamycin was added to a final concentration of 25 μg ml^{-1}. Stocks of strains were maintained in LB medium containing 50% glycerol (v/v) and stored at -70 °C.

2.12.2 Growth on agar plates

Cells were streaked out from glycerol stocks or conjugation plates onto M22 agar (see Appendix). Culture was serially diluted onto sucrose selection plates and subsequently replica plated onto M22 agar.

2.12.3 Semi-aerobic growth

Liquid cultures were grown semi-aerobically at 34 °C. Cultures were shaken continuously at 250 rpm in an orbital shaker. These conditions induced the biosynthesis of pigments and photosystems (Niederman *et al.*, 1976). Single colonies were inoculated in 10 ml of M22 medium and grown for 48 hours. Subsequently, a 125 ml conical flask containing 80 ml of M22 medium was inoculated with the entire 10 ml culture and grown under the same conditions overnight. These cultures were then either used for experimental work or grown further in 1.5 L of M22 medium in a 2 L conical flask or transferred into an appropriate vessel for photosynthetic (Section 2.12.4) or high oxygen growth (Section 2.12.5).

2.12.4 Photosynthetic growth

Anaerobic cultures were grown under photosynthetic conditions and exposed to 20 W MEGAMAN® CFL bulbs, or 116 W Osram Halogen Eco Pro bulbs to achieve the desired light intensity. Light intensity was measured in μmol photons s^{-1} m^{-2} using a LI-250A Light Meter equipped with a LI-190 Quantum Sensor (LI-COR Biosciences). A full 30 ml universal of M22

medium was inoculated with 1 ml of semi-aerobic culture and a small magnetic stir bar. The culture was incubated with gentle agitation and the desired light intensity overnight. This culture was used to inoculate either a 500 ml medical flat or a 1.2 L roux bottle filled to the top with M22 medium and capped with a rubber bung, and with a stir bar for gentle agitation. The terms “high light” and “low light” are used throughout this thesis, where “high light” is approximately 100 µmol photons s^{-1} m^{-2} and “low light” is approximately 10 µmol photons s^{-1} m^{-2}.

2.12.5 High oxygen growth

1 ml of semi-aerobic culture was used to inoculate 200 ml M22 medium in a highly baffled 2 L flask. This was incubated at 30 °C for 16 hours in an orbital shaker at 250-300 rpm to permit maximum aeration. Cells grown in this way lacked the B800, B850 and B875 absorbance peaks associated with the reaction centre, LH1 and LH2 complexes.

2.13 Conjugative transfer of plasmid DNA from *Escherichia coli* S17-1 to *Rhodobacter sphaeroides*

A modified version of the method of Hunter and Turner, 1988 was used to transfer pBBRBB, pIND4 and pK18mob*sacB* constructs from *E. coli* strain S17 to the desired *Rba. sphaeroides* strain. *Rba. sphaeroides* cells were grown in 80 ml M22 medium for 24 hours under semi-aerobic conditions, then pelleted and resuspended in 100 µl LB medium. Freshly transformed *E. coli* S17 colonies containing plasmid DNA were resuspended in 20 µl LB medium. 150 µl of resuspended *Rba. sphaeroides* cells was added to the resuspended *E. coli*. The mixture was pipetted as three drops onto a dried LB agar plate and incubated at 34 °C for less than 16 hours. The cells were scraped off the plate and streaked onto M22 agar containing the appropriate antibiotic and incubated at 34 °C. Transconjugant *Rba. sphaeroides* colonies appeared after 4-6 days.

2.14 Selection of *Rhodobacter sphaeroides* mutants on sucrose

Transconjugant *Rba. sphaeroides* colonies from a conjugative transfer of pK18mob*sacB* plasmid from *E. coli* were selected for on M22 agar plates with 25 µg ml^{-1} kanamycin. Single colonies were grown and scaled up to an 80 ml semi aerobic culture. The cells were serially diluted 10^{-2}, 10^{-3} and 10^{-4} onto M22 agar containing 1 % (w/v) sucrose and incubated until single colonies appeared after 4-6 days. Single colonies were replica plated onto M22 sucrose plates, with and without 25 µg ml^{-1} kanamycin. Colonies that grew on the antibiotic-free plate but not on the kanamycin plate were analysed by PCR. Successful mutants were streaked out

and grown up for further analysis and storage. Further detail on use of the pK18mob*sacB* suicide vector used to produce mutants is available in Sections 3.3.1-2.

2.15 Membrane preparation from *Rhodobacter sphaeroides*

2.15.1 Cell harvesting and breakage

Cells were pelleted at 4,000 x *g* for 30 minutes at 4 °C. Cell pellets were resuspended in the appropriate buffer for downstream analysis. Approximately 5 g of cells were used per 10 ml of buffer.

Lysozyme was added to the cell sample to a final concentration of 0.5 mg ml^{-1} and incubated at 25 °C for 30 min in the dark. A small spatula of deoxyribonuclease 1 from bovine pancreas (Sigma) was added to the cells. The cells were French pressed 3 times at 18,000 psi. The cells were kept on ice throughout. Unbroken cells were removed by centrifugation at 33, 000 x *g* at 4 °C for 25 minutes. The supernatant was transferred to a clean tube prior to loading onto a sucrose gradient.

2.15.2 Standard preparation of mixed intracytoplasmic membranes

High concentrations of intracytoplasmic membranes (ICM) were prepared using a 15/40 % (w/w) discontinuous sucrose gradient. 5-10 ml of broken cells (Section 2.15.1) were layered on top of the 15% sucrose band using a pipette. The gradients were kept on ice throughout. Gradients were centrifuged at 27, 000 rpm (65,000 x *g*) in a Beckman Type 45 Ti rotor at 4 °C for 16 hours. A pigmented band of ICM formed at the 15/40 % interface and was collected using a peristaltic pump. 20 ml of membrane sample and 70 ml of the desired buffer was placed into each of 6 centrifuge tubes. The samples were spun at 40,000 rpm (125,171 x *g*) in a Beckman Type 45 Ti rotor for 2 hours at 4 °C in order to pellet the membrane. The supernatant was discarded. The pellets were homogenised in approximately 3 ml of the desired buffer.

2.15.3 Solubilisation of ICM

Membranes harvested from discontinuous sucrose gradients (Section 2.15.2) were solubilised in 3 % β-DDM at 4 °C for 1 hour with continuous stirring in the dark, unless otherwise stated. The sample was centrifuged for 1 hour in a Beckman Ti 70.1 rotor at 48,000 rpm (160,000 x *g* avg) 4 °C to remove unsolubilised material. The supernatant was collected.

2.15.4 Fractionation of LH1-RC core complexes present in ICM membranes

Solubilised membranes (see Section 2.15.3) were layered on top of discontinuous sucrose gradients containing 20 %, 21.25 %, 22.5 %, 23.75 %, 25 % and 50 % sucrose in 20 mM HEPES, 5 mM EDTA and 0.03 % β-DDM. Gradients were centrifuged in a Beckman SW32 Ti rotor at 27,000 rpm (90,000 x *g* avg) at 4 °C for 40 hours. Pigmented bands were harvested for downstream processing.

2.16 *Synechocystis* strains, growth and fractionation

A glucose tolerant strain of *Synechocystis* sp. PCC6803 was obtained from Professor Wim Vermass (Arizona State University) (Williams, 1988). Strains were grown in BG11 media supplemented with glucose and TES KOH pH 8.2 (See Appendix) and grown under medium (40 µmol photons s^{-1} m^{-2}) or low light conditions (4 µmol photons s^{-1} m^{-2}) with agitation at 150 rpm at 30 °C.

2.16.1 Transformation of *Synechocystis* sp. PCC6803

Synechocystis was grown in 100 ml of liquid medium until an OD_{750} of 0.6 – 0.7 was reached. 1 ml of cell culture was taken and pelleted by centrifugation at 2,656 x *g* for 5 minutes. 900 µl of the supernatant was removed followed by the addition of add 5 µl of miniprep DNA (10 – 50 ng of plasmid or linear DNA). The cells were gently resuspended followed by incubation for 30 min under medium light conditions at 30 °C. The cells were transferred to a small BG11 agar plate and the suspension was allowed to dry. Cells were incubated in medium light at 30 °C for 24 hours. Initial selection was performed by transferring the cells to a BG11 plate containing 5 µg ml^{-1} kanamycin. Plates were incubated under low light at 30 µmol photons s^{-1} m^{-2} at 30 °C until colonies appeared (8 – 12 days). Colonies were transferred to a new agar plate containing double the concentration of antibiotic. This process was repeated with increasing amounts of antibiotic until a fully segregated mutant was obtained, as verified by colony PCR using ACCUZYME™.

Synechocystis contains many copies of its genomic DNA (~ 60 copies per cell). It is therefore vital to ensure the foreign DNA is introduced into every copy of the genome to avoid reversion on removal of antibiotic selection.

2.16.2 Breakage of *Synechocystis* by bead beating

Cultures of *Synechocystis* were harvested by centrifugation at 15,000 x *g* for 10 min. The pellet was resuspended in three volumes of FLAG buffer (25 mM sodium phosphate buffer, 10 mM MgCl2, 10 % glycerol (w/v), 0.05 M NaCl, pH 7.4) containing lysozyme and DNAse. 10

µl of EDTA-free protease inhibitor (BioCompare) was added (1ml to 1 tablet). All subsequent steps were performed in the dark or under a dim green light. The cells were disrupted in a bead beater using 0.1 mm glass beads in a 1:1 ratio of cell suspension to glass beads. 6 X 15 second bead beating cycles were used with 3 min on ice between each cycle. Glass beads were removed from the cell lysate by washing with 2 volumes of FLAG buffer, and unbroken cells were removed by centrifugation for 15 min at 8,000 x *g*.

2.16.3 Separation of thylakoid membrane and soluble cell fractions

Bead beaten cells were centrifuged at 64,000 x *g* for 45 min to separate the insoluble fraction (thylakoid membrane). The supernatant (blue) was removed as the soluble fraction. The thylakoid membrane pellet was resuspended in the minimum volume of FLAG buffer containing EDTA-free protease inhibitor (BioCompare). The membranes were solubilised with β-DDM to a final concentration of 1 % for 1 hour in the dark at 4 °C with gentle stirring. Unsolubilised material was removed by centrifugation at 30,000 x *g* for 30 min. The resulting solubilised material was either loaded onto a FLAG-affinity column or flash frozen in liquid nitrogen and stored at – 80 °C.

2.17 Protein manipulation

2.17.1 Quantification of protein concentration

To measure the total protein concentration, a BCA Protein assay kit (Thermo Scientific) or a Bradford Protein assay (Bio-Rad) was used according to the manufacturers' instructions. To quickly estimate protein concentration, the absorbance at 280 nm was used and the protein concentration in mg ml^{-1} was determined using the following equation:

$$A280\ at\ 1\ mg\ ml^{-1} = \left(5960n_{Trp} + 1280n_{Tyr} + 120n_{Cys}\right) \div Mr$$

Where n_{Trp}, n_{Tyr} and n_{Cys} are the numbers of tryptophan, tyrosine and cysteine residues in the protein, and Mr is the predicted molecular weight of the protein (Gill and Vonhippel, 1989).

For relative LH2 and core complex concentrations the 850 nm and 875 nm absorbance peaks were used respectively.

2.17.2 SDS-polyacrylamide gel electrophoresis

Protein samples were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using 12% Bis-Tris NuPAGE™ pre-cast polyacrylamide gels (Invitrogen). Samples were diluted 1 in 4 with loading buffer (to make 1 ml: 50 µl beta mercaptoethanol, 950 µl Nu PAGE running buffer). Prior to loading, samples were heated at 100 °C for 10 minutes or, for membrane

samples, 37 °C for 30 minutes, followed by centrifugation at 16,600 x *g* in a bench top microcentrifuge for 2 minutes. 5 µl of Precision Plus Protein Dual Colour Standards (Bio-Rad) was added to the gel as a size marker. For gels destined for western blotting, pre-stained markers were used. Protein bands were visualised by staining gels with Coomassie Brilliant Blue R250.

2.17.3 Western blot analysis of proteins

Protein samples were separated by SDS-PAGE as described in Section 2.17.2. After electrophoresis, a sandwich was assembled as follows: a porous sponge pad, 2 sheets of Whatman 3MM paper, the gel, nitrocellulose transfer membrane (Hybond ECL, Amersham), 2 sheets of 3MM paper and a porous sponge pad. This was arranged in a Bio-Rad blotting cassette after saturation in transfer buffer (190 mM glycine, 24 mM Tris, 20 % methanol). The gel was in direct contact with the nitrocellulose transfer membrane. The gel was closer to the cathode (black) and the membrane closer to the anode (red). The transfer was done for 1 hour at 350 mA at 4 °C in a blotting tank (Bio-Rad) submerged completely in transfer buffer. After transfer, the membrane was removed from the cassette and washed in wash buffer (18 mM Tris-HCl pH 7.6, 68 mM NaCl, 0.05 % Tween).

Following transfer, the membrane was blocked for at least 45 minutes in blocking buffer (5 % Marvel Milk Powder, 18 mM Tris-HCl pH 7.6, 68 mM NaCl). The membrane was incubated in primary antibody at the appropriate concentration in wash buffer for 3 - 16 hours at 4 °C. The membrane was then rinsed in 3 quick washes of wash buffer followed by three 5 min 20 ml washes. Secondary antibody at the appropriate concentration in wash buffer was added to the membrane and incubated at room temperature for 1 hour. The membrane was then rinsed in 3 quick washes of wash buffer followed by three 10 min 20 ml washes. Immunodetection was performed using Amersham™ ECL™ Western blotting analysis system (GE Healthcare Life Sciences) according to the manufacturer's instructions.

2.17.4 Purification of His-tagged proteins

His-tagged proteins were purified on a column packed with chelating Sepharose fast-flow resin (GE Healthcare). The column was run on a bench top using gravity. The column was first washed with 20 ml QH_2O in order to wash away the ethanol that the column was stored in. It was then washed with 20 ml charging buffer (70 mM $NiSO_4$, 50 mM sodium acetate, pH 4.5) followed by 20 ml binding buffer (20 mM Tris pH 7.4, 500 mM NaCl, 10 mM imidazole). The sample (often in the form of membranes prepared as in Section 2.15.2) was added to the top of the column. The column was then washed with 20 ml binding buffer, followed by 20 ml binding buffer with 50 mM imidazole, then 20 ml binding buffer with 100 mM imidazole. To

elute protein, the column was washed with 20 ml binding buffer containing 400 mM imidazole. Where required, all buffers contained 0.04 % β-DDM, unless otherwise stated. 1.5 ml fractions were taken. Protein content was analysed spectroscopically and using Bradford assay (Section 2.17.1).

2.17.5 Purification of FLAG tagged proteins

All purification buffers contained 0.04 % β-DDM and the purification was done in the dark. Anti-FLAG-M2-Agarose from mouse (Sigma) was flow packed in a Poly-Prep gravity flow column (BioRad) to the desired bed volume (100 – 300 µl) with 1 ml QH_2O. The resin was equilibrated with 5 ml FLAG buffer (25 mM sodium phosphate buffer, 10 mM $MgCl_2$, 10 % glycerol (w/v), 0.05 M NaCl, pH 7.4) containing EDTA-free protease inhibitor (BioCompare). The sample was loaded onto the column and the flow-through was re-applied once. The column was washed with 10 x 1 ml applications of FLAG buffer. Following washing, the anti-FLAG resin was resuspended in 300 µl wash buffer (containing 150 µg ml^{-1} FLAG peptide (sequence: DTKDDDDKDTKDDDDKDTKDDDDK) (Invitrogen)) followed by incubation on a rocker for 1 hour at 4 °C. The eluate was separated from the FLAG resin by passage through a Costar Spin-X column containing a 0.22 µm cellulose acetate membrane.

2.18 Carotenoid extraction and analysis

2.18.1 Carotenoid extraction from *Rhodobacter sphaeroides*

80 ml of semi-aerobic *Rba. sphaeroides* culture were pelleted by centrifugation and resuspended a minimum volume of 20 mM HEPES pH 7.4. 100 µl of the cell suspension was pelleted by centrifugation and 500 µl of acetone:methanol (7:2 v/v) was used to resuspend the pellet. The sample was centrifuged at 16,600 x *g* in a microcentrifuge at 4°C for 1 min. The supernatant was taken and 200 µl of hexane was added and the tube was inverted to mix. 100 µl of QH_20 was added and the tube was inverted to mix. The sample was centrifuged at 16,600 x *g* in a microcentrifuge at 4°C for 1 min. The top, coloured, layer containing the carotenoids was taken and centrifuged at 16,600 x *g* at 4°C for 5 min to remove any debris and the supernatant was transferred to a glass HPLC sample vial.

Alternatively, the carotenoid sample was dried with nitrogen on ice. The dried samples were sealed with parafilm and frozen at -20 °C overnight. The dried pigment was resuspended in a minimum volume of 80 µl of acetonitrile:H_2O (1:6 v/v), centrifuged at 16,600 x *g* at 4 °C for 1 min to remove any debris and the supernatant was transferred to a glass HPLC sample vial.

2.18.2 Carotenoid extraction from *Escherichia coli*

E. coli cells were harvested by centrifugation and resuspended in a minimum volume of 20 mM HEPES. The culture was sonicated and 0.5 ml of culture was transferred to an Eppendorf. 1 ml of acetone:methanol (7:2 v/v) was added and the sample was centrifuged at 16,600 x *g* for 5 min at 4 °C. The supernatant was taken. The pellet was resuspended in 1 ml acetone: methanol 7:2 and the centrifugation step was repeated. The supernatant was pooled and 1 or 2 ml of hexane was added. The top layer containing the carotenoid pigments was taken and dried on ice with nitrogen. The dried samples were sealed with parafilm and frozen at -20 °C overnight. The dried pigment was resuspended in a minimum volume of 80 µl of acetonitrile:H_2O (1:6 v/v), centrifuged at 16,600 x *g* at 4 °C for 5 min to remove any debris and the supernatant was transferred to a glass HPLC sample vial.

2.18.3 HPLC analysis of extracted carotenoids

HPLC analysis of pigments was performed on an Agilent 1200 high-performance liquid chromatograph using a Phenomenex Aqua C18 reverse phase column (5 µm particle size, 125 Å pore size, 250 x 4.6 mm). The column was washed with 50% each of acetonitrile:H_2O (1:6 v/v) and ethyl acetate. The column was equilibrated in acetonitrile:H_2O (1:6 v/v). The method used was a modified version of that in Garcia-Asua *et al*., 2002 in which a gradient to 100 % ethyl acetate was run over 25 minutes. The pigment sample injection volume was 50 µl. Pigments were eluted at 1 ml min^{-1}, elution of carotenoid species was monitored by scanning from 220 to 950 nm. Pigment elution was detected using a multichannel diode array detector (DAD) (Agilent) set to record absorbance at 442 nm. Pigment retention times were not significantly varied when loaded in either acetonitrile:H_2O (1:6 v/v) or hexane.

2.19 Spectroscopy

2.19.1 Room temperature absorbance spectra

Room temperature Spectra were taken using Cary UV/vis spectrophotometer at wavelengths 950 nm – 260 nm in an ultra violet (UV) cuvette with a 1 cm path length. Dilutions were done using the appropriate buffer or growth medium.

2.19.2 Low temperature absorbance spectra

Absorbance spectra were recorded using a Cary 500 UV/vis spectrophotometer at wavelengths 950- 260 nm in a UV cuvette with a 1 cm path length. Samples were cooled to 77 K in an Optistat DN-V optical cryostat manufactured by Oxford Instruments. Samples were suspended in a cryo-stable buffer consisting of 20 mM Tris-HCl, 80 % glycerol (v/v).

2.19.4 Low temperature fluorescence spectroscopy

All emission and excitation fluorescence spectra were recorded in a cryo-stable buffer consisting of 20 mM HEPES, 80% glycerol. Samples were cooled to 77 K in an Optistat DN-V optical cryostat manufactured by Oxford Instruments. Measurements were recorded on a SPEX FluoroLog spectrofluorimeter (SPEX Industries Inc.). Excitation was provided from a tungsten light source in the visible-IR region of the spectrum. Fluorescence excitation spectra were recorded from 400 nm to 910 nm with an emission of 915 nm. Fluorescence emission spectra were recorded using an excitation wavelength of 590 nm with 5 nm slit widths and emission slit widths of 2.5 mm. An average of 10 individual scans was used to measure excitation and emission spectra.

2.19.5 Fluorescence imaging of *Rhodobacter sphaeroides* cells containing YFP

Cells were washed three times in QH_2O then suspended in QH_2O. 15 µl of cell suspension was dropped on 30 µl of 1.5% agar film on a glass slide and sealed with DPX mountant (Sigma–Aldrich) between the glass slide and a coverslip.

Fluorescence images were taken with an inverted fluorescence microscope (AxioObserverA1m, Zeiss) equipped with a Hal 100 halogen lamp, a high intensity HBO 100 mercury lamp and an ORCA-ER camera (HAMAMATSU). Excitation light was first filtered by a 470/40 nm bandpass filter, then reflected by a 495 nm dichroic beam splitter to the sample through an objective (Plan-Apochromat 63x/1.40 oil objective, Zeiss). Fluorescence emission was filtered by a 520/40 nm bandpass filter before detection by the camera. Each fluorescence image was taken with a 0.1 s exposure time and 50 electron multiplication gain.

2.19.6 Fluorescence spectra and lifetime measurements of *Rhodobacter sphaeroides* cells containing YFP

The fluorescence emission properties of YFP in whole cells were measured on a home-built time-resolved fluorescence microscope. The inverted microscope is equipped with a spectrometer (Acton SP2558, Princeton Instruments), an electron-multiplying charge-coupled device (EMCCD) camera (ProEM 512, Princeton Instruments) and a Hybrid Detector (HPM-100-50, Becker & Hickl). Excitation light source was from a pulse supercontinuum white light laser (SC 480-10, Fianium) with a repetition rate of 40 MHz. The laser beam was focused on the sample surface illuminating a diffraction limited spot using 100 × (PF, NA = 1.4, oil immersion, Olympus). The resulting fluorescence spectral emission was detected through the spectrometer onto the EMCCD camera and the resulting fluorescence lifetime was detected through the spectrometer onto the Hybrid Detector.

During fluorescence spectral and lifetime measurements, the excitation light was filtered by a 472/30 nm bandpass filter, then reflected by a 495 nm dichroic beamsplitter to the sample. The resulting fluorescence emission was filtered by a 596 nm long-pass filter before being detected by the cameras. The fluorescence emission was captured with a slit width of 1500 µm and a grating of 150 line/mm working at a central wavelength of 550 nm in the spectrometer. Multiple measurements were performed on 8 different cells on each sample. Each fluorescence spectrum was detected by EMCCD at an average of 3 frames with a 1 s exposure time and an electron multiplication gain of 80. Analysis was done with OriginPro.

For fluorescence lifetime measurements, the modulation of the laser was synchronized with a time-correlated single-photon counting (TCSPC) module (SPC-150, Becker & Hickl). Fluorescence lifetimes were recorded by parking the focused laser spot over one single cell, selecting a central wavelength of 550nm by use of the monochromator and detected by the Hybrid Detector. SPCM software (Becker&Hickl) was used for the data acquisition. The families of decay curves were analysed with OriginPro and TRI2 software packages by fitting multiexponential decay function:

$$I(t) = \sum_{i=1}^{n} A_i exp\left(\frac{-t}{\tau_i}\right) + B$$

where τ_i is the fluorescence lifetime, A_i is the fractional amplitude contribution of the i^{th} decay component, and B is the background. The quality of fit was judged on the basis of the reduced χ^2 statistic:

$$\chi^2{}_{\mathrm{red}} = \frac{\sum_{\mathrm{k=1}}^{\mathrm{n}} \frac{[\mathrm{I(t_k)} - \mathrm{I_c(t_k)}]^2}{\mathrm{I(t_k)}}}{\mathrm{n - p}} = \frac{\chi^2}{\mathrm{n - p}}$$

where $I(t_k)$ is the data at time point k, $I_c(t_k)$ is the fit at time point k, n is the number of data points and p is the number of variable fit parameters ($n - p$ = degrees of freedom).

The instrument response (IRF) of the system, measured using a mirror, was approximately 0.18 ns, and the convolution of the decay curves with the IRF was taken into account when the fitting was performed.

2.19.7 Extended timescale ultrafast transient absorption measurements

The transient-absorption (TA) studies of the samples (RC and YFP concentrations of ~10-15 µM) utilized an amplified Ti:Sapphire laser system (Spectra Physics) and a Helios or Eos

spectrometer (Ultrafast Systems) with 515 nm, ~0.5 μJ, ~100 fs excitation pulses (at 1 KHz) focused to 1 mm diameter. For these TA studies, any samples containing the RC were spun in a 3-ml, home-built spinning cell at 300-600 rpm (required to prevent re-excitation of P^+); the YFP-only samples used traditional stir bar-containing, 2 mm cells (~500 ul).

Table 2.1 Plasmids

Table 2.1.1 Empty vectors

Plasmid	Resistance	Induction	Source/Reference
pK18mob*sacB*	Km^R Suc^S	N/A	Professor J. Armitage, University of Oxford
pIND4	Km^R	IPTG	Ind *et al*., 2009
pJEXPRESS 414	Amp^R	IPTG	DNA 2.0
pBAD colA	Km^R	Arabinose	Dr. D. Canniffe, University of Sheffield
pBBRBB-*Ppuf* $_{843\text{-}1200}$	Km^R	Constitutive	Tikh *et al*., 2014
pBBRBB-*Ppuf* $_{1\text{-}1200}$	Km^R	Low oxygen	Tikh *et al*., 2014
pET9a	Km^R	IPTG	Novagen
pEXT22	Km^R	IPTG	Dykxhoorn *et al*., 1996
pFLAG	Amp^R Km^R	N/A	Dr P. Davison and Dr D. Canniffe, University of Sheffield

Table 2.1.2 Plasmids used in Chapter 3

Plasmid	Properties	Source/Reference

pK18mob*sacB-crtB*KO	Δ*crtB* construct containing upstream and downstream sequences ligated in frame.	This study
pK18mob*sacB-puhA-syfp2*	C-terminus SYFP2 fusion containing upstream, *syfp2*, and downstream sequences of *puhA* ligated in frame.	E. Martin, University of Sheffield
pBBRBB $_{843\text{-}1200}$ – *syfp2*	Containing *syfp2*.	This study

Table 2.1.3 Plasmids used in Chapter 4

Plasmid	**Properties**	**Source/Reference**
pK18mob*sacB*-Δ*pucBA*	Construct containing upstream and downstream sequences of *puc1BA* to produce a knockout of this gene.	E. Martin, University of Sheffield
pK18mob*sacB*-PR	Construct containing upstream and downstream sequences of *puc1BA* with the PR gene, ligated in frame.	This study
pK18mob*sacB-crtIcrtYblh*	Construct containing upstream and downstream sequences of *crtI*. These sequences flank the *crtI* and *crtY* genes from *Erwinia herbicola* and a codon optimised version of *blh* from uncultured marine bacterium.	This study
pK18mob*sacB-crtYIblh*2	As above but with the ribosome binding site from pIND4.	This study
pIND4-PR	Containing the PR gene	This study
PBBRBB $_{843\text{ - }1200}$ - PR	Containing the PR gene	This study

pIND4-*blh*	Containing a codon optimised version of *blh* from uncultured marine bacterium.	This study
pORANGE	Containing *crt E, B, I, Y* from *Erwinia herbicola*. Arabinose. Cm^R.	Von Lintig and Vogt, 2000
pKJ900	Containing *MBP*, and mouse *β-diox*. Containing *lacUV5* promoter for *MBP*, and *AraC* for *β-diox*. Amp^R.	Kim *et al*., 2008
pK18mob*sacB-ΔccoP*	Construct for the knockout of *Rba. sphaeroides ccoP* gene containing upstream and downstream sequences of *ccoP*.	E. Martin, University of Sheffield

Table 2.1.4 Plasmids used in Chapter 5

Plasmid	**Properties**	**Source/Reference**
pK18mob*sacB*-TM	Δ*1BA* Construct containing upstream and downstream sequences of *puc1BA* with the TM gene, ligated in frame.	This study
pIND4-TM-FLAG	Containing TM with an N-terminal FLAG tag and a C-terminal His tag.	This study
PIND4-TM	Containing TM with a C-terminal His tag.	This study
pBBRBB $_{843 - 1200}$ TM	Containing TM with a C-terminal His tag.	This study
pET9a-TM	TM with a C-terminal His tag.	This study

Table 2.1.5 Plasmids used in Chapter 6

Plasmid	Properties	Source/Reference
pJEXPRESS414-BT6-GK	N terminal His tag	G. Kodali, University of Pennsylvania
pJEXPRESS414-BT6	C terminal His tag	This study
pJEXPRESS414-BT6M-GK	N terminal His tag	G. Kodali, University of Pennsylvania
pJEXPRESS414-BT6M	C terminal His tag	This study
pJEXPRESS414-TorA-BT6	BT6 with a TorA signal peptide on the N-terminus, and a C-terminal His tag.	This study
pJEXPRESS414-TorA-BT6M	BT6M with a TorA signal peptide on the N-terminus, and a C-terminal His tag.	This study
pJEXPRESS414-TorA-KR-BT6	BT6-KR with a TorA signal peptide on the N-terminus, and a C-terminal His tag.	This study
pJEXPRESS414-TorA- KK-BT6	BT6-KK with a TorA signal peptide on the N-terminus, and a C-terminal His tag.	This study
pJEXPRESS414-TorA-BT6-H53A	BT6-H53A with a TorA signal peptide on the N-terminus, and a C-terminal His tag.	This study
pEXT22-BT6	As above	This study; A. Jones, University of Kent
pEXT22-TorA-BT6	As above	This study; A. Jones, University of Kent

pEXT22-BT6M	As above	This study; A. Jones, University of Kent
pEXT22-TorA-BT6M	As above	This study; A. Jones, University of Kent
pEXT22-TorA-KR-BT6	As above	This study; A. Jones, University of Kent
pEXT22-TorA-KK-BT6	As above	This study; A. Jones, University of Kent

Table 2.2 *Rhodobacter sphaeroides* strains

Table 2.2.1 *Rba. sphaeroides* strains used in multiple chapters

Strain	Properties	Source/reference
2.4.1	Wild type	S. Kaplan, University of Texas
Δ*puc1BA*	In frame genomic deletion of *pucBA* operon.	E. Martin, University of Sheffield
Δ*puc1BA* Δ*pufX*	In frame genomic deletion of *pucBA* and *pufX*.	E. Martin, University of Sheffield
Δ*pufX*	In frame genomic deletion of *pufX*.	E. Martin, University of Sheffield
Δ*crtB*	In frame genomic deletion of *crtB*.	This study

Table 2.2.2 *Rba. sphaeroides* strains used in Chapter 3

Strain	Properties	Source/reference
WT *puhA-syfp2* (WT RCH-YFP)	In frame genomic fusion of *puhA* and C-terminus *syfp2*	E. Martin, University of Sheffield

Δ*crtB puhA-syfp2* (Δ*crtB* RCH-YFP)	In frame genomic deletion of *crtB*. In frame genomic fusion of *puhA* and C-terminus *syfp2*	E. Martin, University of Sheffield
WT pBBRBB $_{843\text{-}1,200}$ YFP	Constitutive expression of *syfp2*	This study
Δ*crtB* pBBRBB $_{843\text{-}1,200}$ YFP	Constitutive expression of *syfp2*	This study

Table 2.2.3 *Rba. sphaeroides* strains used in Chapter 4

Strain	**Properties**	**Source/reference**
Δ*crtC*	In frame genomic deletion of *crtC*	E. Martin, University of Sheffield
Δ*crtC* Δ*crtI*:: *crtI crtY*	In frame genomic deletion of *crtC*. *Pantoea agglomerans crtI* and *crtY* replace the native *crtI*.	S. Chi, University of Sheffield
Δ*crtC* Δ*crtI*:: *crtI crtY blh* (*crtIYblh*)	In frame genomic deletion of *crtC*. *Pantoea agglomerans crtI* and *crtY* and a codon optimise variant of *blh* from uncultured marine bacterium replace the native *crtI*.	This study
Δ*crtC* Δ*crtI*:: *crtI crtY blh* 2 (*crtIYblh*2)	As above but using the ribosome binding site from pIND4	This study
ΔpucBA ΔcrtC Δ*crtI*:: *crtI crtY blh* (*crtIYblh*2 Δ*1BA*)	As above. In-frame deletion of the *puc1BA* genes.	This study
ΔpucBA ΔcrtC Δ*crtI*:: *crtI crtY blh PR* (*crtIYblh*2 Δ*1BA*::PR)	As above. The gene for PR in place of the *puc1BA* genes.	This study

ΔpucBA ΔcrtC Δ*crtI ΔccoP:*: *crtI crtY blh* (*crtIYblh*2 Δ*1BA* Δ*ccoP*)	As previously. In-frame deletion of the *ccoP* gene.	This study; E. Martin, University of Sheffield
ΔpucBA ΔcrtC Δ*crtI Δccop:*: *crtI crtY blh PR*	As above. The gene for PR in place of the *puc1BA* genes.	This study

Table 2.2.4 *Rba. sphaeroides* strains used in Chapter 5

Strain	Properties	Source
Δ*bchCFX*	In frame deletion of the *bchCFX* genes.	J. Chidgey, University of Sheffield

Table 2.3 Primers

The primers were synthesised by Sigma. Primer DNA was diluted to a concentration of 125 ng μl^{-1}.

Table 2.3.1 Primers used in Chapter 3

Name	Sequence	Cleavage Site
crtBKOUF	CCGGAATTCCACATCACCATCACCACGGCG	*Eco*RI
crtBKOUR	GCGCTCTAGAGATCTAGGTTCTCATGAAGGTATACCG	*Xba*I
crtBKODF	GCGCTCTAGAGGCAATCATTCCGCGGCAAGC	*Xba*I
crtBKODR	CCCCGCATGCGGCTGTGGCCGAGCCCTA	*Sph*I
crtBseqF	CCCGCAGGCCCGCCCCTC	N/A
crtBseqR	TCGTCAATGCGCCGCGCT	N/A
puhAYFPUF	CCGGAATTCTCGGCCGGCAAGAACCCGATCGG	*Eco*RI
puhAYFPUR	GCTCCTCGCCCTTGCTCACCATGGCGTATTCGGCCAGCATCGCCG	N/A
puhAYFPDF	GCGCTCTAGATCCCCGCATGGCGCGGCCC	*Xba*I
puhAYFPDR	CCCCAAGCTTTAGGGCACCGCATAGGCCACCGC	*Hin*dIII
puhAYFPFor	CGGCGATGCTGGCCGAATACGCCATGGTGAGCAAGGGCGAGGAGCTGTTCAC	N/A

puhAYFPRev	GCGCTCTAGATCATTACTTGTACAGCTCGTCCATGCC GAGAGTGAT	*Xba*I
puhAYFPseqF	GAAGCAGCACGACTTCTTCAAGTC	N/A
puhAYFPseqR	CCACGATCTATTCGATCACCACAGC	N/A
pBBRBB *syfp2* F	GCGCAGATCTATGGTGAGCAAGGGCGAG	*Bgl*II
pBBRBB *syfp2* R	GCGCGGCGGCCGCTTACTTGTACAGCTCGTCCATG	*Not*I

Table 2.3.2 Primers used in Chapter 4

Name	**Sequence**	**Cleavage Site**
crtYI F	GCGCTCTAGAATGAGGGATCTGATTTTAGTCGG	*Xba*I
crtYI R	GCCGTCGACTCATTGCAGATCCTCAATCA	*Sal*I
Blh F	GGCGCCCATGGATATGGGCCTCATGCTCAT	*Nco*I
Blh R	CGCCGAAGCTTTCAGTTCTTGATCTTGATGCG	*Hin*dIII
Blh F rbs	GCGGCGGTCGACGAGGAGAAATTAACCATGGGCCT CATGCTCATCG	*Sal*I
Blh R NdeI	GCGGCGCATATGGCGGCAAGCCTTTC	*Nde*I
PR OE F	GTTGGGAGACGACACAATGGGTAAATTATTACTGAT ATTAGGTAGTG	N/A
PR OE R	GCGCCTTGCGCAGCATCAGTGGTGGTGATGGTG	N/A
PucBA KO F SalI	CCGGTCGACGCCAAGCCATCCTGAAATCTCG	*Sal*I
PR up OE R	CAGTAATAATTTACCCATTGTGTCGTCTCCCAACT	N/A
PR down OE F	CATCACCACCACTGATGCTGCGCAAGGCG	N/A
PucBA KO DR	CCCCAAGCTTGTGTCGGACTTGAACCCGATCAG	N/A
PucBA KO check F	CACGGCCATGTGCTGAAGATC	N/A
PucBA KO check R	CACCGTCTGGATCGTGTGCAC	N/A
PR F FLAG	AAGGAAAAAAGCGGCCGCAGGTAAATTATTACTGA TATTAGG	*Not*I
PR R *BglII*	ATAGATCTTCAACCGGTACGCGTAGAATCGAGACC	*Bgl*II
FLAG check F	CTCTCATTAATCCTTTAGAC	N/A
FLAG check R	GCATTACGCTGACTTGACGG	N/A

Table 2.3.3 Primers used in Chapter 5

Name	**Sequence**	**Cleavage site**

TM no FLAG F	GGCGCCCATGGGATCCGGCCAGATTTGG	*Nco*I
TM R HindIII	CGCCGAAGCTTTCAATGATGATGGTGGTGATG	*Hin*dIII
TM F pET9a	GGCCTCTAGAATGGGATCCGGGCAGATTTG	*Xba*I
TM R pET9a	GGCCGGATCCTCAATGATGATGGTGGTGATGG	*Bam*HI
TM OE F	GTTGGGAGACGACACAATGGGATCCGGGCAG	N/A
TM OE R	GCGCCTTGCGCAGCATCAATGATGATGGTGGTGATG	N/A
PucBA KO F EcoRI	CCGGAATTCGCCAAGCCATCCTGAAATCTCG	*Eco*RI
TM up OE R	CTGCCCGGATCCCATTGTGTCGTCTCCCAACT	N/A
TM down OE F	CACCATCATCATTGATGCTGCGCAAGGCG	N/A
PucBA KO DR	CCCCAAGCTTGTGTCGGACTTGAACCCGATCAG	*Hind*III

Table 2.3.4 Primers used in Chapter 6

Name	**Sequence**	**Cleavage site**
TorA F	GGAATTCCATATGAACAATAACGATCTCTTTCAGG	*Nde*I
TorA-BT6 Rev OE	CTCGCCGTCGCCGCCCGCCGCTTGCGCCGCAGTCGCAC	N/A
TorA-BT6 For OE	GCGGCGCAAGCGGCGGGCGGCGACGGCGAGAATCTG	N/A
BT6 his R	CTGGCTCGAGTTAATGGTGGTGATGATGGTGCAATTGCTTCAGATCTTCAAATTGG	*Xho*I
BT6 F	GGAATTCCATATGGGC GGC GAC GGC GAG AAT C	*Nde*I
TorA KR QC AS	TGTGCCAGAAAACGCCGCTTTGATGCCTGAAAGAGATCGTTATTG	N/A
TorA KR QC S	CAATAACGATCTCTTTCAGGCATCAAAGCGGCGTTTTCTGGCACA	N/A
TorA KK QC S	GATCTCTTTCAGGCATCAAAGAAGCGTTTTCTGGCACAACTC	N/A
TorA KK QC AS	GAGTTGTGCCAGAAAACGCTTCTTTGATGCCTGAAAGAGATC	N/A
TorA-BT6M Rev OE	TTCGCCGTCGCCGCCCGCCGCTTGCGCCGCAGTCGCAC	N/A
TorA-BT6M For OE	GCGGCGCAAGCGGCGGGCGGCGACGGCGAAAACC	N/A

	TC	
BT6M his R	CTGGCTCGAGTTAATGGTGGTGATGATGGTGCAGC TGTTTCAGATCCTCAAACTGA	N/A
BT6M F	GGAATTCCATATGGGCGGCGACGGCGAAAAC	*Nde*I
BT6 H53A QC S	TGCAGCGCATCCTCGGCCTGCTTCCAGATCTC	N/A
BT6 H53A QC AS	GAGATCTGGAAGCAGGCCGAGGATGCGCTGCA	N/A

Chapter 3

Expanding the range of light absorbers for bacterial photosynthesis: YFP-enhanced charge separation at the *Rhodobacter sphaeroides* reaction centre

3.1 Summary

Much of the visible and near-infrared radiation falling on the surface of the Earth is not absorbed by photosynthesising organisms, which occupy particular spectral niches depending on the absorption of the particular pigments they synthesise. While organisms are able to harvest sufficient light for sustenance, for human solar energy needs it is desirable to utilise light across the entire spectrum. Although the incorporation of synthetic chromophores to complement native light-harvesting proteins is promising, the approach generally involves *in vitro* reassembly. Fluorescent proteins such as the *Aequorea victoria* green fluorescent protein (GFP) and its derivatives offer the advantage in that they are well studied and genetically programmable.

In this study, a translational fusion of the *Rhodobacter sphaeroides* reaction centre (RC) H subunit and a yellow variant of GFP, SYFP2, was constructed in *Rba. sphaeroides* wild type and carotenoidless (Δ*crtB*) backgrounds. An increase in the rate of photosynthetic growth was observed in Δ*crtB* RCH-YFP compared to Δ*crtB*. This increased growth rate is due to the transfer of energy from YFP to the *Rba. sphaeroides* LH1 and/or the RC proteins.

The spectral overlap between the emission of YFP and the visible-region (Q_X) absorption bands of the RC allows for energy transfer from YFP to the RC via a Förster mechanism. A number of static and ultrafast time-resolved techniques were used to show that the energy transfer yield from YFP to the RC was 40 ± 10%. This is sufficient to enhance the photosynthetic growth rate in the *Rba. sphaeroides* carotenoidless mutant. This work demonstrates the viability of incorporating new genetically encoded chromophores into existing systems to create new photosynthetic pathways for the augmentation of photosynthesis.

3.2 Introduction

Much of the visible and near-infrared radiation falling on the surface of the Earth is not absorbed by photosynthesising organisms, which occupy particular spectral niches depending on the absorption of the particular chlorophyll, carotenoid, phycobilin and other pigments they synthesise (Blankenship *et al*., 2011; Slouf *et al*., 2012). While organisms are able to harvest sufficient light for sustenance, for human solar energy needs it is desirable to utilise light across the entire spectrum, from the near-UV through to the NIR region. For synthetic biology applications it would be worthwhile to design and construct bacteria for light-powered bioreactors that could utilise a greater range of wavelengths then naturally-evolved photosynthetic bacteria. Thus arises the challenge of constructing antennas that can make use of such a wide range of wavelengths. The light-harvesting antennas found in photosynthetic organisms are flexible and diverse, providing a basis for the design of artificial light-harvesting constructs with a broad spectral coverage and the ability to transfer the resulting energy to a RC-like site with high efficiency.

Challenges arise when creating artificial light-harvesting antenna systems. The designs must emulate the well-defined 3D architecture of natural light-harvesting proteins which contain a large number of pigments and feature rapid, efficient and multistep energy transfer along a "spectroscopic gradient" of a series of pigments.

An emerging platform technology for the creation of artificial antennas is the biohybrid approach, which combines parts of native antennas with synthetic chromophores. Three key design features have been identified that these biohybrid constructs must incorporate: 1, The synthetic and natural chromophores must complement each other to provide pan-chromatic absorption throughout the visible and NIR regions of the spectrum. 2, It must be possible to create a high density of synthetic chromophores within the constructs to maximise absorption. 3, Appropriate energy levels much be achieved in order to funnel the captured light energy in a directional and efficient manner to the downstream chemical processing sites. (Harris *et al*., 2014b)

A combination of proteins analogous to the *Rba. sphaeroides* LH1 complex with a toolbox of bioconjugatable chromophores have been developed. Truncated forms of the *Rba. sphaeroides* LH1 β-polypeptides have been engineered to covalently bind synthetic chromophores that can efficiently transfer energy to the BChl *a* site (Meadows *et al*., 1995; Meadows *et al*., 1998; Kehoe *et al*., 1998; Springer *et al*., 2012; Harris *et al*., 2013; Harris *et al*., 2014a). These short β-peptides bind two BChl *a* molecules each and form homodimeric ββ-dyads. They can also associate to form β-oligomers. Wavelength tuneable bioconjugatable

synthetic bacteriochlorins have been developed that can attach to the peptides to form βn-oligomers which combine a strong NIR absorption with around 90% energy transfer efficiency (Reddy *et al*., 2013).

More recently, native-length bacterial light-harvesting peptides with covalently attached designer chromophores have been created (Harris *et al*., 2013; Harris *et al*., 2014a; Harris *et al*., 2014b). These self-assemble with native BChl *a* to create stable antennas. These structures oligomerize to form biohybrid analogues of the LH1 complex. The designs possess 45 to 60 pigments providing enhanced spectral coverage, and can accommodate pigments at remote sites that contribute to solar light harvesting via an energy transfer cascade. The efficiencies of energy transfer to the BChl *a* target is comparable with native antennas. Free accessory chromophores have been successfully used in these systems within detergent micelles, providing another route to enhance light-harvesting capability. (Harris *et al*., 2014a; Harris *et al*., 2014b).

These biohybrid antennas combine designer chromophores with native scaffolding and are flexible, versatile and tailorable, providing the potential to extend solar coverage beyond natural systems. The absorbed solar energy could then be used for a variety of uses such as charge separation, photocurrent generation or fuel production. Other energy transfer systems have been created that contain diverse pigments and have been used in solar cells (Hardin *et al*., 2009; Bozdemir *et al*., 2011; Ichikawa *et al*., 2013; Botta *et al*., 2013), for clinical diagnostic assays (Algar *et al*., 2012; Fabian *et al*., 2013), and structural probes (Watrob *et al*., 2003; Gehne *et al*., 2012).

There has been much interest in decorating different native light-harvesting and photosynthesis complexes with extra chromophores. However, while biohybrid architectures have great potential for producing artificial light-harvesting architectures *in vitro*, because they utilise synthetic chemistry they are unable to be replicated *in vivo* due to the inability of the bacterium to produce the designer chromophores. In order to create tailor-made light harvesting antennas *in vivo*, we must make use of the toolbox of proteins and pigments available in nature, or create synthetic elements that are able to be created by the host organism.

In *Rba. sphaeroides* light is harvested by the LH2 and LH1 antenna complexes and the resulting excitation energy is used to power the reduction of quinone to quinol prior to the formation of a proton gradient that powers ATP synthesis. *Rba. sphaeroides* transfers the energy from light absorbed by B875 bacteriochlorophylls (BChls) in LH1 and B850 BChls in LH2 to the reaction centre (RC) with near 100% quantum efficiency (Sener *et al*., 2007; Sener

et al., 2009; Sener and Schulten, 2009). The *Rba. sphaeroides* RC is the point at which charge separation takes place when the excitation energy of pigments is converted into biochemical energy. To maximise the efficiency of energy capture the RC was chosen as an attachment point for an additional light-harvesting component. The distance between the donor chromophore and the RC BChl *a* "special pair" is very important for the efficiency of Forster resonance energy transfer (FRET); the efficiency of energy transfer is inversely proportional to the sixth power of distance between donor and acceptor (Piston and Kremers, 2007).

As a pilot study to investigate the possibility of creating artificial light-harvesting antennas *in vivo*, this work investigates the effects of incorporating yellow fluorescent protein, YFP, as a chromophore into the photosynthesis apparatus of *Rba. sphaeroides*. YFP is a well-studied protein and is genetically programmable and has a good spectral overlap with the Q_x bands of the *Rba. sphaeroides* RC (Figure 3.1). Like all fluorescent proteins, YFP absorbs light in a defined range of wavelengths (absorption maximum at 517 nm) and emits it at another wavelength (525 nm). The YFP variant used in this study, SYFP2, exhibits photostability, a high fluorescence quantum yield (~ 70 %) and a large extinction coefficient (~ 100,000 M^{-1} cm^{-1} at 515 nm) (Kremers *et al*., 2006). This study aims to investigate whether the photons that would otherwise be emitted by YFP could instead be funnelled as useful energy and augment the *Rba. sphaeroides* photosynthesis process.

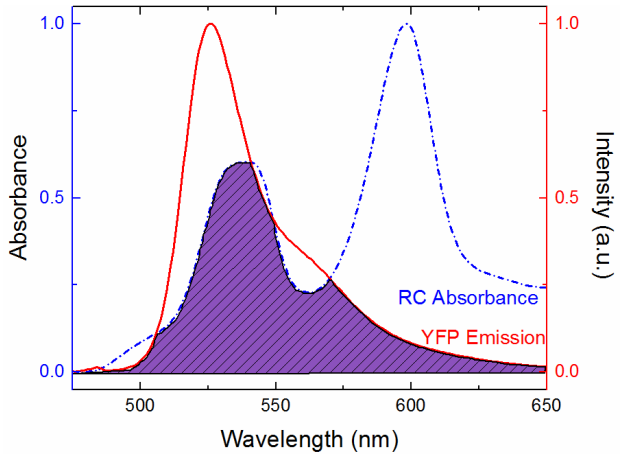


Figure 3.1 The spectral overlap between the *Rba. sphaeroides* reaction centre and SYFP2

Spectral overlap of the normalised YFP energy donor emission (red, solid line) and the RC energy acceptor absorbance (blue, short dashed line). Spectra are taken from the purified protein in Sections 3.3.7 and 3.3.8.

(Figure reproduced with permission from Kaitlyn Faries)

The native absorption spectrum of WT *Rba. sphaeroides* has a large gap between 550-750 nm, with the small Q_x transition at 590 nm the only feature. YFP cannot contribute usefully to the native absorption, because carotenoid pigments of the light-harvesting complexes absorb

in the 400-550 nm range. Accordingly, a carotenoidless mutant of *Rba. sphaeroides* was constructed in order to create a baseline strain to observe the effect of the addition of YFP.

Rba. sphaeroides strains that posess mutations in the *crtB* gene, encoding phytoene synthase, have been important in showing the importance of the role of carotenoids in photosynthetic bacteria. The mutants UV33 and R-26 of *Rba. sphaeroides* are blue-green in colour (Sistrom *et al*., 1956; Clayton and Smith, 1960; Ng *et al*., 2011). *crtB* mutants produce the carotenoid precursor molecule, phytoene, and do not synthesise coloured carotenoids (Clayton, 1962; Lang and Hunter, 1994). They cannot assemble LH2 and are therefore photosynthetically compromised. They are also sensitive to photoxidative damage due to the lack of carotenoids which provide photoprotection (Cogdell and Frank, 1987). These mutants, particularly R-26, have been extremely useful for the study of photosynthesis. However, UV33 and R-26 are prone to reversion and thus the restoration of the production of coloured carotenoids. For this reason the creation of a clean and complete knockout of the *crtB* gene was necessary for this work.

In this chapter, the gene encoding the yellow fluorescent protein (YFP) variant, SYFP2 (Kremers *et al*., 2006), was fused to the 3' end of *puhA*, which encodes the reaction centre H subunit (H). The effects on photosynthetic growth rate in both the wild type and a Δ*crtB* mutant were characterised. A number of static and ultrafast time-resolved techniques were used to characterise the energy transfer processes within the new complex. This project forms the basis of future studies involving the creation of artificial light-harvesting antenna for the increased spectral coverage of *in vivo* photosynthesis.

3.3 Results

3.3.1 Construction of the Δ*crtB* deletion strain

To create a baseline strain for the expression of YFP a *crtB* knockout strain was created in order to remove all carotenoids whose absorbance wavelengths overlap with that of YFP.

To delete the *crtB* ORF a construct containing the upstream and downstream flanking regions of the gene was created. Using the primers "crtBKOUF" and "crtBKOUR" (designed by David Mothersole) a fragment of 361 bp (*Eco*RI and *Xba*I ended) upstream of the *crtB* ORF was amplified, which also included 7 bp of the *crtB* gene. A second fragment of 363 bp (*Xba*I and *Sph*I ended) was amplified using the primers "crtBKODF" and "crtBKODR". This produced the downstream flanking fragment which included 9 bp of the *crtB* gene. Following digestion, the fragments were ligated into the suicide vector pK18mob*sacB*.

The resulting plasmid was transformed into *E. coli* S17 cells and transferred *Rba. sphaeroides* WT via conjugation. The mutant was produced using the method detailed in Section 2.13-14. The recombination events involved are illustrated in Figure 3.2 B.

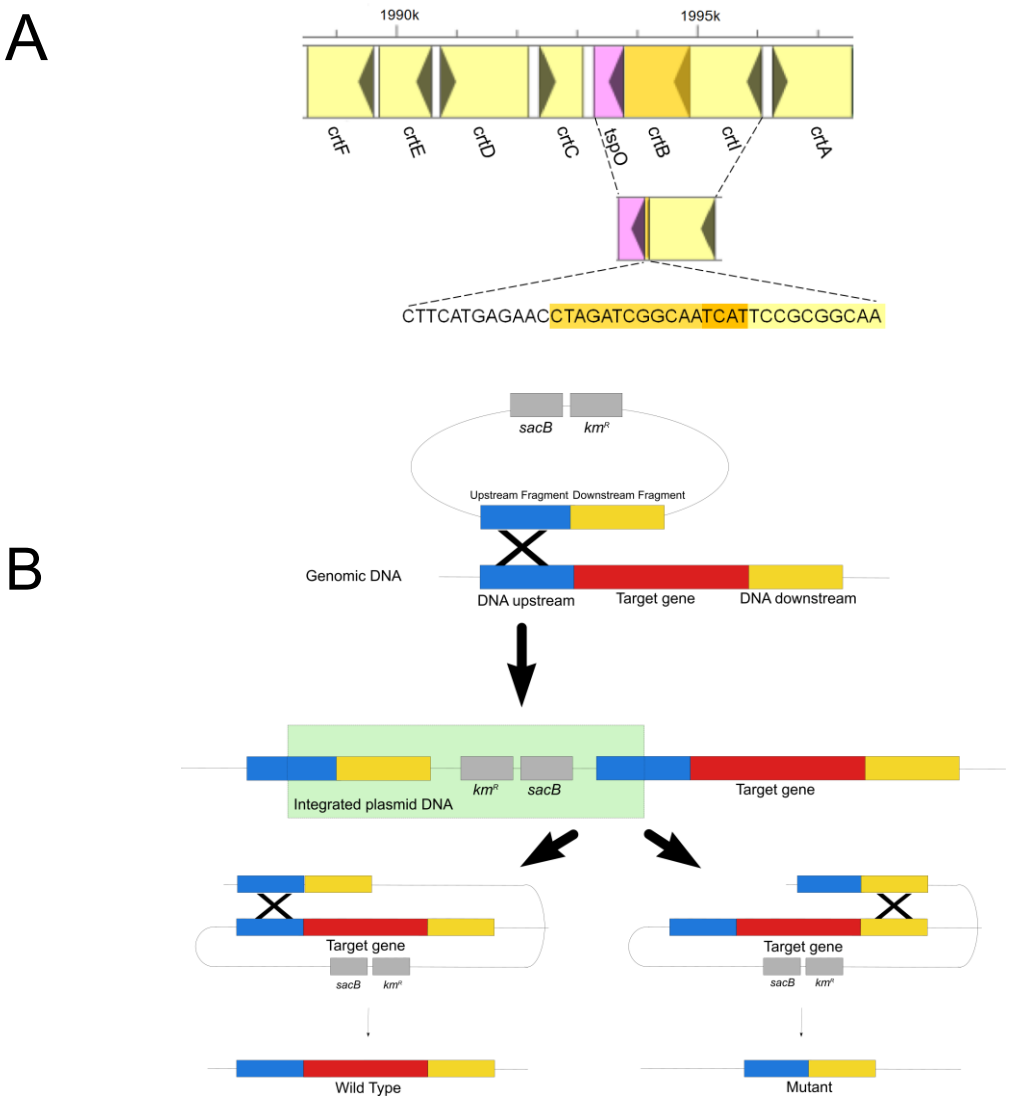


Figure 3.2 Deletion of *crtB* using pK18mob*sacB*

A. Schematic demonstrating the new DNA sequence of the *crtB* deletion strain.

B. The *crtB* gene was deleted using the suicide vector pK18mob*sacB* without the insertion of a resistance cassette. Following conjugation between *Rba. sphaeroides* and *E. coli* S17 cells containing the plasmid construct, a double homologous recombination event occurs. The first event occurs when the plasmid containing the upstream and downstream fragments of the target gene integrates into the genome. When the strain is grown in the presence of sucrose the second event occurs, causing the plasmid to be looped back out. There are two outcomes: either wild type or mutant *Rba. sphaeroides*. Black lines indicate a crossover event. Inserts are shown in blue and yellow. Area of genome corresponding to the DNA to be excised is indicated in red. The *sacB* gene confers sensitivity to sucrose. The km^R gene confers resistance to kanamycin.

(B: Figure reproduced from Mothersole, 2013)

Colonies that grew on a plate containing no antibiotic but failed to grow on the plate containing kanamycin were analysed by colony PCR to detect the deletion of *crtB* using the primers "crtBseqF" and "crtBseqR" designed to specifically amplify a genomic fragment. Sequencing of the fragment confirmed that *crtB* had been successfully and cleanly deleted (Figure 3.2 A).

The theoretical probability of obtaining a deletion strain of *crtB*, or reversal to the original WT strain was 50%. In reality the actual probability of successful mutation was much lower.

3.3.2 Construction of the RCH-YFP fusion strain

The gene encoding the yellow fluorescent protein (YFP) variant, SYFP2 (Kremers *et al*., 2006), was fused to the 3' end of *puhA*, which encodes the reaction centre H subunit (H). This results in a fusion protein with YFP on the C-terminus of H, the cytoplasmic side of the complex. The placement of YFP on the N-terminus of H was ruled out as, although it would bring the YFP within closer distance of the RC special pair and the LH1 BChls, it would have likely blocked the cytochrome c_2 docking site. Hereafter the *puhA-syfp2* mutant is termed RCH-YFP and was constructed by Elizabeth Martin and Sarah Burgess. Primers were designed by David Mothersole.

To create the C-terminal YFP fusion, an upstream fragment was created using the primers "puhAYFPUF" (*Eco*RI) and "puhAYFPUR" (an overlap extension primer). This fragment contained the last 354 bp of the *puhA* ORF excluding the stop codon. A third fragment containing the *syfp2* ORF was amplified using the primers "puhAYFPFor" (an overlap extension primer) and "puhAYFPRev" (*Xba*I). The upstream and YFP fragments were ligated together using overlap extension PCR. A downstream fragment was created using the primers "puhAYFPDF" and "puhAYFPDR" (*Xba*I and *Hin*dIII ended). This fragment contained 354 bp of the DNA immediately downstream of the *puhA* stop codon. Following digestion, the downstream fragment was ligated into the suicide vector pK18mob*sacB* first as after ligation the upstream *Xba*I site is dam methylated and can no longer be cut by *Xba*I. Both fragments (1428 bp in total) were ligated into pK18mob*sacB*.

The resulting plasmid was transformed into *E. coli* S17 cells and transferred *Rba. sphaeroides* WT and Δ*crtB*. The mutant was produced using the method detailed in Sections 2.13-14. The recombination events involved are illustrated in Figure 3.3. Colonies that grew on a plate containing no antibiotic but failed to grow on the plate containing kanamycin were analysed by colony PCR to detect the fusion of *puhA* and *syfp2* using the primers "puhAYFPSeqF" and "puhAYFPSeqF" designed to specifically amplify a genomic fragment. Sequencing of the fragment confirmed that the fusion had been successfully created.

As in Section 3.3.1 the theoretical probability of obtaining a mutant strain, or reversal to the original WT or Δ*crtB* strain was 50%. In reality the actual probability of successful mutation was much lower.

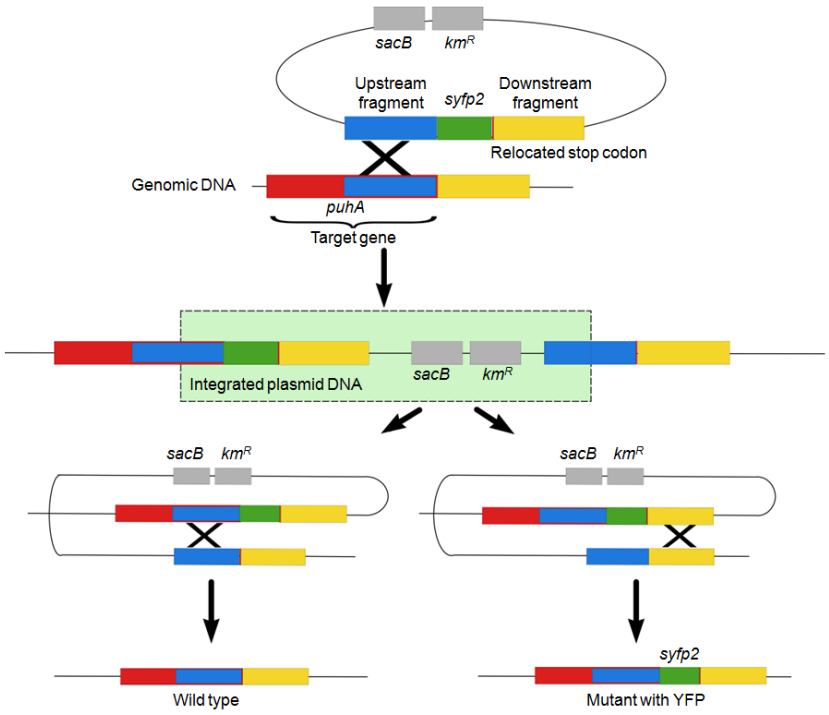


Figure 3.3 Creation of the *Rba. sphaeroides* RCH-YFP fusion strain

Using the suicide vector pK18mob*sacB* the *puhA* gene was fused on the C-terminal end with the fluorescent protein gene, *syfp2*. This was performed by a double homologous recombination event. The first event involves the plasmid containing the upstream and downstream fragments of the target gene integrating into the genome. The second event occurs when the strain is grown in the presence of sucrose causing the plasmid to be excised. This results in two outcomes: either the formation of a mutant or reversion to wild type.

(Modified from Mothersole, 2013)

3.3.3 Biochemical and spectroscopic analysis of the RCH-YFP strains

All strains were grown photosynthetically and membranes were prepared from the WT RCH-YFP and Δ*crtB* RCH-YFP strains as described in Section 2.15. Immunoblotting with antibodies specific to the RC H subunit and, separately, YFP showed the presence of a signal at 54.9 kDa corresponding to the size of a RCH-YFP fusion (Figure 3.4 E). Absorbance spectra of WT RCH-YFP membranes were recorded at room temperature show no isolated YFP peak due to overlap with the 514 nm spheroidenone carotenoid peak (Figure 3.4 B). Room temperature absorbance spectra of Δ*crtB* RCH-YFP membranes show the YFP peak at 517 nm; this peak is shifted to 519 nm at 77 K (Figure 3.4 A, C). Room temperature fluorescence excitation and

emission spectra of Δ*crtB* RCH-YFP whole cells confirm the presence of functional YFP (Figure 3.4 D).

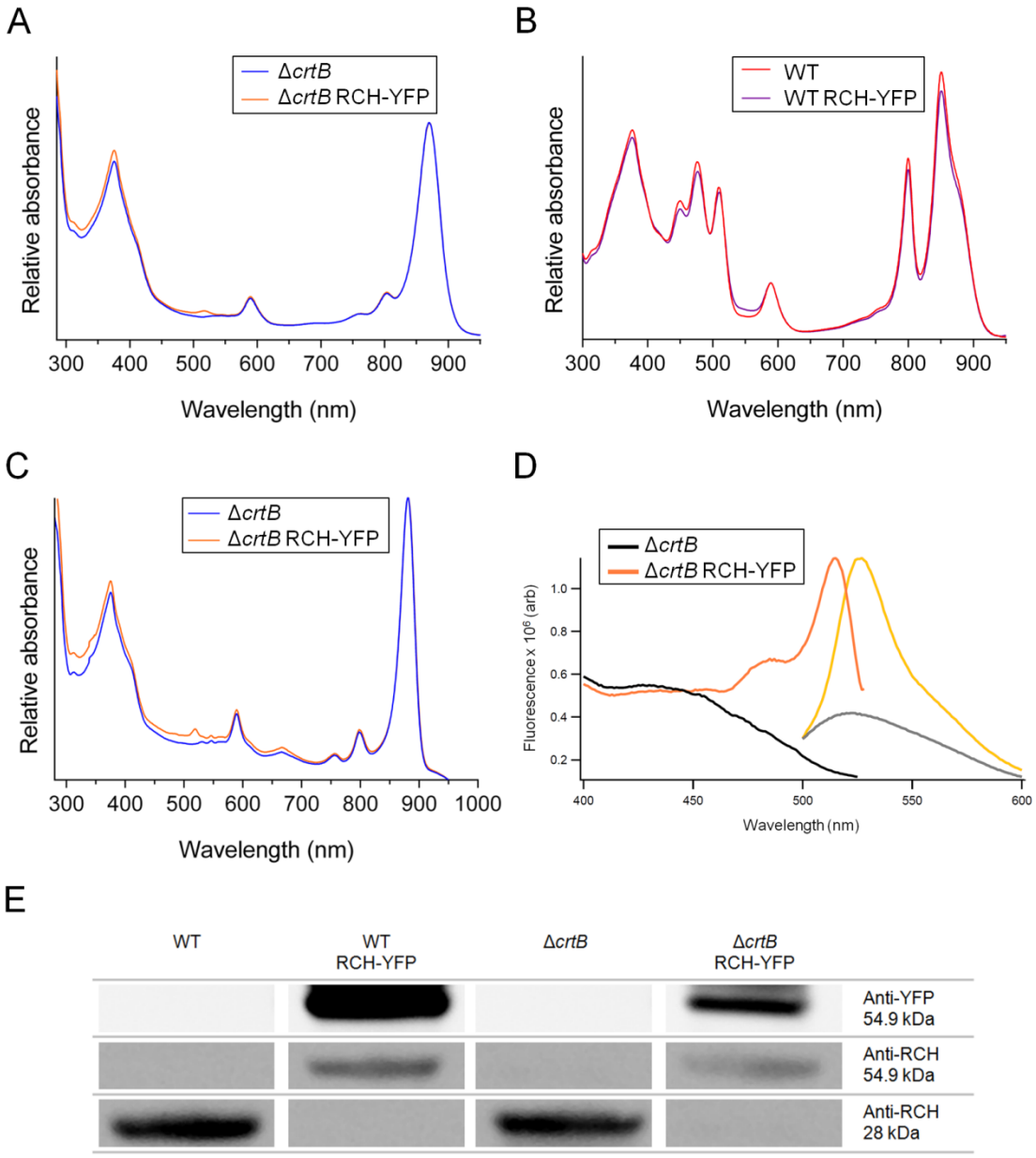


Figure 3.4 Spectroscopic and biochemical analysis of the RCH-YFP fusion

- **A.** Room temperature absorbance spectra of membranes from Δ*crtB* (blue) and Δ*crtB* RCH-YFP (orange) normalised to 870 nm. Δ*crtB* RCH-YFP has a peak at 517 nm corresponding to YFP.
- **B.** Room temperature absorbance spectra of membranes from WT (red) and WT RCH-YFP (purple) normalised to 590 nm.
- **C.** 77 K absorbance spectra of membranes from Δ*crtB* (blue) and Δ*crtB* RCH-YFP (orange) normalised to 880 nm. The region 625 nm – 700 nm is an instrument artefact, spectra in this region should appear as in A.
- **D.** Room temperature fluorescence excitation (ex) and emission (em) spectra of Δ*crtB* (ex: black; em: grey) and Δ*crtB* RCH-YFP (ex: orange; em: light orange) whole cells to check for the presence of functional YFP.
- **E.** Immunoblotting to monitor for the presence of a RCH-YFP fusion.

Absorbance spectra (Figure 3.4 B) of ICM prepared from the WT RCH-YFP strain showed a slight decrease in the LH2 800 and 850 nm absorbance peaks when compared with WT membranes. The Δ*crtB* background produces no LH2; core complex levels remain consistent (Figure 3.4 A and C).

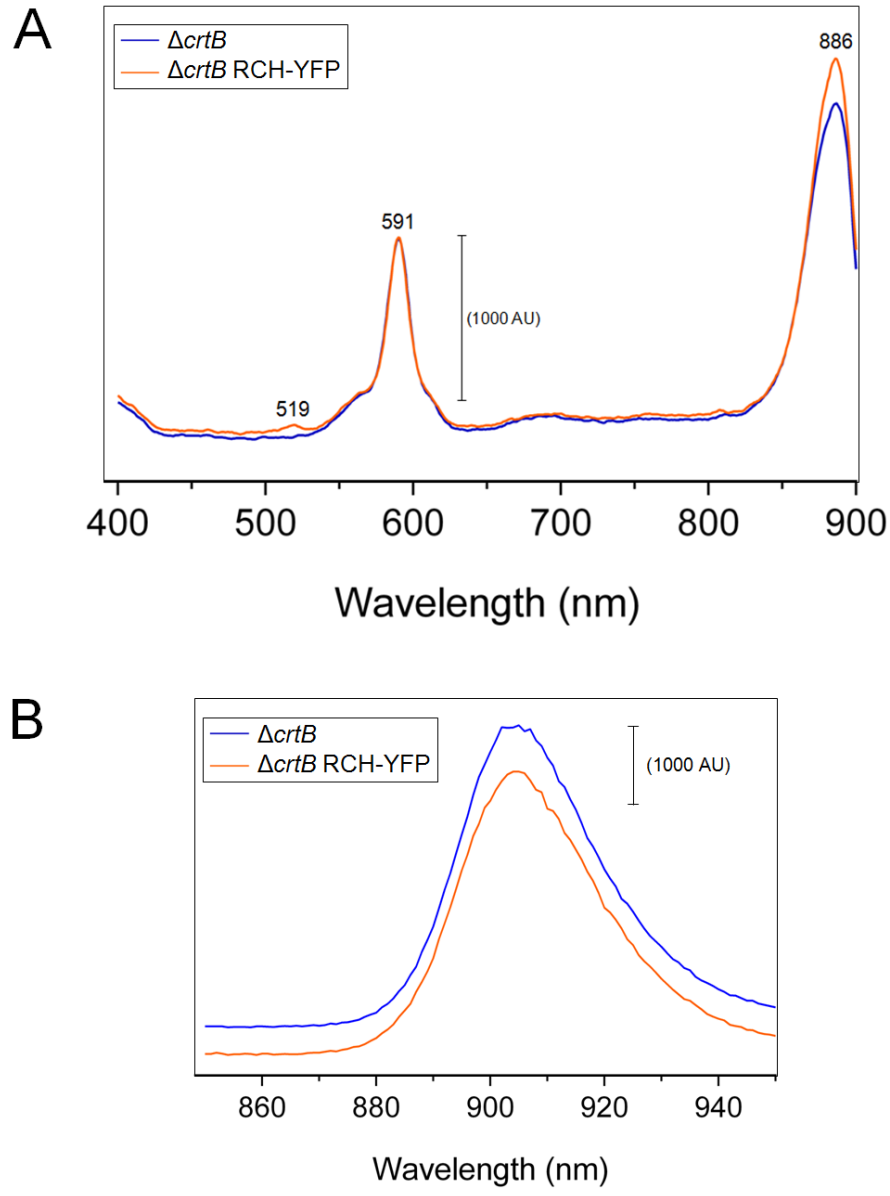


Figure 3.5 Spectroscopic properties of Δ*crtB* RCH-YFP at 77 K

- **A.** Fluorescence excitation spectra. Fluorescence emission was monitored at 910 nm. Excitation and emission slit widths of 5 nm were used. Normalised to 1 AU at 591 nm. Data subject to 2nd order smoothing with 5 neighbours.
- **B.** Fluorescence emission spectra. Samples were normalised to an absorbance of 1.0 at 881 nm and excited at 590 nm. Excitation slit widths of 10 nm and emission slit widths of 5 nm were used. The fluorescence λmax is 905 nm.

All spectra were measured at 77 K and are an average of 10 individual scans.

Fluorescence excitation and emission spectra of ICM prepared from Δ*crtB* RCH-YFP were recorded at 77 K (Figure 3.5 A). Each spectrum presented is an average of 10 individual scans. There is an excitation peak at 519 nm corresponding to YFP. Emission was monitored at 910 nm between excitation wavelengths of 400-910 nm. The BChl *a* Q_x fluorescence peak occurs at 591 nm. The LH1 B875 excitation peak occurs at 886 nm. Compared to the Q_x fluorescence peak at 591 nm, the LH1 B875 peak at 886 nm is higher in Δ*crtB* RCH-YFP compared to Δ*crtB*.

To record emission spectra, the sample was excited at a wavelength of 590 nm and emission recorded between 850 and 950 nm using an average of 10 scans (Figure 3.5 B). The fluorescence λmax was 905 nm in both cases.

High light photosynthetic growth curves of Δ*crtB* RCH-YFP showed an increased growth rate when compared to Δ*crtB* (Figure 3.6). High light photosynthetic growth curves of WT RCH-YFP showed no increase in growth rate when compared to WT (Figure 3.7).

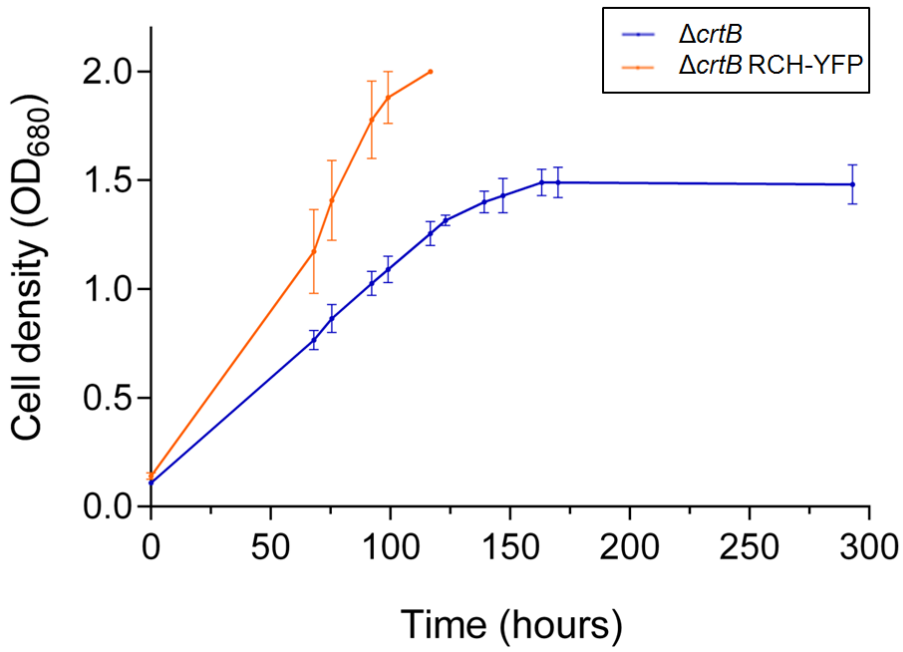


Figure 3.6 Photosynthetic growth of Δ*crtB* RCH-YFP

Photosynthetic growth curves of Δ*crtB* RCH-YFP and Δ*crtB*. Data was obtained in triplicate. Light was provided using Megaman CFL bulbs at an intensity of 100 µmol photons s^{-1} m^{-2}.

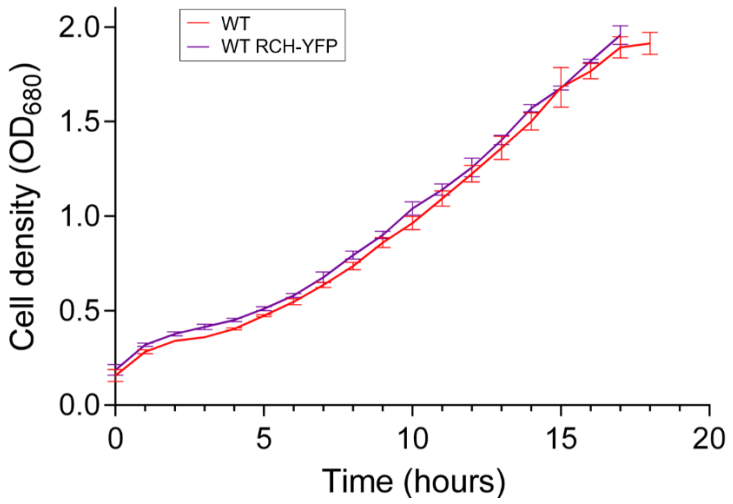


Figure 3.7 Photosynthetic growth of WT RCH-YFP

Photosynthetic growth curves of WT and WT RCH-YFP. Data was obtained in triplicate. Light was provided using Osram Halogen Eco Po bulbs at an intensity of 100 µmol photons s^{-1} m^{-2}.

3.3.4 Analysis of the potential photoprotective effect of YFP

It has been suggested that fluorescent proteins can provide a photoprotective role (Kawaguti, 1944; Salih *et al*., 2000). To investigate how much of a contribution the photoprotective effect has on the increased growth rate seen in Δ*crtB* RCH-YFP, YFP was expressed from the expression plasmid pBBRBB-*Ppuf*$_{843\text{-}1200}$-YFP in Δ*crtB*. High light photosynthetic growth curves show that the expression of YFP from pBBRBB-YFP does not increase the photosynthetic growth rate above that of Δ*crtB* (Figure 3.8).

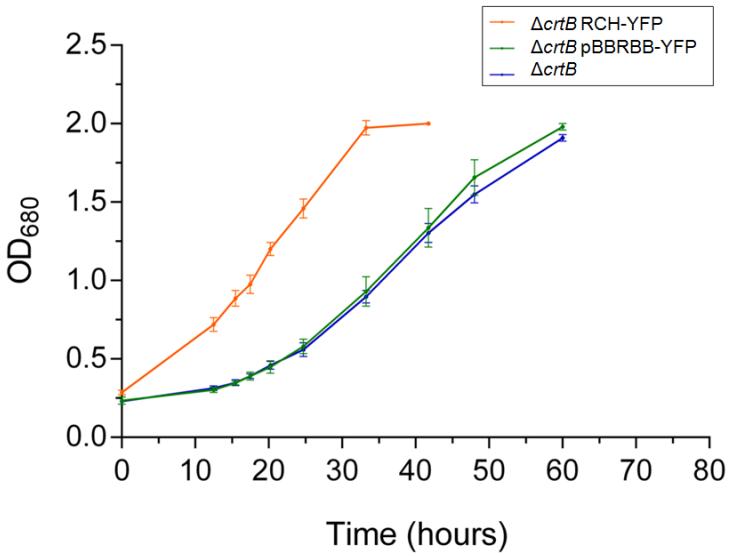


Figure 3.8 Photosynthetic growth curve to investigate the photoprotective effect of YFP

Photosynthetic growth curves of Δ*crtB* pBBRBB-YFP, Δ*crtB* RCH-YFP and Δ*crtB*. Data was obtained in triplicate. Light was provided using Osram Halogen Eco Po bulbs at an intensity of 100 µmol photons s^{-1} m^{-2}.

3.3.5 Fluorescence lifetime measurements of YFP in Δ*crtB* RCH-YFP whole cells

To study the fluorescence lifetime of YFP in Δ*crtB* RCH-YFP and Δ*crtB* pBBRBB-YFP, cells were grown in 10 ml semi-aerobic cultures and fluorescence imaging and lifetime data were obtained by Xia Huang (University of Sheffield). The cells were prepared and data was obtained as described in Section 2.19.6. Images of the cells showed emission from YFP in Δ*crtB* RCH-YFP and Δ*crtB* pBBRBB-YFP but not Δ*crtB*, as expected (Figure 3.9 C).

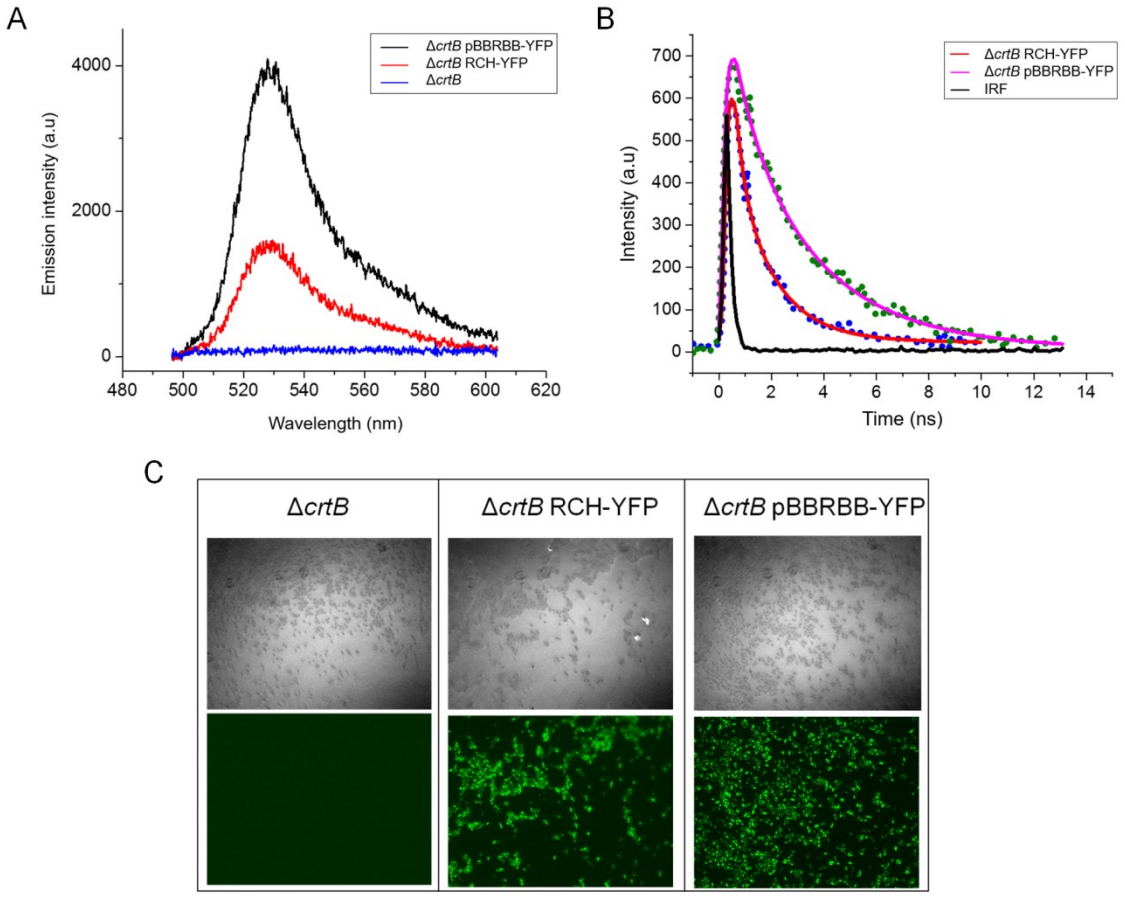


Figure 3.9 YFP lifetimes measured in Δ*crtB* RCH-YFP and Δ*crtB* pBBRBB-YFP whole cells

- **A.** Fluorescence emission spectra of Δ*crtB*, Δ*crtB* RCH-YFP and Δ*crtB* pBBRBB-YFP cells when excited at 495 nm. The spectra are an average of 3 frames with a 1 s exposure time and an electron multiplication gain of 80.
- **B.** Fluorescence lifetime decay curve recorded at a central wavelength of 550 nm. The best fits were achieved using a double-exponent decay function. The measured instrument response (IRF) of the system was approximately 0.18 ns and this was taken in to account during fitting.
- **C.** Images of cells from Δ*crtB*, Δ*crtB* RCH-YFP and Δ*crtB* pBBRBB-YFP with white light (top) and monitoring YFP emission (bottom).

(Data was obtained by Xia Huang.)

Fluorescence spectra of YFP emission were obtained and showed stronger YFP emission from Δ*crtB* pBBRBB-YFP (Figure 3.9 A). The fluorescence lifetimes of YFP in the whole cell samples were measured; decay curves are shown in Figure 3.9 B. The fluorescence decay curves feature a long-lived and a short-lived component, a common feature to fluorescent proteins, each with an associated amplitude contribution (A_1 and A_2 respectively) and a lifetime (τ_1 and τ_2 respectively) (Table 3.1). The Δ*crtB* RCH-YFP cells clearly show a bi-exponential decay (Figure 3.9 B, red trace) in which the lifetimes of both the short- and long-lived components are significantly decreased when compared to Δ*crtB* pBBRBB-YFP.

Sample	A_1	τ_1 [ns]	A_2	τ_2 [ns]	τ_{av} [ns]
Δ*crtB* RCH-YFP	0.49 ± 0.18	1.76 ± 0.28	0.51 ± 0.18	0.57 ± 0.26	1.21
Δ*crtB* pBBRBB-YFP	0.65 ± 0.14	3.15 ± 0.39	0.35 ± 0.14	0.83 ± 0.16	2.32

Table 3.1 YFP lifetimes measured in Δ*crtB* RCH-YFP and Δ*crtB* pBBRBB-YFP whole cells

The fluorescence decay curve shown in Figure 3.8 B was analysed using OriginPro and TRI2 software packages. A1 and A2 are the amplitude contributions of the long- and short-lived components and τ_1 and τ_2 are the lifetimes of each component. The average of these values was taken to give τ_{av}.

Assuming the decrease in YFP fluorescence lifetime in Δ*crtB* RCH-YFP when compared to Δ*crtB* pBBRBB-YFP was due to energy transfer between YFP and the RC-LH1 complex, the average lifetimes (Table 3.1) and Equation 1 were used to calculate an energy transfer yield (ϕ_{EET}) of ~ 47.8 %.

Equation 1. $$\Phi_{\mathrm{EET}} = 1 - \frac{\tau_{YFP-RC}}{\tau_{YFP}}$$

3.3.6 Purification and structural analysis of carotenoidless RCH-YFP-LH1 complexes

Cells were grown photosynthetically as described in Section 2.12.14 to an absorbance of 3 at 680 nm. Intra-cytoplasmic membrane (ICM) preparation was carried out as described in Section 2.15. 224 absorbance units of ICM were solubilised with 3 % β-DDM and the core complexes were fractionated as described in Section 2.15.4. After centrifugation at 27,000 rpm in a SW32 rotor (Beckman Coulter), two bands formed in the gradients (Figure 3.10). The top band consisted of free LH1 subunits and a cytochrome, and the bottom band consisted of monomeric core complexes (Δ*crtB* does not produce dimeric core complexes). The lower band was harvested and loaded on to a DEAE Sepharose column equilibrated with 20 mM

HEPES, 0.03 % β-DDM, pH 7.8 (buffer A) at a flow rate of 0.8 ml min^{-1}. The column was washed with 5 step washes of 50 mM, 100 mM, 150 mM, 200 mM and 250 mM NaCl in 20 mM HEPES, 0.03 % β-DDM, pH 7.8 (buffer B) for 10 minutes each. A gradient of 250 mM NaCl to 400 mM NaCl in buffer B was run over 40 min. The core complex eluted from the column at a concentration of 250 mM NaCl. Fraction sizes of 2 ml were collected and their absorbance spectra recorded to determine their level of purity.

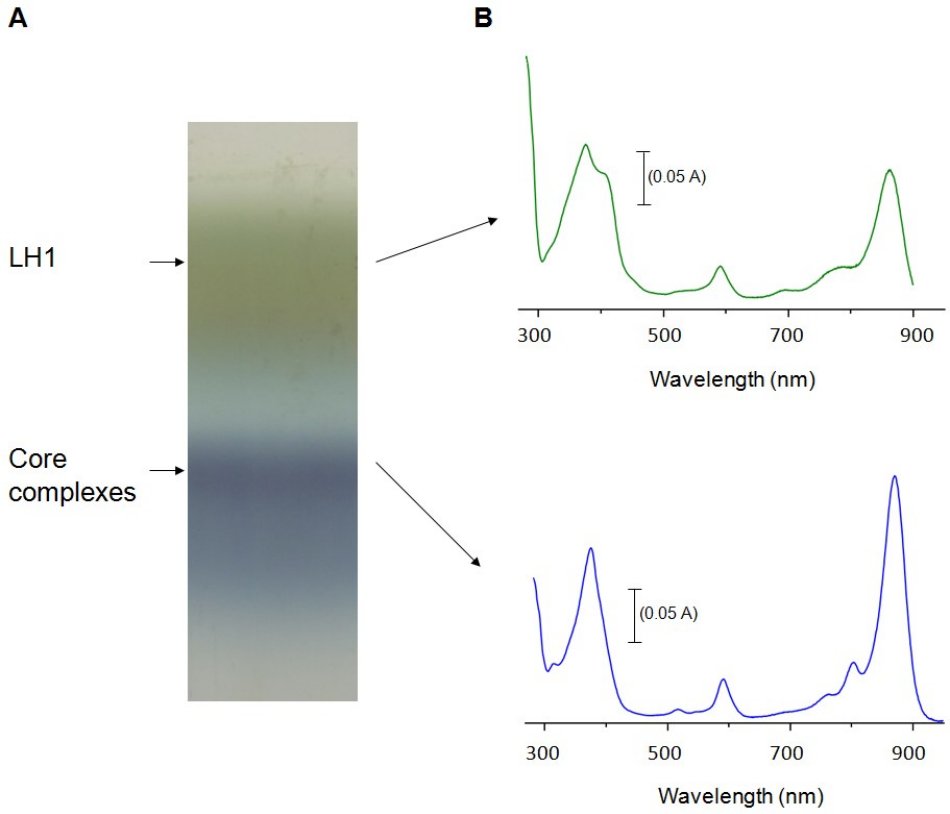


Figure 3.10 Separation of core complexes using discontinuous sucrose gradients

- **A.** Discontinuous sucrose gradient of solubilised Δ*crtB* RCH-YFP ICM
- **B.** Absorbance spectra of the protein harvested from the respective bands in A.

The 875 nm to 280 nm ratio of absorbance of the purified sample provides an indication of the amount of LH1-RC-YFP compared to bulk protein. The absorbance value at 280 nm gives a measure of absolute absorbance by the aromatic residues of all the proteins in the sample. The 875 nm absorbance is specific to LH1 B875 BChl(s). This ratio allows an approximate estimate of the purity of the sample, throughout the various stages of purification. Fractions with an 875:280 nm ratio of 1.7 and above were pooled and concentrated using a Centriprep (Amicon) spin concentrator to a final maximum volume of 1 ml.

The concentrated protein was loaded onto a Superdex 200 (GE Healthcare) gel filtration column equilibrated with 10 mM NaCl in buffer B at a flow rate of 0.4 ml min^{-1}. 1 ml fractions

were collected and analysed spectroscopically to determine purity. Fractions were chosen based on their position within the elution profile, of which those with an 875:280 nm ratio of 1.9 or above were pooled.

An increase in the 875:280 nm ratio from 1 of the starting ICM to 1.92 of the final protein was observed. Absorbance spectra of the pooled fractions from each stage of purification are shown in Figure 3.11 A. SDS-PAGE analysis of the final protein showing the YFP-H, L, and M subunits of the RC, and the α- and β- polypeptides of the LH1 complex is shown in Figure 3.11 B. Single particle electron micrographs of the purified sample were obtained courtesy of Dr. Pu Qian (Figure 3.11 C).

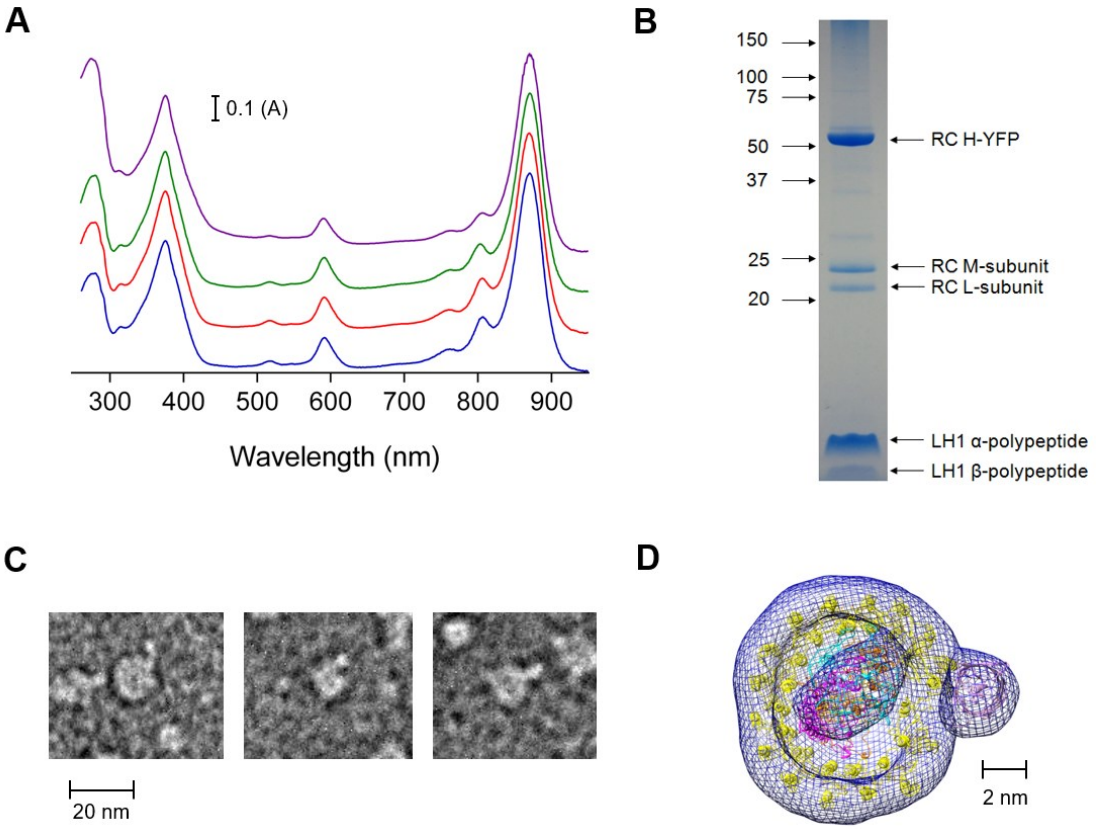


Figure 3.11 Purification of LH1-RC-YFP complexes from *Rba. sphaeroides* Δ*crtB* RCH-YFP

- **A.** Absorbance spectra of stages in purification
  - **Purple** ICM (1)
  - **Green** Core complexes harvested from multistep sucrose gradients (1.63)
  - **Red** Sample after DEAE Sepharose purification (1.73)
  - **Blue** Sample after gel filtration (1.92)

  875:280 nm ratios are indicated in brackets alongside each sample
- **B.** SDS-PAGE analysis of purified sample after gel filtration. Molecular weight in kDa is shown.
- **C.** Electron micrographs of purified sample. YFP can be seen extending outside the LH1 ring.
- **D.** Model of the LH1-RC-YFP complex viewed from the top. Model was made by Dr. Pu Qian and is a low-resolution reconstruction of negatively stained single particles.

A model of the LH1-RC-YFP complex was created by 3D single particle reconstruction of negatively stained electron microscopy of the purified complex at ~25 Å resolution (unpublished data, Dr. Pu Qian) (Figures 3.11 C and D, 3.12). This model is based on an existing structure of the core complex (unpublished data, Dr. Pu Qian). The electron density clearly shows the YFP protruding from the side of the complex, (Figures 3.11 D and 3.12); the distance between the YFP chromophore and the RC special pair BChls was calculated to be ~ 77 Å and to the nearest LH1 B875 BChl was ~ 51 Å.

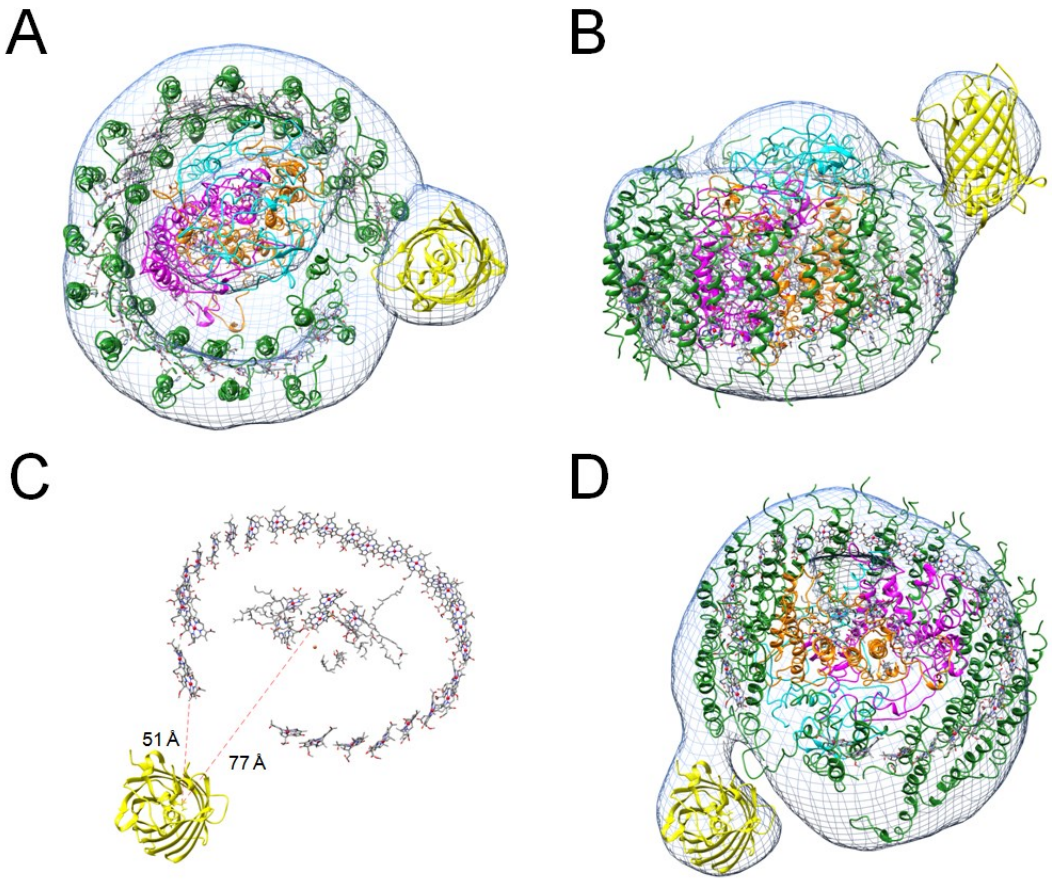


Figure 3.12 Model of LH1-RC-YFP complex from *Rba. sphaeroides* Δ*crtB* RCH-YFP

Model of the LH1-RC-YFP complex created by 3D single particle reconstruction of negatively stained electron microscope images of the purified complex at ~25 Å resolution (unpublished data, Dr. Pu Qian). Showing electron density (blue), LH1 subunits (green), YFP (yellow), RC M (magenta), RC H (cyan) and RC L (orange).

- **A.** Cytoplasmic face.
- **B.** Side view.
- **C.** Periplasmic face showing YFP and pigments of the core complex. Approximate distance between the YFP chromophore and the "special pair" BChls is 77 Å and between YFP chromophore and the nearest B875 BChl is 50.5 Å.
- **D.** Same as C showing electron density and the LH1 and RC subunits.

3.3.7 Purification and spectroscopic analysis of carotenoidless RC-YFP complexes

For experiments to investigate the transfer of energy from YFP to the RC special pair BChls, a purified carotenoidless RC sample without LH1 was required. Reaction centre-YFP complexes were purified from *Rba. sphaeroides* Δ*crtB* RCH-YFP.

Δ*crtB* RCH-YFP cells were grown photosynthetically as specified in Section 2.12.4 to an absorbance of 3 at 680 nm. ICM preparation was carried out as described in Section 2.15. ICM was solublised with 1% LDAO and loaded on to a DEAE Sepharose column equilibrated with 20 mM Tris-HCl, 0.1 % LDAO, pH 7.8 (buffer A) at a flow rate of 1.5 ml min^{-1}. The column was washed with 1 column volume (cv) of buffer A. The column was washed with 2 step washes of 50 mM NaCl for 2 cv and 100 mM NaCl for 5 cv in 20 mM Tris-HCl, 0.1 % LDAO, pH 7.8 (buffer B). RC-YFP was eluted with 200 mM NaCl in buffer B. Fraction sizes of 10 ml were collected and their absorbance spectra recorded to determine their level of purity. Fractions with an 803:280 nm ratio of 0.5 and above were pooled and concentrated using a Centriprep (Amicon) spin concentrator to a final volume of 6 ml maximum.

The concentrated protein was diluted to 50 ml with buffer A to decrease the salt concentration then loaded on to a Q Sepharose (HiTrap QFF) column equilibrated in buffer A at 3 ml min^{-1}. A gradient from 0 mM NaCl to 400 mM NaCl was run over 20 cv and collected in 1.5 ml fractions. RC-YFP was eluted at 300 mM NaCl. Fractions with an 803:280 nm ratio of 0.62 and above were pooled and concentrated using a Centriprep (Amicon) spin concentrator to a final volume of 2 ml maximum.

The concentrated protein was loaded onto a Superdex 200 (GE Healthcare) gel filtration column equilibrated with 10 mM NaCl in buffer B at a flow rate of 0.25 ml min^{-1}. 1 ml fractions were collected and analysed spectroscopically to determine purity. Fractions were chosen based on their position within the elution profile, of which those with an 803:280 nm ratio of 0.73 or above were pooled.

An increase in the 875:280 nm ratio from 0.24 of the starting ICM to 0.73 of the final protein was observed. SDS-PAGE analysis of the pooled fractions from each purification step are shown in Figure 3.13 along with the corresponding absorbance spectra. The YFP-H, L, and M subunits of the RC and the α- and β- polypeptides of the LH1 complex can be seen on the SDS gel and are labelled.

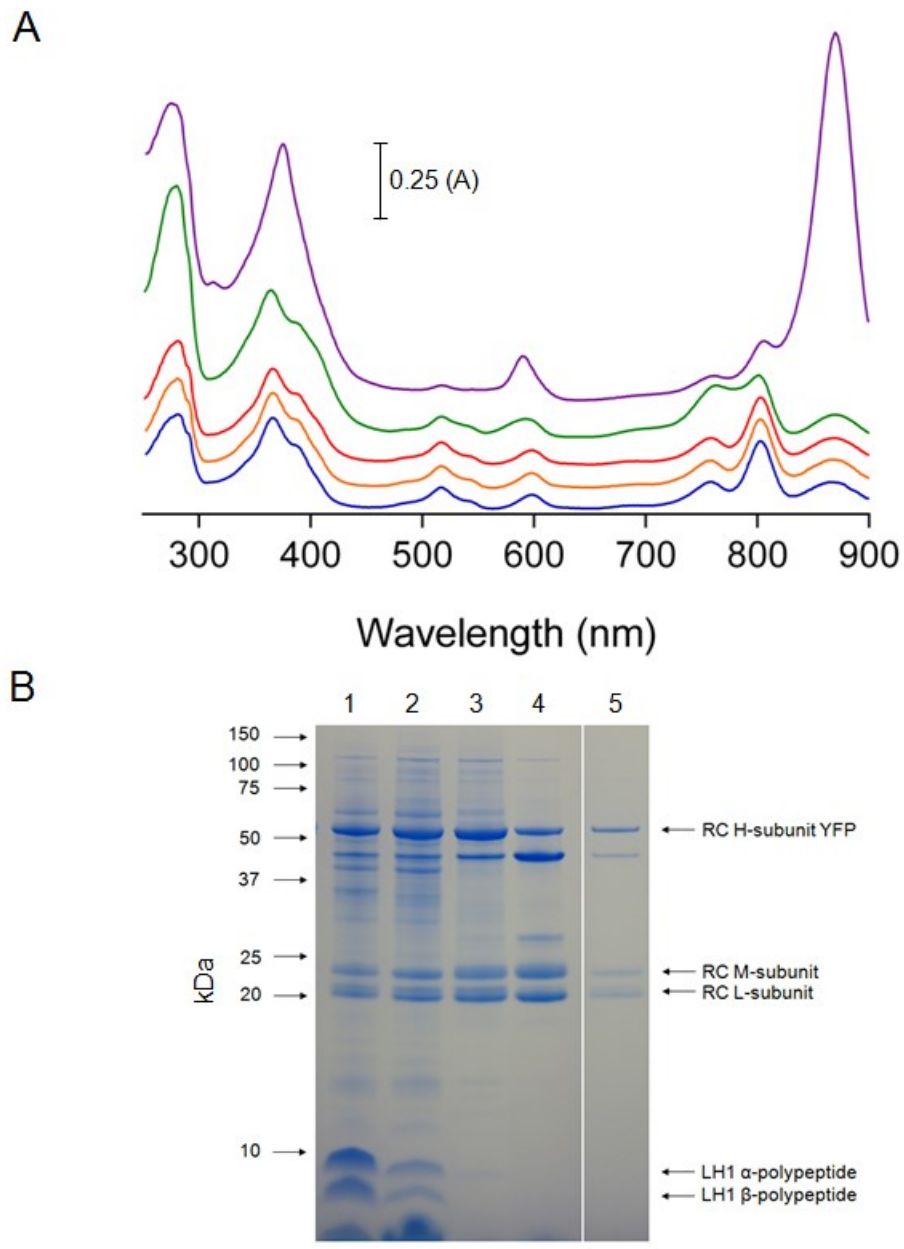


Figure 3.13 Purification of RC-YFP complexes from *Rba. sphaeroides* Δ*crtB* RCH-YFP

A. Absorbance spectra of stages in purification

Purple	ICM	(0.24)
Green	LDAO Solubilised complexes	(0.28)
Red	Sample after DEAE Sepharose purification	(0.57)
Orange	Sample after Q Sepharose purification	(0.64)
Blue	Sample after gel filtration	(0.73)

803:280 nm ratios are indicated in brackets alongside each sample
Data normalised to 803 nm

B. SDS-PAGE analysis of stages in purification

Lane 1 ICM
Lane 2 LDAO solubilised complexes
Lane 3 Sample after DEAE Sepharose purification
Lane 4 Sample after Q Sepharose purification
Lane 5 Sample after gel filtration

Reaction centre complexes were purified from *Rba. sphaeroides* Δ*crtB* to act as a control for further analysis of the RC-YFP complexes purified above. Membranes were prepared, solubilised and two DEAE columns were run as above. The purest fractions were pooled for downstream analysis. Figure 3.14 shows a comparison between the purified RCs from Δ*crtB* and Δ*crtB* RCH-YFP. A model of the purified complex was produced with thanks to Dr. Pu Qian (Figure 3.14). From this model, the distance between the YFP chromophore and the RC special pair BChls could be as short as 40 Å.

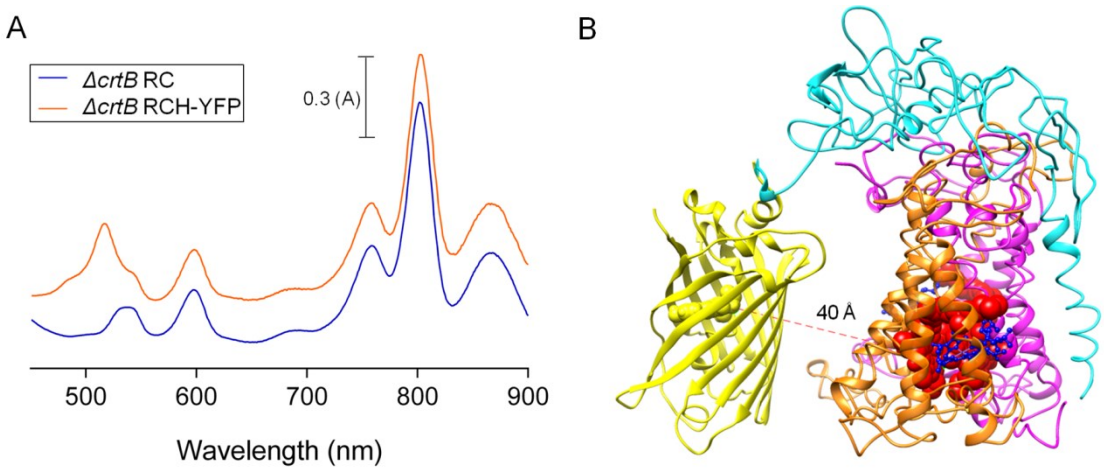


Figure 3.14 Comparison of purified RC from Δ*crtB* and Δ*crtB* RCH-YFP

- **A.** Absorbance spectra of RCs purified from Δ*crtB* and Δ*crtB* RCH-YFP. The peak corresponding to YFP at 517 nm can be seen in the Δ*crtB* RCH-YFP sample. Spectra are normalised to 800 nm.
- **B.** The model was produced in collaboration with Dr. Pu Qian and is based on the LH1-RC-YFP model shown in Figure 3.12. Showing YFP (yellow), RC H (cyan), RC M (magenta), RC L (orange), RC pigments are in blue and the special pair BChls in red. The approximate distance between the YFP chromophore and the RC special pair BChls is shown.

3.3.8 Purification and spectroscopic analysis of YFP from *Rba. sphaeroides* Δ*crtB*

A sample of pure YFP was required to use as a comparison against purified LH1-RC-YFP and RC-YFP.

YFP was expressed from pBBRBB-YFP in *Rba. sphaeroides* Δ*crtB*. Cells were grown semi-aerobically as specified in Section 2.12.3 to an absorbance of 3 at 680 nm. Cells were disrupted by French press as described in Section 2.15.1. The sample was centrifuged at 8200 x *g* (average) for 20 minutes to remove unbroken cells. To remove membranes from the soluble fraction, cells were spun at 125,100 x *g* (average) in a Beckman Type Ti 45 rotor for 2 hours. The supernatant was collected as the soluble fraction, a peak at 515 nm corresponding to YFP was clearly visible (Figure 3.15 A). The sample was loaded on to a DEAE Sepharose

column equilibrated with 20 mM HEPES, pH 7.8 (buffer A) at a flow rate of 0.8 ml min^{-1}. A gradient of 0 mM NaCl to 400 mM NaCl (buffer B) was run over 150 ml. YFP eluted from the column at a concentration of 180 mM NaCl. Fraction sizes of 3 ml were collected and their absorbance spectra recorded to determine their level of purity.

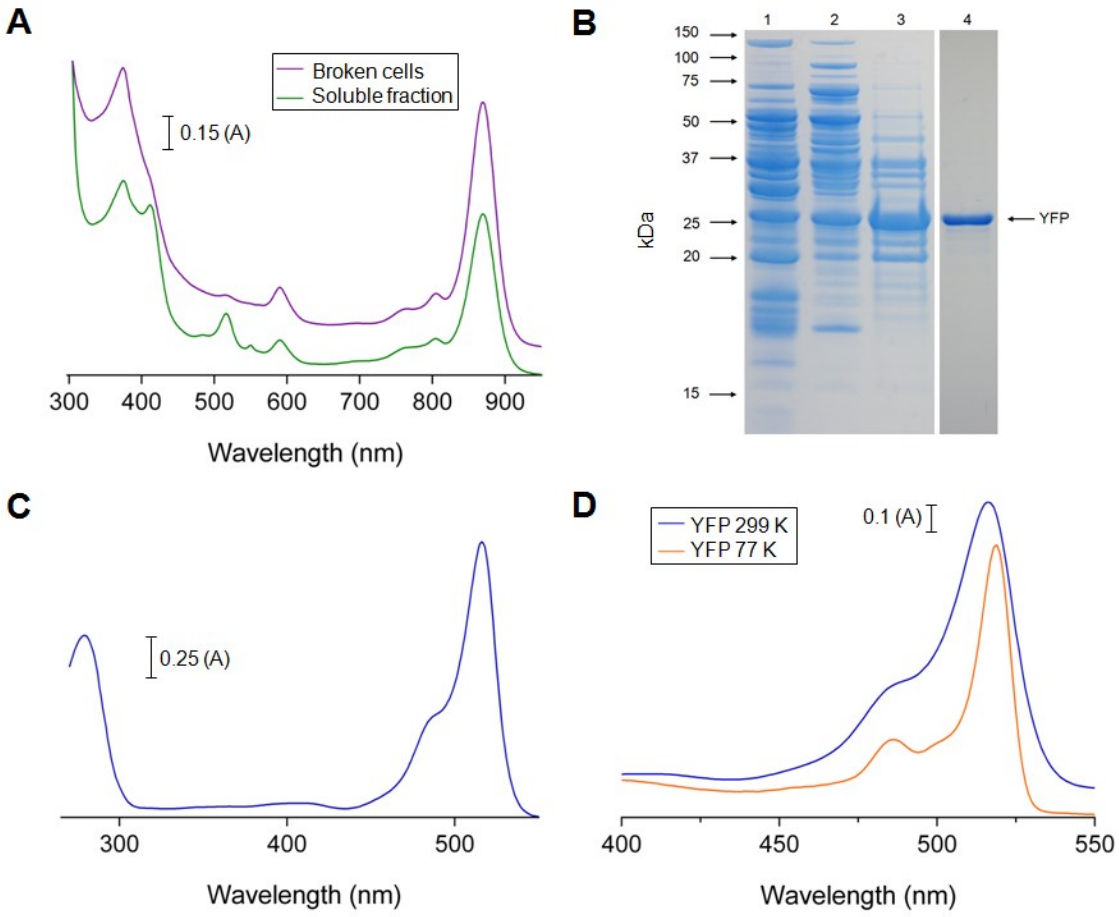


Figure 3.15 Purification of YFP from *Rba. sphaeroides* Δ*crtB* pBBRBB-YFP

- **A.** Absorbance spectra of stages in purification
  - **Purple** Broken cells
  - **Green** Soluble fraction
- **B.** SDS-PAGE analysis of stages in purification
  - **Lane 1** Sample after DEAE Sepharose purification
  - **Lane 2** Sample after Q Sepharose purification
  - **Lane 3** Sample after Gel Filtration purification
  - **Lane 4** Final sample
- **C.** Room temperature absorbance spectrum of purified YFP
- **D.** Room temperature (approximately 299 K) and 77 k absorbance spectra of purified YFP

Yellow fractions were pooled and concentrated using a Centriprep (Amicon) spin concentrator to a final volume of 3 ml maximum. The concentrated protein was diluted to 12 ml with buffer A to decrease the salt concentration then loaded on to a Q Sepharose (HiTrap

QFF) column equilibrated in buffer A 3 ml min^{-1}. A gradient from 0 mM NaCl to 400 mM NaCl was run over 20 cv and collected in 1.5 ml fractions. YFP was eluted at 225 mM NaCl. Fractions were pooled and concentrated using a Centriprep (Amicon) spin concentrator to a final volume of 1 ml maximum.

The concentrated protein was loaded onto a Superdex 200 (GE Healthcare) gel filtration column equilibrated with 1 M NaCl in buffer B at a flow rate of 0.25 ml min^{-1}. 1.5 ml fractions were collected and analysed using SDS PAGE. The most pure fractions were concentrated to 1 ml and were loaded onto a gel filtration column equilibrated with 1 M NaCl in buffer B at a flow rate of 0.25 ml min^{-1}. 0.5 ml fractions were collected and analysed using SDS PAGE. Figure 3.15 shows SDS PAGE and absorbance spectra of various stages in the purification. 77 K absorbance spectra show that the YFP peak maximum shifts from 517 nm at room temperature to 519 nm at 77 K (Figure 3.15 D).

3.3.9 Transient absorbance measurements on RC-YFP complexes

To measure the excitation energy transfer from YFP to the RC special pair BChls (hereafter denoted "P"), ultrafast transient absorbance (TA) measurements were performed on carotenoidless RC-YFP complexes, purified as in Section 3.3.7. The carotenoidless RC complexes purified in Section 3.3.7 were used as the control. The measurements were obtained as described in Section 2.19.7. The data were obtained by Kaitlyn Faries and Dr Christine Kirmaier at the University of Washington in St Louis.

The sample was excited at 515 nm (corresponding to the excitation of YFP) and the change in absorbance was monitored between 830 - 915 nm corresponding to the absorbance of the P BChls. The TA spectra are made up of two components: P-bleaching in the ~ 850 nm region and stimulated emission in the ~ 910 nm region (Figure 3.16 A; see Section 3.4.5). The change in absorbance over time in the 850 nm region was used to produce kinetic fits from which the lifetime of the excited state(s) of RC and RC-YFP could be determined (Figure 3.16 C). These lifetimes are shown in Table 3.2. The state diagrams in Figure 3.16 B show the processes which give rise to the P-bleaching and stimulated emission observed.

Carotenoidless RC complexes show normal charge separation dynamics in which the RC is excited directly at 515 nm (Table 3.2, row 4; Figure 3.16 B, top). Data obtained from the carotenoidless RC-YFP complexes feature the expected charge separation dynamics as for the RC-only sample, but there is an additional component featuring a 1.75 ns increase in bleaching at 850 nm corresponding to the YFP transferring energy to the RC (Table 3.2, row 5; Figure 3.16 B, top and bottom). The excited state of YFP alone was measured as 3.4 ns (Table 3.2, row 3) and the average lifetime of the excited state of RC-YFP is 2.2 ns (Table 3.2, row 5).

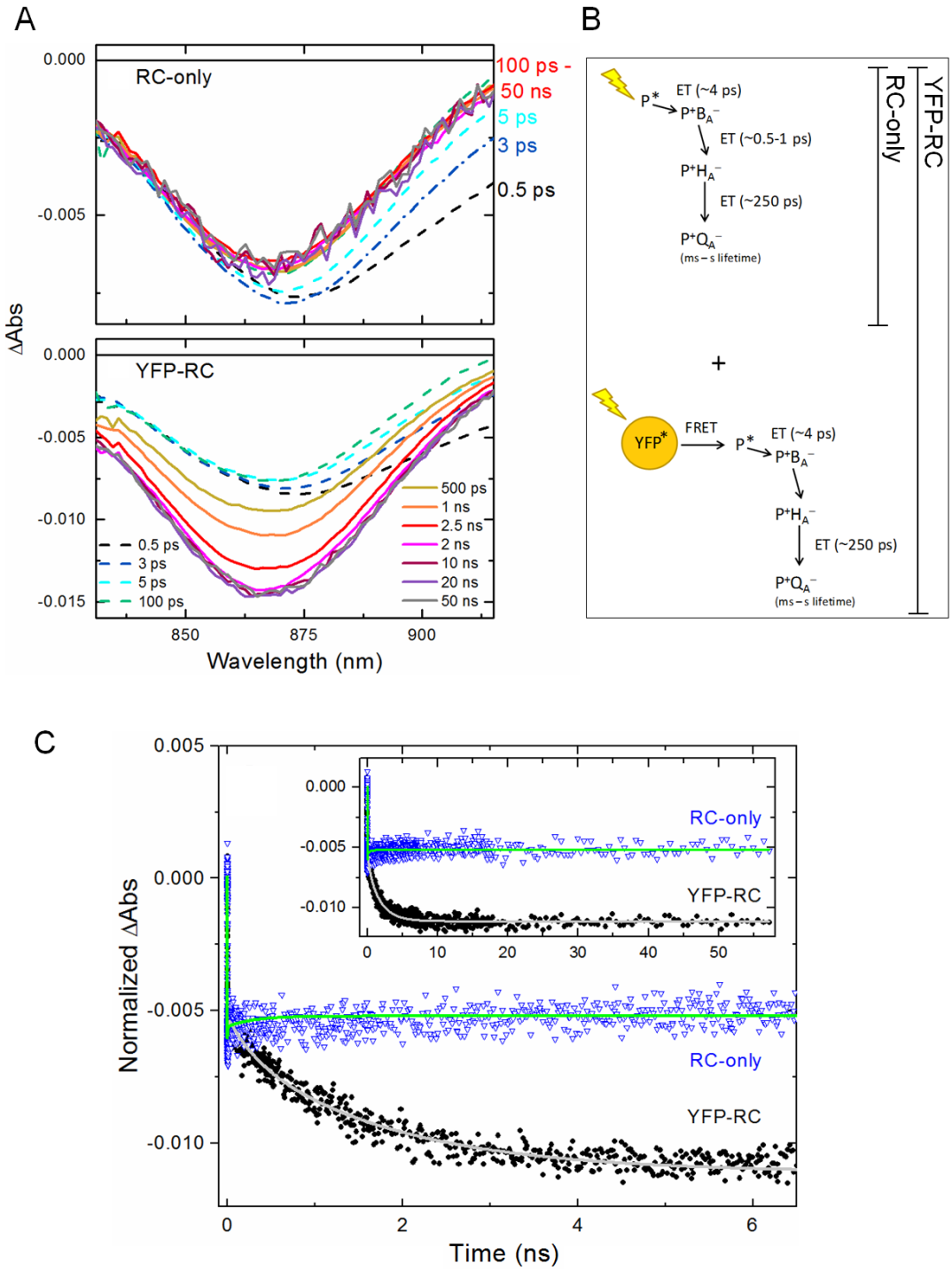


Figure 3.16 Ultrafast transient absorbance measurements on RC-YFP complexes

- **A.** Transient absorption (TA) spectra of P-bleaching (~850 nm region) and stimulated emission (~910 nm region) for the RC-only (top panel) and YFP-RC samples (bottom panel). The dashed lines indicate the P-bleaching and stimulated emission in a portion of RCs in which P* (and subsequently P+) is formed from direct excitation (515 nm) of RC pigments.
- **B.** State diagrams illustrating the process(es) that give rise to the P-bleaching and stimulated emission in the RC-only and RC-YFP samples.
- **C.** Kinetic fits of the 850 nm region of the YFP-RC (black, filled circles) and RC-only (blue, open triangles) samples. The YFP-RC data was fit to three exponentials plus a constant (gray, solid line) and the RC-only data to two exponentials plus a constant (green, solid line). Data were normalized to the absorbance of P_{867}.

(Data obtained and figures produced by Kaitlyn Faries.)

Sample	Transient absorption			
Detection	*560 nm*	*850 nm*	*910 nm*	*Average*
YFP	3.40	n/a	n/a	3.40
RC	n/a	(4.2 ps) (300 ps ^{fixed})	(3.6 ps) (312 ps)	(3.9 ps) (306 ps)
RC-YFP	2.42	1.75 (6.9 ps) (300 ps ^{fixed})	2.10 (3.6 ps) (362 ps)	2.20 (5.3 ps) (331 ps)

Table 3.2 Lifetime (in ns) of the excited state(s) of YFP, RC and RC-YFP complexes as measured by transient absorption

The 850 and 910 nm TA kinetic fits are combined Helios and Eos datasets. Only the Eos was used for the 560 nm wavelength. The RC kinetic data was fit (with the 100-fs excitation pulse) to two exponentials plus a constant; the YFP-RC data was fit (with the pulse) to three exponentials plus a constant. Values in parenthesis with time constants in ps reflect normal charge separation dynamics from P* to BPh_a (~4 ps) and BPh_a to Q_A (~300 ps) in the RC, occurring in a small fraction in which the RC was excited directly and in the major fraction following energy transfer from excited YFP.

Equation 1 was used to calculate the efficiency of energy transfer from YFP to the RC P BChls. Φ_{EET} is the yield of electronic excitation energy transfer, τ is the lifetime of the excited state of the relevant chromophore. The efficiency of energy transfer was determined to be 35 %.

3.4 Discussion

3.4.1 A RCH-YFP fusion been genomically expressed in *Rba. sphaeroides*

A genomic fusion of *puhA* and *syfp2* was successfully created in WT and Δ*crtB* backgrounds. The resulting gene fusion produced a 54.9 kDa RC H subunit-YFP (RCH-YFP) fusion protein (Figure 3.4 E). The YFP is attached to the C terminus of RC H, at the cytoplasmic side of the RC complex.

The YFP is on the opposite side of the complex from the "special pair" BChls P_A and P_B (hereafter denoted "P") and the LH1 B875 BChls. The distance between YFP and P and the B875 BChls affects the efficiency of Forster resonance energy transfer (FRET) as the efficiency of energy transfer is inversely proportional to the sixth power of distance between donor and acceptor (Piston and Kremers, 2007). However, it was unfeasible to place the YFP on the periplasmic side as this would risk interference with or blocking of the cytochrome c_2 docking site.

A model of the LH1-RC-YFP complex showed that the distance between the YFP chromophore and the RC special pair is approximately 77 Å (Figure 3.12 C). The nearest LH1 B875 BChl is approximately 51 Å (Figure 3.12 C). Both values are within FRET distance and that the majority of useful energy transfer from the YFP within native *Rba. sphaeroides* membranes is likely to be to LH1. The region connecting the YFP and RC is flexible, which may allow the YFP to move within closer distance of P and B875 to allow for more efficient energy transfer.

Levels of LH2 and core complexes remained unchanged in RCH-YFP when compared to WT and Δ*crtB*, indicating that YFP does not put significant strain on production of the photosynthetic complexes and does not interfere with complex assembly (Figure 3.4 A - C). YFP is able to fold properly and function as normal, with an absorbance peak visible at 527 nm in Δ*crtB* RCH-YFP ICM (Figure 3.4 A, C), and maintains its typical fluorescence properties (Figure 3.4 D). The peak for YFP is not visible in absorbance spectrum of ICM from WT due to its overlap with the 514 nm spheroidenone carotenoid peak (Figure 3.4 B).

3.4.2 The RCH-YFP fusion results in an increased photosynthetic growth rate in a carotenoidless background

Photosynthetic growth curve analysis of the Δ*crtB* RCH-YFP mutant showed a marked increase in growth rate when compared to Δ*crtB* (Figures 3.6 and 3.8). Under Megaman CFL bulbs, Δ*crtB* RCH-YFP took 100 hours to reach an OD_{680} of 2 while Δ*crtB* which took 175 hours to reach an OD_{680} of 1.5 and could not grow to a higher cell density (Figure 3.6). Δ*crtB* RCH-YFP took 75 hours to reach OD_{680} of 1.5. Under halogen bulbs, Δ*crtB* RCH-YFP took approximately 30 hours to reach an OD_{680} of 1.75 while Δ*crtB* which took 60 hours to reach an OD_{680} of 1.75 (Figure 3.8).

The photosynthetic growth rate of Δ*crtB* RCH-YFP is slower than that of WT which typically takes 17 hours to reach an OD_{680} of 2 (Figure 3.7). The photosynthetic growth rate of Δ*crtB* is much slower than that of WT. WT RCH-YFP did not grow faster than WT. This indicates that any energy transfer from YFP to LH1 or RC is negligible in terms of the total energy yield of WT. It is possible that *Rba. sphaeroides* works at maximum photosynthetic capacity under high light conditions and that processes downstream of light harvesting are a limiting factor to photosynthetic growth rate. The addition of a chromophore whose absorbance wavelengths fill a gap in the absorbance spectrum of WT *Rba. sphaeroides* could provide further insight.

These results indicate that the fusion of YFP to the RC in a Δ*crtB* background causes increased photosynthetic function and suggest that YFP can transfer useful energy to the photosynthesis pathway, most likely to either LH1 or the RC.

3.4.3 YFP does not increase the photosynthetic growth rate through a photoprotective mechanism

It has been previously reported that fluorescent proteins may have a photoprotective function (Kawaguti, 1944; Salih *et al*., 2000). It is known that the carotenoids provide a photoprotective role in *Rba. sphaeroides* and that carotenoidless mutants are susceptible to photo oxidative damage, participating in the decreased photosynthetic ability of *crtB* mutants (Clayton and Smith, 1960; Cogdell and Frank, 1987). To rule out the possibility that the increase in photosynthetic growth rate was purely due to a photoprotective effect, YFP was expressed from the constitutive expression plasmid pBBRBB in Δ*crtB*. The YFP expressed from pBBRBB was free and unattached to the RC; it would most likely not associate with the photosynthetic membrane and thus could not play a role in light harvesting. The photosynthetic growth rate was compared to that of Δ*crtB* and Δ*crtB* RCH-YFP (Figure 3.8). Δ*crtB* pBBRBB-YFP did not grow significantly faster than Δ*crtB*. This indicates that any photoprotective role YFP plays is not significant, and that the majority of the effect of increased photosynthetic growth rate is due to other mechanisms.

3.4.4 YFP transfers energy to LH1 and/or the RC in whole cells and native membranes

77 K fluorescence spectra of ICM from Δ*crtB* RCH-YFP and Δ*crtB* were measured. Excitation of the BChl Q_x transition at 590 nm produces an emission maximum at 905 nm in both cases (Figures 3.5 B). This result is comparable to previous studies on the carotenoidless mutant, R-26 (Ng *et al*., 2011). Fluorescence excitation spectra of Δ*crtB* RCH-YFP showed an excitation maximum at 519 nm corresponding to YFP (Figure 3.5 A). This indicates that YFP contributes to emission at 910 nm and therefore can transfer energy to LH1.

The fluorescence lifetime of YFP was measured in Δ*crtB* RCH-YFP whole cells and was found to be 1.11 ns shorter than the lifetime in Δ*crtB* pBBRBB-YFP (free YFP) (Figure 3.9, Table 3.1). This indicates that YFP can transfer energy to the RC-LH1 complex via a Förster resonance energy transfer (FRET) mechanism, leading to the quenching of the YFP chromophore and thus a shorter lifetime. The efficiency of energy transfer was calculated to be ~ 47.8 %. These results account for the increased photosynthetic growth rate of Δ*crtB* RCH-YFP compared to Δ*crtB* and Δ*crtB* pBBRBB-YFP (Figures 3.6 and 3.8).

3.4.5 YFP transfers energy to the RC in purified RC-YFP complexes

Carotenoidless RC-YFP, RC alone, and YFP alone were purified from *Rba. sphaeroides* and subject to fluorescence lifetime and kinetic measurements to investigate the transfer of energy from YFP to the RC special pair BChls, denoted "P". A model of the RC-YFP complex calculated the distance from the YFP chromophore to P to be ~ 40 Å, within FRET distance.

Ultrafast transient absorbance (TA) measurements on purified carotenoidless RC-YFP complexes probed the formation and decay of the excited special pair BChls, P*. The TA spectra feature two components: P-bleaching at ~ 850 nm and stimulated emission at ~ 910 nm. As the ground state (P) is depleted the excited state is populated (P*); when an incoming photon interacts with an excited molecule, instead of being absorbed it generates another photon through stimulated emission. At a population of 50% ground state and 50% excited state equilibrium is reached in which the amount of absorbed light is equal to the amount of light being produced by stimulated emission. As the ground state is depleted and the excited state is populated the detected amount of absorbed light decreases. In these circumstances it appears that the medium does not absorb light, or absorbs less light, and this feature is termed "bleaching". It is for this reason that the change in absorbance observed during the experiment is negative (Figure 3.16 A and C).

RC-YFP features a ~ 4 ps bleach at 850 nm, which is consistent with that that seen for the RC alone sample (Table 3.2, Figure 3.16). This bleach corresponds to the ground state P dimer absorbing light at ~ 865 nm, forming P* and then bleaching as an electron is donated to the bacteriopheophytin H_A, forming P^+ (Figure 3.16 B). In RC-alone this bleaching feature decays as P^+ returns to the ground state, P (Figure 3.16 A and B). The measurements for the RC-YFP complexes feature a ~ 1.75 ns additional component to the 850 nm bleach not seen in RC-only (Table 3.2). The 850 nm bleaching feature increases as more P* is formed as the YFP transfers energy to the RC special pair BChls (Figure 3.16).

TA data from the purified YFP sample give a fluorescence lifetime of 3.4 ns, whereas the average lifetime of the excited state of RC-YFP is 2.2 ns (Table 3.2). This reflects the quenching of the YFP chromophore and thus a shorter lifetime in the RC-YFP sample.

The efficiency of energy transfer from YFP to the RC in purified RC-YFP complexes was calculated to be ~ 35 %. Additional stroboscopic and time-correlated single photon counting (TCSPC) florescence decay data obtained by Kaitlyn Faries revised this value to 40 $\pm$ 10%. This value is comparable to that determined from whole cells, ~ 47.8 % (Section 3.3.5).

3.4.6 Additional work

This work demonstrates that YFP has a positive effect on photosynthetic growth rate in a carotenoidless mutant of *Rba. sphaeroides*. This result paves the way for the use of chromophores that absorb light of wavelengths that *Rba. sphaeroides* WT does not absorb, between 640 nm and 750 nm. Possible candidates are the fluorescent proteins mApple, mCherry and mKate2 (Figure 1.17) (Shaner *et al*., 2004; Shaner *et al*., 2008; Shcherbo *et al*., 2009).

3.4.7 Conclusion

A genomic fusion of *puhA*, encoding the RC H subunit and *syfp2*, encoding the YFP variant SYFP2, was successfully created in WT and Δ*crtB Rba. sphaeroides* backgrounds to produce a RCH-YFP fusion protein. Photosynthetic growth analysis, fluorescence spectroscopy, fluorescence lifetime studies, and ultrafast TA measurements have indicated that YFP can pass energy to the *Rba. sphaeroides* LH1 and RC complexes and increase photosynthetic growth rate in a carotenoidless background. This opens up new possibilities for the creation of new photosynthetic pathways by expanding the range of light absorbers for the augmentation of photosynthesis.

Chapter 4

The assembly of proteorhodopsin and the biosynthesis of its chromophore retinal in *Rhodobacter sphaeroides*

4.1 Summary

Proteorhodopsin (PR) is a light-driven proton pump found in marine bacterioplankton. PR has potential as an additional component for insertion into the photosynthetic pathways of *Rba. sphaeroides* to convert light into a proton gradient to drive downstream metabolism. This chapter details the expression and assembly of PR in the photosynthetic membrane of *Rba. sphaeroides*.

PR requires the pigment all-*trans* retinal for complete folding and full functionality. In this study, the carotenoid biosynthesis pathway of *Rba. sphaeroides* was engineered through the introduction of three foreign genes for the biosynthesis of retinal. Small amounts of retinal were produced in *Rba. sphaeroides* and several strategies were used to attempt to increase the retinal yield. Data presented in this chapter indicate that sufficient retinal is produced for the assembly of a small amount of PR, although higher levels of expression are achieved through the addition of retinal to the growth medium. PR was expressed in the cyanobacterium *Synechocystis*, which is naturally capable of retinal biosynthesis. Only a very small amount of PR could be purified from membranes obtained from this strain, and it is likely that natural retinal biosynthesis capability of *Synechocystis* is too low for the complete assembly of useful amounts of PR.

Retinal-bound PR is able to assemble in the membrane of *Rba. sphaeroides*, and work done in this study demonstrates that PR prefers to sit in a more planar membrane environment. However, further work is required to improve expression levels of PR and the biosynthesis of retinal. Ultimately, PR could be used to augment the proton motive force that drives downstream metabolism.

4.2 Introduction

The assembly of artificial photosynthetic pathways through the use of diverse pigments and proteins to expand spectral coverage is an emerging field (see Sections 1.8 and 3.2). A pilot study to explore the challenge of creating new light-harvesting pathways in *Rba. sphaeroides* is detailed in Chapter 3; the yellow fluorescent protein (YFP) can pass energy to the reaction centre in amounts sufficient to increase the photosynthetic growth rate in a carotenoidless background. The work presented in this chapter explores the potential to assemble another foreign light-harvesting protein, proteorhodopsin, in the *Rba. sphaeroides* photosynthetic membrane.

Proteorhodopsin (PR) is a light-driven proton pump found in marine bacterioplankton. PR is a transmembrane protein and in its functional form it is complexed with an all-*trans* retinal pigment molecule (for further reading, see Section 1.9). Due to its simplicity PR has potential as an additional component for insertion into existing photosynthetic pathways to convert light into a proton gradient to drive downstream metabolism.

Rba. sphaeroides has two light-harvesting proteins, LH1 and LH2, which transfer electronic excitation energy to the reaction centre where charge separation takes place. This initiates electron transfer which ultimately leads to pumping of proteins across the photosynthetic membrane by the cytochrome bc_1 complex. The resulting proton gradient is used to drive the ATP synthase to produce ATP. While the harvesting of light to form a transmembrane proton gradient in *Rba. sphaeroides* is highly productive and efficient, organisms containing PR are able to harvest light and produce a proton gradient using this single protein.

There has been debate as to whether PR alone is sufficient to enable cells to survive photosynthetically (see Section 1.9.2) however studies expressing PR in *E. coli* have shown that PR is capable of having a significant impact on cellular energy production. PR has been found to embed in the *E. coli* cell membrane and function as a proton pump to drive ATP synthesis (Beja *et al*., 2000; Martinez *et al*., 2007). PR has also been used in *E. coli* as the sole cellular energy source for flagellar rotation (Walter *et al*., 2007). PR therefore has the potential to augment photosynthesis by contributing to the production of the transmembrane proton gradient along with the native cytochrome bc_1 complex.

Functional PR requires the pigment all-*trans* retinal. Retinal is bound to PR via a lysine residue to form a protonated Schiff base. On excitation with light retinal undergoes an isomerisation from all-*trans* to 13-*cis* which triggers conformational changes in PR leading to proton transfer across the protein (see Section 1.9.3). Previous studies of the expression of PR in *E.*

coli have involved the supplementation of the growth medium with retinal or the expression of genes necessary for the biosynthesis of retinal, and its precursor beta-carotene (Beja *et al*., 2000; Martinez *et al*., 2007). There has been interest in producing retinal and other retinoids in *E. coli* as they are an important active ingredient in many cosmetics and pharmaceuticals (Jang *et al*., 2011).

Previous studies have made modifications to the *Rba. sphaeroides* carotenoid biosynthesis pathway to produce foreign pigments. The *crtI*, *crtY*, *crtB* and *crtZ* genes from *Pantoea (P.) agglomerans* (previously known as *Erwinia herbicola*) have been expressed from a plasmid to produce lycopene, beta-carotene, beta-cryptoxanthin, zeaxanthin and zeaxanthin glucoside. Some of these carotenoids were able to be incorporated into the photosynthetic complexes and are able to transfer energy to the BChl molecules within the light harvesting complexes (Hunter *et al*., 1994). The biosynthesis of 2,2' diketospirilloxanthin in *Rba. sphaeroides* was achieved with the expression of the four-step phytoene desaturase *crtI* gene from *P. agglomerans* (Garcia-Asua *et al*., 2002; Chi *et al*., 2015). When the native *crtC* gene was deleted in this strain the spirilloxanthin pathway was truncated resulting in the accumulation of lycopene.

Retinal can be produced from beta-carotene by the enzyme beta-carotene-15',15' mono(di)oxygenase (BCM(D)O). The gene encoding this protein is found in many species including mouse, fruit fly, and humans and is involved in vitamin A biosynthesis. Several studies have used various BCM(D)O genes to produce retinal in *E. coli* (von Lintig and Vogt, 2000; Kim *et al*., 2008; Jang *et al*., 2011). The *blh* gene used in this study is from the uncultured marine bacterium 66A03 and produces BCDO (Sabehi *et al*., 2005; Kim *et al*., 2009). A codon optimised version of *blh* has previously been used to produce retinal in *E. coli* and was found to have high beta-carotene cleavage activity (Jang *et al*., 2011).

In a recent study by Tikh *et al* PR was produced in *Rba. sphaeroides* from an expression vector, and the growth medium was supplemented with retinal. Just under 1 mg ml^{-1} of retinal-bound PR was purified from wild type (WT) *Rba. sphaeroides* whole cell lysate, and 3 mg ml^{-1} from a *Rba. sphaeroides* mutant lacking the reaction centre (Tikh *et al*., 2014). The aim of the work presented in this chapter was to produce a *Rba. sphaeroides* strain with the genes for PR and the biosynthesis of all-*trans* retinal incorporated into the genome to determine whether functional PR is able to assemble in the *Rba. sphaeroides* photosynthetic membrane.

In this chapter the native carotenoid biosynthesis pathway of *Rba. sphaeroides* was modified through the addition of three foreign genes: *crtI* and *crtY* from *P. agglomerans* and *blh* from

uncultured marine bacterium 66A03 to produce all-*trans* retinal (Figure 4.1). The pigment content of this mutant strain was characterised and different strategies were attempted with the aim to improve the retinal yield. The gene for PR was expressed in *Rba. sphaeroides* and membranes were analysed to detect the presence of functional PR. The effect of membrane curvature on PR incorporation was studied. This work forms the basis of future studies involving the creation of artificial light-harvesting antenna *in vivo* using non-native proteins and pigments for increased spectral coverage and augmentation of photosynthesis.

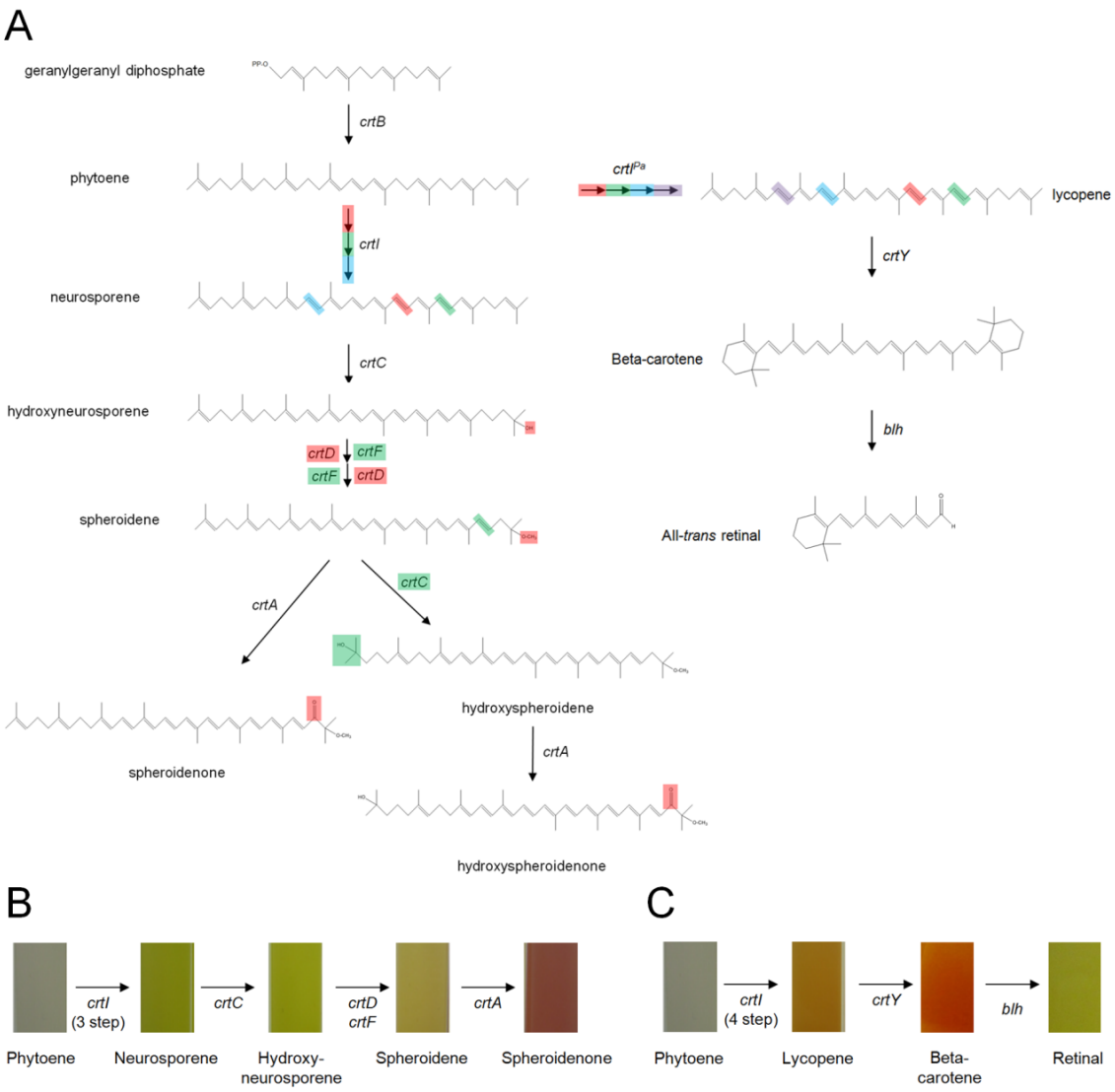


Figure 4.1 Carotenoid biosynthesis pathways of *Rba. sphaeroides*

- **A.** Overview of the carotenoid biosynthesis pathways of *Rba. sphaeroides* and the modifications required to make retinal.
- **B.** Overview of the spheroidenone biosynthesis pathway consisting of photographs of cell cultures.
- **C.** Overview of the retinal biosynthesis pathway consisting of photographs of cell cultures.

(Figure modified from an earlier version by Dr D. Mothersole.)

4.3 Results

4.3.1 Production of retinal in *Rba. sphaeroides*

Proteorhodopsin requires the pigment all-*trans* retinal to fully assemble. *Rba. sphaeroides* does not naturally produce retinal so three foreign genes were used to modify the existing carotenoid biosynthesis pathway (Figures 4.1 and 4.2 A). The genes were *crtI* and *crtY* from *P. agglomerans* and *blh* from uncultured marine bacterium 66A03. *blh* encodes bacteriorhodopsin-related protein-like homolog protein, a beta-carotene 15,15'-dioxygenase (BCDO) and produces retinal from beta-carotene (Kim *et al*., 2009). The *blh* gene encoding BCDO was chosen as it cleaves beta-carotene to produce retinal with a high efficiency (Jang *et al*., 2011).

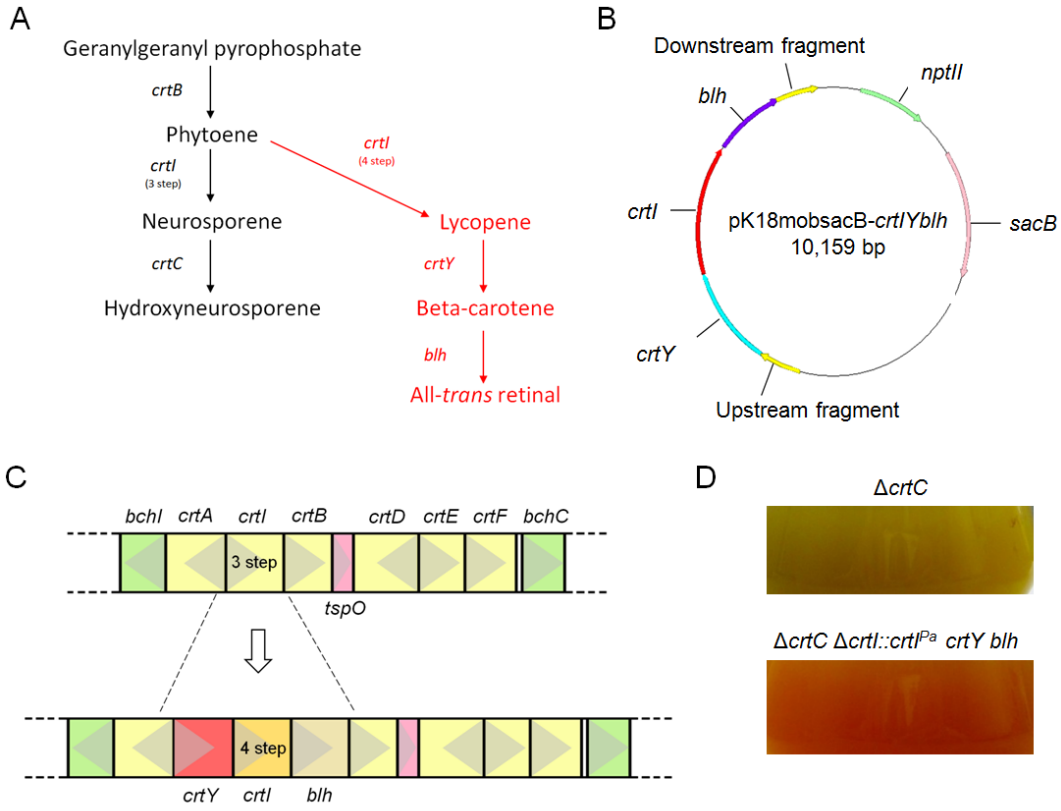


Figure 4.2 Construction of *Rba. sphaeroides crtIYblh*

- **A.** Part of the native *Rba. sphaeroides* carotenoid biosynthesis pathway is shown in black, with the genes that encode the enzymes performing each step indicated. The three new genes that have been introduced to produce retinal are shown in red.
- **B.** The pK18mob*sacB* construct created to replace the native 3-step phytoene desaturase *crtI* gene with the 4 step *crtI* and *crtY* from *P. agglomerans* and *blh* from uncultured marine bacterium 66A03.
- **C.** The *Rba. sphaeroides* carotenoid gene cluster before and after the gene substitutions. These genes were introduced into a *crtC* knockout strain to direct pigment biosynthesis down the new pathway. The grey arrows indicate the direction of transcription.
- **D.** Semi-aerobically grown cultures of Δ*crtC* (top) and *crtIYblh* (bottom) strains. The orange colour is due to the presence of beta-carotene.

The *blh* gene was codon optimised for expression in *Rba. sphaeroides*. A DNA fragment was synthesised containing: the first 348 bp of the native *crtA* gene plus the next 146 bp, an *XbaI* site and a *SalI* site, a predicted ribosome binding site for the native *crtI* gene, the codon optimised *blh*, 17 bp upstream of the native *crtB* gene and 483 bp of the native *crtB* gene. The fragment was flanked by the restriction sites *Hin*dIII and *Eco*RI. The fragment was cut with *Hin*dIII and *Eco*RI and ligated into the plasmid pK18mob*sacB*. *CrtI* and *crtY* were amplified from the *Rba. sphaeroides* strain DM21 containing the plasmid pERWIY (unpublished data) using the primers "crtYI F" and "crtYI R". The *crtI* and *crtY* fragment was digested with *Xba*I and *Sal*I and ligated into pK18mob*sacB-blh* to create the construct pK18mobsacB-*crtYIblh* (Figure 4.2 B).

The *crtI*, *crtY* and *blh* genes were integrated into the *Rba. sphaeroides* genome using the pK18mob*sacB* system of homologous recombination, as described previously in Sections 2.13-14 and 3.3.2. The genes were integrated into a *crtC* deletion mutant of *Rba. sphaeroides* to direct pigment production down the new pathway. The genes were integrated into the carotenoid biosynthesis gene cluster between *crtA* and *crtB*, replacing the native *crtI* (Figure 4.2 C). The resulting strain, *Rba. sphaeroides* Δ*crtC* Δ*crtI*::$crtI^{Pa}$ $crtY^{Pa}$ *blh*, hereafter termed *crtIYblh*, is orange in colour (Figure 4.1 D)

4.3.2 Extraction and HPLC analysis of the pigments produced by *crtIYblh*

The cellular pigments produced by *Rba. sphaeroides crtIYblh* were extracted with acetone:methanol (7:1 v/v) according to the method described in Section 2.18.1 and analysed using HPLC. A Phenomenex C18 column was used to separate the extracted pigments using a method described in Section 2.18.3. The DAD absorbance spectra were recorded and elution was monitored at 442 nm, corresponding to the absorbance of carotenoids (Figure 4.3). Cell pigment extracts from both *ΔcrtC* and *crtIYblh* strains contain a peak at 13.7 minutes corresponding to neurosporene, which is produced by the *Rba. sphaeroides* 3-step phytoene desaturase (*crtI*) and as an intermediate by the *P. agglomerans* 4-step phytoene desaturase. *crtIYblh* contains an additional peak at 15.8 minutes which corresponds to beta-carotene. No lycopene is observed in *crtIYblh* possibly indicating that the enzyme encoded by *crtY*, lycopene cyclase, (lycopene to beta-carotene) is very efficient. No retinal was observed in *crtIYblh*.

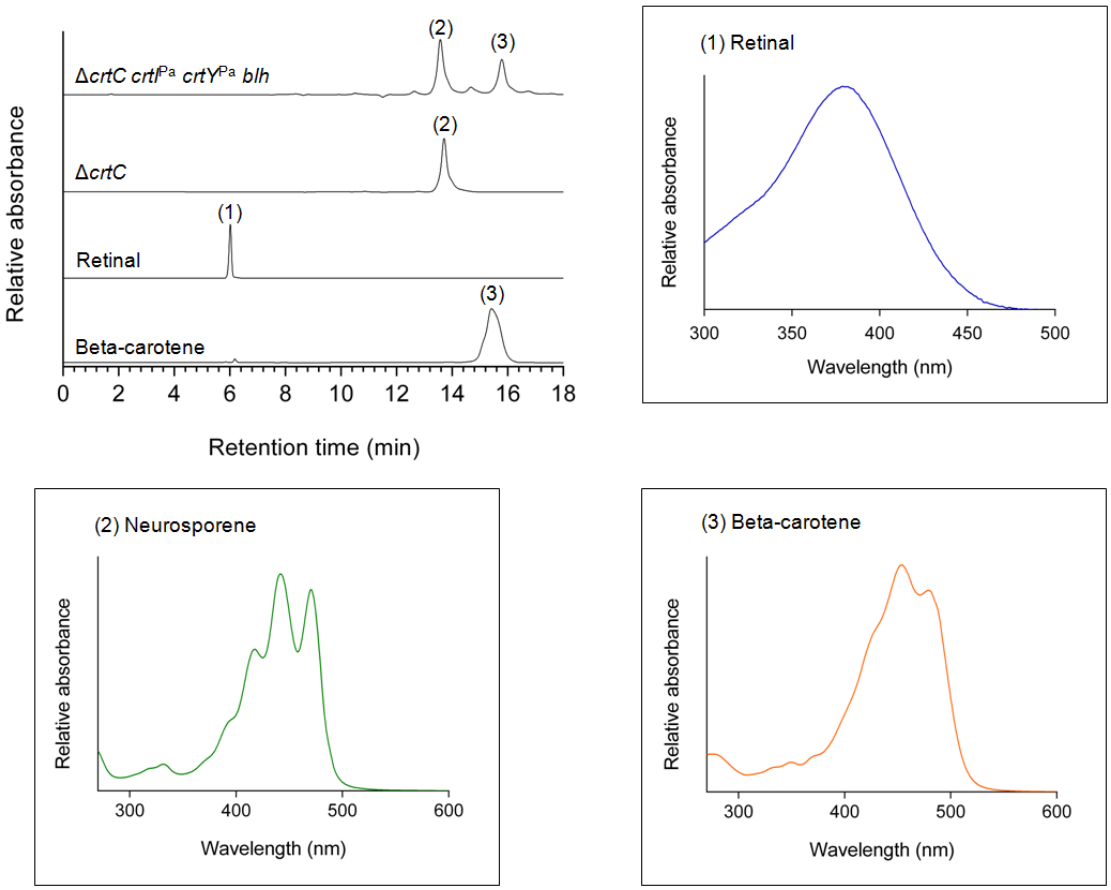


Figure 4.3 HPLC analysis of *crtIYblh*

Pigments were extracted from *Rba. sphaeroides* Δ*crtC* and Δ*crtC* Δ*crtI*::*crtI*^{Pa} *crtY*^{Pa} *blh* (*crtIYblh*) using acetone:methanol (7:2) and analysed on a Phenomenex C18 column. Beta-carotene and retinal pigment standards were used as a comparison. Separation of pigments was detected by a diode array detector (DAD) set to 442 nm, spectra are normalised according to the highest peak. The large negative peak contributed by bacteriochlorophyll which occurs at a retention time of between 7 and 9 minutes has been removed for clarity. The observed pigments and retention times are (1) retinal 6 min, (2) neurosporene, 13.7 min, and (3) beta-carotene, 15.8 min. The respective absorbance spectra are shown. Other peaks are thought to be pigment degradation products.

The absence of retinal observed in *crtIYblh* indicates either a problem with the *blh* gene, the environmental conditions affecting the functionality of the BCDO protein, the degradation of retinal, or the availability of beta-carotene for conversion to retinal.

To investigate the possibility that the *blh* gene is not being expressed from the genome, it was overexpressed from the plasmid pIND4. The *blh* gene was amplified using the primers "Blh F" and "Blh R", digested with *Nco*I and *Hin*dIII and ligated into pIND4 to produce pIND4-*blh*. pIND4-*blh* was transferred via conjugation to a beta-carotene producing strain, *Rba. sphaeroides ΔcrtC crtI*^{Pa} *crtY*^{Pa} made by Shuang Chi. 1.5 L of semi-aerobic culture was grown and protein expression was induced at an OD_{680} of 0.8 with 1 mM IPTG and grown for 7 hours. Pigments were extracted and HPLC analysis was performed, but no retinal was

observed (data not shown). Retinal biosynthesis requires oxygen as a substrate for the Blh monooxygenase; accordingly the cells were grown under oxygenated conditions, 500 ml M22 medium in a 2 L flask shaken at 300 rpm, but no retinal was observed (data not shown).

4.3.3 Production of retinal in *E. coli*

Retinal has been previously produced in *E. coli* (Kim *et al*., 2008; Jang *et al*., 2011). To test the functionality of the codon optimised *blh* gene used in this study, it was expressed from pIND4-*blh* in *E. coli* JM109. The equivalent gene from mouse, *β-diox*, was expressed from pKJ900 (Kim *et al*., 2008). The genes for the biosynthesis of beta-carotene were also expressed from the plasmid pORANGE: *P. agglomerans crtE, crtB, crtI* and *crtY* (von Lintig and Vogt, 2000). Cells were grown and protein expression was induced. Pigments were extracted and HPLC analysis was performed as described in Sections 2.18.2-3. Retinal was seen in both cases; BCDO (*blh*) is more efficient at converting beta-carotene to retinal than β-diox (Figure 4.4). These results are consistent with the findings of Jang *et al*., 2011.

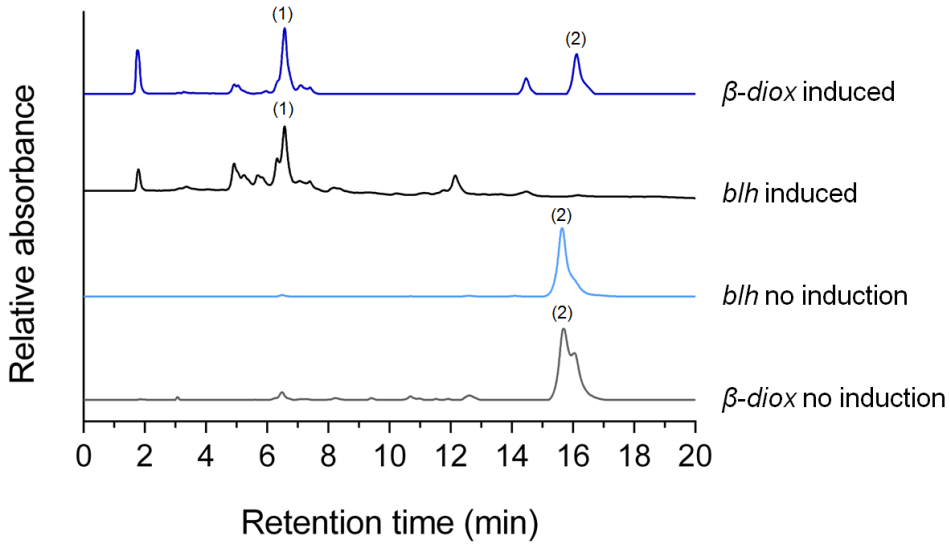


Figure 4.4 Production of retinal in *E. coli*

E. coli JM109 cells containing the plasmids pORANGE to produce beta-carotene (2) and either pKJ900 (*βdiox*; blue traces) or pIND4-*blh* (black traces) to produce retinal (1). 20 ml of overnight culture was used to inoculate 500 ml of LB containing the appropriate antibiotics. Cells were grown and induced with 0.2% L-arabinose at an OD_{600} of 0.4 (pORANGE and pKJ900) and 1 mM IPTG at an OD_{600} of 0.8 (pIND4-*blh*). After induction the cells were grown for 24 hours at 20 °C. Cellular pigments were extracted and separated on a Phenomenex C18 column using HPLC. Separation of pigments was detected by a diode array detector (DAD) set to 442 nm. Spectra are normalised according to the highest peak. Traces are shown with and without induction of the retinal producing genes. Other observed peaks are thought to be degradation products.

While retinal was the major peak seen at the retention time of 6 minutes, many smaller peaks were seen at a similar retention time. These peaks correspond to various degradation products of retinal. This raises the possibility of retinal degradation in *Rba. sphaeroides*.

4.3.4 Improved expression of *blh* in *Rba. sphaeroides*

To improve expression of *blh*, a new ribosome binding site (RBS) was introduced before the *blh* gene. The previously used RBS was the predicted site for the *Rba. sphaeroides crtI* gene (GAGTTCGC); instead the RBS from the pIND4 vector was used (Ind *et al*., 2009). The new RBS was incorporated into the primer "Blh F rbs", which was used to amplify *blh* from pK18mob*sacB-crtYIblh* with the reverse primer "Blh R NdeI". Both the resulting DNA fragment and pK18mob*sacB-crtYIblh* were digested with *Sal*I and *Nde*I. Digested pK18mob*sacB-crtYIblh* was run on an agarose gel and the cut plasmid was gel extracted to remove the original *blh* insert. The two fragments were ligated together to produce pK18mob*sacB-crtYIblh*2. The genes were re-integrated into *Rba. sphaeroides* Δ*crtC* as described previously to produce a strain termed *crtYIblh*2. Pigment extraction and HPLC analysis was performed as before and a small amount of retinal was observed (Figure 4.5).

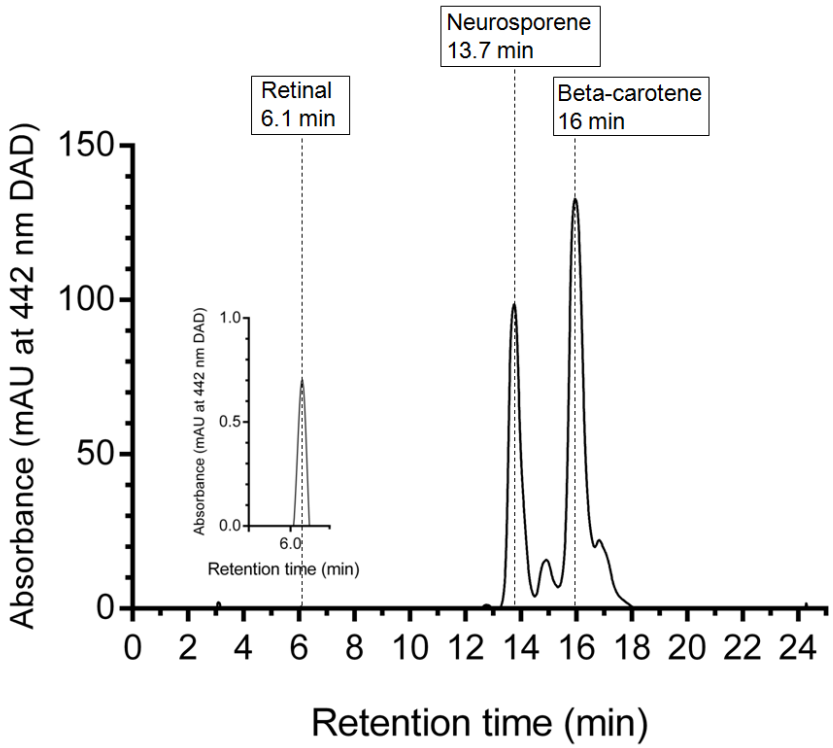


Figure 4.5 HPLC analysis of pigments produced by *Rba. sphaeroides crtIYblh*2

Pigments were extracted from *Rba. sphaeroides crtIYblh*2 using acetone:methanol (7:2) and analysed on a Phenomenex C18 column. Separation of pigments was detected by a diode array detector (DAD) set to 442 nm. The large negative peak contributed by bacteriochlorophyll which occurs at a retention time of between 7 and 9 minutes has been removed for clarity. The unlabelled peaks are thought to be degradation products.

4.3.5 The deletion of LH2 in *Rba. sphaeroides crtIYblh*

BCDO, encoded by *blh*, is hydrophobic and likely an intrinsic membrane protein (Figure 4.6) (Kim *et al*., 2009). It is possible that the crowded membrane environment of *Rba. sphaeroides* cannot accommodate BCDO. To free up membrane space, LH2 was removed through the knockout of the *puc1BA* (hereafter referred to as *1BA*) genes in *crtIYblh*2. It is known that neurosporene can be incorporated into the light harvesting complexes of *Rba. sphaeroides*, but beta-carotene can not (Garcia-Asua 1999; Garcia-Asua *et al*., 2002). This deletion also aimed to increase the availability of neurosporene for conversion into beta-carotene by preventing it being sequestered by LH2. The *puc1BA* genes were knocked out using the plasmid pK18mob*sacB*-Δ*pucBA* constructed by E. Martin following the protocol described previously.

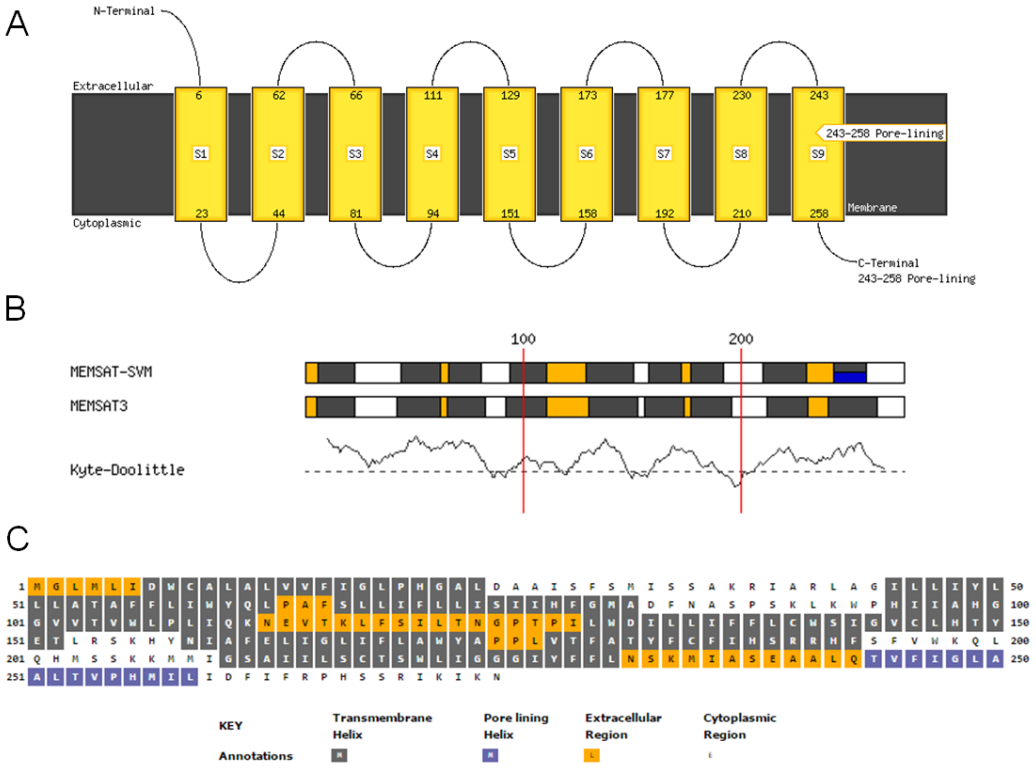


Figure 4.6 Membrane topology prediction for BCDO encoded by *blh*

The amino acid sequence of BCDO encoded by blh from uncultured marine bacterium 66A03 was analysed using the transmembrane topology prediction software MEMSAT3 and MEMSAT-SVM.

- **A.** MEMSAT-SVM representation of the predicted 9 transmembrane helices of BCDO. The N-terminus is predicted to be extracellular (or periplasmic), and the C-terminus is predicted to be cytoplasmic.
- **B.** Schematic of the transmembrane model predicted by both MEMSAT-SVM and MEMSAT3. Kyte-Doolittle hydropathy plot demonstrating the presence of regions of high hydrophobicity. Nine transmembrane helices are observed, the final one of which is predicted to be a pore-lining helix.
- **C.** Amino acid sequence of BCDO showing the presence of the transmembrane helices and location of the inter-helix loops. The pore lining helix is highlighted in blue.

Semi-aerobically grown cells of *crtIYblh*2 Δ*1BA* are more yellow than *crtIYblh* (Figure 4.7 A). This could be due to the balance of neurosporene (yellow) and beta-carotene (orange) compared to bacteriochlorophyll (blue) and/or could indicate the presence of retinal, which is yellow. Figure 4.7 B shows absorbance spectra of intracytoplasmic membranes (ICM) from Δ*crtC*, Δ*crtC* Δ*crtI*::*crtI*^{Pa} and *crtIYblh*2 Δ*1BA*. Neurosporene is the major carotenoid of Δ*crtC* with the absorbance peaks at 430, 457 and 490 nm. A strain containing the *P. agglomerans crtI* gene (Δ*crtC* Δ*crtI*::*crtI*^{Pa} created by Shuang Chi) accumulates lycopene with absorbance maxima at 459, 490 and 523 nm.

As with *crtIYblh* and *crtIYblh*2, *crtIYblh*2 Δ*1BA* contains no lycopene, indicating efficient conversion into beta-carotene by the *crtY* gene product, lycopene cyclase (Figure 4.7 C). Small amounts of retinal were observed in *crtIYblh*2 Δ*1BA* but the amount was 10 fold lower than seen for *crtIYblh*2. A much higher proportion of beta-carotene compared to neurosporene was observed in this strain - 2.5:1 compared to 1.3:1 in *crtIYblh*2 (Figures 4.5 and 4.7 C). The retention times of the pigments are different than recorded previously, likely due to the use of a brand new Phenomenex C18 column; the retention times of pigments varies with the age of the column.

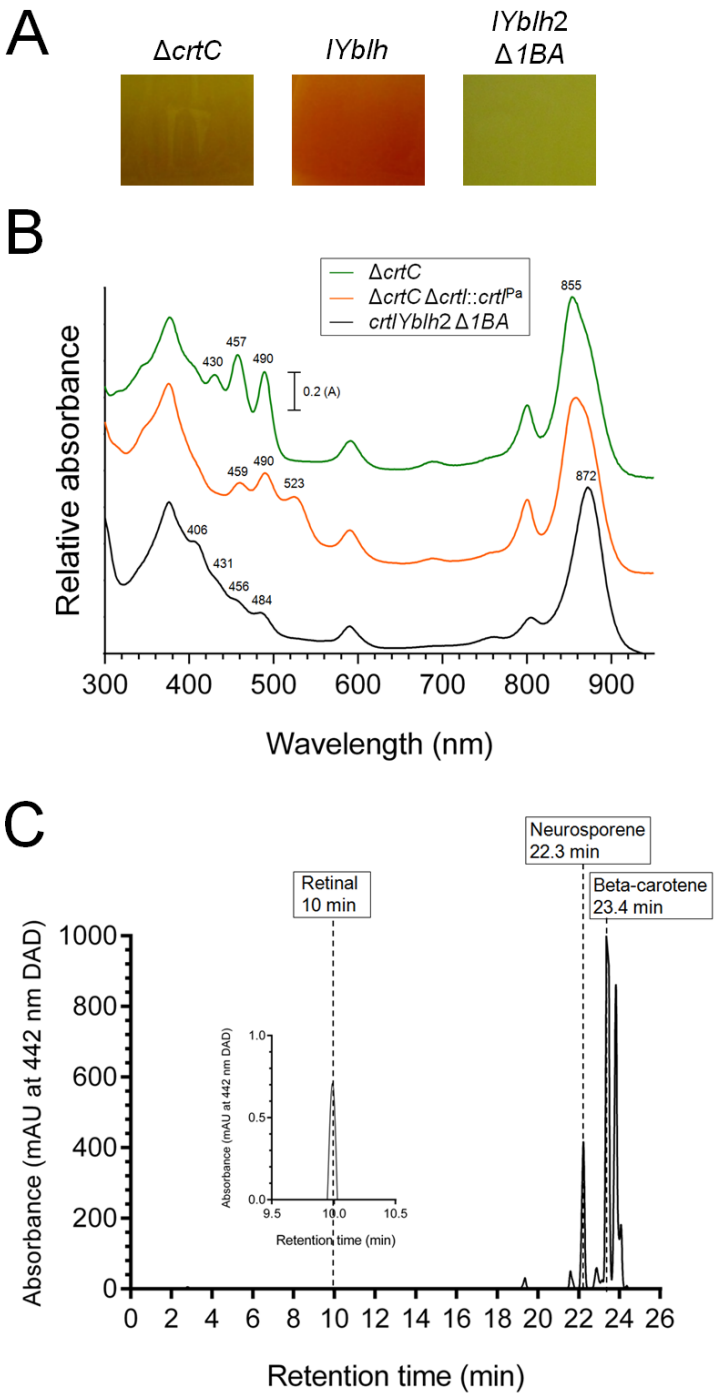


Figure 4.7 Analysis of pigments produced by *Rba. sphaeroides crtIYblh2 Δ1BA*

Semi-aerobically grown cells of the strains Δ*crtC*, Δ*crtC* Δ*crtI*::*crtI*^{Pa}, Δ*crtC* Δ*crtI*::*crtI*^{Pa} *crtY blh* (*crtIYblh*), Δ*crtC* Δ*crtI*::*crtI*^{Pa} *crtY blh* Δ*puc1BA* (*crtIYblh2 Δ1BA*) were analysed.

- **A.** Photographs of cell cultures from Δ*crtC* (neurosporene), IYblh (beta-carotene), and *crtIYblh2 Δ1BA*.
- **B.** Absorbance spectra of ICM of Δ*crtC*, Δ*crtC* Δ*crtI*::*crtI*^{Pa} (strain production and spectrum collection by Shuang Chi) and *crtIYblh2 Δ1BA*.
- **C.** HPLC analysis of *crtIYblh2 Δ1BA*. Pigments were extracted using acetone:methanol (7:2) and analysed on a Phenomenex C18 column. Separation of pigments was detected by a diode array detector (DAD) set to 442 nm. The large negative peak contributed by bacteriochlorophyll which occurs at a retention time of between 18 and 19 minutes has been removed for clarity. The unlabelled peaks are thought to be degradation products.

4.3.6 Increasing the intracellular molecular oxygen levels in *crtIYblh*2 and *crtIYblh*2 Δ*1BA*

BCDO requires molecular oxygen for the conversion of beta-carotene to retinal (Kim *et al*., 2009). As the intracellular environment of *Rba. sphaeroides* is largely anaerobic there may not be sufficient oxygen for the BCDO reaction. *crtIYblh*2 was grown under different oxygen conditions. This was achieved by varying the total culture volume in 125 ml flasks: 20, 40, 50, 60 and 80 ml. No improvement in the retinal yield was observed (data not shown).

The *Rba. sphaeroides* cbb_3 oxidase participates in keeping the intracellular environment anaerobic by reducing oxygen to water. By deleting a subunit of the cbb_3 oxidase, encoded by the *ccoP* gene, cellular levels of oxygen are increased. The *ccoP* gene was knocked out of *crtIYblh*2 Δ*1BA* by E. Martin to produce *crtIYblh*2 Δ*1BA* Δ*ccoP*. The *crtIYblh*2 Δ*1BA* strain was chosen as a higher proportion of beta-carotene is accumulated in comparison to neurosporene.

Cells of *crtIYblh*2 Δ*1BA* Δ*ccoP* appeared green on the plate and yellow in semi-aerobic culture. Figure 4.8 A shows the absorbance spectra of ICM from *crtIYblh*2 Δ*1BA* and *crtIYblh*2 Δ*1BA* Δ*ccoP*. Significantly reduced amounts of LH1 and RC are seen and the BChl Soret band at 370 nm is absent. A peak at 632 nm is observed that is not normally seen in WT *Rba. sphaeroides*. No retinal was observed by HPLC analysis of this strain (Figure 4.8 B).

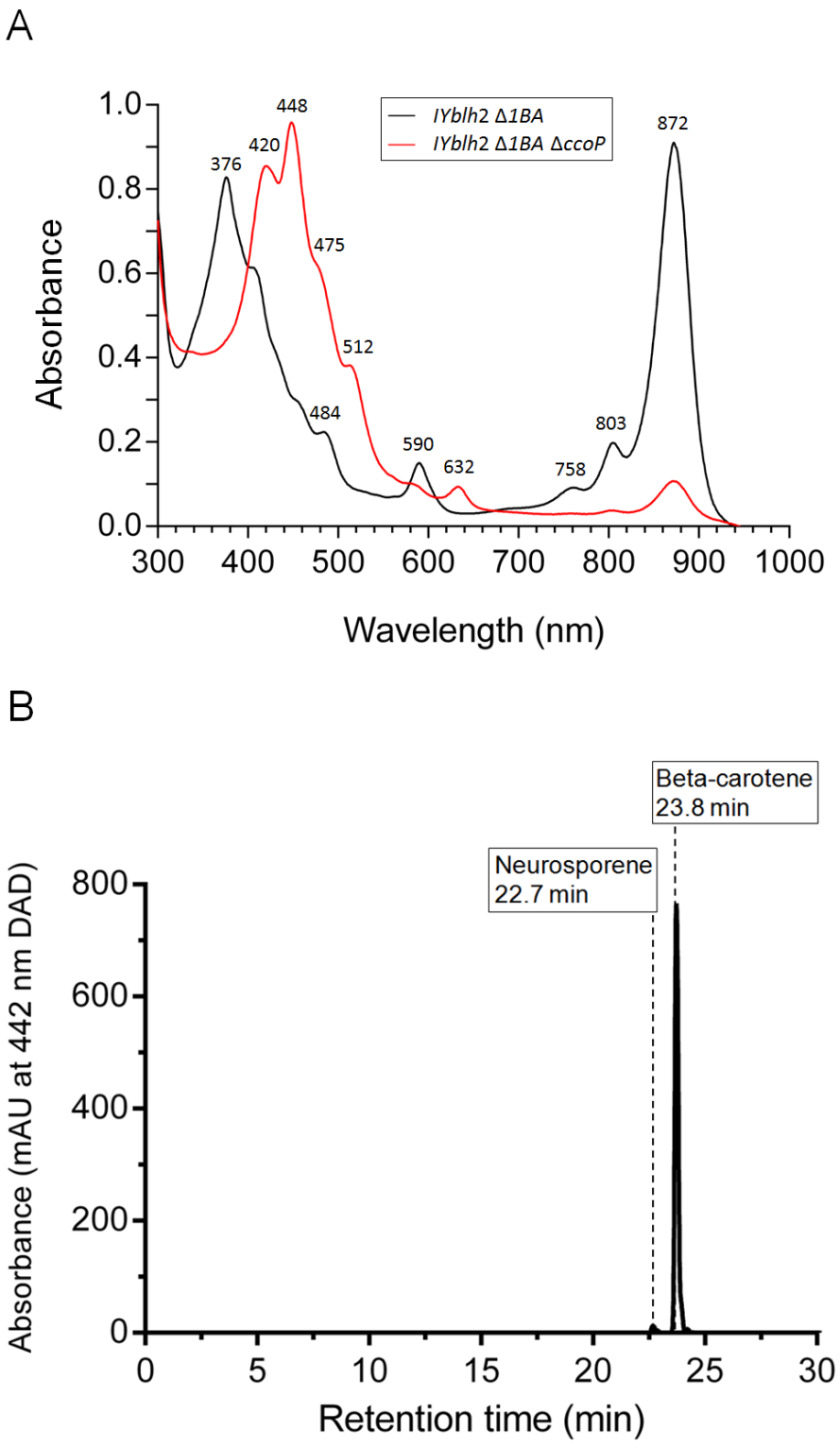


Figure 4.8 Analysis of *Rba. sphaeroides crtIYblh2 Δ1BA ΔccoP*

A. Absorbance spectra of *crtIYblh2 Δ1BA* (black) and *crtIYblh2 Δ1BA ΔccoP* (red). Peak maxima are labelled.

B. HPLC pigment analysis of *crtIYblh2 Δ1BA ΔccoP*. Pigments were extracted using acetone:methanol (7:2) and analysed on a Phenomenex C18 column. Separation of pigments was detected by a diode array detector (DAD) set to 442 nm. The large negative peak contributed by bacteriochlorophyll which occurs at a retention time of between 18 and 19 minutes has been removed for clarity.

4.3.7 Production of proteorhodopsin in *Rba. sphaeroides*

The gene for proteorhodopsin (PR) was introduced in the genome of *crtIYblh*2 in place of the *puc1BA* genes encoding the LH2 complex subunits. The PR variant used is the SAR86 γ-proteobacterial PR variant (Geneart, Toronto, ON, Canada). A pK18mob*sacB* construct was created containing a 1740 bp fragment containing the region upstream of the *puc1BA* genes, the PR gene and the region downstream of the *puc1BA* genes. These fragments were fused together using overlap extension PCR. PR, along with a C-terminal His tag, was amplified from the plasmid pFLAG-PR (Section 4.3.10) using the primers "PR OE F" and "PR OE R" to create a 904 bp fragment. A 404 bp region of DNA upstream of the *puc1BA* genes was amplified using the primers "PucBA KO F SalI" and "PR up OE R". A 463 bp region of DNA downstream of the *puc1BA* genes was amplified using the primers "PR down OE F" and "PucBA KO DR". The fragment was cut with *Sal*I and *Hin*dIII and ligated into pK18mob*sacB* to create the plasmid pk18mob*sacB*-PR. The plasmid was transferred to *crtIYblh*2 and the mutant *crtIYblh*2 Δ*1BA*::PR was produced using the method as described previously.

*crtIYblh*2 Δ*1BA*::PR was grown in 1 L semi-aerobic cultures in 2 L flasks with and without supplementation with 10 µM retinal. ICM were prepared according to the method described in Section 2.15.2. No difference was observed in the colour of the membranes with and without retinal and the absorbance spectra were very similar (Figure 4.9 A). There was a shoulder in the absorbance spectrum of both at 520 nm, which is more pronounced with retinal supplementation. This feature does not appear in the absorbance spectrum of ICM from *crtIYblh*2 Δ*1BA*, so it is likely that this corresponds to PR which absorbs at 520 nm.

Western blot analysis showed that proteorhodopsin is present in the membrane but that there is more proteorhodopsin present in membranes when retinal is added to the media (Figure 4.9 B). As has been seen in previous studies, the actual molecular weight of PR is 27 kDa but it runs at 22 kDa (Tikh *et al*., 2014).

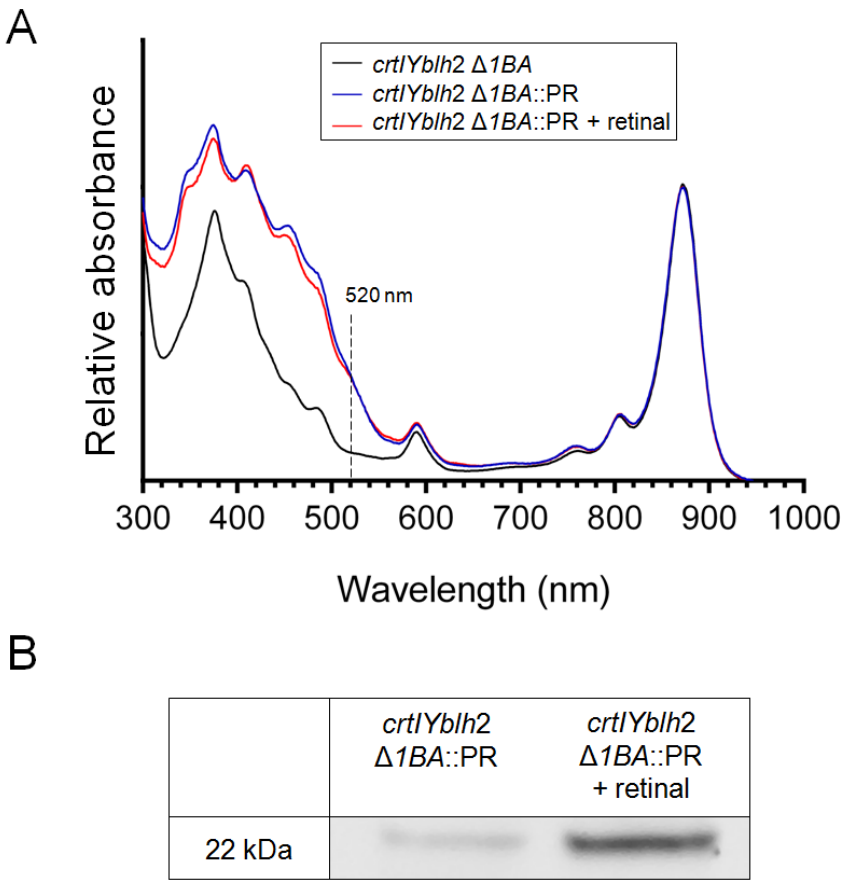


Figure 4.9 Proteorhodopsin in membranes of *Rba. sphaeroides crtIYblh*2 Δ*1BA*::PR

Cells were grown semi-aerobically with and without supplemented 10 µM retinal. ICM was prepared by French pressing the cells followed by centrifugation on a 15% 40% sucrose step gradient.

- **A.** Room temperature absorbance spectra of ICM.
- **B.** Western blot of ICM. The membrane was immunoblotted with an anti-His antibody.

PR was expressed from the plasmid pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR, a plasmid that has been previously shown to produce large amounts of PR in *Rba. sphaeroides* (Tikh *et al.*, 2014). pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR contains a truncated version of the *Rba. sphaeroides puf* promoter and protein expression is constitutive. Cells were grown and membranes prepared as described above. The membranes appeared redder when retinal was supplemented (Figure 4.10 B). There is a clear peak at 520 nm in the ICM absorbance spectrum of membranes both with and without supplemented retinal (Figure 4.10 A). These peaks are more pronounced than when PR is expressed from the genome (Figure 4.9 A), likely due to the increased expression of PR. The 520 nm peak is larger when retinal is supplemented. Western blot analysis showed large amount of PR in the membranes with and without supplementation with retinal (Figure 4.10 C).

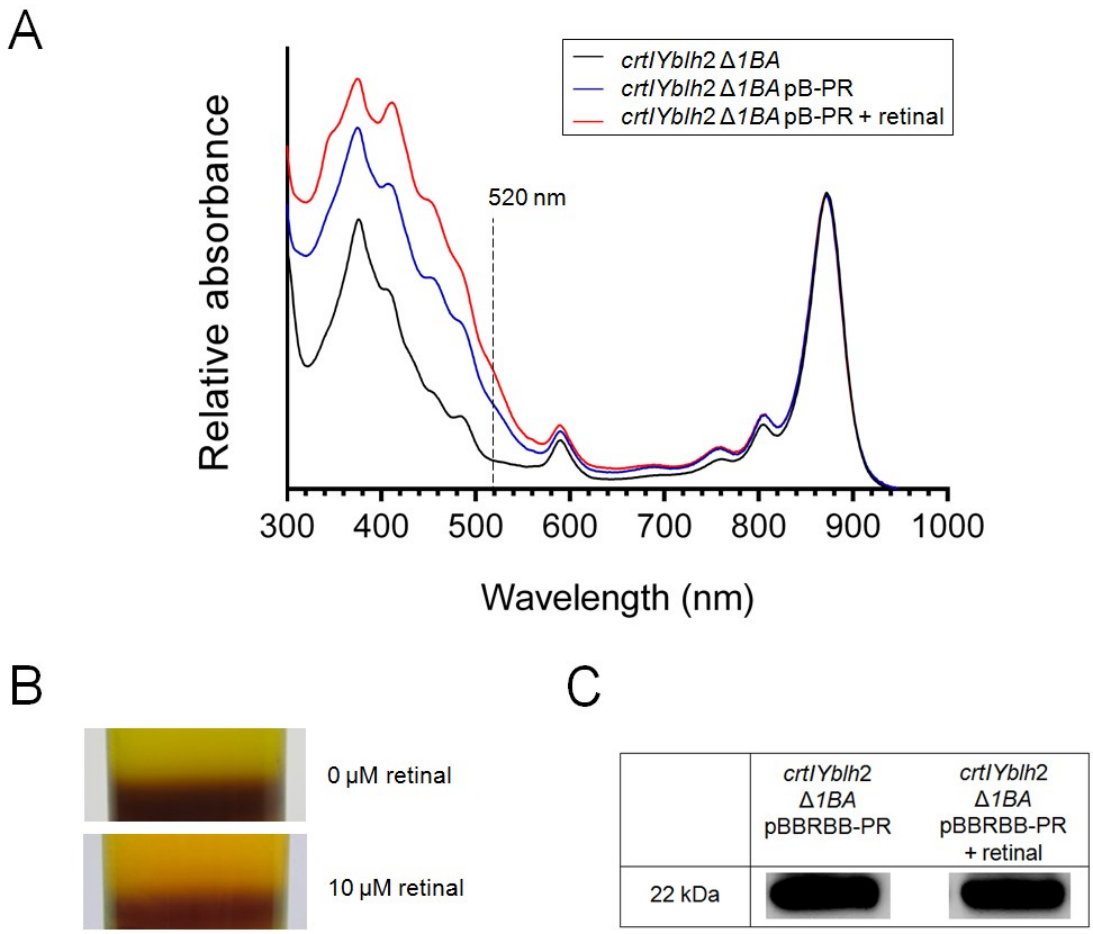


Figure 4.10 Proteorhodopsin in membranes of *Rba. sphaeroides crtIYblh*2 Δ*1BA* pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR

Cells were grown semi-aerobically with and without supplemented 10 µM retinal. ICM was prepared by French pressing the cells and centrifugation on a 15% 40% sucrose step gradient. PB-PR refers to pBBRBB-843-1200-PR.

- **A.** Room temperature absorbance spectra of ICM.
- **B.** 15/40% sucrose gradients with and without supplemented retinal.
- **C.** Western blot of ICM. The membrane was immunoblotted with an anti-His antibody.

PR was expressed from pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR in *Rba. sphaeroides crtIYblh*2 Δ*1BA* Δ*ccoP* (see Section 4.3.6). The cells were grown semi-aerobically in 750 ml LB in 2 L flasks for 32 hours at 30 ^{o}C. 10 µM retinal was added to the growth medium. The cells were spun down and ICM was prepared as described previously. The ICM appeared greener than ICM prepared from the same strain without the pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR and grown in 1.5 L M22 medium in a 2 L flask, conditions of lower aeration (Figure 4.11 A). Although PR was expressed in this strain and under these conditions, there was no peak in the ICM absorbance spectrum at 520 nm corresponding to the absorbance of PR (Figure 4.11 B and C). This indicates that PR is able to associate or assemble into the ICM without bound retinal. Lower levels of photosynthetic

complexes were seen and the carotenoid portion of the absorbance spectrum appeared significantly changed.

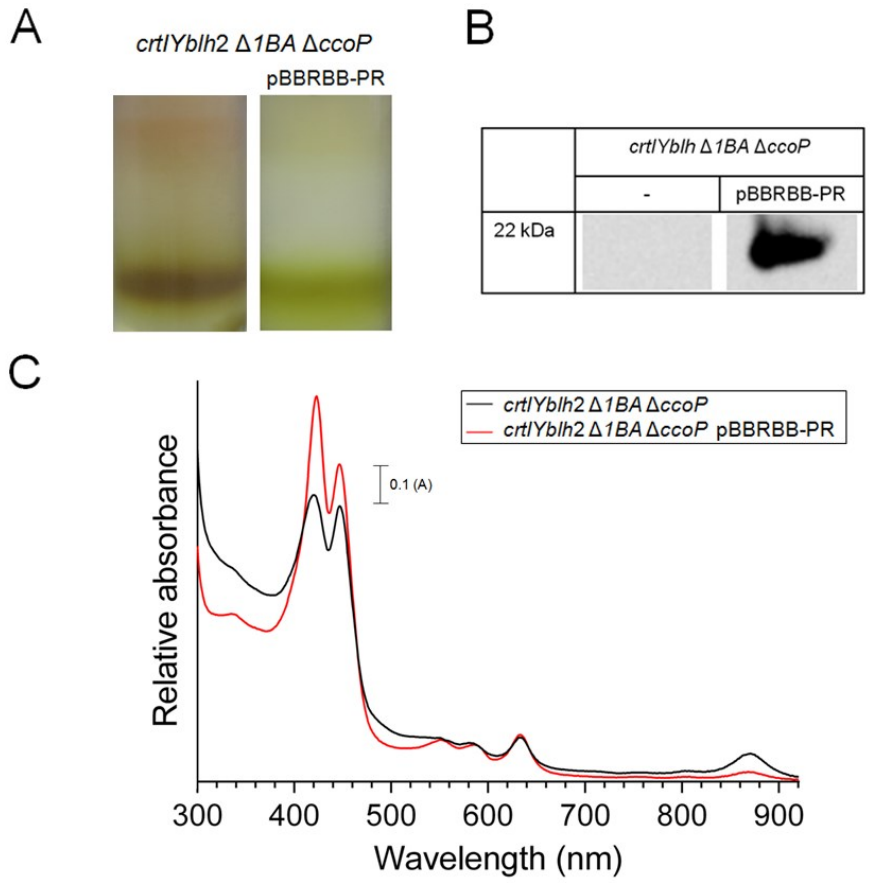


Figure 4.11 Production of proteorhodopsin in *Rba. sphaeroides crtIYblh*2 Δ*1BA* Δ*ccoP*

*crtIYblh*2 Δ*1BA* Δ*ccoP* and *crtIYblh*2 Δ*1BA* Δ*ccoP* pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR were grown in 750 ml of LB in 2 L flasks supplemented with 10 µM retinal, according to the protocol for expression of high levels of PR detailed in Tikh *et al*., 2014.

- **A.** Cells were French pressed then applied to sucrose gradients. The dark band is the membrane fraction at the 40 % / 15 % w/w sucrose boundary.
- **B.** Western blot of membranes isolated from sucrose gradients immunoblotted with an anti-His antibody.
- **C.** Absorbance spectra of ICM from both strains.

4.3.8 The effect of membrane curvature on proteorhodopsin expression

The photosynthetic membrane (ICM) of wild type (WT) *Rba. sphaeroides* is highly curved. The photosynthetic protein complexes induce curvature and are shaped in such a way as to best fit this environment (Figure 4.12 A and B). AFM of PR arrays in *E coli* membranes shows that they have only a small degree of curvature (Dr. Michael Cartron, unpublished data) which is consistent with the shape of PR (Figure 4.12 C). Previous studies have indicated that the

knockout of various photosynthetic complexes, including LH2, results in higher expression of foreign membrane proteins (Laible *et al*., 2009; Tikh *et al*., 2014). To test the hypothesis that PR prefers to sit in flatter membranes, it was expressed in different *Rba. sphaeroides* mutants with varying levels of curvature.

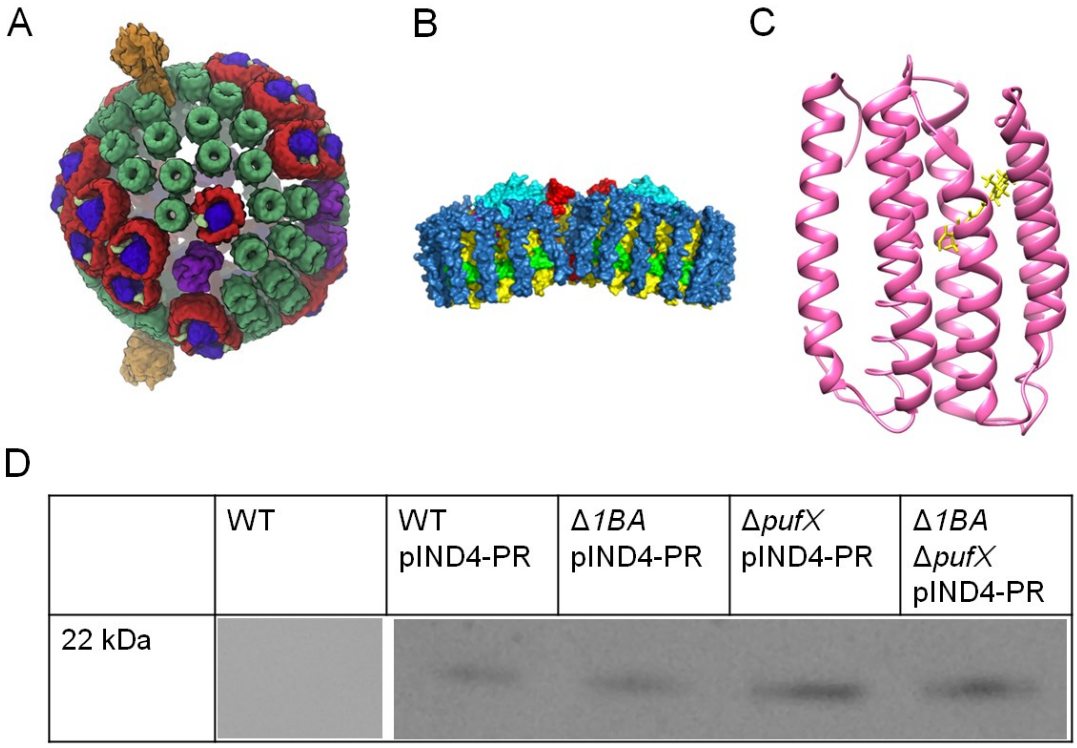


Figure 4.12 Expression of proteorhodopsin in membranes of varying curvature

- **A.** Model of the *Rba. sphaeroides* photosynthetic chromatophore vesicle featuring highly curved membrane. Figure reproduced with permission from Elsevier from Cartron *et al*., 2014.
- **B.** Side view of the dimeric *Rba. sphaeroides* core complex, shaped to induce membrane curvature. Figure reproduced with permission from Qian *et al*., 2013. Copyright 2013 American Chemical Society.
- **C.** Structure of PR, figure from the pdb file of the NMR solution structure of green PR submitted by Reckel *et al*., 2011.
- **D.** PR was expressed from pIND4-PR in *Rba. sphaeroides* mutants of varying membrane curvature. Equal amounts of whole cell extract were subject to Western blotting using an anti-His antibody to the C-terminal His tag on PR.

PR was expressed from pIND4-PR (construct made by Dr. Michael Cartron). pIND4-PR was transferred via conjugation to the *Rba. sphaeroides* strains: WT, Δ*1BA*, Δ*pufX*, and Δ*1BA* Δ*pufX*. *Rba. sphaeroides* mutants lacking LH2 as a result of the knockout of the *puc1BA* genes feature tubular membranes (Hunter *et al*., 1988; Golecki *et al*., 1991). The dimeric core complex induces membrane curvature due to its shape (Figure 4.12 B). When the *pufX* gene is knocked out the core complexes become monomeric. In Δ*pufX* mutants the ICM is larger and less curved, resulting in slower LH2 assembly (Adams *et al*., 2011). The membranes of Δ*1BA*

Δ*pufX* are flattened sheets (Siebert *et al*., 2004). The cells were grown in 80 ml semi-aerobic cultures. Protein expression was induced at an OD_{680} of 0.6 for 16 hours. Cells were spun down and an equal amount of cells standardised by OD_{680} was analysed using western blot (Figure 4.12 D). PR was expressed in higher levels in Δ*pufX* and Δ*1BA* Δ*pufX* than in WT and Δ*1BA*.

4.3.9 Purification of proteorhodopsin from *Rba. sphaeroides* Δ*pufx*

The results presented in Section 4.3.8 indicate that expression of plasmid-borne PR genes in a *Rba. sphaeroides* strain with less curved ICM yields higher levels of the PR protein. Next, the gene encoding PR was introduced into the genome of *Rba. sphaeroides ΔpufX* according to the same method as described in Sections 4.3.1 and 4.3.5.

The resulting strain, *ΔpufX* Δ*1BA*::PR was grown photosynthetically in a 1.5 L culture with and without supplemented retinal. ICM was prepared and solubilised as described in Sections 2.15.1-3. Spectra of the solubilised ICM (Figure 4.13 A) do not show a peak for PR, as the carotenoid absorbance bands are in the 520 nm region. The solubilised membranes were applied to a Ni NTA column equilibrated in 20 mM Tris pH 7.4 and 0.04 % β-DDM. After wash steps of 10 and 20 mM imidazole, protein was eluted with 400 mM imidazole. The fractions were not coloured, so the elution fractions were analysed using western blot. PR was not visible on a Coomassie stained SDS gel. A band at 22 kDa was observed for both with and without supplanted retinal, corresponding to the expected observed size of PR (Figure 4.13 B). The same band was observed for a purification performed on cells that did not have retinal supplemented. This indicates that PR is able to associate with the membrane even without retinal bound.

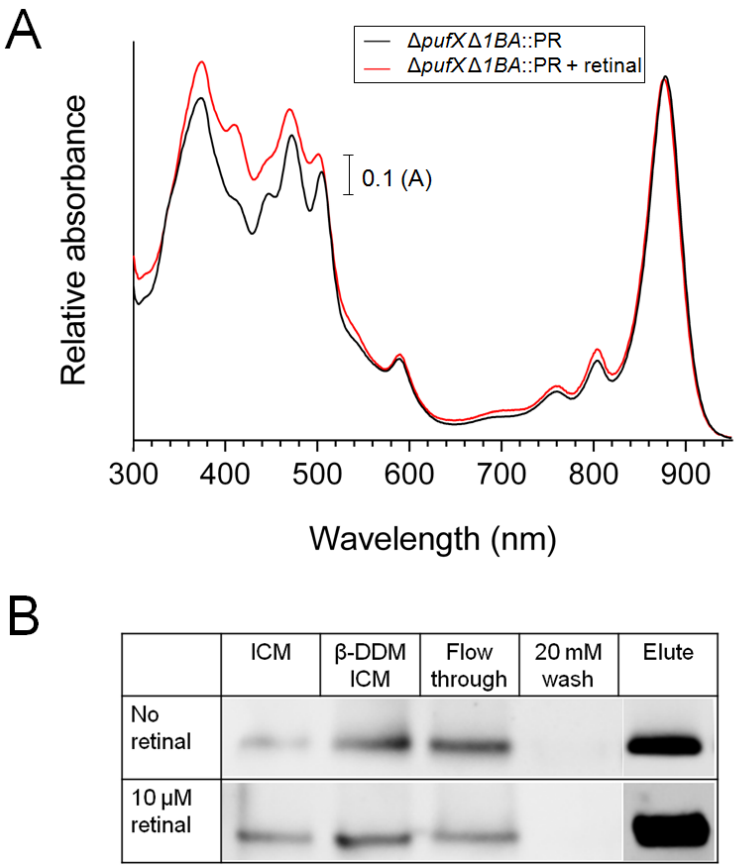


Figure 4.13 Purification of proteorhodopsin from *Rba. sphaeroides* Δ*pufX* Δ*1BA*::PR

Cells were grown photosynthetically with and without supplementation with 10 µM retinal. Cells were spun down and resuspended in membrane buffer (20 mM HEPES, 5 mM EDTA pH 7.4), French pressed and the ICM was isolated on 15/40% sucrose step gradients. The membranes were solubilised in 3 % β-DDM for 1 hour at 4 °C followed by centrifugation at 48,000 rpm (158, 054 x *g*) in a Beckman Ti 70.1 rotor for 1 hour to remove unsolubilised material.

A. Absorbance spectra of ICM solubilised with 3 % β-DDM.

B. Western blot analysis of the various stages in purification using an anti-His antibody to the C-terminal His tag on PR. The bands are at 22 kDa, the expected running size of PR. Membranes were harvested from sucrose gradients and solubilised with 3 % β-DDM (β-DDM ICM). Solubilised membranes were applied to a Ni NTA column, which was washed with 20 mM imidazole. Protein was eluted with 400 mM imidazole.

4.3.10 Production and purification of proteorhodopsin from *Synechocystis* PCC6803

Synechocystis PCC6803 is a cyanobacterium that is able to biosynthesise retinal. *Synechocystis* is naturally transformable and can integrate foreign plasmid and linear DNA into its genome by homologous recombination allowing targeted gene replacement. *Synechocystis* contains approximately 10 genome copies per cell (Herdman *et al.*, 1979), therefore to avoid the reversion of a mutant back to a base strain genotype, a fully-segregated mutant must be isolated.

A plasmid was constructed for the introduction of PR into the genome of *Synechocystis* at the *slr1311* (*psbAII*) locus. The photosystem II D1 protein of *Synechocystis* is encoded by three genes, *psbAI*, *psbAII* and p*sbAIII*. Expression of these genes is enhanced under high light conditions (Hihara *et al*., 2001) and the deletion of the *psbAII* gene (*slr1311*) does not result in a mutant phenotype (Dr. Roman Sobotka, personal communication). It is therefore a good location for the introduction of the gene encoding PR. The plasmid pFLAG was created by Dr. Daniel Canniffe (University of Sheffield) and contains the sequence for a 5' 3X FLAG tag flanked by the sequences found up- and downstream of *slr1311*. The gene for PR was amplified from pBAD-PR (created by Dr. M. Cartron) using the primers "PR F FLAG" and "PR R *BglII*". The resulting fragment was digested using *Not*I and *Bgl*II and ligated into pFLAG to create the plasmid pFLAG-PR.

Synechocystis was transformed with pFLAG-PR according to the method described in Section 2.16.1. To identify fully segregated mutant strains, colony PCR was carried out using the primers "FLAG Check F" and "FLAG Check R", designed by Dr. D. Canniffe to amplify the ORF *slr1311* (Figure 4.14 A). Replacement of *slr1311* with the PR gene in mutant colonies was confirmed by DNA sequencing.

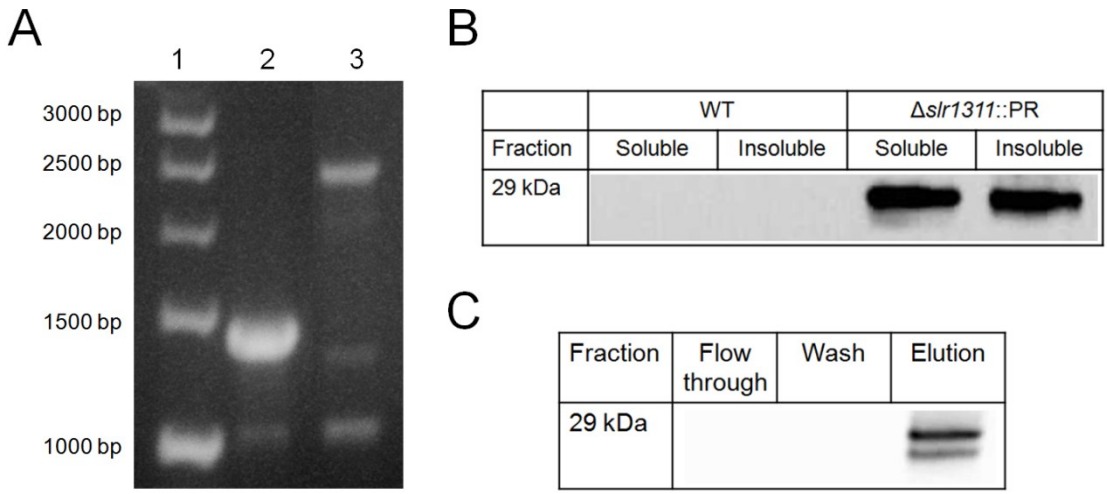


Figure 4.14 Production of proteorhodopsin in *Synechocystis* PCC6803

- **A.** Agarose gel showing the products of colony PCR of *Synechocystis* WT (lane 2) and Δ*slr1311*::PR (lane 3). The analytical PCR was carried out using primers designed to amplify the *slr1311* locus.
- **B.** Western blot analysis of WT and Δ*slr1311*::PR insoluble (membrane) and soluble fractions. Using anti-FLAG antibody to the N-terminal FLAG tag on PR.
- **C.** Western blot analysis of purification of PR from Δ*slr1311*::PR using anti-FLAG antibody to the N-terminal FLAG tag on PR.

To determine whether PR was expressed and assembled in the membrane fraction, a 100 ml liquid culture was grown without retinal supplementation and separated into the insoluble (membrane) and soluble fractions as described in Section 2.16.3. PR ran on the gel at approximately 29 kDa, the expected running size of PR with an N-terminal FLAG tag. PR was observed in both the soluble and insoluble (membrane) fractions (Figure 4.14 B).

An 8 L culture of *Synechocystis* Δ*slr1311*::PR was grown up as described in Section 2.16. PR was purified according to the method described in Section 2.17.5. The growth medium was supplemented with retinal to maximise the yield of fully-assembled PR. A very small amount of PR was purified, but the concentration was so low that it could not be seen on a silver stained gel and could only be detected by western blot (Figure 4.14 C).

4.4 Discussion

4.4.1 Small amounts of retinal are produced in *Rba. sphaeroides*

Functional proteorhodopsin requires the pigment all-*trans* retinal. Three foreign genes, *crtI* and *crtY* from *P. agglomerans* and *blh* from uncultured marine bacterium 66A03, were introduced into the carotenoid gene cluster of *Rba. sphaeroides* replacing the native *crtI* gene (Figures 4.1 and 4.2). These genes were introduced into a *crtC* knockout strain to direct pigment biosynthesis down the new pathway; the strain was termed *crtIYblh*. The *P. agglomerans crtI* gene encodes a 4-step phytoene desaturase which produces lycopene from phytoene. The *P. agglomerans crtY* gene encodes lycopene cyclase and produces beta-carotene from lycopene. *P. agglomerans crtI* and *crtY* have both been previously expressed from an expression vector in *Rba. sphaeroides* to produce large amounts of beta-carotene (Hunter *et al*., 1994). This work demonstrates that these genes are capable of producing beta-carotene when expressed from the *Rba. sphaeroides* genome (Figure 4.3). No lycopene is observed in any of the strains expressing *crtY* in this work indicating efficient conversion into beta-carotene.

The gene used in this study to convert beta-carotene to all-*trans* retinal was the *blh* gene from the uncultured marine bacterium 66A03 as it has been found to convert beta-carotene to retinal with a high efficiency in *E. coli* (Jang *et al*., 2011). *blh* encodes a beta-carotene 15, 15' dioxygenase (BCDO) which cleaves beta-carotene at its central double bond to produce retinal; this reaction requires molecular oxygen (Kim *et al*., 2009). *blh* was codon optimised for *Rba. sphaeroides*. Initially, no retinal was observed in the *crtIYblh* strain (Figure 4.3).

To test that the *blh* gene with codons optimised for *Rba. sphaeroides* was functional it was expressed in *E. coli* along with the genes from *P. agglomerans* to produce beta-carotene. The efficiency of retinal production was compared to a strain with the *β-diox* gene from mouse. BCDO expressed from the codon optimised *blh* was found to convert beta-carotene to retinal with a higher efficiency than β-diox (Figure 4.4). These results are consistent with the findings of Jang *et al*., 2011. This result eliminates the possibility that the codon optimised gene cannot produce a functional BCDO.

The predicted RBS for the native *crtI* gene was placed before the *blh* gene in *crtIYblh*. The introduction of the RBS from the pIND4 expression vector before the *blh* gene in *Rba. sphaeroides crtIYblh* enabled the production of small amounts of retinal in the strain *crtIYblh2* (Figure 4.5). The new RBS contains a GGAGA sequence 8 base pairs prior to the ATG start codon of *blh*, in line with the optimal RBS position for prokaryotic protein expression (Isaacs *et al*., 2004; Pfleger *et al*., 2006; Salis *et al*., 2009). It can therefore be concluded that the predicted *crtI* RBS was either wrong or insufficient for *blh* expression. Several strategies were used to attempt to improve the retinal yield of the *crtIYblh*2 strain without success, detailed below.

4.4.2 Genetic removal of the LH2 antenna does not result in increased retinal biosynthesis

BCDO, encoded by *blh*, is a putative membrane protein (Figure 4.6). The *puc1BA* genes were knocked out in *crtIYblh*2 to abolish LH2 expression. The knockout of LH2 may provide more space in the photosynthetic membrane for BCDO to sit, and Δ*1BA* mutants are known to result in higher expression of membrane proteins (Laible *et al*., 2009). However, no increase in retinal levels was observed (Figure 4.7 C). Beta-carotene is not soluble in an aqueous environment; it is not freely diffusible *in vivo* and associates with lipophilic structures such as membranes or binding proteins. Although it is incapable of being assembled into the light harvesting complexes, it is possible that beta-carotene associates with another protein which protects it from degradation or conversion to retinal (Garcia-Asua, 1999).

4.4.3 Increasing the level of intracellular oxygen does not increase retinal biosynthesis

BCDO requires molecular oxygen to produce retinal from beta-carotene, however the interior of *Rba. sphaeroides* cells is a low oxygen environment. The native carotenoid spheroidenone is produced from spheroidene, a reaction that requires molecular oxygen (Shneour, 1962). Small amounts of spheroidenone are found when *Rba. sphaeroides* is grown under anaerobic conditions but significantly higher levels are produced when *Rba. sphaeroides* is grown in the

presence of oxygen (Shneour, 1962). To investigate the possibility that retinal biosynthesis requires an increased amount of available molecular oxygen *crtIYblh2 Δ1BA* cells were grown in conditions of high aeration, but no retinal was observed.

The *ccoP* gene encodes a subunit of the *Rba. sphaeroides* cbb_3 oxidase, the deletion of which results in higher levels of oxygen within the cell (Ogara and Kaplan, 1997). The cbb_3 oxidase catalyses the reduction of O_2 to H_2O and uses the liberated free energy to pump protons across the periplasmic membrane (Pitcher and Watmough, 2004). The strain *crtIYblh2* Δ*1BA* Δ*ccoP* was produced but, although increased levels of beta-carotene were observed in comparison to neurosporene, no retinal was observed (Figure 4.8 B). It is possible that retinal is degraded by the increased levels of oxygen within the cells.

The knockout of *ccoP* in wild type *Rba. sphaeroides* results in the de-repression of the *puc* and *puf* operons under high oxygen conditions (Ogara and Kaplan, 1997). The biosynthesis of BChl is repressed by high oxygen conditions but the knockout of *ccoP* does not result in increased BChl production (Ogara and Kaplan, 1997). However, in the *crtIYblh2* Δ*1BA* Δ*ccoP* background significantly reduced amounts of LH1 and RC are seen and the 370 nm BChl Soret peak is not resolvable in the ICM absorbance spectrum (Figure 4.8 A). A peak at 632 nm that is not present in wild type ICM is observed in the Δ*ccoP* strain (Figure 4.8 A), this may correspond to a BChl biosynthetic intermediate such as protochlorophyllide. It is therefore possible that the growth conditions used in this study result in high enough oxygen tension to significantly repress BChl synthesis and therefore the assembly of the reaction centre and LH1 complexes. The reduced level of LH1 may account for the decreased amount of neurosporene seen in this strain; if neurosporene is not sequestered by LH1 it becomes available for conversion into beta-carotene.

4.4.4 Retinal is sensitive to degradation

The native *crtI*/*crtB* operon of *Rba. sphaeroides* is repressed by oxygen, which may explain the lack of observed retinal under increased oxygen conditions. Retinoids, such as retinal, are readily degraded and isomerised by heat, oxygen, light and biological degradation via retinoic acid (Jang *et al*., 2011). This effect can be seen in the HPLC analysis of pigments produced in *E. coli* in which many smaller peaks are observed around the retinal peak; these peaks correspond to various degradation species of retinal (Figure 4.4). In an attempt to find the balance between sufficient oxygen for retinal production but not retinal degradation *crtIYblh2 Δ1BA* was grown under varying oxygen conditions, but no improvement in retinal yield was seen.

Retinal is degraded to retinoic acid by retinal dehydrogenase. *Salmonella enterica* has a retinal dehydrogenase (Accession No. CBY96723), a possible homologue of which is found in *E. coli* (Jang *et al*., 2011). A blastp search found no homologues of the *S. enterica* retinal dehydrogenase in *Rba. sphaeroides*, although it shares 26 % sequence identity with an aldehyde dehydrogenase (Accession No. WP_011339575) which could participate in the degradation of retinal.

4.4.5 Proteorhodopsin assembles in the membrane of *Rba. sphaeroides*

The gene for the SAR86 γ-proteobacterial PR variant (a green PR) was introduced to the genome of *Rba. sphaeroides crtIYblh*2 Δ*1BA* in place of the *puc1BA* genes encoding the LH2 subunits. This location was chosen as the knockout of LH2 results in increased expression of foreign membrane proteins, possibly due to increasing the space in the photosynthetic membrane for new complexes to sit (Laible *et al*., 2009). The *puc* promoter is strong and inducible by light and decreased oxygen (Kiley and Kaplan, 1987), although good expression occurs when the cells are grown semi-aerobically. *crtIYblh*2 Δ*1BA*::PR was grown semi-aerobically to reduce the possibility of light- and oxygen- induced degradation of retinal, and to ensure that there was enough molecular oxygen present for the conversion of beta-carotene into retinal by BCDO. A peak at 520 nm corresponding to the absorbance maximum of PR is visible in the absorbance spectrum of ICM prepared from these cells (Figure 4.9 A). This peak becomes more pronounced on supplementation of the growth medium with retinal. Western blot analysis confirmed that PR is present in the ICM and in larger quantities when retinal is added to the growth medium (Figure 4.9 B). These data suggest that PR is able to assemble into the photosynthetic membrane of *Rba. sphaeroides* and that the small amount of retinal produced by *crtIYblh*2 Δ*1BA* is sufficient for the expression of some PR, although supplementation is required to produce increased amounts of this protein.

To increase the amount of PR in the membrane, it was produced in the *crtIYblh*2 Δ*1BA* strain through expression from the plasmid pBBRBB-$Ppuf_{843\text{-}1200}$-PR. This plasmid has been previously used to produce large amounts of PR in a *Rba. sphaeroides* ΔRCLH (reaction centre minus) background (Tikh *et al*., 2014). A peak at 520 nm is observed in the absorbance spectrum of ICM isolated from this strain (Figure 4.10 A) and is more pronounced than when PR is expressed genomically. This peak is increased when the cells were supplemented with retinal, and ICM appeared redder in colour (Figure 4.10 B). Western blot analysis of ICM showed the presence of PR in membranes both with and without supplementation with retinal (Figure 4.10 C). The same result was observed in the *crtIYblh*2 Δ*1BA* Δ*ccoP* background in which no retinal is present (Figure 4.11), which indicates that PR is able to assemble into

the ICM in the absence of retinal. Studies on the folding of bacteriorhodopsin show that retinal binds at a late stage in the folding process (Booth, 2000). Therefore, it is possible that the same is the case for PR and that it is able to associate or assemble in the membrane without bound retinal.

4.4.6 Decreased membrane curvature results in increased expression of proteorhodopsin

The photosynthetic complexes of *Rba. sphaeroides* are shaped in such a way as to induce membrane curvature of the photosynthetic membrane (Figure 4.12 A and B). Membrane proteins from other organisms may be shaped to fit into flatter membranes and the highly curved membrane environment of *Rba. sphaeroides* may not favour the expression of these proteins. Previous studies have shown that the knockout of various photosynthetic complexes results in increased expression of foreign membrane proteins in *Rba. sphaeroides* (Laible *et al*., 2009; Tikh *et al*., 2014), although this was suggested to arise from making more membrane area available rather than making the membrane more planar. To test the curvature hypothesis, PR was expressed in different *Rba. sphaeroides* backgrounds. Increased PR expression was seen in Δ*pufX* and Δ*pufX* Δ*1BA* backgrounds when compared to WT (spherical ICM) and Δ*1BA* (tubular membranes) (Figure 4.12 D). A Δ*pufX* Δ*1BA* background was therefore used for the expression of PR with the aim to purify the protein. However, while PR was found in the membrane fraction of these cells, the amount of purified proteorhodopsin was too low to obtain an absorbance spectrum to confirm full assembly (Figure 4.13).

4.4.7 Low amounts of proteorhodopsin can be purified from *Synechocystis* PCC6803

The cyanobacterium *Synechocystis* PCC6803 is able to synthesise retinal. The gene for PR was therefore introduced onto the genome of this strain to determine whether it is able to assemble functional PR. PR was detected in both the soluble and insoluble (membrane) fractions of this strain (Figure 4.14 B). Only a very small amount of PR could be purified from this strain, and there was evidence of protein degradation (Figure 4.14 C). The appearance of PR in the soluble fraction indicates that a very large amount of PR is produced or that PR cannot fully insert into the membrane, possibly due to incomplete binding of retinal. It is possible that PR requires a specific leader sequence for correct insertion into the thylakoid membrane.

Synechocystis does not encode a polypeptide with homology to a known opsin, therefore the retinal it produces may have a role in signalling or regulation (Ruch *et al*., 2005). As a result, only very low amounts may be synthesised and found within the cell. In organisms that

produce PR, the genes necessary for the biosynthesis of retinal are often found downstream of the PR gene indicating that for functional relevance, the target cell should have the desired capability of retinal production (McCarren and DeLong, 2007).

The gene for proteorhodopsin was positioned under the light-driven *slr1311* (*psbAII*) locus. There are two possible explanations associated with this gene placement for the low amount of PR that could be purified. Firstly, as the cell culture grows, self-shading prevents optimal transcription form the light-driven promoter. Secondly, the proteorhodopsin-producing mutant appeared to be light sensitive as cells grown in high light levels did not grow or died. This necessitated the use of low light levels for cell growth. The light-sensitivity of this strain raises the possibility that the proton pumping ability of PR could have been damaging to the cells. PR is able to vary the direction in which it pumps protons depending on the pH of the environment. Depending on the orientation in which PR is inserted into the membrane and the pH of its surroundings, it may deplete the transmembrane proton gradient used to drive ATP synthesis.

4.4.8 Additional work

Native operons feature the genes necessary for retinal biosynthesis in reverse order. The placement of *blh* as the first of the three added genes in *Rba. sphaeroides* may result in increased retinal production. It may also be necessary to place these genes under the same promoter as PR.

Further purification strategies must be attempted to verify the correct assembly of PR in *Rba. sphaeroides*. This includes the purification of PR from *crtIYblh*2 Δ*1BA*::PR. The use of pBBRBB-*Ppuf*$_{843\text{-}1200}$-PR in this strain will result in increased yields to facilitate purification.

To further probe the effect of membrane curvature on the expression of PR, membranes should be prepared to discount any signal contribution from PR that had not assembled in the membrane. There is currently little research on the distribution of PR in native membranes. AFM and EM could be performed on *Rba. sphaeroides* chromatophores vesicles to determine the arrangement of the protein in this system.

Photosynthetic growth curve analysis of *crtIYblh*2 Δ*1BA*::PR will determine whether PR has the potential to augment photosynthesis. In this strain, the cytochrome bc_1 complex could be deleted or inhibited to determine whether the PR can replace its proton-pumping abilities. Ultimately a membrane packed with PR could replace the light-harvesting, reaction centre and cytochrome bc_1 complexes. It may be possible to enhance the performance of this new

photosynthetic unit by creating an artificial antenna system for the PR in place of the *Rba. sphaeroides* LH1 and LH2 molecules.

If PR it is to be useful for photosynthesis in *Synechocystis*, it must be located in the thylakoid membrane and in the correct orientation. This could be achieved by attaching a signal sequence to PR, for example the Tat signal sequence which targets the petC1 and petC2 proteins to the thylakoid membrane (Aldridge *et al*., 2008).

4.4.9 Conclusion

The aim of this work was to explore the potential for the introduction of foreign pigments and proteins into the photosynthetic membrane of *Rba. sphaeroides* with the eventual aim of augmenting photosynthesis. This work has established the principle that the native carotenoid biosynthesis pathway can be engineered through the inclusion of three foreign genes to produce small amounts of retinal. Proteorhodopsin has been successfully expressed in the photosynthetic membrane, but the levels of this protein may have to be increased significantly before it will be possible to augment photosynthesis by contributing to the transmembrane proton gradient used to drive the ATP synthase.

Chapter 5

Design and expression of a transmembrane maquette

5.1 Summary

This chapter explores the potential for the use of synthetic peptide maquettes as additional photosystem components in *Rhodobacter sphaeroides*. The gene for a transmembrane maquette (TM) was designed using design principles from natural membrane proteins and is based on existing amphiphilic maquettes. The maquette features a hydrophilic, cytoplasmic, domain and a hydrophobic membrane-spanning domain. Incorporated into the design are six histidine residues which are capable of non-covalently binding three haem molecules.

The TM is able to assemble in to the intracytoplasmic membrane of *Rba. sphaeroides* and the membranes of *Escherichia coli*. The net charge of the cytoplasmic N-terminus was shown to be important for expression of the TM, with the introduction of negatively charged residues leading to decreased expression. It was determined that the extent of membrane curvature is important for TM expression in *Rba. sphaeroides*; decreased membrane curvature resulted in increased expression of TM.

Strategies used to determine the pigment binding properties of TM included the stimulation of haem production in *E. coli* and introduction of the maquette into a *Rba. sphaeroides* mutant that accumulates bacteriochlorophyll precursors. The data in this chapter provide evidence for pigment-binding by the maquette although further work is required to explore this. This work forms the basis of the bottom-up redesign of photosystem components with the eventual aim to augment photosynthesis in *Rba. sphaeroides* and create new photosynthetic pathways.

5.2 Introduction

At the turn of the 20th century Emil Fischer "foresaw a time in which physiological chemistry will not only make greater use of natural enzymes but will actually resort to creating synthetic ones" (Fischer, 1902, 1905). Recently, the field of *de novo* protein design has expanded rapidly. However, translating any particular protein function into a successful synthetic protein remains a challenge, even more so to express function synthetic protein *in vivo*. One approach pioneered by the laboratory of Professor Les Dutton at the University of Pennsylvania is to create synthetic self-assembling peptide structures known as maquettes. Maquettes aim to emulate the function of natural oxidoreductases whilst avoiding the complexity of natural proteins arising from Darwin's principle of multiple-utility (Darwin, 1862). Maquettes are reviewed in Section 1.11. The 'bottom-up' design of maquettes provides a simple, modular approach for tailoring the structure and function of components that transfer excitation energy and electrons in photosynthetic systems.

The key structural characteristics and functional elements of natural membrane proteins have been reproduced in artificial peptide maquettes. If a maquette is to replicate the function of, or become, a photosystem component it must be not only membrane-bound but have a vectorial orientation (Discher *et al*., 2003). This vectorial character was conferred on early hydrophilic maquettes through the addition of palmitoylates or cholesterols to the loop regions of the 4-helix bundles; these were able to form stable oriented Langmuir monolayer films (Chen, 1999; Chen *et al*., 1998). More recently, amphiphilic α-helical bundles have been formed which are a fusion of sequences from previously designed hydrophilic and lipophilic maquettes and incorporate lipophilic domains from natural proteins (Discher *et al*., 2005). These amphiphilic (AP) maquettes consist of two distinct continuous hydrophilic and lipophilic domains. Figure 5.1 shows the evolution of AP maquettes from an original hydrophilic maquette.

The original AP maquette is designated AP0, the structure of which is shown in Figure 5.1 B. The hydrophilic domain of AP0 consists of 4-α-helices which are patterned with polar and non polar amino acid residues which assemble in such a way as to present a hydrophilic face to the exterior and a hydrophobic face to the interior of the maquette (Ye *et al*., 2004). The α-helices of the lipophilic domain are extensions of the hydrophilic helices and are patterned in the opposite way. AP0 has haem-binding histidine residues on the interior of the hydrophilic domain.

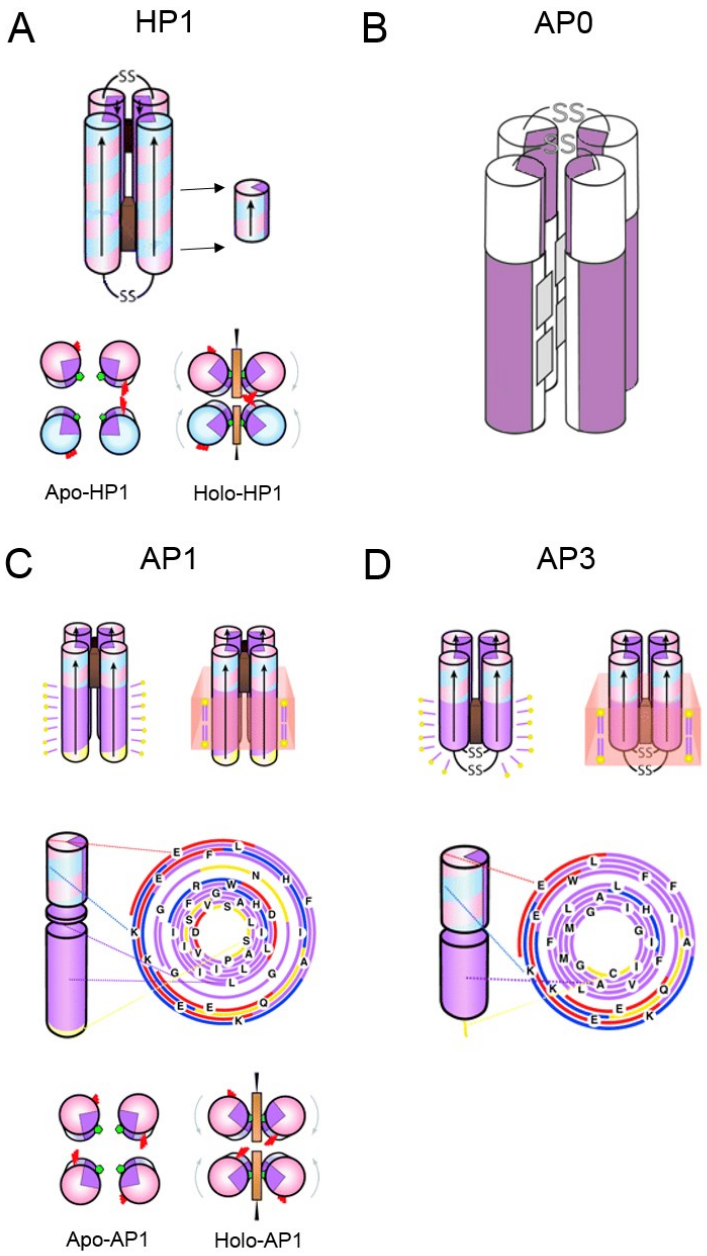


Figure 5.1 Schematic drawings of amphiphilic maquettes

De novo designed maquettes are built from α-helices. In all cases except B, positive residues are coloured blue, negative are red, polar uncharged are yellow and nonpolar residues are purple. In B, hydrophobic regions are coloured purple and hydrophilic regions are white. The positions of haem-ligating histidines are shown in green and the redox-coupled glutamates are shown in red. AP maquettes readily co-assemble with detergents to form micelles or with lipids to form membranes (C and D, top).

- **A.** The hydrophilic domains of the amphiphilic (AP) maquettes are based on the hydrophilic maquette, HP1, which consists of two antiparallel dihelices; each dihelix is connected by a disulphide bond.
- **B.** The original AP maquette, AP0.
- **C.** The lipophilic and hydrophilic domains of AP1 are connected by a flexible linker.
- **D.** The lipophilic and hydrophilic domains of AP3 are connected directly.

(Figure adapted with permission from Discher *et al*., 2005 and from Ye *et al*., 2004. Copyright American Chemical Society, 2004 and 2005)

There are now several members of the AP maquette family (see Figure 5.1). In the latest designs, porphyrin and chlorin cofactors are able to bind in either the hydrophilic or lipophilic domains (Discher *et al*., 2005; Noy *et al*., 2005). In order to ligate highly hydrophobic metal porphyrins such chlorins within the lipophilic domain, it was necessary to provide a hydrophobic environment around the cofactor binding site. Chlorophylls and bacteriochlorophylls (BChls) (and their metal-substituted analogues) are unable to bind hydrophilic maquettes due to the tendency of these pigments to aggregate or precipitate in aqueous or polar solvents (Noy *et al*., 2005). Haem, and Zn- and Ni- bacteriochlorophylls have been bound within the liphilic domains of AP maquettes, although the binding affinities for BChls are much lower than for haem, which is in the nanomolar range (Noy *et al*., 2005).

More recently, the maquette AP6 was designed to emulate the proton and electron transfer reactions of the *Rba. sphaeroides* cytochrome bc_1 complex (Chobot *et al*., 2010; Hokanson, 2010; Fry *et al*., 2011). AP6 combines the engineering principles learned from previous maquettes to perform quinol-cytochrome *c* oxidoreductase activity with a near natural turnover rate.

There are various design features included in AP family maquettes that facilitate folding, membrane insertion and cofactor binding. Asparagine confers stability when placed at the membrane-solvent interface (Lear *et al*., 2003). Hydrophobic mismatch between the membrane interface and the maquettes can be overcome through the use of charged residues with long and/or aromatic side chains such as lysine or phenylalanine (Strandberg and Killian, 2003). In addition phenylalanine residues are often found near haem binding sites in natural proteins (Koder *et al*., 2009) and lysines are often found in cytoplasmic loops due to their positive charge (von Heijne, 1986). Residues such as lysine, with straight (unbranched) side chains, have high α-helix forming propensities (Chou and Fasman, 1978). Glycine has high conformational flexibility and therefore has a very low helix forming propensity, as such glycine residues are found in the loops connecting the maquette α-helices. Tryptophan residues are incorporated to facilitate optical detection using UV-visible spectroscopy.

AP maquettes have been designed for membrane insertion and to perform transmembrane electron transfer (Noy *et al*., 2005; Discher *et al*., 2005). Amphiphilic maquettes expand the potential of maquettes towards emulating natural membrane proteins, for example man-made terminal oxidase proteins that could create a proton motive force within a living organism. The majority of the published research involving AP maquettes has involved their synthesis and assembly with cofactors *in vitro*. The single chain amphiphilic maquette, AM1 has been expressed in *E. coli* in inclusion bodies, and refolding is necessary after purification

(Bohdana Discher, personal communication). The work performed in this chapter used a modified version of AM1 and aimed to determine how bacteria respond when these synthetic proteins are expressed and whether they can assemble in the native lipid bilayer membrane. This study forms the basis of exploring the possibility of building maquettes into the photosynthetic machinery of *Rba. sphaeroides* as a working component.

5.3 Results

5.3.1 Design of the maquette

The amphiphilic haem-binding maquette (denoted TM) used in this study was based on the single-chain amphiphilic maquette, AM1. AM1 was designed through many iterations by the group of Bodhana Discher and Les Dutton; TM was produced by modifying the net charges of the loops of AM1. Membrane proteins typically feature positively charged cytoplasmic loops and negatively charged periplasmic or extracellular loops and TM was designed by Neil Hunter to satisfy these principles (Figure 5.2 A) (von Heijne, 1986). The differences in sequence between TM and AM1 are shown in Figure 5.2 D. TM is a 21 kDa protein and features a C-terminal His-tag for immunodetection and purification. As with the AP-family maquettes, TM features distinct continuous hydrophilic and lipophilic domains (Figure 5.2 B and C). TM is a 21 kDa protein. The sequence was analysed using the transmembrane topology prediction software MEMSAT3 and MEMSAT-SVM to confirm that the maquette is likely to feature four transmembrane helices (Figure 5.3).

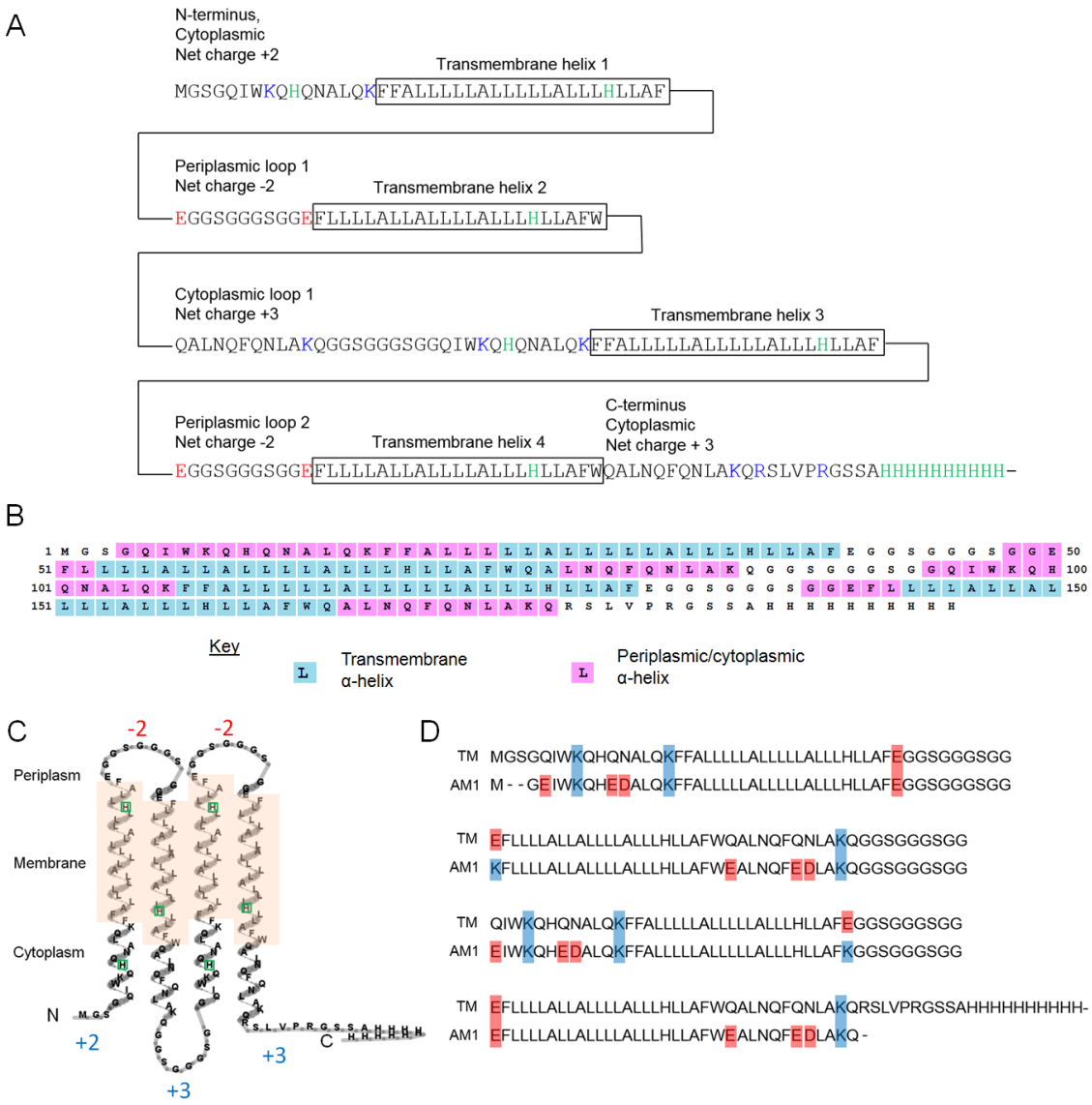


Figure 5.2 The sequence of TM

- **A.** The amino acid sequence of TM with predicted secondary and tertiary structure indicated. Negatively charged residues are in red, positive in blue and the histidines are in green.
- **B.** The amino acid sequence of TM was analysed using the topology prediction software MEMSAT3 and DISOPRED.
- **C.** Schematic of TM showing alpha-helical and membrane (beige) regions with net charges on the periplasmic and cytoplasmic regions. Haem binding histidine residues are highlighted in green.
- **D.** ClustalW2 alignment of TM with AM1. Positively charged residues are in blue, negative in red.

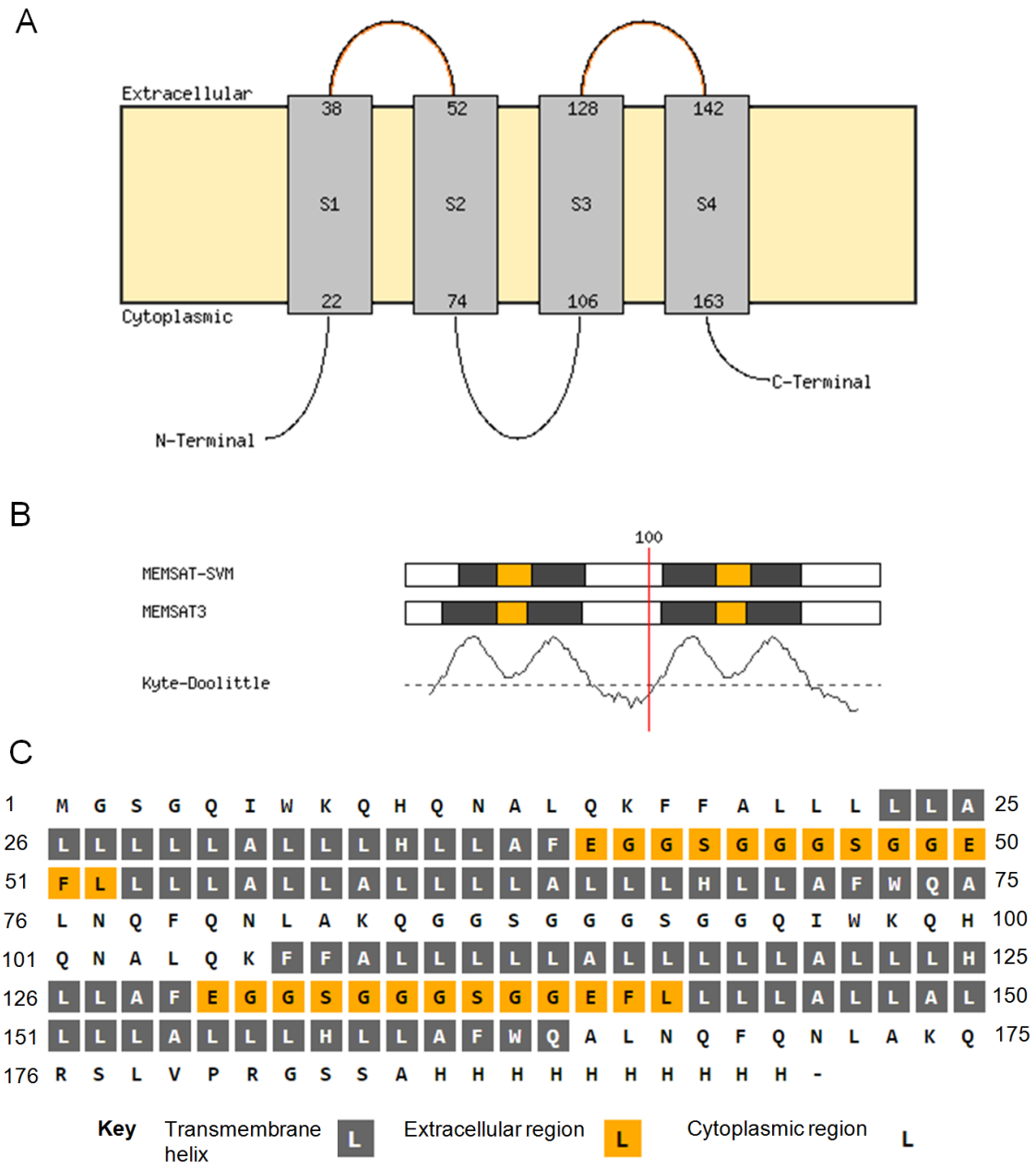


Figure 5.3 Membrane topology prediction for TM

The amino acid sequence of TM was analysed using the transmembrane topology prediction software MEMSAT3 and MEMSAT-SVM.

- **A.** MEMSAT-SVM representation of the predicted 4 transmembrane helices of TM. Both the N- and C-termini are predicted to be cytoplasmic.
- **B.** Schematic of the transmembrane model predicted by both MEMSAT-SVM and MEMSAT3. Kyte-Doolittle hydropathy plot demonstrating the presence of regions of high hydrophobicity. Four transmembrane helices are observed.
- **C.** Amino acid sequence of TM showing the presence of the transmembrane helices and location of the inter-helix loops.

The gene for the expression of TM was designed with codons optimised for expression in *Rba. sphaeroides*. However, due to constraints imposed by the DNA synthesis company, Bio Basic

Inc., compromises had to be made to lower the GC content and repetition of the gene. This was not anticipated to be a problem as *Rba. sphaeroides* is reasonably tolerant of rare codons (Laible *et al*., 2009).

5.3.2 Expression of TM in *Rba. sphaeroides* membranes

To determine to what extent the charge of the cytoplasmic loops affects expression of TM, two variants were expressed in *Rba. sphaeroides*. The first variant contained only a C-terminal His tag and the second variant, denoted TM-FLAG, also contained an N-terminal FLAG tag. The FLAG tag conferred a significant negative charge onto the cytoplasmic N-terminus of the protein (Figure 5.4).

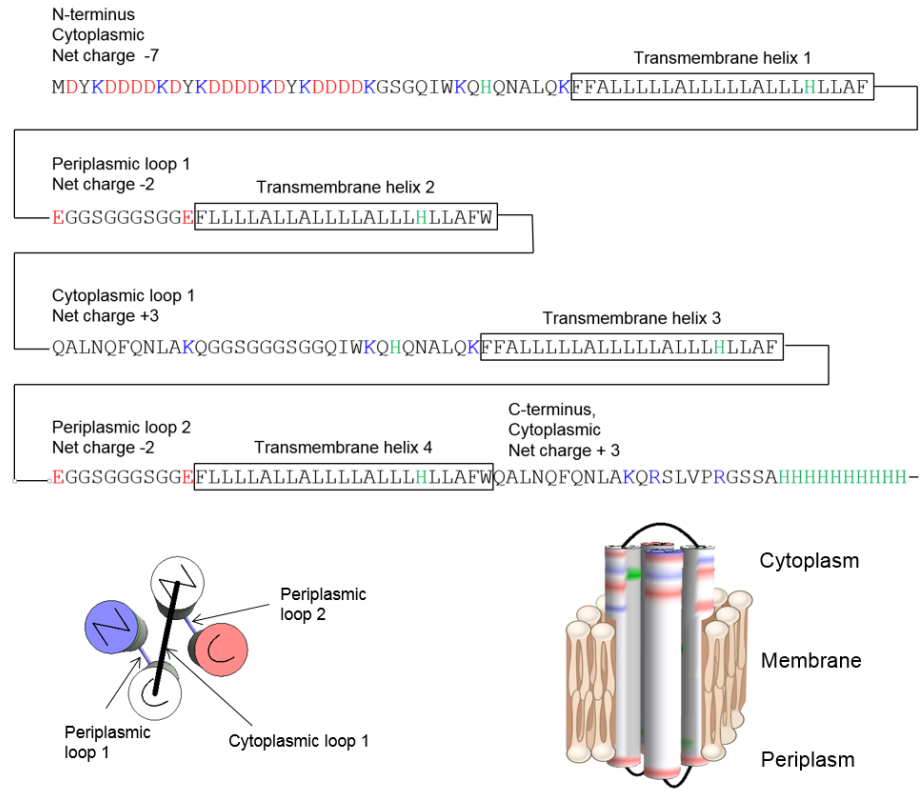


Figure 5.4 The sequence of TM with an N-terminal FLAG tag

The amino acid sequence of TM with predicted secondary and tertiary structure is indicated. Negatively charged residues are in red, positive in blue and the histidines are in green.

The synthesised DNA contained the TM gene including a region encoding an N-terminal FLAG tag and was cut with *Nco*I and *Hin*dIII. The DNA fragment was ligated into pIND4 to produce pIND4-TM-FLAG. The plasmid was transferred to *Rba. sphaeroides* Δ*puc1BA* Δ*pufX* via conjugation. This strain was chosen due to the findings of Chapter 4 (Section 4.3.8) in which proteorhodopsin expressed best in this strain. The expression of pIND4-TM-FLAG in *Rba.*

sphaeroides yielded only low levels of protein (Figure 5.5, pIND4-TM FLAG). The FLAG tag was removed by amplifying TM from pIND4-TM-FLAG using the primers "TM no FLAG F" and "TM R *Hin*dIII" to create pIND4-TM. It was found that expression of TM was increased when the FLAG tag was removed (Figure 5.5).

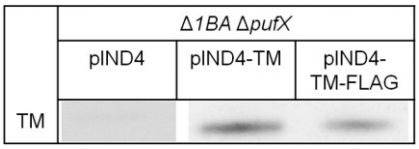


Figure 5.5 Expression of TM with and without the N-terminal FLAG tag

A 10 ml starter culture of each was grown for 2 days and used to inoculate 80 ml of M22 in a 120 ml flask. The 80 ml cultures were grown semi-aerobically for 6 hours at 34 °C before induction with 1 mM IPTG for 16 hours. Equal amounts of whole cells were separated by SDS-PAGE followed by transfer to a nitrocellulose membrane. Immunoblotting was performed using an anti-His antibody to detect the C-terminal His tag found on both TM versions.

5.3.3 Expression of TM in *E. coli*

Prior to further expression and analysis in *Rba. sphaeroides*, TM was expressed in *E. coli* to ensure that the gene product assembled in the membrane with haem bound as designed.

E. coli BL21(DE3) pLysS cells were transformed with pIND4-TM and induced with 1 mM IPTG, but only very low levels of TM were detected by western blot (Figure 5.6). However, good levels of TM expression were achieved through growth in autoinduction media (Figure 5.6).

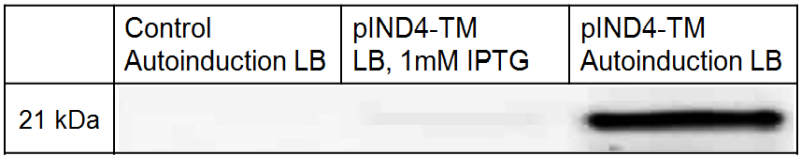


Figure 5.6 Expression of TM in *E. coli* BL21

E. coli BL21 cells were transformed with pIND4-TM. 20 ml of overnight starter culture were used to inoculate 500 ml of LB or autoinduction media in a 1.5 L flask. Cells were grown at 37 °C and agitated at 250 rpm for 1 hour followed by the addition of 1 mM IPTG where relevant and growth at 20 °C with shaking at 150 rpm for 16 hours. Cells were pelleted by centrifugation and resuspended in 20 mM Tris, 0.5 mM EDTA, pH 7.4. Equal amounts of whole cells were separated by SDS page, transferred to a nitrocellulose membrane and immunoblotted using an anti-His antibody to detect the His tag on TM.

5.3.4 Purification of TM from *E. coli* membranes

To determine whether the maquette was located in the membrane fraction, membranes were prepared from *E. coli* cells producing TM using sucrose gradients according to the

method detailed in Section 2.11 (Figure 5.7 B). Western blot analysis of the membrane fractions showed a stronger signal corresponding to TM in the outer membrane fraction (Figure 5.7 A). This is an unexpected result and it is likely that the outer membrane fraction is contaminated with inner membrane.

To maximise TM yield during purification the inner and outer membranes were pooled, pelleted and solubilised according to the method described in Section 2.11. The clarified membrane sample was loaded onto an Ni NTA column, washed to remove any contaminating proteins and eluted with 400 mM imidazole. All purification buffers contained 0.1 % LDAO. The elution fractions were analysed using SDS-PAGE with the gel stained with Coomassie; a band corresponding in size to TM was observed (Figure 5.7 C). The elution fractions were pooled and concentrated but the sample was not coloured indicating that TM did not have any pigment bound.

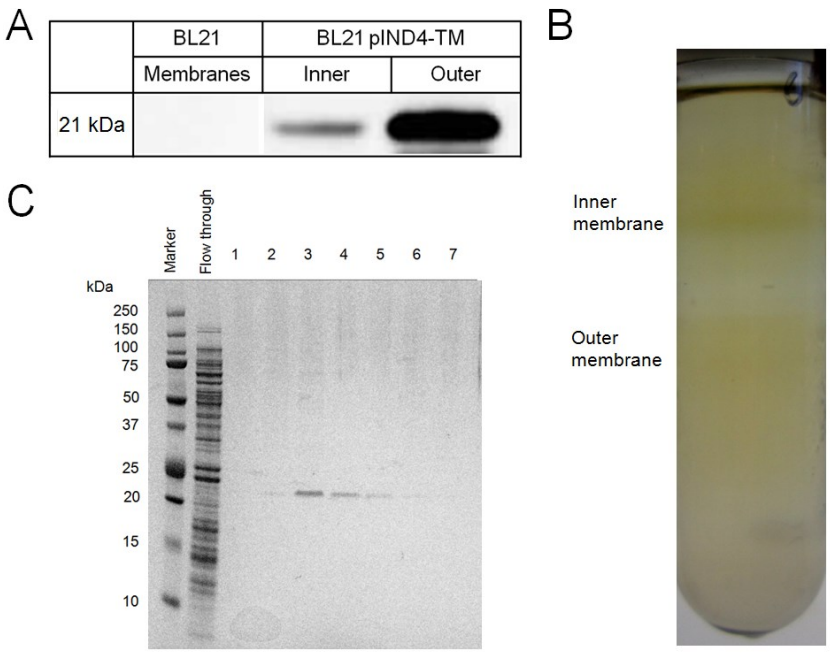


Figure 5.7 Purification of TM from *E. coli* membranes

E. coli BL21(DE3) pLysS cells were transformed with pIND4-TM. Protein expression was induced through the use of autoinduction media at 20 °C with shaking at 150 rpm for 16 hours. Membranes were prepared on a 55-30% sucrose step gradient. TM was purified using an Ni-NTA affinity column.

- **A.** Western blot analysis of total membranes purified from BL21(DE3) pLysS and inner and outer membranes purified from BL21(DE3) pLysS expressing TM. Equal amounts of protein were loaded onto each.
- **B.** Sucrose gradient separation of *E. coli* inner and outer membranes.
- **C.** SDS-PAGE analysis of material that did not bind to the Ni-NTA column and elution fractions 1-7, showing the maquette at 21 kDa.

5.3.5 Supplementation of *E. coli* with ALA and iron

To increase the amount of haem available in the cell for TM to bind, *E. coli* cells were supplemented with 500 µM D-aminolevulinic acid (ALA, a precursor in the haem biosynthesis pathway) and 50 µM ferrous chloride. To achieve higher expression levels of TM, the gene was amplified using the primers "TM F pET9a" and "TM R pET9a" and ligated into pET9a to create pET9a-TM. pET9a-TM was used to transform *E. coli* BL21(DE3) cells. The cells were grown, harvested and membranes were prepared as described in Section 2.11. As a control BL21(DE3) cells were grown without pET9a-TM, with and without supplementation with ALA and iron.

Whole cells expressing TM appear slightly pink, and when ALA and Fe are supplemented the cells appear dark red (Figure 5.8 A, bottom panels). Cells not expressing TM appear a similar colour with and without supplementation. The outer membranes were darker in the supplemented cells, and free haem was present at the top of the sucrose gradient in both cases (Figure 5.8 A).

The absorbance spectra of membranes obtained from cells without TM are very similar and feature absorbance maxima between 415-422 nm in all cases (Figure 5.8 B). The absorbance spectrum of the inner membrane from cells expressing TM shows peaks characteristic of free haem (Figure 5.8 D). Spectra of the outer membrane fraction appear similar in all cases (Figure 5.8 B-D). Western blot analysis of the membrane fractions showed that TM is present only in the outer membrane fraction of cells both with and without supplementation (Figure 5.8 C and D, inset).

After harvesting, the membranes collected from cells expressing TM and supplemented with ALA and iron were pelleted by centrifugation at 40,000 rpm for 2.5 hours in a Beckman Ti 45 rotor. The inner membrane pellet was very small with weak colouration, indicating that the majority of the red colour came from free haem. The outer membrane pellet was larger and dark brown (data not shown).

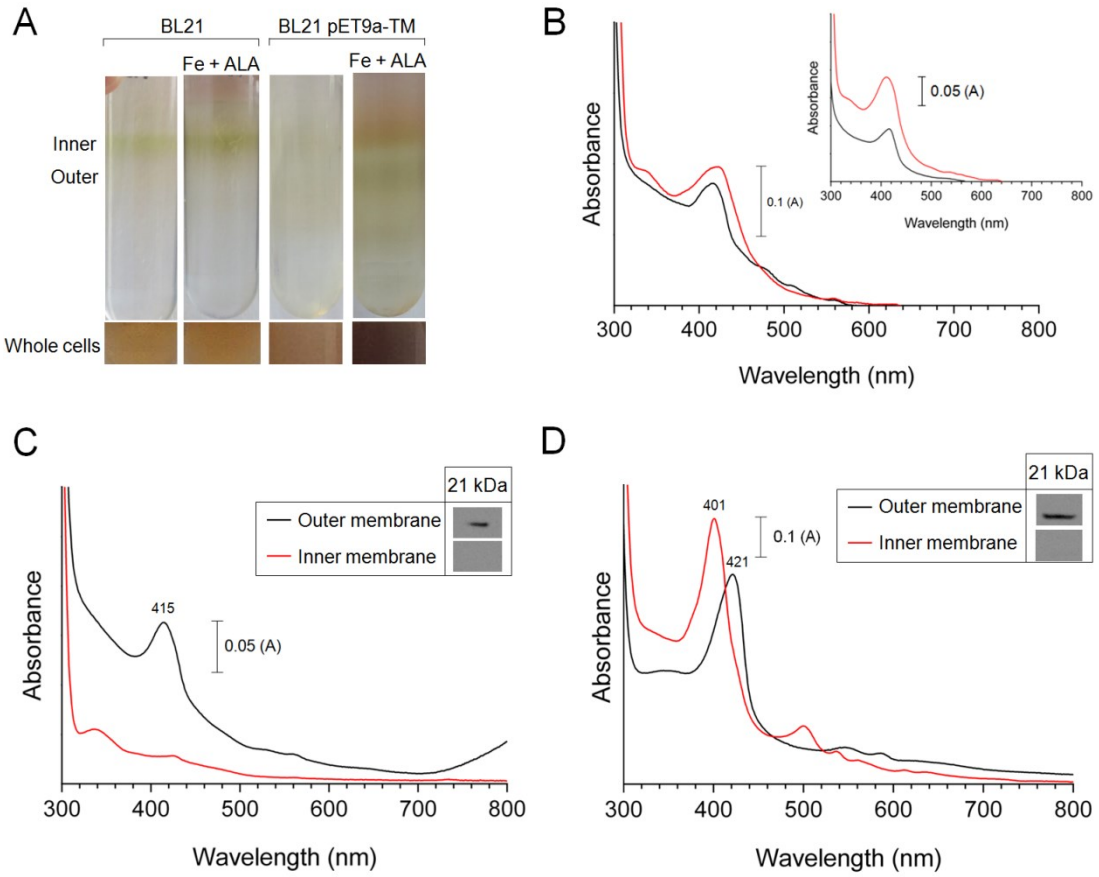


Figure 5.8 Expression of TM in *E. coli* supplemented with ALA and iron

TM was expressed from pET9a-TM in BL21(DE3). 5 ml of an overnight starter culture was used to inoculate 500 ml of LB. Cells were induced with 1 mM IPTG for 48 hours at 30 °C with shaking at 250 rpm. Where indicated, cells were supplemented with 500 µM D-aminolevulinic acid (ALA) and 50 µM ferrous chloride.

- **A.** Photographs of whole cells, and sucrose gradients onto which the broken cell material was loaded. The locations of the inner and outer membrane fractions are indicated.
- **B.** Absorbance spectra of the inner and outer membranes prepared from *E. coli* not expressing TM. Without supplementation, and (inset) with ALA and iron.
- **C.** *E. coli* membranes expressing TM without supplementation. The inset shows western blot analysis of membrane fractions, immunoblotted with an anti-His primary antibody.
- **D.** *E. coli* membranes expressing TM with supplementation with ALA and iron. The inset shows western blot analysis of membrane fractions, immunoblotted with an anti-His primary antibody.

5.3.6 Expression in *Rba. sphaeroides* strains that contain membranes of varying curvature

Previous studies have found that *Rba. sphaeroides* strains unable to assemble various photosynthetic complexes expressed higher levels of foreign membrane proteins (Laible *et al.*, 2009; Tikh *et al.*, 2014). The data obtained in Chapter 4 regarding the expression of proteorhodopsin (Section 4.3.8) supported these findings and indicated that a more planar

membrane environment facilitated protein expression. TM, being made up of only of 4 transmembrane α-helices, may also prefer to sit in flat membranes. TM was expressed from pIND4-TM in various *Rba. sphaeroides* backgrounds with differing membrane morphologies. Expression was significantly better in a *ΔpufX* background (Figure 5.9).

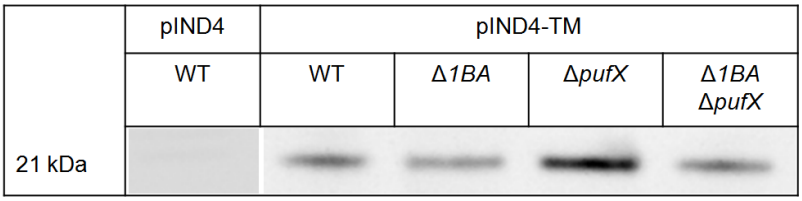


Figure 5.9 Expression of TM in *Rba. sphaeroides* strains with membranes of varying curvature

A 10 ml starter culture of each strain was grown for 2 days and used to inoculate 80 ml of M22 in a 120 ml flask. The 80 ml cultures were grown semi-aerobically for 6 hours at 34 °C before induction with 1mM IPTG for 16 hours. Equal amounts of whole cells were separated by SDS-PAGE, blotted onto a nitrocellulose membrane and immunoblotted with an anti-His antibody to detect the C-terminal His tag on TM.

5.3.7 Expression of TM from the genome of *Rba. sphaeroides*

The results presented in Section 5.3.6 indicate that the expression of the plasmid-borne TM gene yields higher levels of TM protein in a *Rba. sphaeroides* Δ*pufX* strain that makes larger, less curved membranes. In order to investigate whether genomic integration of the TM gene could improve expression levels, TM was introduced into the *Rba. sphaeroides* Δ*pufX* genome in place of the *puc1BA* genes, encoding the LH2 subunits. Despite the fact that TM appeared to express from pIND4-TM at lower levels in Δ*1BA* Δ*pufX* compared to Δ*pufX*, the placement of TM at the *puc1BA* locus was chosen as it has a strong promoter.

A pK18mob*sacB* construct was created containing a 1527 bp fragment containing the region upstream of the *puc1BA* genes, the TM gene and the region downstream of the *puc1BA* genes. These fragments were fused together using overlap extension PCR. TM, along with a C-terminal his tag, was amplified from the plasmid pIND4-TM using the primers "TM OE F" and "TM OE R" to create a 619 bp fragment. A 410 bp region of DNA upstream of the *puc1BA* genes was amplified using the primers "PucBA KO F *Eco*RI" and "TM up OE R". A 469 bp region of DNA downstream of the *puc1BA* genes was amplified using the primers "TM down OE F" and "PucBA KO DR". The fragment was cut with *Eco*RI and *Hin*dIII and ligated into pK18mob*sacB* to create the plasmid pK18mob*sacB*-TM. The plasmid was transferred to *Rba. sphaeroides* Δ*pufX* via conjugation and the mutant *Rba. sphaeroides* Δ*pufX* Δ*1BA*::TM was created according to the method described in Sections 2.13-14 and 3.3.2.

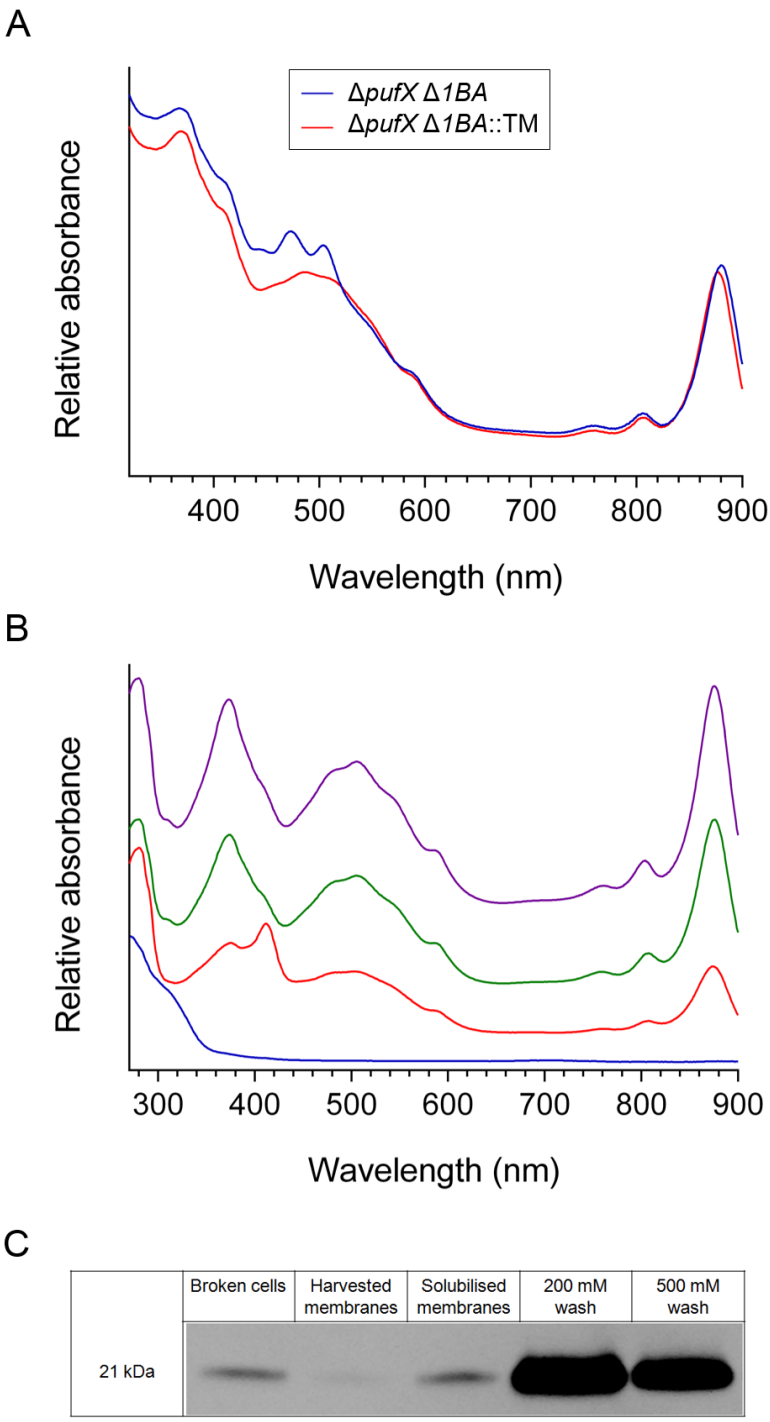


Figure 5.10 Purification of TM from Δ*pufX* Δ*1BA*::TM

- **A.** Absorbance spectra of ICM purified from Δ*pufX* Δ*1BA* and Δ*pufX* Δ*1BA*::TM
- **B.** Absorbance spectra of various stages of the purification process.
  Purple - Solubilised ICM
  Green - Flow through
  Red - 200 mM imidazole wash
  Blue - 500 mM imidazole elution
- **C.** Western blot analysis of the stages in the purification process.

An 8 L photosynthetic culture of *Rba. sphaeroides* Δ*pufX* Δ*1BA*::TM was grown, pelletted and French pressed. The membranes were prepared as described in Section 2.15 and resuspended in 20 mM HEPES pH 7.4. Other than differences in the 450-500 nm region, the absorbance spectrum of ICM prepared from Δ*pufX* Δ*1BA*::TM contains no peaks that are not seen in ICM prepared from Δ*pufX* Δ*1BA* (Figure 5.10 A). The membranes were solubilised as in Section 2.15.3. The solubilised membranes were applied to a Ni NTA column, which was washed to remove any contaminating proteins and eluted with 500 mM imidazole. The elution fractions were concentrated but appeared colourless, but a peak at 420 nm was enriched in the 200 mM wash (Figure 5.10 B). Western blot analysis showed that TM was present in all fractions including a wash with 200 mM imidazole prior to elution (Figure 5.10 C). Protein levels were too low to be seen on a Coomassie stained SDS-PAGE gel (data not shown).

5.3.8 Expression of TM in a *Rba. sphaeroides* strain that accumulates bacteriochlorophyll precursors

TM may be able to bind porphyrins other than haem. To test this hypothesis, TM was expressed in *Rba. sphaeroides* Δ*bchCFX* (Δ*CFX*) which accumulates the BChl precursors Mg-monovinyl protochlorophyllide *a* and chlorophyllide *a* (see Section 1.3.2). Δ*CFX* cannot grow photosynthetically as it lacks BChl and cannot produce the LH1, LH2 or RC complexes.

The TM gene was placed onto the genome of this strain according to the method described previously to produce Δ*CFX* Δ*1BA*::TM. This strain was grown semi-aerobically and ICM was prepared on sucrose gradients as described previously in buffer containing 20 mM Tris, 500 mM NaCl, 20 mM imidazole. The membranes from Δ*CFX* Δ*1BA*::TM appeared more red/orange than those from Δ*CFX*, the membranes also appeared aggregated (Figure 5.11 A). A peak appeared in the absorbance spectrum at 388 nm that is absent in the absorbance spectrum of ICM from Δ*CFX* (Figure 5.11 B).

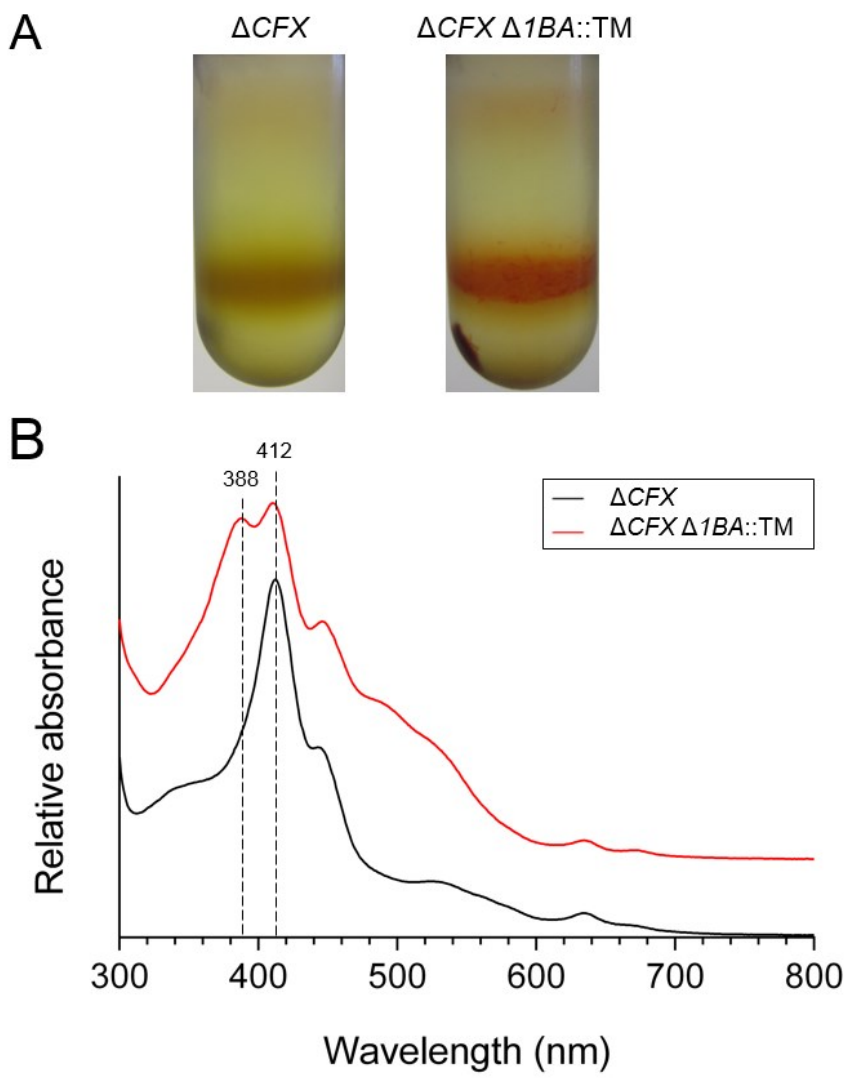


Figure 5.11 ICM prepared from Δ*CFX* and Δ*CFX* Δ*1BA*::TM

- **A.** Sucrose gradients to isolate ICM from Δ*CFX* and Δ*CFX* Δ*1BA*:TM grown semi-aerobically.
- **B.** Absorbance spectra of ICM purified from Δ*CFX* (black) and Δ*CFX* Δ*1BA*:TM (red). Peak maxima in nm are indicated.

Membranes were solubilised as described in Section 2.15.3. The solubilised membranes were applied to a Ni NTA column which was washed with increasing concentrations of imidazole to remove any contaminating proteins and protein eluted with 500 mM imidazole. On solubilisation, the peak that was not observed in Δ*CFX* ICM at 388 nm shifted to 384 nm (Figure 5.12 A). This peak was present in the flow through fraction that did not bind to the column (Figure 5.12 A); western blot with an anti-His antibody did not show a signal in this fraction, suggesting that the 384 nm peak does not correspond to TM (Figure 5.12 B). A peak at 414 nm was seen in the elution fraction, and western blot analysis confirmed that TM was present in this fraction (Figure 5.12).

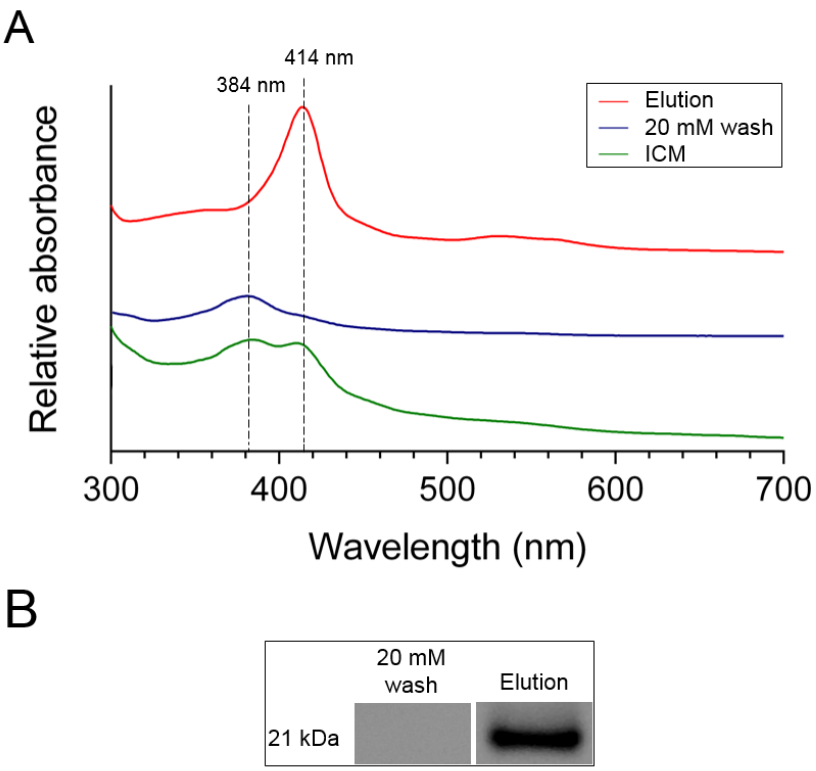


Figure 5.12 Purification of TM from Δ*CFX* Δ*1BA*::TM

ICM was prepared from semi-aerobically grown cells, solubilised in 3 % β-DDM and applied to an Ni NTA column.

- **A.** Absorbance spectra of the various stages of purification
- **B.** Western blot analysis of the flow through and elution fractions using an anti-His antibody.

5.4 Discussion

5.4.1 Charge patterning across the maquette is important for insertion into membranes

The design for TM incorporates many of the design principles found in natural membrane proteins. One of the most important of these features is that for correct insertion and orientation in the membrane, any cytoplasmic portions must be predominantly positively charged and any periplasmic or extracellular regions must be negatively charged (von Heijne, 1986). The introduction of an N-terminal FLAG tag to TM violated this rule and introduced a large negative charge to the cytoplasmic N-terminus. On removal of the FLAG tag, maquette expression levels in *Rba. sphaeroides* increased (Figure 5.5). It is possible that any maquette that did not sit in the membrane properly was degraded, resulting in the decreased appearance of TM when expressed from pIND4-TM-FLAG.

5.4.2 TM sits in the membrane of *E. coli*

TM was found predominantly in the outer membrane fraction of *E. coli* (Figures 5.7 A and 5.8 C and D, inset). This was unexpected as the maquette has no targeting sequence to the outer membrane. It is possible that the band harvested as the outer membrane also contained unseparated cell envelope material (Spencer and Guest, 1974), or was contaminated with inner membrane. To test this, future work should involve western blotting of each membrane fraction with an antibody to a protein unique to the inner or outer membrane. TM was purified from the pooled membranes, however it appeared that no pigment was bound (Figure 5.7). It is likely that the pigment was lost during purification as haem B is ligated non-covalently. Until recently, maquettes have been synthesized and assembled *in vitro*, or expressed as apoproteins in *E. coli* requiring *in vitro* addition of the selected cofactor (Robertson *et al*., 1994; Farid *et al*., 2013). Currently, there is no maquette expressed in *E. coli* that 100% binds haem B, despite nanomolar binding affinities *in vitro*. This is possibly due to the unfolding of the maquette in the crowded *E. coli* cytoplasm, or insufficient stimulation of the haem B biosynthetic pathway (Farid *et al*., 2013; Watkins *et al*., 2014). With the supplementation of the growth medium with ALA and iron to stimulate the haem biosynthesis pathway it is possible to purify soluble maquettes with some haem bound although much haem is lost during the purification process and it is necessary to reconstitute the protein by adding haem *in vitro* (data not shown, and Goutham Kodali, personal communication).

It was not possible to definitively conclude that membrane-bound TM binds pigment in *E. coil* as absorbance spectra of the inner and outer membranes feature a peak possibly corresponding to a native cytochrome in the region that haem-bound TM would be expected to appear (Figures 5.8 B-D and 5.13). The red colour of cells expressing TM, particularly when supplemented with ALA and Fe, is evidence to suggest that TM can bind haem (Figure 5.8 A, bottom panels).

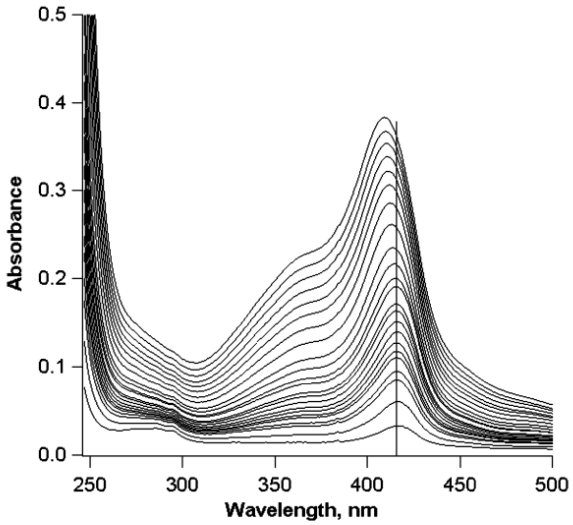


Figure 5.13 Absorbance spectra of the amphiphilic maquette AP0

Titration of hemin into a 1.0 µM solution of AP0 recorded in a 1 cm path length cuvette. The spectra shown contain increasing amounts of added hemin. The vertical line at peak 412 nm indicates the blue shift of the peak, due to contribution of free haem (400 nm) in the solution.

(Figure reproduced with permission from Ye *et al*., 2004. Copyright American Chemical Society, 2004.)

5.4.3 TM prefers to insert into a flatter membrane environment

The photosynthetic membrane of *Rba. sphaeroides* is highly curved, and although the LH2 complex preferentially inserts into this membrane (Adams *et al*., 2011) this environment may not favour the expression of foreign membrane proteins with no intrinsic curvature. There is also the issue of crowding to consider; membrane vesicles from wild-type cells appear to have no empty areas to accommodate foreign membrane proteins (Cartron *et al*., 2014). Previous studies have shown that foreign membrane protein expression is improved with the knockout of various native complexes (Laible *et al*, 2009; Tikh *et al*., 2014). This was also found to be the case in Chapter 4 in which proteorhodopsin expressed in greater amounts in Δ*pufX* Δ*1BA* and Δ*pufX* backgrounds (Section 4.3.8). Figure 5.9 shows evidence for increased TM expression in a Δ*pufX* background which features large, flat membranes (Siebert *et al*., 2004), although the absence of LH2 complexes appeared to have no beneficial effect. One explanation could arise from the work of Adams *et al*., 2011; the absence of PufX already lowers LH2 levels because of the more planar membranes are unfavourable for LH2 insertion, so there is ample room for inserting TM, especially given the low expression levels. Creating even more membrane area by removing LH2 altogether would not improve TM levels.

5.4.4 TM is found in the ICM of *Rba. sphaeroides*

TM was successfully expressed from the genome of *Rba. sphaeroides* Δ*pufX* Δ*1BA*::TM (Figure 5.10). Overlap between the expected absorbance of haem-bound TM (Figure 5.13) and the BChl Soret absorbance peak obscures differences in the absorbance spectrum of ICM with and without TM, but they can be seen in 400-500 nm region in Figure 5.10 A. TM was found in the membrane fraction of these cells (Figure 5.10 C). The 200 mM elution fraction was contaminated with native *Rba. sphaeroides* proteins and pigments, making it difficult to assign a peak to TM (Figure 5.10 B), but the strong peak at ~420 nm is similar to that seen in the *E. coli* experiment in Figure 5.8 D. TM was present in the 500 mM elution fraction and the absorbance spectrum revealed that it was free of contaminants, however no peak was observed that was likely to correspond to pigment-bound maquette (Figure 5.10 B and C). Possibly, the very high concentration of imidazole used for elution displaced haems from their binding sites within TM.

5.4.5 TM pigment binding in *Rba. sphaeroides*

The fact that TM could not be purified with bound pigment could either indicate that any bound pigment is lost during the purification process or that the maquette was unable to bind pigment. As well as haem, amphiphilic maquettes have been shown to be able to bind Zn- and Ni- BChls (Discher *et al*., 2005; Noy *et al*., 2005). The central ion of BChl is a coordinately unsaturated Mg^{2+}. Chlorophyll-binding proteins use a polar protein side chain, often from histidine, to provide an extra ligand to coordinate the Mg^{2+}, making it penta-coordinate (although hexa-coordinate pigments have also been observed) (Haehnel *et al*., 2009). Therefore, it is possible that TM may be able to bind this pigment via its histidine residues. However, there is evidence to suggest that chlorophyll requires a specific sequence motif in order to bind a maquette (Eggink and Hoober, 2000).

WT *Rba. sphaeroides* may not accumulate sufficient BChl for TM to bind, and there could be complications arising from the presence of the long phytyl tail, so TM was introduced into a strain with a truncated chlorophyll biosynthesis pathway. The Δ*bchCFX* mutant accumulates the BChl precursors Mg-monovinyl protochlorophyllide *a* and chlorophyllide *a*. The ICM prepared from Δ*CFX* Δ*1BA*::TM was significantly redder in colour than Δ*CFX*, which provides compelling evidence for TM expression and pigment binding (Figure 5.11 A). The ICM fraction on the sucrose gradient appeared aggregated, which could indicate contamination with other cell debris, or it could be that TM caused membrane stacking due to interaction between the positively charged cytoplasmic regions and negatively charged periplasmic regions. A peak in the ICM absorbance spectrum of Δ*CFX* Δ*1BA*::TM at 388 nm was not visible in ICM prepared

from Δ*CFX* (Figure 5.11 B). However, data from an attempted purification of the maquette suggested that the 388 nm peak was not TM (Figure 5.12). It is possible that any bound pigment is lost during the purification process, or that the maquette absorbance spectrum overlaps with the native 412 nm peak.

5.4.6 Additional work

It has not been possible to definitively determine whether TM is able to bind pigment *in vivo*. Native gel analysis of membranes prepared from strains expressing TM may provide an answer to this. Titration of hemin into the purified maquette may determine whether haem is able to bind. Future purification protocols should involve gentle washes and elution with a gradient of increasing imidazole concentration.

For increased understanding of the effect of membrane curvature and the negative charge introduced by the FLAG tag on the expression of TM, membranes should be prepared to discount any signal contribution from apo-maquette that had not assembled in the membrane.

The original TM design features extensions of the transmembrane helices into the aqueous phase, into the cytoplasmic side of the membrane. This domain is hydrophilic and is a feature common to the AP family of maquettes. This hydrophilic domain provided an amphiphilic character to the maquettes which was important for vectorial orientation of the maquette within a membrane. This domain also features a third haem binding site, designed to participate in electron transfer across the membrane; it is also important for stability and assembly (Bodhana Discher, personal communication). This hydrophilic domain has been removed from TM (with thanks to Andrew Hitchcock). Future work involves investigating whether the charges on the periplasmic and cytoplasmic regions of the maquette would be sufficient for membrane insertion in the correct orientation.

Ultimately, the design of TM could be modified to replicate light harvesting and electron transfer roles *in vivo*.

5.4.7 Conclusion

The gene for an amphiphilic transmembrane maquette (TM) was designed and expressed in *E. coli* and *Rba. sphaeroides*. The data presented in this chapter show that TM is able to assemble in to the membrane fraction of both organisms. The design of the maquette takes into consideration the design principles of natural membrane proteins and demonstrates that the charge on the cytoplasmic N-terminus is very important for protein expression. There is evidence to suggest that TM is able to bind pigment although, when expressed in *Rba.*

sphaeroides, it was not possible to determine whether the maquette bound haem or a chlorophyll precursor. Further work is required to investigate the pigment-binding properties of the maquette. This work forms the basis of the bottom-up redesign of photosystem components which could be used to replace existing proteins in vivo or emulate their function in artificial membrane systems.

Chapter 6

Tat-mediated recognition of the folded state of a wholly synthetic protein and its export by the *Escherichia coli* Tat transporter

6.1 Summary

The twin arginine translocation (Tat) system will only export fully, or largely, folded proteins from the cytoplasm to the periplasm. It is currently unclear how the quality control mechanism of the Tat pathway works. The plasticity of the *Escherichia coli* Tat system was investigated through the use of synthetic protein maquettes. A Tat signal peptide, TorA, was fused to the N-terminus of the soluble bis-haem binding maquette, BT6. The Tat system recognises and exports BT6 to the periplasm, during translocation the TorA signal peptide is cleaved.

Maquette residues were selectively mutated to modify the cofactor binding, and thus the folding properties, of the maquette to test the Tat quality control system. The Tat system will not export the maquette without the haem cofactors bound, likely due to loss of ordered tertiary structure. This work demonstrates that the Tat system is able to export wholly synthetic proteins never seen before by nature. The quality control property of the Tat pathway makes it a desirable system for efficient large-scale protein production. Ultimately, maquettes could be exported into the growth medium, facilitating purification. This strategy offers the potential of maximising protein yield as build-up of protein and possible degradation inside the cell are avoided.

6.2 Introduction

The transport of proteins across membranes is one of the great challenges faced by the cell. In prokaryotes, two major pathways are used to achieve protein translocation across the cytoplasmic (inner) membrane: the secretory (Sec) pathway and the twin-arginine translocation (Tat) pathway. The Sec pathway transports proteins in an unfolded configuration using energy provided by ATP hydrolysis and the transmembrane proton gradient. In contrast, the Tat pathway transports fully folded proteins using only the transmembrane proton gradient. The Tat pathway is found in bacteria, archaea and plant chloroplasts.

A range of proteins are transported by the Tat pathway, including those that fold too quickly for Sec transport (Palmer and Berks, 2012), those that assemble into oligomeric complexes (Rodrigue *et al*., 1999), and those that bind a series of redox cofactors in the cytoplasm (Berks 1996). The final group, proteins containing redox cofactors, constitute the majority of Tat substrates. Any pathway involved in aerobic respiration or photosynthesis which involves the cytochrome bc_1 complex (the b_6f complex in plastids) is dependent on the Tat system to export the cofactor-containing Rieske subunit (Molik *et al*., 2001; Bachmann *et al*., 2006; De Buck *et al*., 2007). Cytoplasmic enzymes insert a variety of cofactors into these complex proteins which then need to be transported in a folded or largely folded form. How the Tat pathway knows whether substrates are folded correctly prior to export is currently not well understood.

Proteins translocated by the Tat pathway have an N-terminal signal sequence characterised by a twin arginine motif. The signal sequence is cleaved from the precursor protein during or immediately after translocation, liberating the mature protein into the periplasm. One of the most commonly used Tat signal peptides for investigating the Tat transporter is that of trimethylamine N-oxide reductase (TorA) from *E. coli*. Use of the TorA signal peptide often results in high efficiency of export of its substrate (Matos *et al*., 2012).

In *E. coli* the minimum set of components for Tat translocase assembly comprises TatA, TatB and TatC. The general consensus is that the substrate binds initially to TatBC, independent of other Tat components (Mori and Cline, 2002). The TatA complex then associates with the TatBC-substrate complex in the presence of a pH gradient across the membrane (Cline and Mori, 2001). There are two current models for the mechanism of translocation: the translocation pore, or trap door, model and the membrane destabilisation model; these are discussed further in Section 1.12.6 (Hauer *et al*., 2013).

The ability of the Tat system to transport fully folded proteins raises the question of whether a quality control or proofreading mechanism exists to prevent futile export of misfolded or misassembled proteins. Many studies have tried to address this question. DeLisa *et al*. have shown that several proteins that feature disulphide bonds in their native state are not exported by Tat in the absence of cytoplasmic disulphide bond formation (DeLisa *et al*., 2003). *E. coli* strains, termed CyDisCo, that oxidise disulphide bonds in the cytoplasm permit the Tat export of disulphide-bond containing proteins (Matos *et al*., 2014). In contrast, it has been found that some proteins that acquire their disulphide bonds in the periplasm can be exported via Tat, suggesting that the system can export reduced or unfolded proteins if they adopt near-native structures (Alanen *et al*., 2015). Richter *et al*. found that the presence of hydrophobic surface patches aborts transport (Richter *et al*., 2007). Various studies have shown that there are chaperone or binding molecules that prevent export of a protein until cofactor binding has taken place (Oresnik *et al*., 2001; Ray *et al*., 2003; Jack *et al*., 2004). Even though advances have been made in understanding the quality control mechanism of the Tat system, it is not currently clear how this mechanism would work.

In order to help clarify the function of the Tat translocase in proofreading of its substrates in this study the export of a synthetic protein maquette was investigated. While it is known that the *E. coli* Tat machinery can process foreign proteins, such as GFP (Thomas *et al*., 2001), it is a new challenge to investigate whether the system will export synthetic proteins never before seen by nature.

Maquettes are man-made proteins designed from the bottom up while avoiding natural amino acid sequences or motifs. Undesirable complexity is minimised to increase engineering freedom; their simplicity facilitates the easy modification of residues to confer new properties or function. Maquettes are discussed further in Section 1.11. Maquettes have been engineered to bind a wide range of redox-active cofactors. The maquette used in this work is termed BT6 (Figure 6.1). BT6 is a bundle of 4 α-helices and binds two haem *b* molecules, each coordinated by two histidine residues within the helices. BT6 can be expressed in *E. coli*; expression can be boosted through the addition of 5-aminolevulinic acid (ALA) and iron to increase haem production, although without supplementation the majority of expressed maquette is able to scavenge enough haem to bind at both sites (G. Kodali, personal communication). Like many Tat substrates, maquettes must assemble in the cytoplasm with the cofactor and be exported without the loss of the bound cofactor.

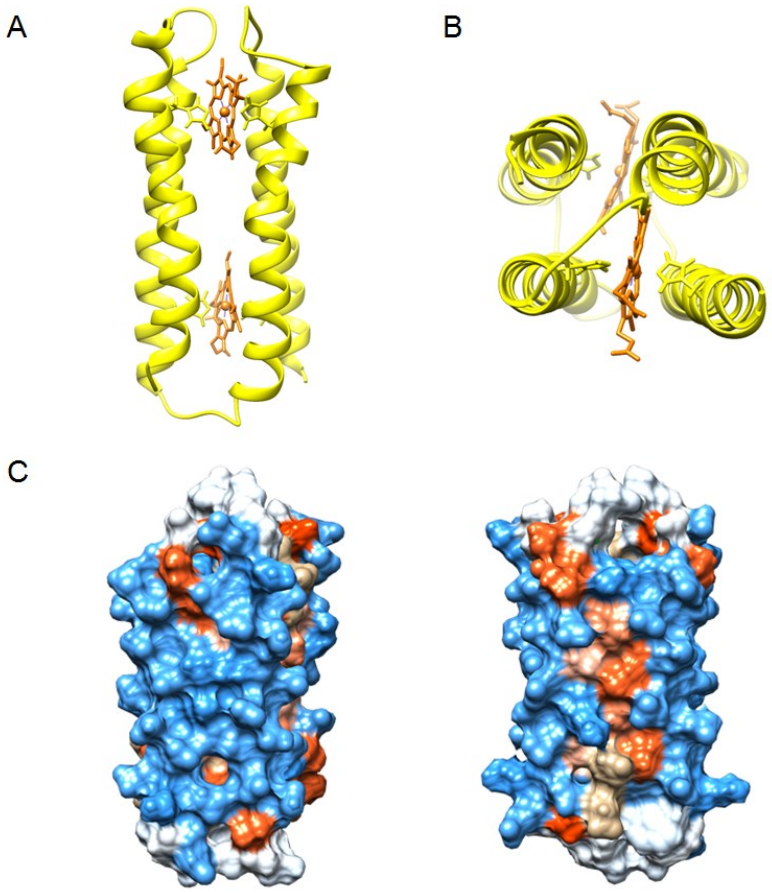


Figure 6.1 The structure of BT6

Model of BT6 based molecular dynamics trajectory simulation, provided by Bryan Fry, University of Pennsylvania.

- **A.** Side view showing protein backbone in yellow with the haem coordinating histidines shown. The two haems are in orange.
- **B.** Top view.
- **C.** Surface hydrophobicity plots determined using the Kyte-Doolittle scale (blue – most hydrophilic, red – most hydrophobic).

As well as the opportunity to gain new insights into the Tat system, the ability to export maquettes has advantages for purification and downstream study. Export of recombinant protein to the growth medium outside the cell would simplify purification. The quality control property of the Tat pathway makes it a desirable system for efficient large-scale protein production. Maquette export into the growth medium offers the potential of maximising protein yield as build-up of protein and possible degradation inside the cell is avoided. Export of the maquette to the periplasm is the first step to achieving this goal.

In this chapter, the *E. coli* Tat export machinery was challenged to process the completely synthetic protein maquette, BT6. The TorA signal peptide was fused to the N-terminus of BT6 and expressed in *E. coli* BL21(DE3) cells. The TorA signal peptide was also fused to the N-

terminus of a BT6 variant that cannot bind haem in order to determine the fate of Tat substrate molecules that are unable to assemble the cofactor correctly.

6.3 Results

6.3.1 Construction of the TorA-BT6 strain and expression in *E. coli*

The Tat signal sequence peptide from the *E. coli* trimethylamine N-oxide reductase protein (this signal sequence is hereafter referred to as TorA) was fused to the N-terminus of the haem-binding maquette, BT6 (Figure 6.2 A). The primers "TorA F" and "TorA-BT6 Rev OE" were used to amplify the TorA sequence from *E. coli* genomic DNA. The primers "TorA-BT6 For OE" and "BT6 his R" were used to amplify BT6 from the plasmid pJExpress414-BT6 (provided by G. Kodali, University of Pennsylvania). The primer "BT6 his R" introduced a C-terminal His tag for immunodetection on Western blots. The two fragments were fused by overlap extension PCR. The fragment was cloned into the expression plasmid pJExpress414 using the restriction sites *Nde*I and *Xho*I. As a control, BT6 was amplified without the TorA signal sequence using the primers "BT6 F" and "BT6 his R" and cloned into pJExpress414 to produce the plasmids pJExpress414-BT62 and pJExpress414-TorA-BT6 (Figure 6.2 A). The constructs were transformed into *E. coli* BL21 (DE3) cells.

Cells were grown and protein expression was induced with 0.5 mM IPTG for 5 hours. Total cell samples were collected and subjected to SDS-PAGE and subsequent western blot analysis using anti-His antibodies (Figure 6.2 B).

In the absence of IPTG, no signal is seen on the western blot. For the TorA-BT6 sample there is a signal at 23 kDa corresponding in size to precursor (pre) TorA-BT6 and a signal at 17 kDa, corresponding to mature (mat) BT6. This suggests the processing of TorA-BT6 through the Tat system and the cleavage of the TorA signal peptide. The band seen at 20 kDa is likely a degradation product, additional bands such as this are routinely seen in Tat assays with diverse Tat substrates (personal communication, Colin Robinson). For BT6 with no signal sequence, only a band at 17 kDa is observed, corresponding to the molecular weight of BT6.

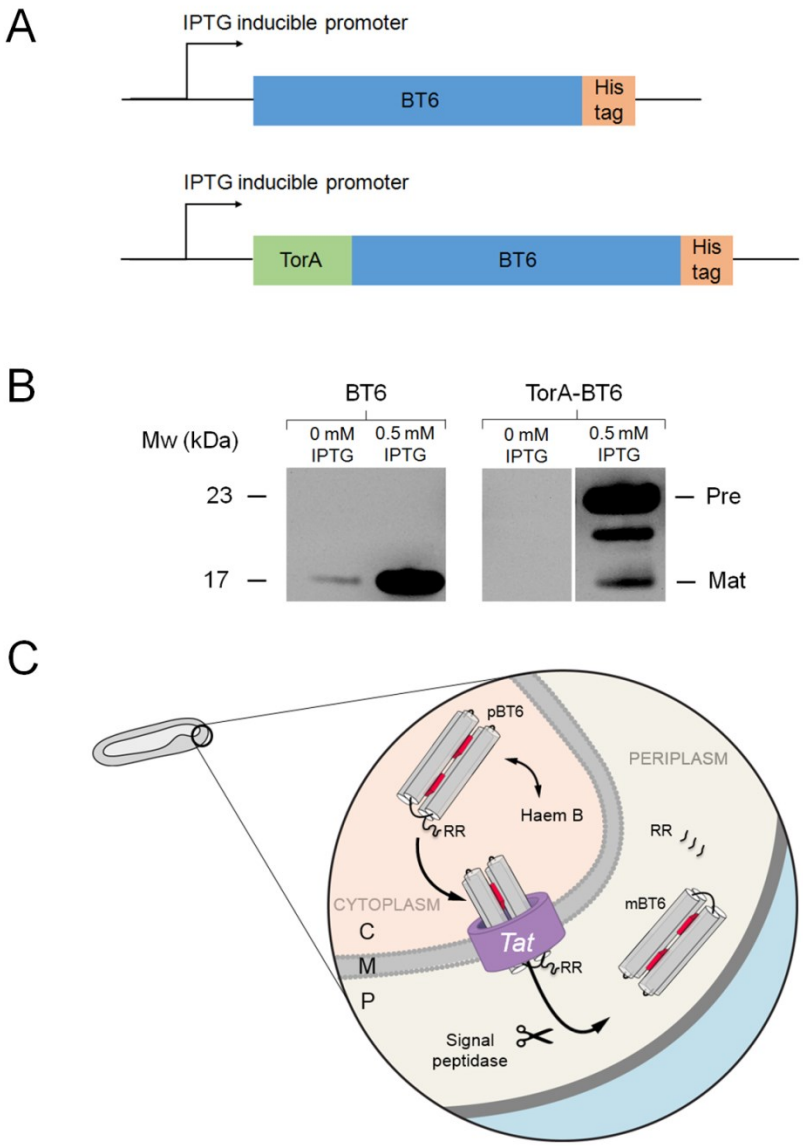


Figure 6.2 TorA-BT6 can be processed by the Tat system

Single colonies were used to inoculate 5 ml LB with 100 µg ml^{-1} ampicillin and grown overnight at 37 ^{o}C. 0.5 ml of the overnight culture was transferred to 9.5 ml LB plus ampicillin in a 50 ml Falcon tube. Expression of protein was induced with 0.5 mM IPTG and the cells were grown for 5 hours at 37 ^{o}C.

- **A.** Schematic representation of the IPTG-inducible constructs, TorA-BT6 and BT6.
- **B.** Western blot analysis of *E. coli* BL21 (DE3) whole cells expressing BT6 and TorA-BT6.
- **C.** Schematic of processing of TorA-BT6 through the Tat system. Figure adapted from an earlier version by Les Dutton.

To further establish that TorA-BT6 is processed by the Tat system, two commonly used control mutations of the TorA signal peptide were fused to the N terminus of BT6. The mutation of the first arginine of the twin arginine motif to a lysine (KR) allows the normal export of the Tat substrate (Figure 6.3 C, left). The mutation of both arginines to lysines (KK) prevents the export of many proteins (Figure 6.3 C, right) (Cristobal *et al*., 1999; Stanley *et al*., 2000).

pJExpress414-TorA-KR-BT6 was produced using QuikChange mutagenesis (Section 2.2.7) of the plasmid pJExpress414-TorA-BT6 using the primers “TorA KR QC AS” and “TorA KR QC S”. pJExpress414-TorA-KK-BT6 was produced using QuikChange mutagenesis of the plasmid pJExpress414-TorA-KR-BT6 using the primers “TorA KK QC AS” and “TorA KK QC S”. After transformation into BL21 (DE3) cells followed by protein induction, whole cell extracts were analysed. The 17 kDa signal corresponding to mat-BT6 is seen for KR but only a faint signal for KK is seen, possibly corresponding to a degradation product.

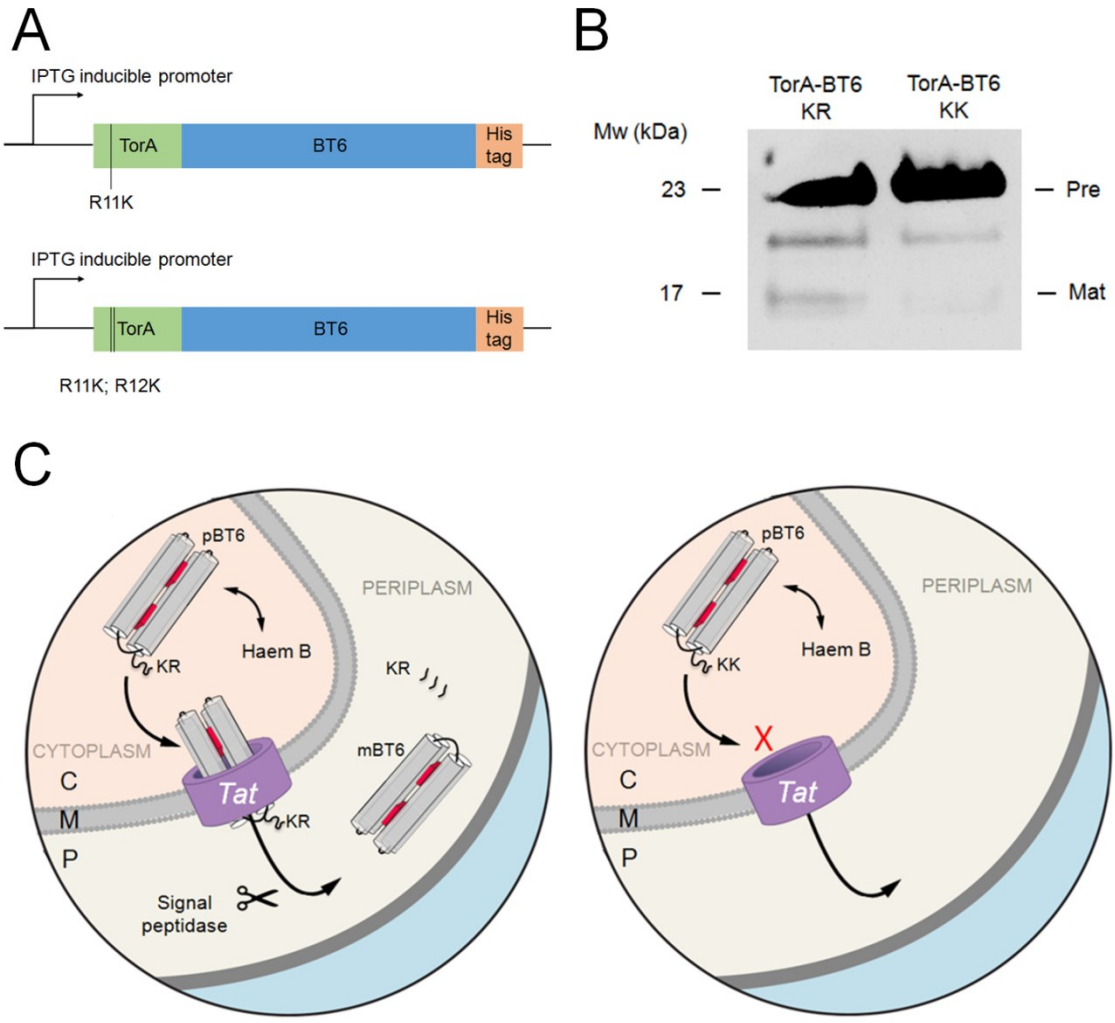


Figure 6.3 Expression of TorA-BT6 using KR and KK mutants of the TorA signal peptide

Single colonies were used to inoculate 5 ml LB with 100 µg ml^{-1} ampicillin and grown overnight at 37 ^{o}C. 0.5 ml of the overnight culture was transferred to 9.5 ml LB plus ampicillin in a 50 ml Falcon tube. Expression of protein was induced with 0.5 mM IPTG and the cells were grown for 5 hours at 37 ^{o}C.

- **A.** Schematic representation of the IPTG-inducible constructs, TorA-KR-BT6 and TorA-KK-BT6.
- **B.** Western blot analysis of whole cells *E. coli* BL21 (DE3) expressing TorA-KR-BT6 and TorA-KK-BT6.
- **C.** Schematic of processing of TorA-KR-BT6 (left) and TorA-KK-BT6 (right) through the Tat system. Figure adapted from an earlier version by Les Dutton.

6.3.2 Sub-cellular fractionation of cells expressing TorA-BT6

Cells were grown as before and the media was supplemented with 200 µM $FeSO_4$ and 200 µM 5-aminolevulinic acid (ALA – a precursor in the haem biosynthesis pathway) to stimulate haem biosynthesis. In order to confirm that the processed mat-BT6 is localised to the periplasm, a sub-cellular fractionation was carried out according to the method described in Section 2.10, followed by immunoblotting (Figure 6.4). A signal at 17 kDa corresponding to mat- BT6 was found only in the periplasmic fraction. No mat-BT6 was detected in the membrane or cytoplasmic fractions, where only the 23 kDa pre-BT6 is visible. This indicates that processing, likely in the form of cleavage of the signal peptide, occurs exclusively in the periplasm. The same result was observed for the TorA-KR mutant. As expected, no mat-BT6 was visible in the periplasm for the TorA-KK mutant.

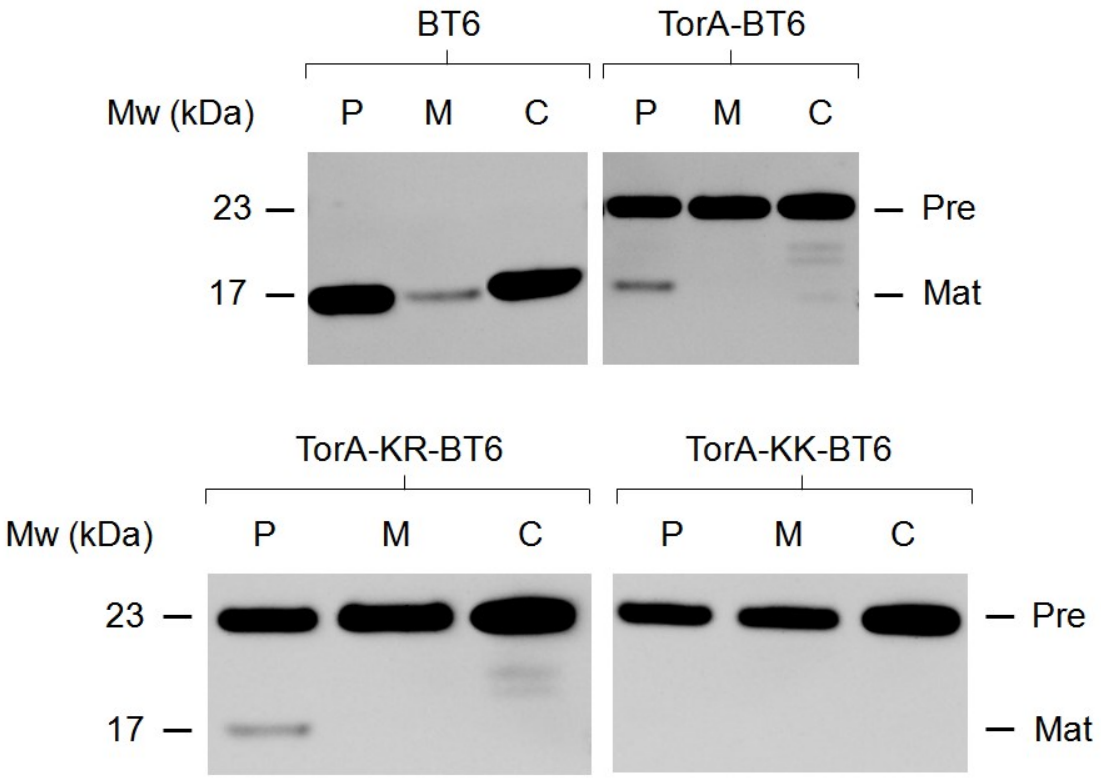


Figure 6.4 Tat export of BT6

Expression was induced with 0.5 mM IPTG in the presence of 200 µM $FeSO_4$ and 200 µM ALA for 2 hours at 37°C. Cells were fractionated into cytoplasmic (C), membrane (M) and periplasmic (P) samples. Samples were immunoblotted using antibodies to the C-terminal 6-His tag on BT6. Pre-BT6 (23 kDa) appears to be exported to the periplasm where the TorA signal sequence is cleaved, producing mat-BT6 (17 kDa). TorA-KR-BT6 exports to the periplasm; TorA-KK-BT6 does not. Contamination of the periplasmic fraction by the precursor form is observed in all cases.

In samples expressing BT6 alone, a 17 kDa signal was seen in all fractions. In the case of the TorA-BT6 samples, Pre-BT6 at 23 kDa is observed in all fractions. This "contamination" could be due to issue with the fractionation procedure or leakage from the cytoplasm to the periplasm. BT6 overexpresses from pJExpress414 extremely well (Figure 6.5) which could cause some membrane destabilisation. The leakage of precursor protein to the periplasm has been observed in previous studies when large amounts of the protein, often GFP, are produced and saturate the Tat machinery (Barrett *et al*., 2003) (Alex Jones, personal communication).

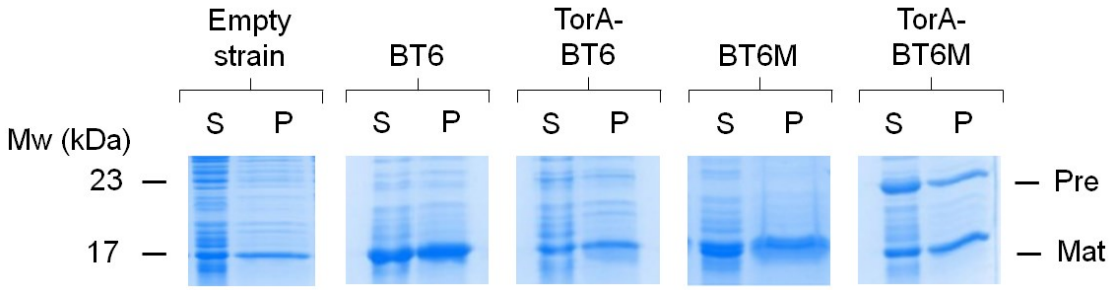


Figure 6.5 Relative expression levels of BT6 and BT6M

Expression was induced with 0.5 mM IPTG for 3 hours at 30 ^{o}C. Cells were fractionated into spheroplast (S) and periplasmic (P) samples. Samples were standardised to OD_{600} and the same amount of material loaded into each lane. The band at 17 kDa is native to *E. coli* but the presence of BT6, mat-BT6 and BT6M causes this band to become more intense. Higher expression levels were found for TorA-BT6M relative to TorA-BT6, as evidenced by the stronger band at 23 kDa.

To attempt to reduce the amount of BT6 produced by the cells, a shorter induction time of 1 hour was used but this did not decrease leakage of pre-BT6 from the cytoplasm to the periplasm. Decreasing the IPTG concentration to 0.1 mM did not improve the situation (personal communication, Alex Jones, University of Kent).

6.3.3 Analysis of growth medium

To determine whether BT6 is present in the growth medium, which could indicate outer membrane destabilisation, the medium was tested for presence of maquette. Cells were induced with 0.5 mM IPTG at 30° C for 3 hours in the presence 50 µM $FeSO_4$ and 500 µM ALA. The cells were spun down, and separated into spheroplast and periplasm fractions followed by immunoblotting (Figure 6.6). No signal was observed in the medium fraction, indicating that BT6 is not exported and does not leak outside the cell. The decrease in induction temperature did not prevent leakage of pre-BT6 to the periplasm. The presence of a signal corresponding in size to mat-BT6 in the spheroplast fraction is either a degradation product,

mat-BT6 in the membrane prior to release in the periplasm, or indicative of inefficient fractionation.

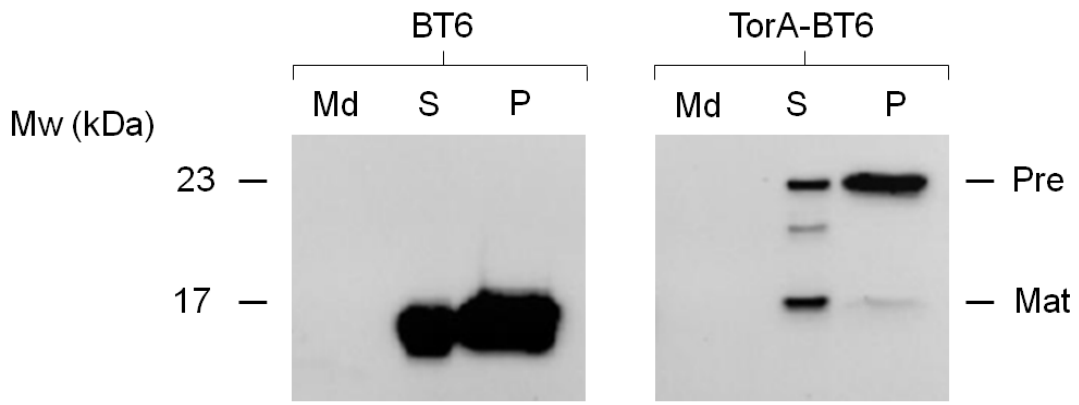


Figure 6.6 BT6 is not present in the growth medium

Expression was induced with 0.5 mM IPTG in the presence of 50 µM $FeSO_4$ and 500 µM ALA for 3 hours at 30°C. Cells were fractionated into spheroplast (S) and periplasmic (P) samples. The growth medium was also analysed (Md). Samples were immunoblotted using antibodies to the C-terminal 6-His tag on BT6. No signal is seen for the medium fraction.

6.3.4 An alternate protocol for sub-cellular fractionation of *E. coli*

It is possible that the expression of BT6 disrupts the inner membrane. A more gentle fractionation procedure was used in the hope of minimising the leakage of BT6 and TorA-BT6 (pre-BT6) into the periplasm. The procedure used was developed by Quan *et al*. in which the release of envelope proteins occurs in the presence of sucrose and EDTA (Quan *et al*., 2013).

After induction with 0.5 mM IPTG for 1 hour, 10 ml of cells were pelleted by centrifugation at 3,000 rpm for 10 min at 4 °C. The supernatant was discarded and the pellet gently resuspended with a wire loop in 100 µl cold TSE buffer (20 mM Tris-HCl pH8, 0.5 M sucrose, 1 mM EDTA, 1 EDTA-free protease inhibitor tablet per 100 ml buffer). The cells were incubated on ice for 30 minutes followed by centrifugation at 16, 000 x *g* in a microcentrifuge at 4 °C for 30 minutes. The supernatant was carefully removed and designated as the periplasmic fraction. The spheroplast pellet was resuspended in lysis buffer (50 mM Tris-HCl pH8, 0.1 mg ml^{-1} lysozyme, 20 µg ml^{-1} DNase). The samples were analysed by western blot (Figure 6.7).

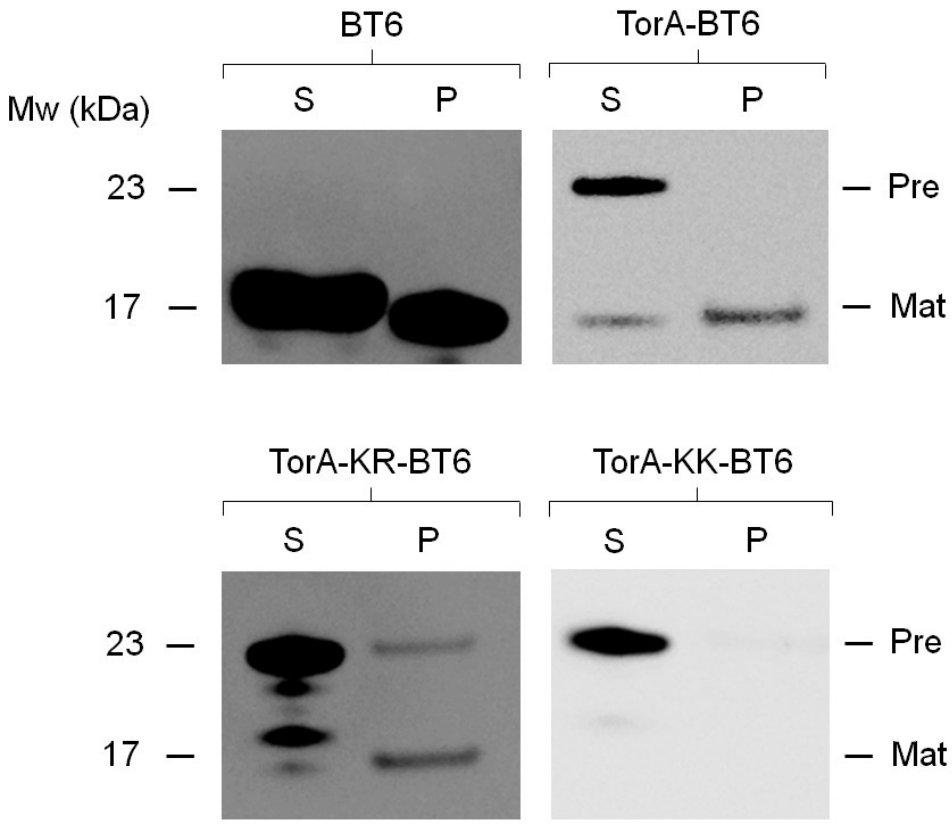


Figure 6.7 Fractionation of *E. coli* expressing BT6 using a more gentle procedure

Expression was induced with 0.5 mM IPTG for 1 hour. Cells were fractionated into spheroplast (S) and periplasmic (P) samples using the method developed by Quan *et al*. Samples were immunoblotted using antibodies to the C-terminal 6-His tag on BT6. The contamination of Pre-BT6 (23 kDa) in the periplasm appears to be decreased in all TorA-BT6 cases. BT6 leaks into the periplasm even using this more gentle fractionation technique.

In the case of TorA-BT6, and largely TorA-KR-BT6 and TorA-KK-BT6, this method proved successful in eliminating pre-BT6 from the periplasmic fraction. However fractionation was inefficient and some periplasmic contents remained in the spheroplast fraction (as indicated by a signal at 17 kDa). When expressed without the signal peptide, BT6 was seen in both the periplasmic and spheroplast fractions. BT6 expresses more than TorA-BT6 under the same induction conditions (Figure 6.5) and therefore may have more potential to disrupt the membrane.

6.3.5 The effect of the addition of ALA and iron on BT6 export

The addition of ALA and iron boosts haem production and increases the expression of BT6. To determine whether the addition of ALA and Fe increases export and/or the periplasmic contamination of pre-BT6, TorA-BT6 expressing cells were induced with and without the addition of 50 µM $FeCl_2$ and 500 µM ALA. The cells were induced for 16 hours to allow the haem biosynthesis pathway to be sufficiently stimulated. The cells that had been

supplemented were a reddish/brown colour (Figure 6.8). The cells were standardised to the same OD_{600} and fractionated into spheroplast and periplasm samples using the method developed by Quan *et al*. The samples were analysed by western blot (Figure 6.8). The presence of ALA and iron did not appear to increase TorA-BT6 export. Periplasmic contamination was observed in both cases.

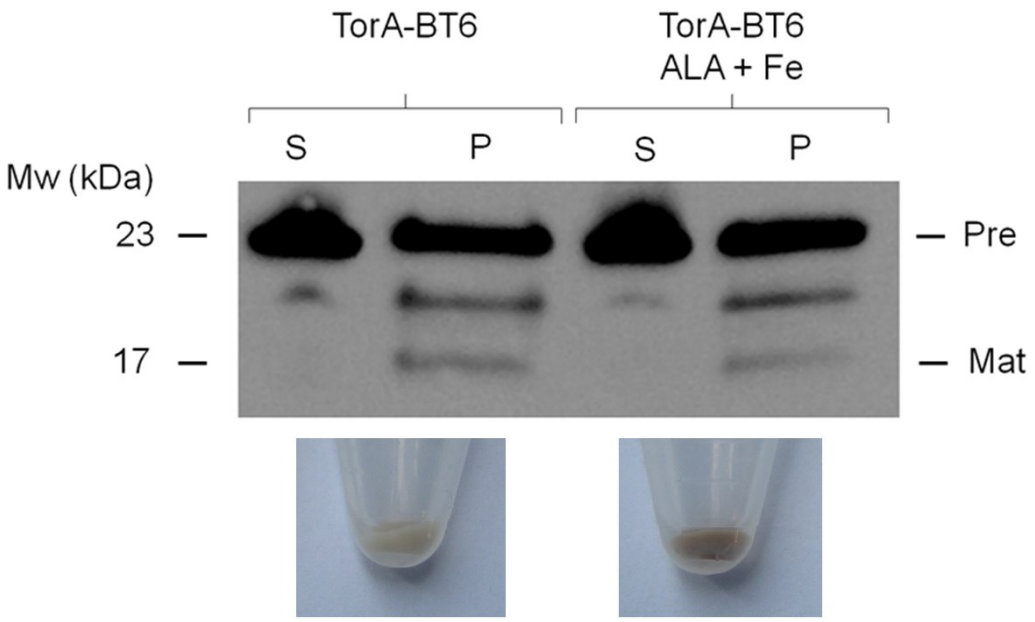


Figure 6.8 The effect of the addition of ALA and iron on TorA-BT6 export

Expression was induced with 0.5 mM IPTG in the presence of 50 µM $FeCl_2$ and 500 µM ALA for 16 hours at 37°C. Cell pellets are shown. Cells were fractionated using the method developed by Quan *et al* into spheroplast (S) and periplasmic (P) samples. Samples were immunoblotted using antibodies to the C-terminal 6-His tag on BT6. Contamination of the periplasmic fraction by pre-BT6 is observed. No difference in the level of contamination is seen with and without supplementation with ALA and iron. Supplementation does not improve efficiency of BT6 export.

6.3.6 Induction of TorA-BT6 from pEXT22

It is possible that the high expression level of BT6 disrupts the inner membrane and causes contamination of the periplasm with pre-BT6. The BT6 and TorA-BT6 variants were cloned into pEXT22 by Alex Jones (University of Kent), an expression plasmid that produces lower levels of the maquette. The constructs were transformed into *E. coli* W3110 cells and induced with 0.5 mM IPTG for 3 hours at 30 °C. The cells were fractionated according to the original protocol (Section 2.10) and analysed by western blot (Figure 6.9). No leakage of the cytoplasmic precursor to the periplasmic fraction was observed. The signal for pre-BT6 is either faint or non-existent in the cytoplasmic and membrane fractions indicating

degradation. A band at 17 kDa in the periplasmic fraction corresponding to mat-BT6 is observed for the TorA-BT6 and TorA-KR-BT6 variants.

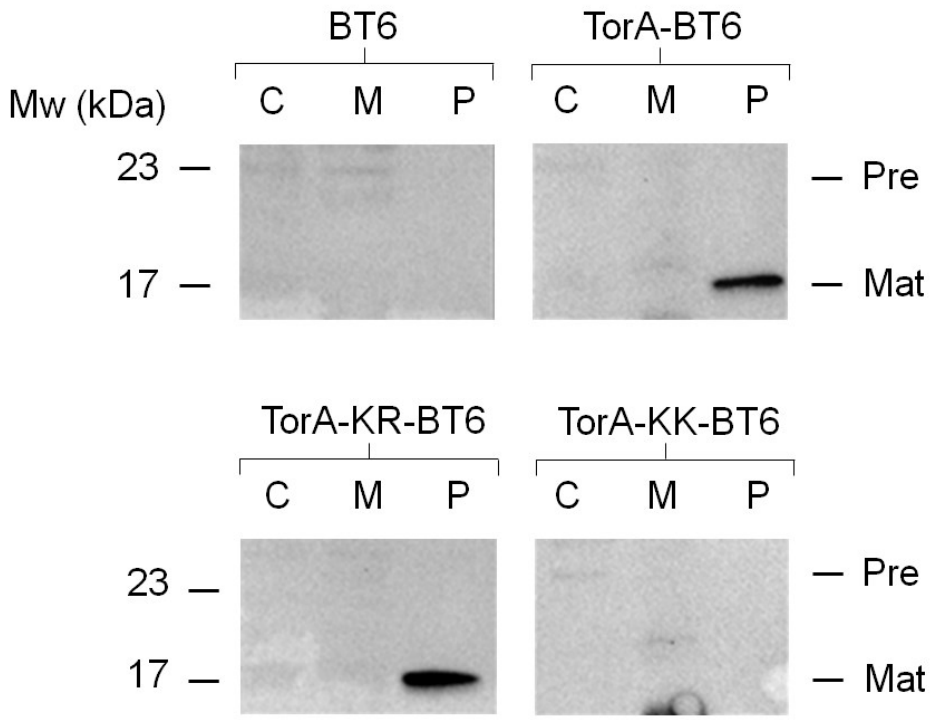


Figure 6.9 Expression of BT6 from pEXT22

Expression was induced with 0.5 mM IPTG for 3 hours at 30 ^{o}C. Cells were fractionated into cytoplasmic (C), membrane (M) and periplasmic (P) samples. Samples were immunoblotted using antibodies to the C-terminal 6-His tag on BT6.

6.3.7 Export of maquette with 1 haem bound

A mutant version of BT6 with one of its haem-binding histidine residues mutated to an alanine was created using QuikChange mutagenesis on the plasmid pJExpress414-TorA-BT6 using the primers “BT6 H53A QC S” and “BT6 H53A QC AS”. The resulting maquette is only capable of binding one haem but remains largely structured (Farid *et al*., 2013; and Figure 6.12 B). The construct, termed pJExpress414-TorA-BT6 H53A was transformed into *E. coli* BL21 (DE3) cells and induced with 0.5 mM IPTG for 3 hours at 30 °C. The cells were fractionated according to the original protocol (Section 2.10) and analysed by western blot (Figure 6.10). The one-haem binding maquette was exported to the periplasm, but in lower amounts than BT6. TorA-BT6 H53A appears to have leaked from the spheroplasts into the apparent periplasmic fraction.

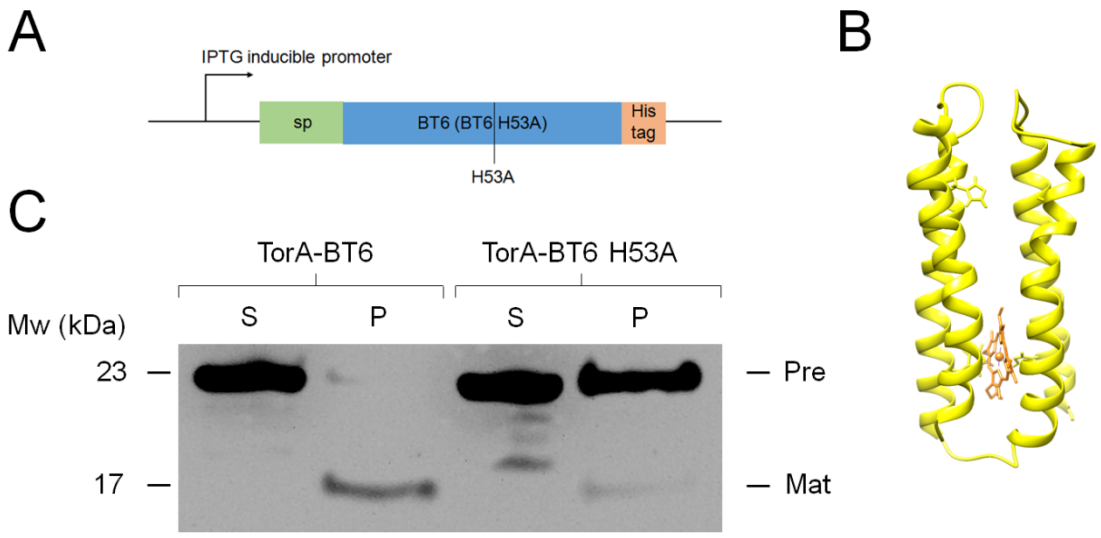


Figure 6.10 Export of 1 haem bound maquette

- **A.** Schematic representation of the IPTG-inducible construct, TorA-BT6 H53A.
- **B.** Model of BT6 H53A based on the structure shown in Figure 6.1.
- **C.** Western blot analysis of fractionated *E. coli* BL21 (DE3) expressing TorA- BT6 and TorA-BT6 H53A. Expression was induced with 0.5 mM IPTG for 1 hour. Cells were fractionated into spheroplast (S) and periplasm (P) fractions using the protocol described in Section 2.10.

6.3.8 Expression of a BT6 mutant incapable of binding haem

The Tat pathway exports only fully folded proteins, and many Tat substrates contain a cofactor. A mutant version of BT6, BT6M, was constructed in which two of its haem-binding histidine residues were mutated to alanines, preventing haem binding. As a result, BT6M is significantly less structured in comparison to BT6 (Farid *et al*., 2013; and Figure 6.13). To analyse the relationship between haem incorporation (and thus extent of maquette folding) and protein export via the Tat pathway, the TorA signal sequence was fused to the N terminus of BT6M (Figure 6.11 B i). This was done as in Section 6.3.1 using the primers "TorA F", "TorA-BT6M rev OE", "TorA-BT6M For OE" and "BT6M his R". BT6M was amplified from pJExpress414-BT6M (G. Kodali, University of Pennsylvania). The resulting construct was termed pJExpress414-TorABT6M. A C-terminal His tag was included for immunodetection. As a control, BT6M with no signal peptide and a C-terminal His tag was amplified using the primers "BT6M F" and "BT6M his R" and was cloned into pJExpress414 to create pJExpress414-BT6M2 (Figure 6.11 A i). TorA-BT6M and BT6M were cloned into pEXT22 by Alex Jones, University of Kent.

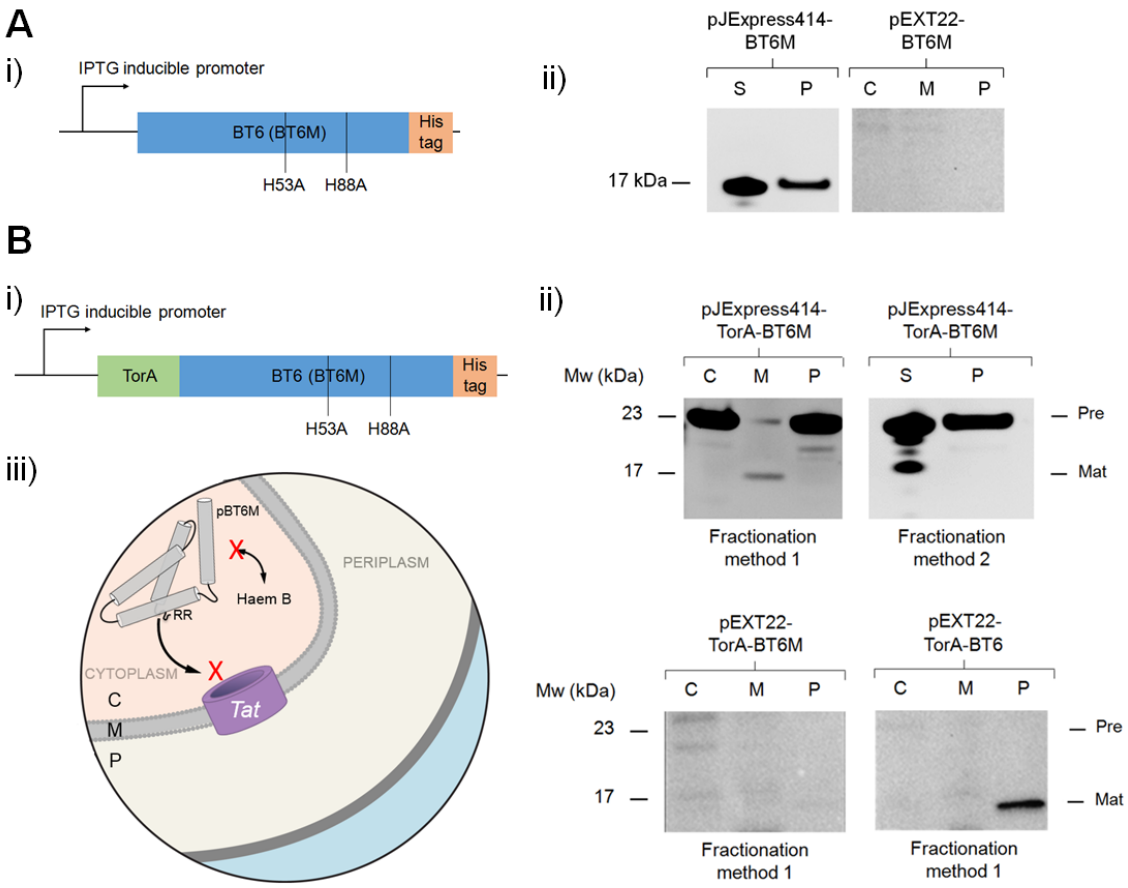


Figure 6.11 Expression of BT6M and TorA-BT6M

A. **i)** Schematic representation of IPTG-inducible BT6M.
ii) Western blot analysis of fractionated *E. coli* expressing BT6M. Cells were fractionated into spheroplast (S), periplasm (P), membrane (M) and cytoplasm (C) fractions.

B. **i)** Schematic representation of the IPTG-inducible TorA-BT6M construct.
ii) Western blot analysis of fractionated *E. coli* expressing BT6M. Top panels – TorA-BT6M expressed from pJExpress414; bottom panels – TorA-BT6M and TorA-BT6 expressed from pEXT22. Cells were fractionated into spheroplast (S), periplasm (P), membrane (M) and cytoplasm (C) fractions. Fractionation method 1 is that detailed in Section 2.10. Method 2 is the more gentle technique developed by Quan *et al* (Section 6.3.4).
iii) The export of TorA-BT6M is blocked by the Tat system. Figure adapted from an earlier version by Les Dutton.

Protein expression was induced and the cells fractionated and analysed by western blot as described previously. When expressed from pJExpress414 and fractionated using the gentle method of Quan *et al*, BT6M is seen in both the spheroplast and periplasmic fractions (Figure 6.11 A ii, left panel). BT6M expresses more than BT6, and possibly causes membrane destabilisation to a greater extent. No protein expression of BT6M is observed when expressed from pEXT22 (Figure 6.11 A ii, right panel).

Neither pre-BT6M nor mat-BT6M is observed on the western blot of the fractions when TorA-BT6M is expressed from pEXT22 (Figure 6.11 B ii, bottom left). For both fractionation techniques, contamination of the periplasm of pre-BT6M is seen when TorA-BT6M is expressed from pJExpress414 (Figure 6.11 B ii, top panels).

The loss of the ability to bind haem has a dramatic inhibitory effect on export of TorA-BT6M; no mat-BT6M is seen in the periplasmic fraction (Figure 6.11 B ii, top panels). These data provide strong evidence that the Tat pathway efficiently assesses the folding state of BT6 and completely blocks the export of non haem binding forms.

A 17 kDa band corresponding in size to mat-BT6M can be seen in the membrane fraction when TorA-BT6M is expressed from pJExpress414 (Figure 6.11 B ii, top left), which could indicate that TorA-BT6M is recognised by the Tat apparatus but is not released into the periplasm.

6.4 Discussion

6.4.1 The *E. coli* Tat system can process the synthetic protein maquette, BT6

The gene sequence encoding the TorA signal peptide for export through the Tat pathway was fused to the start of the gene for the maquette BT6 and cloned into the expression plasmid pJExpress414. The resulting construct was expressed in *E. coli* BL21(DE3) to produce a TorA-BT6 fusion (Figure 6.2 A).

Western blot analysis of whole cells showed two signals at 23 kDa and 17 kDa, corresponding respectively in size to precursor TorA-BT6 and mat-BT6 with the TorA signal peptide cleaved (Figure 6.2 B). The result shown in Figure 6.2 B alone is inconclusive so to test if there is any genuine Tat export, further experiments involved the mutation of the TorA signal peptide.

The TorA signal sequence contains a characteristic twin arginine motif, RR. When the first arginine is mutated to a lysine, KR, the protein is targeted to the Tat machinery and exported (Figure 6.3). When both arginines are mutated to lysines, KK, the Tat apparatus will not process the protein (Figure 6.3). TorA-KR-BT6 featured a 23 kDa and 17 kDa bands indicating processing by the Tat machinery. TorA-KK-BT6 only featured the 23 kDa band, indicating that processing by the Tat machinery was unable to take place and the signal sequence was cleaved. These results show that a functional TorA sequence is required for processing of pre-BT6 to the mature form, mat-BT6. The decrease in molecular weight of the observed signals from 23 kDa to 17 kDa is most likely due to the cleavage of the TorA signal sequence

occurring during translocation of the Tat machinery and release of mat-BT6 into the periplasm.

6.4.2 Tat-processed BT6 is released to the periplasm

The Tat pathway releases processed proteins to the periplasm. To confirm that the mature form of BT6 is localised to the periplasm, the cells were fractionated into cytoplasmic, membrane and periplasmic fractions. Western blot results showed the presence of a 17 kDa mature protein in the periplasm for TorA-BT6 and TorA-KR-BT6, but not for TorA-KK-BT6 (Figure 6.4). This indicates that processing, likely in the form of cleavage of the signal peptide, occurs causes the release of BT6 into the periplasm.

6.4.3 Expression of BT6 causes periplasmic contamination of the precursor form

For BT6 lacking TorA, a 17 kDa signal was seen in all fractions. In the case of the TorA-BT6 samples, pre-BT6 at 23 kDa is observed in all fractions (Figure 6.4). This indicates contamination due either to an issue with the fractionation procedure or leakage from the cytoplasm to the periplasm. BT6 overexpresses from pJExpress414 extremely well which could cause some membrane destabilisation and lack of inner membrane integrity (Figure 6.5; personal communication, Colin Robinson, University of Kent). This leakage of precursor protein to the periplasm has been observed in previous studies when large amounts of the protein, often GFP, are produced and saturate the Tat machinery (Barrett *et al*., 2003; Alex Jones, personal communication). The maquette is not exported to the growth medium, indicating that outer membrane integrity is maintained (Figure 6.6).

Shorter induction times and lack of supplementation of ALA and iron to the medium did not decrease the appearance of pre-BT6 in the periplasm (Figures 6.8 and 6.10). When a more gentle fractionation technique was used to minimise this contamination the situation was improved but not eliminated (Figure 6.7). Expression of BT6 without the TorA signal sequence continued to give large amounts of contamination of the periplasmic fraction with the protein; export of BT6 alone to the periplasm should be impossible as there is no signal sequence. This indicates that the maquette itself is disrupting the inner membrane, rather than the maquette having an effect on the Tat machinery during export and allowing leakage.

The TorA-BT6 variants and BT6 were cloned into a vector, pEXT22, which produces lower expression levels. Almost no signal was seen in any fraction, other than a strong signal at 17 kDa in the periplasmic fractions of TorA-BT6 and TorA-KR-BT6 (Figure 6.9). These results indicate cytoplasmic degradation of the maquette when expressed at low levels and it is not

exported (Alex Jones, personal communication). This result also indicates that it is the high expression of BT6 that causes leakage of the precursor form into the periplasm.

6.4.4 BT6 must bind haem for transport through the Tat pathway

A version of BT6 that could only bind one haem (BT6 H53A) was found to export through the Tat pathway (Figure 6.10), although to a lesser extent than BT6. A previous study has shown that on binding only one haem, BT6 is relatively structured (Figure 6.12) (Farid *et al*., 2013), sufficient to avoid rejection by the Tat quality control mechanism. The low export efficiency could indicate that BT6 H53A is a dynamic structure that does not always appear sufficiently folded, perhaps revealing its hydrophobic interior causing rejection by the Tat pathway.

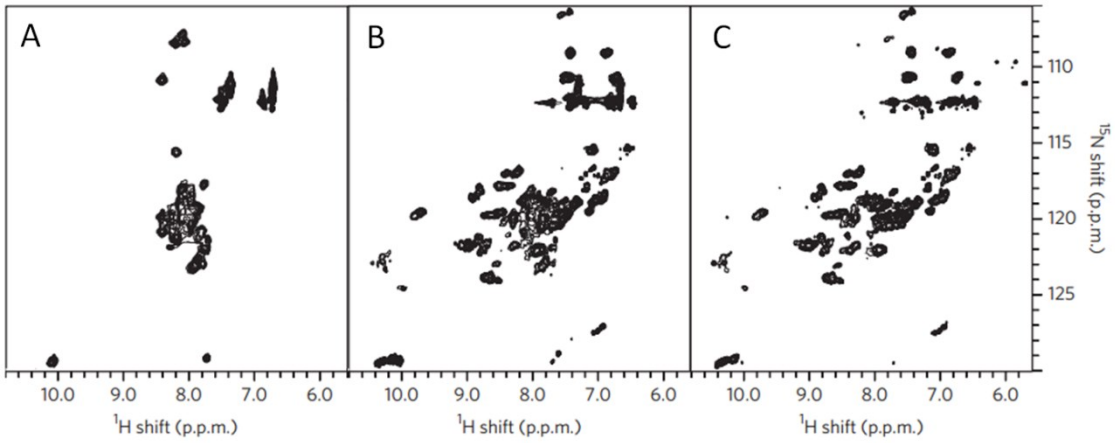


Figure 6.12 NMR spectral dispersion of BT6

750 MHz 15N-HSQC showing NMR spectral dispersion. In the absence of haem, NMR resonances of BT6 are relatively undispersed indicating the absence of unique tertiary structure (A). On the addition of one equivalent of haem resonances of BT6 become partially dispersed indicating stable structure (B). Addition of a second equivalent of haem induces further structuring (C).

The TorA signal peptide was attached to a variant of BT6, termed BT6M. BT6M has two of its histidines mutated to alanines and cannot bind any haem. As a result, BT6M is unstructured compared to haem-bound BT6. A previous study has shown this lack of structure using NMR (Figure 6.13) (Farid *et al*., 2013).

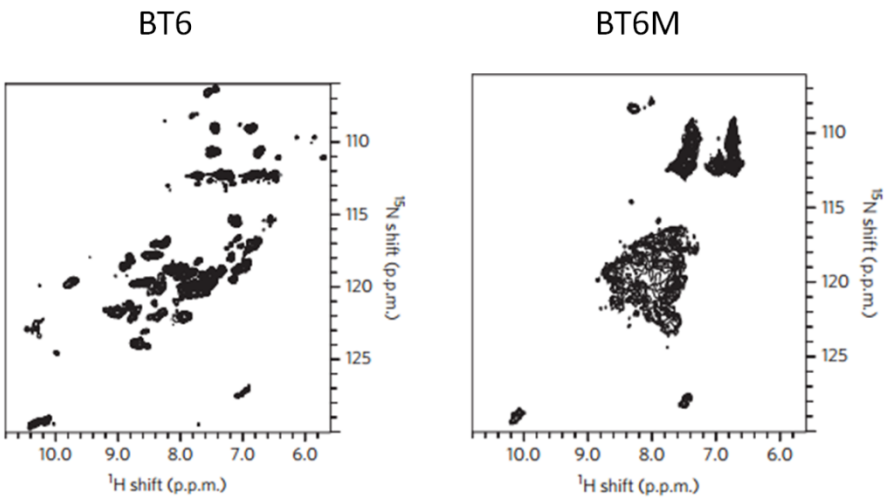


Figure 6.13 NMR spectral dispersion of BT6 and BT6M in the presence of haem

750 MHz 15N-HSQC showing NMR spectral dispersion. NMR resonances of BT6M are relatively undispersed indicating the absence of unique tertiary structure. Resonances of BT6 relatively dispersed indicating stable structure.

(Figure adapted with permission from Macmillan Publishers Ltd, Nature, Farid *et al* 2013, Copyright 2013)

TorA-BT6M is not processed by the Tat apparatus with no 17 kDa mature protein seen in the periplasm when expressed from either pJExpress414 or pEXT22 (Figure 6.11 B ii, top panels, bottom left). These data provide strong evidence that the Tat pathway efficiently assesses the folding state of BT6 and completely blocks the export of non haem binding forms. This could be because the hydrophobic maquette interior becomes concealed on haem binding.

A 17 kDa band corresponding in size to mat-BT6M has been observed in the membrane fraction for the pJExpress414 experiment (Figure 6.11 B ii; top left panel). This could indicate that TorA-BT6M is recognised by the Tat apparatus but is not released into the periplasm. This provides a possible location for the quality control step.

6.4.5 Increased haem biosynthesis does not increase Tat export efficiency of TorA-BT6

The supplementation of the growth medium with ALA and iron increases maquette expression and the proportion of maquette with two haems bound (G. Kodali, University of Pennsylvania, personal communication). Supplementation with ALA and iron does not increase the efficiency of export of BT6 (Figure 6.8). This suggests that BT6 can scavenge enough haem to stabilise its tertiary structure sufficient for Tat export and that the Tat apparatus exports BT6 to the optimum efficiency without the need for supplementation.

6.4.6 Additional work

Initial experiments have involved the export of maquettes to the periplasm, but ultimately they could be exported outside the cell to the growth medium, simplifying purification. Maquette export into the medium offers the potential of maximising protein yield as build-up of protein and possible degradation inside the cell are avoided.

Due to the easy mutation of the maquette sequence, more subtle mutations to perturb folding or introduce hydrophobic patches on the surface of the maquette could be introduced to determine the "tipping point" between acceptance and rejection by the Tat quality control apparatus. Maquettes with other cofactors such as zinc porphyrins could be tested.

6.4.7 Conclusion

The Tat system, which exports fully or largely folded proteins, recognises and exports the completely man-designed haem-binding synthetic protein maquette BT6. On selective mutation of the maquette to prevent cofactor binding, and thus the folding properties, the Tat apparatus will not export the maquette.

7. Concluding remarks

Attempts to recreate photosynthesis-like light-harvesting processes using purely chemical means have fallen far short of the efficiency and versatility of natural systems (Harvey, 2003; Aratani and Osuka 2010; Harvey *et al*., 2011; Lindsey *et al*., 2011; Jiang *et al*., 2014). A major challenge is to make use of the diverse spectral range of incoming solar radiation.

Ongoing biohybrid antenna research has integrated elements found in natural photosynthetic pathways with components produced by synthetic chemistry. Native-like bacterial photosynthetic antenna peptides have been used as a scaffold to attach synthetic chromophores. The use of synthetic chromophores enables the broadening of spectral coverage of the resulting antenna beyond that provided by native antennas. (Harris *et al*., 2014a; Jiang *et al*., 2014). However while these biohybrid antennas show great promise *in vitro* to augment photosynthesis *in vivo* we must make use of the proteins and pigments that are available in nature. Thus arises the challenge of creating bacteria with new photosynthetic pathways to utilise a greater range of wavelengths than naturally-evolved photosynthesis.

Chapter 3 details a pilot study to explore the potential of using new light harvesters to augment bacterial photosynthesis. YFP was fused to the reaction centre of the purple photosynthetic bacterium *Rhodobacter sphaeroides*. YFP was able to funnel sufficient energy to the native photosynthesis pathway to increase the photosynthetic growth rate in a carotenoidless background.

Future work following on from this should explore new chromophores that can absorb light of wavelengths that wild type *Rba. sphaeroides* does not absorb (between 640 nm and 750 nm). Possible candidates include orange and red fluorescent proteins such as mApple, mCherry or mKate2 (Shaner *et al*., 2004; Shaner *et al*., 2008; Shcherbo *et al*., 2009). Another possible candidate is the Cyanobacterial CpcA protein. CpcA is a fluorescent phycobiliprotein with potential as a biological label. CpcA can be expressed in *E. coli* along with genes encoding enzymes to modify the native haem biosynthesis pathway to produce various bilin chromophores (Alvey *et al*., 2011).

Additional chromophores need not be confined to one per native complex. A ring of multiple chromophores around the reaction centre may be achieved through the fusion of the gene(s) for a native LH1 subunit with the gene for a new chromophore. The use of two different FRET partner fluorescent proteins expressed on the LH1 protein and the RC would enhance charge separation by funnelling additional energy.

Chapter 4 explored the potential of proteorhodopsin (PR) as an additional photosystem component. PR is a light-driven proton pump and may be able to augment the *Rba. sphaeroides* transmembrane proton gradient used to drive the ATP synthase. PR is an attractive candidate as, in contrast to *Rba. sphaeroides*, it performs light harvesting and proton pumping roles in a single protein. Although PR has been successfully expressed, the challenge of assembling sufficient functional protein in the photosynthetic membrane to augment photosynthesis is one that is yet to be overcome. Organisms that natively express proteorhodopsin typically fill the membrane with many copies of the protein, and there is debate as to whether PR alone is sufficient to power metabolism with light in a heterotroph (see Section 1.9.2).

The biosynthesis of the PR chromophore, retinal, proved challenging. Large-scale production of this pigment may be lucrative as retinal and its derivatives are high value components of various beauty products. However, only very small amounts of retinal were made and this research highlighted various hurdles that may arise when trying to biosynthesise foreign pigments. It is likely that striking the balance between the availability of sufficient oxygen for retinal production without risking degradation is difficult in *Rba. sphaeroides*. In addition, it is possible that *Rba. sphaeroides* possesses unknown or poorly characterised proteins that are able to degrade retinal.

The results presented in this thesis confirmed the findings of previous studies that the deletion of photosystem components improved the expression of foreign membrane proteins (Labile *et al*., 2009; Tikh *et al*., 2014). Furthermore, the results indicated that a more planar membrane environment aided the expression of proteorhodopsin and a transmembrane maquette. For future studies involving the expression of foreign membrane proteins a Δ*pufX* mutant of *Rba. sphaeroides* may therefore provide good yields.

Beyond the use of foreign proteins to augment photosynthesis, completely synthetic proteins such as maquettes have potential for the creation of custom photosynthetic pathways due to their stability and versatility. Artificial peptide maquettes can have diverse functions conferred on them, such as light-harvesting and reaction centre-like properties, and may be capable of replacing the native components of bacterial photosynthetic pathways (Farid *et al*., 2013; Anderson *et al*., 2014; Watkins *et al*., 2014). The work presented in Chapter 5 forms the basis of the bottom-up redesign of photosystem components which could be used to replace or augment existing proteins *in vivo* or emulate their function in artificial membrane systems.

This study involved the design and expression of a gene for a transmembrane maquette. The amino acid sequence of the maquette made use of the design principles of natural membrane proteins in which the cytoplasmic regions are positively charged and the periplasmic or extracellular regions are negatively charged. The results of this study highlighted the importance of correct charge patterning, as the introduction of a negatively charged FLAG tag to the cytoplasmic N terminus of the maquette significantly decreased expression.

While the transmembrane maquette was found to sit in the membrane of *E. coli* and *Rba. sphaeroides*, purification and characterisation was difficult. This was in part due to the low amounts of protein that were able to be purified and difficulty in obtaining purified maquette with pigment bound. Future purification strategies may include a more gentle elution gradient when using nickel affinity columns. Instead of solubilisation with a traditional detergent which may disrupt cofactor binding, styrene maleic acid (SMA) has potential for the purification of maquettes. SMA has been used for the purification of *Rba. sphaeroides* reaction centres and acts as a "cookie cutter" to remove proteins from the membrane with a ring of native lipids around them (Swainsbury *et al*., 2014b). In this way the functional properties of the native membrane are maintained and facilitate stability of the isolated protein. Another strategy that may be useful for the purification of transmembrane maquettes is GraDeR, a method which results in mono-disperse and stable complexes that facilitate structural characterisation such as single-particle electron microscopy (Hauer *et al*., 2015).

Accelerated evolution in a photosynthetically compromised mutant of *Rba. sphaeroides* may result in the maquette acquiring a functional role, such as a light harvester. Alternatively, the maquette may play an electron transfer role. Recent work in the lab of Les Dutton has involved the attachment of CpcA to soluble maquettes to act as a light harvester. Artificial bilin attachment sites have been conferred on maquettes to make them photoactivatable (Les Dutton, personal communication).

Beyond transmembrane maquettes, a soluble *c*-type maquette has been designed which is capable of electron transfer and oxygen transport (Anderson *et al*., 2014). If the redox potential of this maquette is tailored it may replace the function of the *Rba. sphaeroides* native cytochrome c_2 protein. To do this the maquette must be transported to the photosynthetic membrane vesicles, either through the Tat or Sec systems. As shown in Chapter 6, export of maquettes has been successful in *E. coli* and this is something that may be transferable to *Rba. sphaeroides*.

Appendix

Growth medium for *Rhodobacter sphaeroides*

M22 10 x Stock (4 L)

Potassium dihydrogen orthophosphate	KH_2PO_4	122.4 g
Dipotassium hydrogen orthophosphate	K_2HPO_4	120.0 g
DL – Lactic acid (fridge)	Na lactate solution	100.0 g
Ammonium sulphate	$(NH_4)_2SO_4$	20 g
Sodium chloride	NaCl	20 g
Sodium succinate		173.7 g
Sodium glutamate	L – Glutamic acid	10.8 g
Aspartic acid	DL – Aspartic acid	1.6 g
Solution C		800ml

Make up with QH_2O to 2-3 L, pH to 6.8 and top up to 4 L

Solution C (4 L)

Nitrilotriacetic acid (NTA)		40 g
Magnesium chloride	$MgCl_2$	96 g
Calcium chloride	$CaCl_2$	13.36 g
EDTA		0.5 g
Zinc chloride	$ZnCl_2$	1.044 g
Ferrous chloride	$FeCl_2$	1.0 g
Manganese chloride	$MnCl_2$	0.36 g
Ammonium heptamolybdate	$(NH_4)_6Mo_7O_{24}(H_2O)_4$	0.037 g
Cupric chloride	$CuCl_2$	0.031 g

Cobaltous nitrate	$Co(NO_3)_2$	0.0496 g
Boric acid (orthoboric acid)		0.0228 g

Do not autoclave. Freeze at -20 ^{o}C in 800 ml aliquots

1 x M22 (4 L)

10 x M22 stock	400 ml
5 % CAA	80 ml
water	3520 ml

1 x M22 Agar (200 ml)

1 x M22	200 ml
Bacto Agar	3 g

1 x M22 10 % Sucrose Agar (200 ml)

1 x M22	180 ml
Sucrose	20 g
Bacto Agar	3 g

Ensure sucrose is dissolved prior to autoclaving, sucrose can be dissolved in batch but Bacto Agar must be added to Durans individually. CAA is not included in M22 agar.

Vitamins (10 K Vits)	**10 000 x**
Nicotinic acid (poison)	1 g
Thiamine (poison)	0.5 g
pABA (4-aminobenzoic acid) (fridge)	0.1 g
Biotin (d-Biotin) (fridge)	10 mg
water	100 ml

Growth media for *Escherichia coli*

LB/ LB agar and autoinduction LB

LB/ LB agar	Ready mixed LB media from FORMEDIUM prepared following the manufacturer's instructions.
Autoinduction LB	Ready mixed LB media from FORMEDIUM prepared following the manufacturer's instructions. Grows IPTG-inducible expression strains without induction to a high cell density, and induces production of the target protein automatically. Inclusion of a small amount of glucose and lactose in the media regulates expression from IPTG inducible promoters.

Terrific broth (1 L)

Tryptone	12 g
Yeast extract	24 g
Glycerol	4 ml

Make up to 900 ml with QH_2O, autoclave then add 100 ml sterile phosphate solution.

Phosphate solution for terrific broth (250 ml)

KH_2PO_4	5.78 g
K_2HPO_4	31.35 g

Top up to 250 ml with QH_2O. Filter sterilise.

Growth medium for *Synechocystis*

Trace minerals (1 L)

Boric acid	H_3BO_3	2.86 g
Manganese chloride	$MnCl_2$	1.81 g
Zinc sulphate	$ZnSO_4$	0.22 g
Sodium molybdate	$Na_2MoO_4 \cdot 2H_2O$	0.39 g
Copper sulphate	Cu_2SO_4	0.079 g
Cobaltous nitrate	$Co(NO_3)_2 \cdot 6H_2O$	100 ml

Make up to 1 L with QH_2O. Autoclave

BG11 100 X stock (1 L)

Sodium nitrate	$NaNO_3$	149.6 g
Manganese sulphate	$MnSO_4(H_2O)$	7.49 g
Calcium chloride	$CaCl_2$	3.60 g
Citric acid		0.60 g
EDTA (disodium salt)		0.10 g
Trace minerals		100 ml

1000 X Iron stock (1 L)

Ferric ammonium citrate	6 g

Make up to 1 L with QH_2O. Filter sterilise.

1000 X Phosphate stock (1 L)

Dipotassium hydrophosphate	KH_2PO_4	30.5 g

Make up to 1 L with QH_2O. Filter sterilise.

1000 X Carbonate stock (1 L)

Sodium carbonate	Na_2CO_3	20 g

Make up to 1 L with QH_2O. Filter sterilise.

1 M glucose stock (1 L)

Glucose	180 g

Make up to 1 L with QH_2O. Dissolve in water bath. Filter sterilise.

1 M TES stock (1 L)

TES	$C_6H_{15}NO_6S$	229.2 g

Make up to 1 L with QH_2O. pH to 8.2 with KOH. Filter sterilise.

1 X BG11 (1 L)

100 X BG11	10 ml
1000 X Iron	1 ml
1000 X Phosphate	1 ml
1000 X Carbonate	1 ml

Autoclave. Before use add TES to a final concentration of 10 mM and 1 M glucose to a final concentration of 5 mM

References

Aagaard, J. and Sistrom, W. R. (1972). Control of synthesis of reaction center bacteriochlorophyll in photosynthetic bacteria. *Photochemistry and Photobiology* 15, 209.

Abrahams, J. P., Leslie, A. G. W., Lutter, R. and Walker, J. E. (1994). Structure at 2.8-angstrom resolution of f1-atpase from bovine heart-mitochondria. *Nature* 370, 621-628.

Adams, P. G. and Hunter, C. N. (2012). Adaptation of intracytoplasmic membranes to altered light intensity in *Rhodobacter sphaeroides*. *Biochimica Et Biophysica Acta-Bioenergetics* 1817, 1616-1627.

Adams, P. G., Mothersole, D. J., Ng, I. W., Olsen, J. D. and Hunter, C. N. (2011). Monomeric RC-LH1 core complexes retard LH2 assembly and intracytoplasmic membrane formation in PufX-minus mutants of *Rhodobacter sphaeroides*. *Biochimica Et Biophysica Acta-Bioenergetics* 1807, 1044-1055.

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

Alanen, H. I., Walker, K. L., Suberbie, M. L. V., Matos, C. F. R. O., Boenisch, S., Freedman, R. B., Keshavarz-Moore, E., Ruddock, L. W. and Robinson, C. (2015). Efficient export of human growth hormone, interferon alpha 2b and antibody fragments to the periplasm by the *Escherichia coli* Tat pathway in the absence of prior disulfide bond formation. *Biochimica Et Biophysica Acta-Molecular Cell Research* 1853, 756-763.

Aldridge, C., Spence, E., Kirkilionis, M. A., Frigerio, L. and Robinson, C. (2008). Tat-dependent targeting of Rieske iron-sulphur proteins to both the plasma and thylakoid membranes in the cyanobacterium *Synechocystis* PCC6803. *Molecular Microbiology* 70, 140-150.

Algar, W. R., Ancona, M. G., Malanoski, A. P., Susumu, K. and Medintz, I. L. (2012). Assembly of a concentric Forster resonance energy transfer relay on a quantum dot scaffold: characterization and application to multiplexed protease sensing. *Acs Nano* 6, 11044-11058.

Allen, J. P., Feher, G., Yeates, T. O., Komiya, H. and Rees, D. C. (1987a). Structure of the reaction center from *Rhodobacter sphaeroides* R-26 - the cofactors .1. *Proceedings of the National Academy of Sciences of the United States of America* 84, 5730-5734.

Allen, J. P., Feher, G., Yeates, T. O., Komiya, H. and Rees, D. C. (1987b). Structure of the reaction center from *Rhodobacter sphaeroides* R-26 - the protein subunits. *Proceedings of the National Academy of Sciences of the United States of America* 84, 6162-6166.

Alvey, R. M., Biswas, A., Schluchter, W. M. and Bryant, D. A. (2011). Attachment of Noncognate Chromophores to CpcA of *Synechocystis* sp PCC 6803 and *Synechococcus* sp PCC 7002 by Heterologous Expression in *Escherichia coli*. *Biochemistry*, 50, 4890-4902.

Anderson, J. L. R., Armstrong, C. T., Kodali, G., Lichtenstein, B. R., Watkins, D. W., Mancini, J. A., Boyle, A. L., Farid, T. A., Crump, M. P., Moser, C. C. and Dutton, P. L. (2014). Constructing a man-made c-type cytochrome maquette in vivo: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. *Chemical Science*, 5, 507-514.

Anderson, J. L. R., Koder, R. L., Moser, C. C. and Dutton, P. L. (2008). Controlling complexity and water penetration in functional de novo protein design. *Biochemical Society Transactions*, 36, 1106-1111.

Aratani N and Osuka A (2010) Synthetic strategies toward multiporphyrinic arrays. *In*: Kadish KM, Smith KM, Guilard R (eds) Handbook of porphyrin science, Vol 1. *World Scientific Publishing Co*., Singapore, pp 1–132

Ashby, M. K., Coomber, S. A. and Hunter, C. N. (1987). Cloning, nucleotide-sequence and transfer of genes for the b800-850 light harvesting complex of *Rhodobacter sphaeroides. Febs Letters*, 213, 245-248.

Atamna-Ismaeel, N., Sabehi, G., Sharon, I., Witzel, K.-P., Labrenz, M., Juergens, K., Barkay, T., Stomp, M., Huisman, J. and Beja, O. (2008). Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *Isme Journal*, 2, 656-662.

Bachmann, J., Bauer, B., Zwicker, K., Ludwig, B. and Anderka, O. (2006). The Rieske protein from *Paracoccus denitrificans* is inserted into the cytoplasmic membrane by the twin-arginine translocase. *Febs Journal*, 273, 4817-4830.

Bahatyrova, S., Frese, R. N., Siebert, C. A., Olsen, J. D., van der Werf, K. O., van Grondelle, R., Niederman, R. A., Bullough, P. A., Otto, C. and Hunter, C. N. (2004a). The native architecture of a photosynthetic membrane. *Nature*, 430, 1058-1062.

Bahatyrova, S., Frese, R. N., van der Werf, K. O., Otto, C., Hunter, C. N. and Olsen, J. D. (2004b.) Flexibility and size heterogeneity of the LH1 light harvesting complex revealed by atomic force microscopy - Functional significance for bacterial photosynthesis. *Journal of Biological Chemistry*, 279, 21327-21333.

Balashov, S. P., Imasheva, E. S., Boichenko, V. A., Anton, J., Wang, J. M. and Lanyi, J. K. (2005). Xanthorhodopsin: A proton pump with a light-harvesting carotenoid antenna. *Science*, 309, 2061-2064.

Balashov, S. P., Petrovskaya, L. E., Lukashev, E. P., Imasheva, E. S., Dioumaev, A. K., Wang, J. M., Sychev, S. V., Dolgikh, D. A., Rubin, A. B., Kirpichnikov, M. P. and Lanyi, J. K. (2012). Aspartate-histidine interaction in the retinal schiff base counterion of the light-driven proton pump of *Exiguobacterium sibiricum*. *Biochemistry*, 51, 5748-5762.

Bamann, C., Bamberg, E., Wachtveitl, J. and Glaubitz, C. (2014). Proteorhodopsin. *Biochimica Et Biophysica Acta-Bioenergetics*, 1837, 614-625.

Barber, J. (2009). Photosynthetic energy conversion: natural and artificial. *Chemical Society Reviews*, 38, 185-196.

Barrett, C. M. L., Ray, N., Thomas, J. D., Robinson, C. and Bolhuis, A. (2003). Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in *Escherichia coli. Biochemical and Biophysical Research Communications*, 304, 279-284.

Beck, D., Vasisht, N., Baglieri, J., Monteferrante, C. G., van Dijl, J. M., Robinson, C. and Smith, C. J. (2013). Ultrastructural characterisation of *Bacillus subtilis* TatA complexes suggests they are too small to form homooligomeric translocation pores. *Biochimica Et Biophysica Acta-Molecular Cell Research*, 1833, 1811-1819.

Beekman, L. M. P., Vanmourik, F., Jones, M. R., Visser, H. M., Hunter, C. N. and van Grondelle, R. (1994). Trapping kinetics in mutants of the photosynthetic purple bacterium *Rhodobacter sphaeroides* - influence of the charge separation rate and consequences for the rate-limiting step in the light-harvesting process. *Biochemistry*, 33, 3143-3147.

Beja, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., Jovanovich, S., Gates, C. M., Feldman, R. A., Spudich, J. L., Spudich, E. N. and DeLong, E. F. (2000). Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science*, 289, 1902-1906.

Beja, O., Spudich, E. N., Spudich, J. L., Leclerc, M. and DeLong, E. F. (2001). Proteorhodopsin phototrophy in the ocean. *Nature,* 411, 786-789.

Belasco, J. G., Beatty, J. T., Adams, C. W., Vongabain, A. and Cohen, S. N. (1985). Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic rxca transcript. *Cell*, 40, 171-181.

Bergo, V. B., Sineshchekov, O. A., Kralj, J. M., Partha, R., Spudich, E. N., Rothschild, K. J. and Spudich, J. L. (2009). His-75 in Proteorhodopsin, a Novel Component in Light-driven Proton Translocation by Primary Pumps. *Journal of Biological Chemistry*, 284, 2836-2843.

Berks, B. C. (1996). A common export pathway for proteins binding complex redox cofactors? *Molecular Microbiology*, 22, 393-404.

Berry, E. A., Huang, L. S., Saechao, L. K., Pon, N. G., Valkova-Valchanova, M. and Daldal, F. (2004). X-ray structure of *Rhodobacter capsulatus* cytochrome bc_1: comparison with its mitochondrial and chloroplast counterparts. *Photosynthesis Research*, 81, 251-275.

Bieszke, J. A., Braun, E. L., Bean, L. E., Kang, S. C., Natvig, D. O. and Borkovich, K. A. (1999). The *nop-1* gene of *Neurospora crassa* encodes a seven transmembrane helix retinal-binding protein homologous to archaeal rhodopsins. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 8034-8039.

Birge, R. R. (1990). Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin. *Biochimica Et Biophysica Acta*, 1016, 293-327.

Bixon, M. and Jortner, J. (1999). Electron transfer - from isolated molecules to biomolecules. *Advances in chemical physics, Pt 1*, 106, 35-202.

Blankenship, R. E. (2002). Genetics, Assembly and Regulation of Photosynthetic Systems. Molecular mechanisms of photosynthesis (Blankenship, R. E., ed.) *Blackwell Science Ltd*, Oxford, U.K., 204-219.

Blankenship, R. E., Madigan, M. T. and Bauer, C. E. (Eds.). (1995). Anoxygenic photosynthetic bacteria (Vol. 2). *Springer*.

Blankenship, R. E., Tiede, D. M., Barber, J., Brudvig, G. W., Fleming, G., Ghirardi, M., Gunner, M. R., Junge, W., Kramer, D. M., Melis, A., Moore, T. A., Moser, C. C., Nocera, D. G., Nozik, A. J., Ort, D. R., Parson, W. W., Prince, R. C. And Sayre, R. T. (2011). Comparing Photosynthetic and Photovoltaic Efficiencies and Recognizing the Potential for Improvement. *Science*, 332, 805-809.

Bogsch, E., Brink, S. and Robinson, C. (1997). Pathway specificity for a Delta pH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein. *Embo Journal*, 16, 3851-3859.

Bohorquez, L. C., Ruiz-Perez, C. A. and Mercedes Zambrano, M. (2012). Proteorhodopsin-Like Genes Present in Thermoacidophilic High-Mountain Microbial Communities. *Applied and Environmental Microbiology*, 78, 7813-7817.

Bokman, S. H. and Ward, W. W. (1981). Renaturation of *Aequorea* green-fluorescent protein. *Biochemical and Biophysical Research Communications*, 101, 1372-1380.

Bollivar, D. W., Wang, S. J., Allen, J. P. And Bauer, C. E. (1994). Molecular-genetic analysis of terminal steps in bacteriochlorophyll a biosynthesis - characterization of a *Rhodobacter capsulatus* strain that synthesizes geranylgeraniol-esterified bacteriochlorophyll *a*. *Biochemistry*, 33, 12763-12768.

Booth, P. J. (2000). Unravelling the folding of bacteriorhodopsin. *Biochimica Et Biophysica Acta-Bioenergetics*, 1460, 4-14.

Borland, C. F., Cogdell, R. J., Land, E. J. And Truscott, T. G. (1989). Bacteriochlorophyll alpha-triplet state and its interactions with bacterial carotenoids and oxygen. *Journal of Photochemistry and Photobiology B-Biology*, 3, 237-245.

Botta, C., Betti, P. and Pasini, M. (2013). Organic nanostructured host-guest materials for luminescent solar concentrators. *Journal of Materials Chemistry A*, 1, 510-514.

Bowyer, J. R., Hunter, C. N., Ohnishi, T. and Niederman, R. A. (1985). Photosynthetic membrane-development in *Rhodopseudomonas sphaeroides* - spectral and kinetic characterization of redox components of light-driven electron flow in apparent photosynthetic membrane growth initiation sites. *Journal of Biological Chemistry*, 260, 3295-3304.

Bozdemir, O. A., Erbas-Cakmak, S., Ekiz, O. O., Dana, A. and Akkaya, E. U. (2011). Towards unimolecular luminescent solar concentrators: bodipy-based dendritic energy-transfer cascade with panchromatic absorption and monochromatized emission. *Angewandte Chemie-International Edition*, 50, 10907-10912.

Bradforth, S. E., Jimenez, R., Vanmourik, F., Vangrondelle, R. and Fleming, G. R. 1995. Excitation transfer in the core light-harvesting complex (LH-1) of *Rhodobacter sphaeroides* - an ultrafast fluorescence depolarization and annihilation study. *Journal of Physical Chemistry*, 99, 16179-16191.

Breton, J., Martin, J. L., Migus, A., Antonetti, A. and Orszag, A. (1986). Femtosecond spectroscopy of excitation-energy transfer and initial charge separation in the reaction center of the photosynthetic bacterium *Rhodopseudomonas viridis*. *Proceedings of the National Academy of Sciences of the United States of America*, 83, 5121-5125.

Brune, D. C. (1995). Sulfur compounds as photosynthetic electron donors. *Anoxygenic Photosynthetic Bacteria*, 847-870.

Bruser, T. and Sanders, C. (2003). Hypothesis-review - An alternative model of the twin arginine translocation system. *Microbiological Research*, 158, 7-17

Bryant, D. A., Costas, A. M. G., Maresca, J. A., Chew, A. G. M., Klatt, C. G., Bateson, M. M., Tallon, L. J., Hostetler, J., Nelson, W. C., Heidelberg, J. F. and Ward, D. M. (2007). *Candidatus Chloracidobacterium thermophilum*: An aerobic phototrophic acidobacterium. *Science*, 317, 523-526.

Buchanan, G., Sargent, F., Berks, B. C. and Palmer, T. (2001). A genetic screen for suppressors of *Escherichia coli* Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif. *Archives of Microbiology*, 177, 107-112.

Burton, A. J., Thomas, F., Agnew, C., Hudson, K. L., Halford, S. E., Brady, R. L. and Woolfson, D. N. (2013). Accessibility, Reactivity, and Selectivity of Side Chains within a Channel of de Novo Peptide Assembly. *Journal of the American Chemical Society*, 135, 12524-12527.

Cartron, M. L., Olsen, J. D., Sener, M., Jackson, P. J., Brindley, A. A., Qian, P., Dickman, M. J., Leggett, G. J., Schulten, K. and Hunter, C. N. (2014). Integration of energy and electron transfer processes in the photosynthetic membrane of *Rhodobacter sphaeroides*. *Biochimica Et Biophysica Acta-Bioenergetics*, 1837, 1769-1780.

Chaddock, A. M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R. G., Klosgen, R. B. and Robinson, C. (1995). A new-type of signal peptide - central role of a twin-arginine motif in transfer signals for the delta-ph-dependent thylakoidal protein translocase. *Embo Journal*, 14, 2715-2722.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene-expression. *Science*, 263, 802-805.

Chang, C. H., Elkabbani, O., Tiede, D., Norris, J. And Schiffer, M. (1991). Structure of the membrane-bound protein photosynthetic reaction center from *Rhodobacter sphaeroides*. *Biochemistry*, 30, 5352-5360.

Chang, C. H., Tiede, D., Tang, J., Smith, U., Norris, J. and Schiffer, M. (1986). Structure of *Rhodopseudomonas sphaeroides* R26 reaction center. *Febs Letters*, 205, 82-86.

Chang, H. C., Kaiser, C. M., Hartl, F. U. And Barral, J. M. (2005). De novo folding of GFP fusion proteins: High efficiency in eukaryotes but not in bacteria. *Journal of Molecular Biology*, 353, 397-409.

Channon, K., Bromley, E. H. C. and Woolfson, D. N. (2008). Synthetic biology through biomolecular design and engineering. *Current Opinion in Structural Biology*, 18, 491-498.

Chen, X. X. (1999) Ph.D. thesis, Biochemistry and Biophysics. University of Pennsylvania, Philadelphia, PA.

Chen, X. X., Moser, C. C., Pilloud, D. L. and Dutton, P. L. (1998). Molecular orientation of Langmuir-Blodgett films of designed heme protein and lipoprotein maquettes. *Journal of Physical Chemistry B*, 102, 6425-6432.

Chi, S. C., Mothersole, D. J., Dilbeck, P., Niedzwiedzki, D. M., Zhang, H., Qian, P., Vasilev, C., Grayson, K. J., Jackson, P. J., Martin, E. C., Li, Y., Holten, D. and Hunter, C. N. (2015). Assembly of functional photosystem complexes in *Rhodobacter sphaeroides* incorporating carotenoids from the spirilloxanthin pathway. *Biochimica Et Biophysica Acta-Bioenergetics*, 1847, 189-201.

Chobot, S. E., Wiedman, G., Moser, C. C., Discher, B. M. and Dutton, P. L. (2010). Design and Characterization of an Enzymatically Active Amphiphilic Maquette Protein. *Biophysical Journal*, 98, 639A-639A.

Chou, P. Y. and Fasman, G. D. (1978). Empirical predictions of protein conformation. *Annual Review of Biochemistry*, 47, 251-276.

Clark, S. A. and Theg, S. M. (1997). A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Molecular Biology of the Cell*, 8, 923-934.

Clayton, R. K. (1962). Primary reactions in bacterial photosynthesis. 3. Reactions of carotenoids and cytochromes in illuminated bacterial chromatophores. *Photochemistry and Photobiology*, 1, 313-323.

Clayton, R. K. And Smith, C. (1960). *Rhodopseudomonas spheroides* - high catalase and blue-green double mutants. *Biochemical and Biophysical Research Communications*, 3, 143-145.

Cline, K. and Mori, H. (2001). Thylakoid Delta ph-dependent precursor proteins bind to a cptatc-Hcf106 complex before Tha4-dependent transport. *Journal of Cell Biology*, 154, 719-729.

Cline, K., Ettinger, W. F. and Theg, S. M. (1992). Protein-specific energy-requirements for protein-transport across or into thylakoid membranes - 2 lumenal proteins are transported in the absence of ATP. *Journal of Biological Chemistry*, 267, 2688-2696.

Cogdell, R. J. and Frank, H. A. (1987). How carotenoids function in photosynthetic bacteria. *Biochimica Et Biophysica Acta*, 895, 63-79.

Cohen-Bazire, G. and Kunisawa, R. (1963). Fine structure of *Rhodospirillum rubrum*. *Journal of Cell Biology*, 16, 401-.

Cooper, D. E., Rands, M. B. and Woo, C. P. (1975). Sulfide reduction in fellmongery effluent by red sulfur bacteria. *Journal Water Pollution Control Federation*, 47, 2088-2100.

Cormack, B. P., Valdivia, R. H. and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173, 33-38.

Crameri, A., Whitehorn, E. A., Tate, E. and Stemmer, W. P. C. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotechnology*, 14, 315-319.

Cristobal, S., de Gier, J. W., Nielsen, H. and von Heijne, G. (1999). Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *Embo Journal*, 18, 2982-2990.

Crouch, L. I. and Jones, M. R. (2012). Cross-species investigation of the functions of the *Rhodobacter* pufx polypeptide and the composition of the RC-LH1 core complex. *Biochimica Et Biophysica Acta-Bioenergetics*, 1817, 336-352.

Cubitt, A. B., Woollenweber, L. A. and Heim, R. (1999). Understanding structure-function relationships in the *Aequorea victoria* green fluorescent protein. *Methods in Cell Biology*, Vol 58, 58, 19-.

Currin, A., Swainston, N., Day, P. J. and Kell, D. B. (2015). Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chemical Society Reviews,* 44, 1172-239

Darwin, C., (1862). The origin of species.

Day, R. N. and Davidson, M. W. (2009). The fluorescent protein palette: tools for cellular imaging. *Chemical Society Reviews*, 38, 2887-2921.

De Buck, E., Vranckx, L., Meyen, E., Maes, L., Vandersmissen, L., Anne, J. and Lannnertyn, E. (2007). The twin-arginine translocation pathway is necessary for correct membrane insertion of the Rieske Fe/S protein in *Legionella pneumophila. Febs Letters*, 581, 259-264.

De Leeuw, E., Porcelli, I., Sargent, F., Palmer, T. and Berks, B. C. (2001). Membrane interactions and self-association of the tata and tatb components of the twin-arginine translocation pathway. *Febs Letters*, 506, 143-148.

Deisenhofer, J. And Michel, H. (1989). The photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis*. *Science,* 245, 1463-1473.

Deisenhofer, J. And Michel, H. (1991). High-resolution structures of photosynthetic reaction centers. *Annual Review of Biophysics and Biophysical Chemistry*, 20, 247-266.

Deisenhofer, J., and Norris, J. R. (1993). The photosynthetic reaction center, volumes I and II. *San Diego Academic Press*

Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. (1985). Structure of the protein subunits in the photosynthetic reaction center of *Rhodopseudomonas viridis* at 3 Å resolution. *Nature*, 318, 618-624.

Deisenhofer, J., Epp, O., Sinning, I. And Michel, H. (1995). Crystallographic refinement at 2.3-angstrom resolution and refined model of the photosynthetic reaction center from *Rhodopseudomonas viridis*. *Journal of Molecular Biology*, 246, 429-457.

DeLisa, M. P., Lee, P., Palmer, T. and Georgiou, G. (2004). Phage shock protein pspa of *Escherichia coli* relieves saturation of protein export via the Tat pathway. *Journal of Bacteriology*, 186, 366-373.

Delisa, M. P., Samuelson, P., Palmer, T. and Georgiou, G. (2002). Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *Journal of Biological Chemistry*, 277, 29825-29831.

DeLisa, M. P., Tullman, D. and Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 6115-6120.

DeLong, E. F. and Beja, O. (2010). The Light-Driven Proton Pump Proteorhodopsin Enhances Bacterial Survival during Tough Times. *Plos Biology*, 8.

Dioumaev, A. K., Brown, L. S., Shih, J., Spudich, E. N., Spudich, J. L. and Lanyi, J. K. (2002). Proton transfers in the photochemical reaction cycle of proteorhodopsin. *Biochemistry*, 41, 5348-5358.

Dioumaev, A. K., Wang, J. M., Balint, Z., Varo, G. and Lanyi, J. K. (2003). Proton transport by proteorhodopsin requires that the retinal Schiff base counterion Asp-97 be anionic. *Biochemistry*, 42, 6582-6587.

Discher, B. M., Koder, R. L., Moser, C. C. and Dutton, P. L. (2003). Hydrophillic to amphiphilic design in redox protein maquettes. *Current Opinion in Chemical Biology*, 7, 741-748.

Discher, B. M., Noy, D., Strzalka, J., Ye, S. X., Moser, C. C., Lear, J. D., Blasie, J. K. and Dutton, P. L. 2005. Design of amphiphilic protein maquettes: Controlling assembly, membrane insertion, and cofactor interactions. *Biochemistry*, 44, 12329-12343.

Donohue, T. J., Hoger, J. H. and Kaplan, S. (1986). Cloning and expression of the *Rhodobacter sphaeroides* reaction center-h gene. *Journal of Bacteriology*, 168, 953-961.

Drews, G. (1960). Untersuchungen zur substruktur der chromatophoren von *Rhodospirillum rubrum* und *Rhodospirillum molischianum. Archiv Fur Mikrobiologie*, 36, 99-108.

Dutton, P. L. and Moser, C. C. (2011). Engineering enzymes. *Faraday Discussions*, 148.

Dykxhoorn, D. M., Stpierre, R. and Linn, T. (1996). A set of compatible tac promoter expression vectors. *Gene*, 177, 133-136.

Eggink, L. L. and Hoober, J. K. (2000). Chlorophyll binding to peptide maquettes containing a retention motif. *Journal of Biological Chemistry*, 275, 9087-9090.

Ermler, U., Fritzsch, G., Buchanan, S. K. and Michel, H. (1994). Structure of the photosynthetic reaction-center from *Rhodobacter-sphaeroides* at 2.65-angstrom resolution - cofactors and protein-cofactor interactions. *Structure*, 2, 925-936.

Esser, L., Elberry, M., Zhou, F., Yu, C.-A., Yu, L. and Xia, D. (2008). Inhibitor-complexed structures of the cytochrome *bc*(1) from the photosynthetic bacterium *Rhodobacter sphaeroides. Journal of Biological Chemistry*, 283, 2846-2857.

Esser, L., Gong, X., Yang, S., Yu, L., Yu, C.-A. and Xia, D. (2006). Surface-modulated motion switch: Capture and release of iron-sulfur protein in the cytochrome *bc*(1) complex. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 13045-13050.

Fabian, A., Horvath, G., Vamosi, G., Vereb, G. and Szoellosi, J. (2013). Triplefret measurements in flow cytometry. *Cytometry Part A*, 83A, 375-385.

Farid, T. A., Kodali, G., Solomon, L. A., Lichtenstein, B. R., Sheehan, M. M., Fry, B. A., Bialas, C., Ennist, N. M., Siedlecki, J. A., Zhao, Z., Stetz, M. A., Valentine, K. G., Anderson, J. L. R., Wand, A. J., Discher, B. M., Moser, C. C. and Dutton, P. L. (2013). Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nature Chemical Biology*, 9, 826-.

Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H. and Phillips, G. J. (2000). Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli. Journal of Bacteriology*, 182, 4068-4076.

Feniouk, B. A. and Junge, W. (2009). Proton Translocation and ATP Synthesis by the F0F1-atpase of Purple Bacteria. *Advances in Photosynthesis and Respiration*, 28, 475-493.

Finkel, O. M., Beja, O. and Belkin, S. (2013). Global abundance of microbial rhodopsins. *Isme Journal,* 7, 448-451.

Fischer, E., Synthesen in der Purin- und Zuckergruppe. In Les Prix Nobel en 1902, 1905.

Fleming, G. R. and van Frondelle, R. (1994). The primary steps of photosynthesis. *Physics Today*, 47, 48-55.

Follenius-Wund, A., Bourotte, M., Schmitt, M., Iyice, F., Lami, H., Bourguignon, J. J., Haiech, J. and Pigault, C. (2003). Fluorescent derivatives of the GFP chromophore give a new insight into the GFP fluorescence process. *Biophysical Journal*, 85, 1839-1850.

Foote, C. S. and Denny, R. W. (1968). Chemistry of singlet oxygen .7. Quenching by beta-carotene. *Journal of the American Chemical Society*, 90, 6233-.

Foster, K. W., Saranak, J., Patel, N., Zarilli, G., Okabe, M., Kline, T. and Nakanishi, K. (1984). A rhodopsin is the functional photoreceptor for phototaxis in the unicellular eukaryote *chlamydomonas*. *Nature*, 311, 756-759.

Francia, F., Wang, J., Venturoli, G., Melandri, B. A., Barz, W. P. and Oesterhelt, D. (1999). The reaction center-LH1 antenna complex of *Rhodobacter sphaeroides* contains one pufx molecule which is involved in dimerization of this complex. *Biochemistry*, 38, 6834-6845.

Freer, A. A., Prince, S. M., Sauer, K., Papiz, M. Z., Hawthornthwaite-Lawless, A. M., mcDermott, G., Cogdell, R. J. and Isaacs, N. W. (1996) Pigment-pigment interactions and energy transfer in the antenna complex of the photosynthetic bacterium *Rhodopseudomonas acidophila. Structure* 4, 449-462.

Friedrich, T., Geibel, S., Kalmbach, R., Chizhov, I., Ataka, K., Heberle, J., Engelhard, M. and Bamberg, E. (2002). Proteorhodopsin is a light-driven proton pump with variable vectoriality. *Journal of Molecular Biology*, 321, 821-838.

Froebel, J., Rose, P. and Mueller, M. (2011). Early Contacts between Substrate Proteins and tata Translocase Component in Twin-arginine Translocation. *Journal of Biological Chemistry*, 286, 43679-43689.

Froebel, J., Rose, P. and Mueller, M. (2012). Twin-arginine-dependent translocation of folded proteins. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367, 1029-1046.

Fry, B. A., Wiedman, G. R., Moser, C. C., Dutton, P. L. and Discher, B. M. (2011). Design of Transmembrane Electron Transport Chain within Amphiphilic Protein Maquettes. *Biophysical Journal*, 100, 131-132.

Fujii, R., Adachi, S., Roszak, A. W., Gardiner, A. T., Cogdell, R. J., Isaacs, N. W., Koshihara, S. and Hashimoto, H. (2010). 3I4D: Photosynthetic reaction center from *Rhodobacter sphaeroides* 2.4.1. Worldwide Protein Data Bank.

Garcia-Asua, G., (1999) PhD. Thesis. University of Sheffield. Studies on the biosynthesis and functions of carotenoids in photosynthesis, using *Rhodobacter sphaeroides* as a model organism, Chapter 3

Garcia-Asua, G., Cogdell, R. J. and Hunter, C. N. (2002). Functional assembly of the foreign carotenoid lycopene into the photosynthetic apparatus of *Rhodobacter sphaeroides*, achieved by replacement of the native 3-step phytoene desaturase with its 4-step counterpart from *Erwinia herbicola. Molecular Microbiology*, 44, 233-244.

Gehne, S., Flehr, R., Kienzler, A. A. N., Berg, M., Bannwarth, W. and Kumke, M. U. (2012). Dye Dynamics in Three-Color FRET Samples. *Journal of Physical Chemistry B*, 116, 10798-10806.

Gibney, B. R., Mulholland, S. E., Rabanal, F. and Dutton, P. L. (1996). Ferredoxin and ferredoxin-heme maquettes. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 15041-15046.

Gill, S. C. and Von Hippel, P. H. (19890. Calculation of protein extinction coefficients from amino-acid sequence data. *Analytical Biochemistry*, 182, 319-326.

Giovannoni, S. J., Bibbs, L., Cho, J. C., Stapels, M. D., Desiderio, R., Vergin, K. L., Rappe, M. S., Laney, S., Wilhelm, L. J., Tripp, H. J., Mathur, E. J. and Barofsky, D. F. (2005). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature*, 438, 82-85.

Glaeser, J. and Klug, G. (2005). Photo-oxidative stress in *Rhodobacter sphaeroides*: protective role of carotenoids and expression of selected genes. *Microbiology-Sgm*, 151, 1927-1938.

Gohlke, U., Pullan, L., mcdevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H. R. and Berks, B. C. (2005). The tata component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 10482-10486.

Golecki, J. R., Ventura, S. and Oelze, J. (1991). The architecture of unusual membrane tubes in the b800-850 light-harvesting bacteriochlorophyll-deficient mutant 19 of *Rhodobacter sphaeroides. Fems Microbiology Letters*, 77, 335-340.

Gomez-Consarnau, L., Akram, N., Lindell, K., Pedersen, A., Neutze, R., Milton, D. L., Gonzalez, J. M. and Pinhassi, J. (2010). Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation. *Plos Biology*, 8.

Gomez-Consarnau, L., Gonzalez, J. M., Coll-Llado, M., Gourdon, P., Pascher, T., Neutze, R., Pedros-Alio, C. and Pinhassi, J. (2007). Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature*, 445, 210-213.

Gopta, O. A., Cherepanov, D. A., Junge, W. and Mulkidjanian, A. Y. (1999). Proton transfer from the bulk to the bound ubiquinone Q(B) of the reaction center in chromatophores of *Rhodobacter sphaeroides*: Retarded conveyance by neutral water. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 13159-13164.

Gorbunov, M. Y., Kolber, Z. S., Lesser, M. P. and Falkowski, P. G. (2001). Photosynthesis and photoprotection in symbiotic corals. *Limnology and Oceanography*, 46, 75-85.

Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. and Henderson, R. (1996). Electron-crystallographic refinement of the structure of bacteriorhodopsin. *Journal of Molecular Biology*, 259, 393-421.

Grosset, A. M., Gibney, B. R., Rabanal, F., Moser, C. C. and Dutton, P. L. (2001). Proof of principle in a de novo designed protein maquette: An allosterically regulated, charge-activated conformational switch in a tetra-alpha-helix bundle. *Biochemistry*, 40, 5474-5487.

Haehnel, W., Noy, D., and Scheer, H. (2009) De novo designed bacteriochlorophyll-binding helix-bundle proteins. *In*: The Purple Phototrophic Bacteria. *Springer.*

Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R. and Sprenger, G. A. (1999). The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *European Journal of Biochemistry*, 263, 543-551.

Hardin, B. E., Hoke, E. T., Armstrong, P. B., Yum, J.-H., Comte, P., Torres, T., Frechet, J. M. J., Nazeeruddin, M. K., Graetzel, M. and Mcgehee, M. D. (2009). Increased light harvesting in dye-sensitized solar cells with energy relay dyes. *Nature Photonics*, 3, 406-411.

Harris, M. A., Jiang, J., Niedzwiedzki, D. M., Jiao, J., Taniguchi, M., Kirmaier, C., Loach, P. A., Bocian, D. F., Lindsey, J. S., Holten, D. and Parkes-Loach, P. S. (2014a). Versatile design of biohybrid light-harvesting architectures to tune location, density, and spectral coverage of attached synthetic chromophores for enhanced energy capture. *Photosynthesis Research*, 121, 35-48.

Harris, M. A., Parkes-Loach, P. S., Springer, J. W., Jiang, J., Martin, E. C., Qian, P., Jiao, J., Niedzwiedzki, D. M., Kirmaier, C., Olsen, J. D., Bocian, D. F., Holten, D., Hunter, C. N., Lindsey, J. S. and Loach, P. A. (2013). Integration of multiple chromophores with native photosynthetic antennas to enhance solar energy capture and delivery. *Chemical Science*, 4, 3924-3933.

Harris, M. A., Sahin, T., Jiang, J., Vairaprakash, P., Parkes-Loach, P. S., Niedzwiedzki, D. M., Kirmaier, C., Loach, P. A., Bocian, D. F., Holten, D. and Lindsey, J. S. (2014b). Enhanced Light-Harvesting Capacity by Micellar Assembly of Free Accessory Chromophores and LH1-like Antennas. *Photochemistry and Photobiology*, 90, 1264-1276.

Hartz, A. J., Sherr, B. F. and Sherr, E. B. (2011). Photoresponse in the Heterotrophic Marine Dinoflagellate *Oxyrrhis marina. Journal of Eukaryotic Microbiology*, 58, 171-177.

Harvey, P. D., (2003) Recent advances in free and metalated multiporphyrin assemblies and arrays; a photophysical behavior and energy transfer perspective. *In*: Kadish KM, Smith KM, Guilard R (eds) The porphyrin handbook, vol 18. *Academic Press*, San Diego, pp 63–250

Harvey, P. D., Stern, C. and Guilard, R. (2011). Bio-inspired Molecular Devices Based on Systems Found in Photosynthetic Bacteria. Handbook of Porphyrin Science with Applications to Chemistry, Physics, Materials Science, Engineering, Biology and Medicine, Vol 11: Catalysis and Bio-Inspired Systems, Pt Ii, 11, 1-179.

Hauer, F., Gerle, C., Fischer, N., Oshima, A., Shinzawa-Itoh, K., Shimada, S., Yokoyama, K., Fujiyoshi, Y. and Stark, H. (2015). GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. *Structure* (London, England : 1993), 23, 1769-75.

Hauer, R. S., Schlesier, R., Heilmann, K., Dittmar, J., Jakob, M. and Kloesgen, R. B. (2013). Enough is enough: tata demand during Tat-dependent protein transport. *Biochimica Et Biophysica Acta-Molecular Cell Research*, 1833, 957-965.

Hawthornthwaite, A. M. and Cogdell, R. J. (1993). Bacteriochlorophyll-binding proteins. *In*: The Chlorophylls (Scheer, H., ed.), pp. 493–528, CRC, Boca Raton, Florida.

Hempelmann, F., Hoelper, S., Verhoefen, M.-K., Woerner, A. C., Koehler, T., Fiedler, S.-A., Pfleger, N., Wachtveitl, J. and Glaubitz, C. (2011). His75-Asp97 Cluster in Green Proteorhodopsin. *Journal of the American Chemical Society*, 133, 4645-4654.

Herdman, M., Janvier, M., Waterbury, J.B., Rippka, R., Stanier, R.Y. and Mandel, M. (1979) Deoxyribonucleic acid base composition of cyanobacteria. *Journal of General Microbiology* 111, 63-71

Hess, S., Akesson, E., Cogdell, R. J., Pullerits, T. and Sundstrom, V. (1995a). Energy transfer in spectrally inhomogeneous light-harvesting pigment-protein complexes of purple bacteria. *Biophysical Journal*, 69, 2211-2225.

Hess, S., Chachisvilis, M., Timpmann, K., Jones, M. R., Fowler, G. J. S., Hunter, C. N. and Sundstrom, V. (1995b). Temporally and spectrally resolved subpicosecond energy transfer within the peripheral antenna complex (LH2) and from LH2 to the core antenna complex in photosynthetic purple bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 12333-12337.

Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., Ikeuchi, M. (2001). DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *Plant Cell* 13, 793-806.

Hoff, A. J. and Deisenhofer, J. (1997). Photophysics of photosynthesis. Structure and spectroscopy of reaction centers of purple bacteria. *Physics Reports-Review Section of Physics Letters*, 287, 1-247.

Hoff, W. D., Jung, K. H. and Spudich, J. L. (1997). Molecular mechanism of photosignaling by archaeal sensory rhodopsins. *Annual Review of Biophysics and Biomolecular Structure*, 26, 223-258.

Hokanson, Sarah C., (2010). Deconvoluting the engineering and assembly instructions for complex iii activity. Publicly accessible Penn Dissertations. Paper 105.

Holzapfel, W., Finkele, U., Kaiser, W., Oesterhelt, D., Scheer, H., Stilz, H. U. and Zinth, W. (1990). Initial electron-transfer in the reaction center from *Rhodobacter sphaeroides*. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 5168-5172.

Hopf, M., Gohring, W., Ries, A., Timpl, R. and Hohenester, E. (2001). Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1. *Nature Structural Biology*, 8, 634-640.

Huang, S. S., Gibney, B. R., Stayrook, S. E., Dutton, P. L. and Lewis, M. (2003). X-ray structure of a Maquette scaffold. *Journal of Molecular Biology*, 326, 1219-1225.

Huang, S. S., Koder, R. L., Lewis, M., Wand, A. J. and Dutton, P. L. (2004). The HP-1 maquette: From an apoprotein structure to a structured hemoprotein designed to promote redox-coupled proton exchange. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5536-5541.

Hunte, C., Solmaz, S., Palsdottir, H. and Wenz, T. (2008). A Structural Perspective on Mechanism and Function of the Cytochrome *bc*(1) Complex. *Results and Problems in Cell Differentiation*, 45, 253-278.

Hunter, C. N. and Turner, G. (1988). Transfer of genes-coding for apoproteins of reaction center and light-harvesting LH1 complexes to *Rhodobacter sphaeroides*. *Journal of General Microbiology*, 134, 1471-1480

Hunter, C. N., Holmes, N. G., Jones, O. T. G. and Niederman, R. A. (1979a). Membranes of *Rhodopseudomonas sphaeroides*. 7. Photo-chemical properties of a fraction enriched in newly synthesized bacteriochlorophyll alpha-protein complexes. *Biochimica Et Biophysica Acta*, 548, 253-266.

Hunter, C. N., Hundle, B. S., Hearst, J. E., Lang, H. P., Gardiner, A. T., Takaichi, S. and Cogdell, R. J. (1994). Introduction of new carotenoids into the bacterial photosynthetic apparatus by combining the carotenoid biosynthetic pathways of *Erwinia herbicola* and *Rhodobacter sphaeroides. Journal of Bacteriology*, 176, 3692-3697.

Hunter, C. N., mcglynn, P., Ashby, M. K., Burgess, J. G. and Olsen, J. D. (1991). DNA sequencing and complementation deletion analysis of the bcha-puf operon region of *Rhodobacter sphaeroides* – in vivo mapping of the oxygen-regulated *puf* promoter. *Molecular Microbiology*, 5, 2649-2661.

Hunter, C. N., Pennoyer, J. D., Sturgis, J. N., Farrelly, D. and Niederman, R. A. (1988). Oligomerization states and associations of light-harvesting pigment protein complexes of *Rhodobacter sphaeroides* as analyzed by lithium dodecyl-sulfate polyacrylamide-gel electrophoresis. *Biochemistry*, 27, 3459-3467.

Hunter, C. N., Tucker, J. D. and Niederman, R. A. (2005). The assembly and organisation of photosynthetic membranes in *Rhodobacter sphaeroides*. *Photochemical and Photobiological Sciences*, 4, 1023-1027.

Hunter, C. N., van Grondelle, R., Holmes, N. G., Jones, O. T. G. and Niederman, R. A. (1979b). Fluorescence yield properties of a fraction enriched in newly synthesized bacteriochlorophyll a-protein complexes from *Rhodopseudomonas sphaeroides*. *Photochemistry and Photobiology*, 30, 313-316.

Ichikawa, M., Takekawa, D., Jeon, H.-G. and Banoukepa, G. D. R. (2013). Cascade-type excitation energy relay in organic thin-film solar cells. *Organic Electronics*, 14, 814-820

Imasheva, E. S., Balashov, S. P., Choi, A. R., Jung, K.-H. and Lanyi, J. K. (2009). Reconstitution of *Gloeobacter violaceus* Rhodopsin with a Light-Harvesting Carotenoid Antenna. *Biochemistry*, 48, 10948-10955.

Imhoff, J. F. (1995). Taxonomy and physiology of phototrophic purple bacteria and green sulfur bacteria. *Anoxygenic Photosynthetic Bacteria*, 1-15.

Inamine, G. S., Vanhouten, J. and Niederman, R. A. (1984). Intracellular-localization of photosynthetic membrane growth initiation sites in *Rhodopseudomonas sphaeroides*. *Journal of Bacteriology*, 158, 425-429.

Ind, A. C., Porter, S. L., Brown, M. T., Byles, E. D., de Beyer, J. A., Godfrey, S. A. and Armitage, J. P. (2009). Inducible-Expression Plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans. Applied and Environmental Microbiology*, 75, 6613-6615

Isaacs, F. J., Dwyer, D. J., Ding, C. M., Pervouchine, D. D., Cantor, C. R. and Collins, J. J. (2004). Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotechnology*, 22, 841-847.

Ishida, M., Dohmae, N., Shiro, Y., Oku, T., Iizuka, T. and Isogai, Y. (2004). Design and synthesis of de novo cytochromes c. *Biochemistry*, 43, 9823-9833.

Jack, R. L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T. and Sargent, F. (2004). Coordinating assembly and export of complex bacterial proteins. *Embo Journal*, 23, 3962-3972.

Jamieson, S. J., Wang, P. Y., Qian, P., Kirkland, J. Y., Conroy, M. J., Hunter, C. N. and Bullough, P. A. (2002). Projection structure of the photosynthetic reaction centre-antenna complex of *Rhodospirillum rubrum* at 8.5 angstrom resolution. *Embo Journal*, 21, 3927-3935.

Jang, H.-J., Yoon, S.-H., Ryu, H.-K., Kim, J.-H., Wang, C.-L., Kim, J.-Y., Oh, D.-K. and Kim, S.-W. (2011). Retinoid production using metabolically engineered *Escherichia coli* with a two-phase culture system. *Microbial Cell Factories*, 10.

Janke, C., Scholz, F., Becker-Baldus, J., Glaubitz, C., Wood, P. G., Bamberg, E., Wachtveitl, J. and Bamann, C. (2013). Photocycle and Vectorial Proton Transfer in a Rhodopsin from the Eukaryote *Oxyrrhis marina. Biochemistry*, 52, 2750-2763.

Jensen, R. B., Kelemen, B., Ward, D. E., Asato, A. E. and Kelemen, B. R. (2013). Photochromic material useful for optical applications and as optical data storage carrier, security ink, comprises proteorhodopsin apoprotein and azulenic retinoid compound, or proteorhodopsin apoprotein and retinal analog. Patent.

Jiang, J., Reddy, K. R., Pavan, M. P., Lubian, E., Harris, M. A., Jiao, J., Niedzwiedzki, D. M., Kirmaier, C., Parkes-Loach, P. S., Loach, P. A., Bocian, D. F., Holten, D. and Lindsey, J. S. (2014). Amphiphilic, hydrophilic, or hydrophobic synthetic bacteriochlorins in biohybrid light-harvesting architectures: consideration of molecular designs. *Photosynthesis research*, 122, 187-202.

Jimenez, R., Dikshit, S. N., Bradforth, S. E. and Fleming, G. R. (1996). Electronic excitation transfer in the LH2 complex of *Rhodobacter sphaeroides. Journal of Physical Chemistry*, 100, 6825-6834.

Joo, T. H., Jia, Y. W., Yu, J. Y., Jonas, D. M. and Fleming, G. R. (1996). Dynamics in isolated bacterial light harvesting antenna (LH2) of *Rhodobacter sphaeroides* at room temperature. *Journal of Physical Chemistry*, 100, 2399-2409.

Kamran, M., Friebe, V. M., Delgado, J. D., Aartsma, T. J., Frese, R. N. and Jones, M. R. (2015). Demonstration of asymmetric electron conduction in pseudosymmetrical photosynthetic reaction centre proteins in an electrical circuit. *Nature Communications*, 6.

Karrasch, S., Bullough, P. A. and Ghosh, R. (1995). The 8.5-angstrom projection map of the light-harvesting complex-i from *Rhodospirillum rubrum* reveals a ring composed of 16 subunits. *Embo Journal*, 14, 631-638.

Katona, G., Andreasson, U., Landau, E. M., Andreasson, L. E. and Neutze, R. (2003). Lipidic cubic phase crystal structure of the photosynthetic reaction centre from *Rhodobacter sphaeroides* at 2.35 angstrom resolution. *Journal of Molecular Biology*, 331, 681-692.

Kawaguti, S. (1944). On the Physiology of Reefs Corals VI. Study on the Pigments. *Palao Tropical Biological Station Studies Tokyo*, 2, 617-673.

Kawaguti, S. (1969). Effect of the green fluorescent pigment on the productivity of the reef corals. *Micronesica*, 5, 313-313.

Kehoe, J. W., Meadows, K. A., Parkes-Loach, P. S. and Loach, P. A. (19980. Reconstitution of core light-harvesting complexes of photosynthetic bacteria using chemically synthesized polypeptides. 2. Determination of structural features that stabilize complex formation and their implications for the structure of the subunit complex. *Biochemistry*, 37, 3418-3428.

Kelmanson, I. V. and Matz, M. V. (2003). Molecular basis and evolutionary origins of color diversity in great star coral *Montastraea cavernosa* (Scleractinia : Faviida). *Molecular Biology and Evolution*, 20, 1125-1133.

Kiley, P. J. and Kaplan, S. (1987). Cloning, dna-sequence, and expression of the *Rhodobacter sphaeroides* light-harvesting B800-850-alpha and B800-850-beta genes. *Journal of Bacteriology*, 169, 3268-3275.

Kim, D.-S., Song, H.-N., Nam, H. J., Kim, S.-G., Park, Y.-S., Park, J.-C., Woo, E.-J. and Lim, H.-K. (2014). Directed Evolution of Human Heavy Chain Variable Domain (V-H) Using In Vivo Protein Fitness Filter. *Plos One*, 9.

Kim, S. Y., Waschuk, S. A., Brown, L. S. and Jung, K.-H. (2008). Screening and characterization of proteorhodopsin color-tuning mutations in *Escherichia coli* with endogenous retinal synthesis. *Biochimica Et Biophysica Acta-Bioenergetics*, 1777, 504-513.

Kim, Y.-S., Kim, N.-H., Yeom, S.-J., Kim, S.-W. and Oh, D.-K. (2009). In Vitro Characterization of a Recombinant Blh Protein from an Uncultured Marine Bacterium as a beta-Carotene 15,15 '-Dioxygenase. *Journal of Biological Chemistry*, 284, 15781-15793.

Klyszejko, A. L., Shastri, S., Mari, S. A., Grubmueller, H., Muller, D. J. and Glaubitz, C. (2008). Folding and assembly of proteorhodopsin. *Journal of Molecular Biology*, 376, 35-41.

Koblizek, M., Shih, J. D., Breitbart, S. I., Ratcliffe, E. C., Kolber, Z. S., Hunter, C. N. and Niederman, R. A. (2005). Sequential assembly of photosynthetic units in *Rhodobacter sphaeroides* as revealed by fast repetition rate analysis of variable bacteriochlorophyll a fluorescence. *Biochimica Et Biophysica Acta-Bioenergetics*, 1706, 220-231.

Koder, R. L., Anderson, J. L. R., Solomon, L. A., Reddy, K. S., Moser, C. C. and Dutton, P. L. (2009). Design and engineering of an O-2 transport protein. *Nature*, 458.

Koder, R. L., Valentine, K. G., Cerda, J., Noy, D., Smith, K. M., Wand, A. J. and Dutton, P. L. (2006). Native like structure in designed four alpha-helix bundles driven by buried polar interactions. *Journal of the American Chemical Society*, 128, 14450-14451.

Koepke, J., Hu, X. C., Muenke, C., Schulten, K. and Michel, H. (1996). The crystal structure of the light-harvesting complex II (B800-850) from *Rhodospirillum molischianum*. *Structure*, 4, 581-597.

Koh, E. Y., Atamna-Ismaeel, N., Martin, A., Cowie, R. O. M., Beja, O., Davy, S. K., Maas, E. W. and Ryan, K. G. (2010). Proteorhodopsin-Bearing Bacteria in Antarctic Sea Ice. *Applied and Environmental Microbiology*, 76, 5918-5925.

Kremers, G. J., Goedhart, J., van Munster, E. B. and Gadella, T. W. J. (2006). Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius. *Biochemistry*, 45, 6570-6580.

Kudva, R., Denks, K., Kuhn, P., Vogt, A., Mueller, M. and Koch, H.-G. (2013). Protein translocation across the inner membrane of Gram-negative bacteria: the Sec and Tat dependent protein transport pathways. *Research in Microbiology*, 164, 505-534.

Labas, Y. A., Gurskaya, N. G., Yanushevich, Y. G., Fradkov, A. F., Lukyanov, K. A., Lukyanov, S. A. and Matz, M. V. (2002). Diversity and evolution of the green fluorescent protein family. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 4256-4261.

Laible, P. D., Mielke, D. L. and Hanson, D. K. (2009) Foreign gene expression in photosynthetic bacteria. *In*: The Purple Phototrophic Bacteria. *Springer.*

Lang, H. P. and Hunter, C. N. (1994). The relationship between carotenoid biosynthesis and the assembly of the light-harvesting LH2 complex in *Rhodobacter sphaeroides*. *Biochemical Journal*, 298, 197-205.

Lang, H. P., Cogdell, R. J., Takaichi, S. and Hunter, C. N. (1995). Complete DNA-sequence, specific tn5 insertion map, and gene assignment of the carotenoid biosynthesis pathway of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, 177, 2064-2073.

Lange, C., Mueller, S. D., Walther, T. H., Buerck, J. and Ulrich, A. S. (2007). Structure analysis of the protein translocating channel tata in membranes using a multi-construct approach. *Biochimica Et Biophysica Acta-Biomembranes*, 1768, 2627-2634.

Lanyi, J. K. (1998). Understanding structure and function in the light-driven proton pump bacteriorhodopsin. *Journal of Structural Biology*, 124, 164-178.

Lanyi, J. K. (2004). Bacteriorhodopsin. *Annual Review of Physiology*, 66, 665-688.

Lear, J. D., Gratkowski, H., Adamian, L., Liang, J. and Degrado, W. F. (2003). Position-dependence of stabilizing polar interactions of asparagine in transmembrane helical bundles. *Biochemistry*, 42, 6400-6407.

Lee, J. K., Dehoff, B. S., Donohue, T. J., Gumport, R. I. and Kaplan, S. (1989). Transcriptional analysis of *puf* operon expression in *Rhodobacter sphaeroides* 2.4.1 and an intercistronic transcription terminator mutant. *Journal of Biological Chemistry*, 264, 19354-19365.

Leibly, D. J., Arbing, M. A., Pashkov, I., DeVore, N., Waldo, G. S. And Terwilliger, T. C. (2015). A suite of engineered GFP molecules for oligomeric scaffolding. *Structure*, 23, 1754-1768.

Lichtenstein, B. R. (2010). A new approach to understanding biological control of quinone electrochemistry. University of Pennsylvania. *Publicly accessible Penn dissertations.*

Lichtenstein, B. R., Farid, T. A., Kodali, G., Solomon, L. A., Anderson, J. L. R., Sheehan, M. M., Ennist, N. M., Fry, B. A., Chobot, S. E., Bialas, C., Mancini, J. A., Armstrong, C. T., Zhao, Z., Esipova, T. V., Snell, D., Vinogradov, S. A., Discher, B. M., Moser, C. C. and Dutton, P. L. (2012). Engineering oxidoreductases: maquette proteins designed from scratch. *Biochemical Society Transactions*, 40, 561-566

Lindsey, J. S., Mass, O. and Chen, C.-Y. (2011). Tapping the near-infrared spectral region with bacteriochlorin arrays. *New Journal of Chemistry*, 35, 511-516.

Lo, S. M. and Theg, S. M. (2012). Role of Vesicle-Inducing Protein in Plastids 1 in cptat transport at the thylakoid. *Plant Journal*, 71, 656-668.

Lundberg, K. S., Shoemaker, D. D., Adams, M. W. W., Short, J. M., Sorge, J. A. and Mathur, E. J. (1991). High-fidelity amplification using a thermostable DNA-polymerase isolated from *Pyrococcus furiosus*. *Gene*, 108, 1-6.

Ma, Y. Z., Cogdell, R. J. and Gillbro, T. (1997). Energy transfer and exciton annihilation in the B800-850 antenna complex of the photosynthetic purple bacterium *Rhodopseudomonas acidophila* (Strain 10050). A femtosecond transient absorption study. *Journal of Physical Chemistry B*, 101, 1087-1095.

Malo, G. D., Pouwels, L. J., Wang, M., Weichsel, A., Montfort, W. R., Rizzo, M. A., Piston, D. W. and Wachter, R. M. (2007). X-ray structure of cerulean GFP: A tryptophan-based chromophore useful for fluorescence lifetime imaging. *Biochemistry*, 46, 9865-9873.

Man, D. L., Wang, W. W., Sabehi, G., Aravind, L., Post, A. F., Massana, R., Spudich, E. N., Spudich, J. L. and Beja, O. (2003). Diversification and spectral tuning in marine proteorhodopsins. *Embo Journal*, 22, 1725-1731.

Mao, J., Nhu-Nguyen, D., Scholz, F., Reggie, L., Mehler, M., Lakatos, A., Ong, Y.-S., Ullrich, S. J., Brown, L. J., Brown, R. C. D., Becker-Baldus, J., Wachtveitl, J. and Glaubitz, C. (2014). Structural Basis of the Green-Blue Color Switching in Proteorhodopsin as Determined by NMR Spectroscopy. *Journal of the American Chemical Society*, 136, 17578-17590.

Martinez, A., Bradley, A. S., Waldbauer, J. R., Summons, R. E. and DeLong, E. F. (2007). Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 5590-5595.

Matos, C. F. R. O., Branston, S. D., Albiniak, A., Dhanoya, A., Freedman, R. B., Keshavarz-Moore, E. and Robinson, C. (2012). High-yield export of a native heterologous protein to the periplasm by the tat translocation pathway in *Escherichia coli*. *Biotechnology and Bioengineering*, 109, 2533-2542.

Matos, C. F. R. O., Robinson, C., Alanen, H. I., Prus, P., Uchida, Y., Ruddock, L. W., Freedman, R. B. and Keshavarz-Moore, E. (2014). Efficient export of prefolded, disulfide-bonded recombinant proteins to the periplasm by the Tat pathway in *Escherichia coli* cydisco strains. *Biotechnology Progress*, 30, 281-290.

Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L. and Lukyanov, S. A. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology*, 17, 969-973.

Maurer, C., Panahandeh, S., Jungkamp, A.-C., Moser, M. and Mueller, M. (2010). Tatb functions as an oligomeric binding site for folded tat precursor proteins. *Molecular Biology of the Cell*, 21, 4151-4161.

Mazel, C. H., Lesser, M. P., Gorbunov, M. Y., Barry, T. M., Farrell, J. H., Wyman, K. D. and Falkowski, P. G. (2003). Green-fluorescent proteins in Caribbean corals. *Limnology and Oceanography*, 48, 402-411.

McAuley, K. E., Fyfe, P. K., Cogdell, R. J., Isaacs, N. W. and Jones, M. R. (2000). X-ray crystal structure of the YM210W mutant reaction centre from *Rhodobacter sphaeroides. Febs Letters*, 467, 285-290.

McCarren, J. and DeLong, E. F. (2007). Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environmental Microbiology*, 9, 846-858.

McDermott, G., Prince, S. M., Freer, A. A., Hawthornthwaite-Lawless, A. M., Papiz, M. Z., Cogdell, R. J. and Isaacs, N. W. (1995). Crystal-structure of an integral membrane light-harvesting complex from photosynthetic bacteria. *Nature*, 374, 517-521.

McIsaac, R. S., Engqvist, M. K. M., Wannier, T., Rosenthal, A. Z., Herwig, L., Flytzanis, N. C., Imasheva, E. S., Lanyi, J. K., Balashov, S. P., Gradinaru, V. and Arnold, F. H. (2014). Directed evolution of a far-red fluorescent rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 13034-13039.

Meadows, K. A., Iida, K., Tsuda, K., Recchia, P. A., Heller, B. A., Antonio, B., Nango, M. and Loach, P. A. (1995). Enzymatic and chemical cleavage of the core light-harvesting polypeptides of photosynthetic bacteria - determination of the minimal polypeptide size and structure required for subunit and light-harvesting complex-formation. *Biochemistry*, 34, 1559-1574.

Meadows, K. A., Parkes-Loach, P. S., Kehoe, J. W. and Loach, P. A. (1998). Reconstitution of core light-harvesting complexes of photosynthetic bacteria using chemically synthesized polypeptides. 1. Minimal requirements for subunit formation. *Biochemistry*, 37, 3411-3417.

Mehler, M., Scholz, F., Ullrich, S. J., Mao, J., Braun, M., Brown, L. J., Brown, R. C. D., Fiedler, S. A., Becker-Baldus, J., Wachtveitl, J. and Glaubitz, C. (2013). The EF Loop in Green Proteorhodopsin Affects Conformation and Photocycle dynamics. *Biophysical Journal*, 105, 385-397.

Miller, K. R. (1979). Structure of a bacterial photosynthetic membrane. *Proceedings of the National Academy of Sciences of the United States of America*, 76, 6415-6419.

Molik, S., Karnauchov, I., Weidlich, C. E., Herrmann, R. G. and Klosgen, R. B. (2001). The Rieske Fe/S protein of the cytochrome b(6)/f complex in chloroplasts - Missing link in the evolution of protein transport pathways in chloroplasts? *Journal of Biological Chemistry*, 276, 42761-42766.

Moltke, S. and Heyn, M. P. (1995). Photovoltage kinetics of the acid-blue and acid-purple forms of bacteriorhodopsin: Evidence for no net charge transfer. *Biophysical Journal*, 69, 2066-2073.

Monshouwer, R., Dezarate, I. O., Vanmourik, F. and van Grondelle, R. (1995). Low-intensity pump-probe spectroscopy on the B800 to B850 transfer in the light-harvesting 2 complex of *Rhodobacter sphaeroides*. *Chemical Physics Letters*, 246, 341-346.

Mori, H. and Cline, K. (2002). A twin arginine signal peptide and the ph gradient trigger reversible assembly of the thylakoid Delta ph/Tat translocase. *Journal of Cell Biology*, 157, 205-210.

Mothersole, D. (2013). Assembly, structure and organisation of photosynthetic membranes. University of Sheffield.

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research*, 1, 2-9.

Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K. and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology*, 20, 87-90.

Nagarajan, V. and Parson, W. W. (1997). Excitation energy transfer between the B850 and B875 antenna complexes of *Rhodobacter sphaeroides. Biochemistry*, 36, 2300-2306.

Naylor, G. W., Addlesee, H. A., Gibson, L. C. D. and Hunter, C. N. (1999). The photosynthesis gene cluster of *Rhodobacter sphaeroides. Photosynthesis Research*, 62, 121-139.

Ng, I. W., Adams, P. G., Mothersole, D. J., Vasilev, C., Martin, E. C., Lang, H. P., Tucker, J. D. and Hunter, C. N. (2011). Carotenoids are essential for normal levels of dimerisation of the RC-LH1-pufx core complex of *Rhodobacter sphaeroides*: Characterisation of R-26 as a crtb (phytoene synthase) mutant. *Biochimica Et Biophysica Acta-Bioenergetics*, 1807, 1056-1063.

Nguyen, D., Maranger, R., Balagué, V., Coll-Lladó, M., Lovejoy, C., and Pedrós-Alió, C., (2015). Winter diversity and expression of proteorhodopsin genes in a polar ocean. The ISME Journal, 9, 1835-1845

Niederman, R. A., Mallon, D. E. and Langan, J. J. (1976). Membranes of R*hodopseudomonas* sphaeroides .4. Assembly of chromatophores in low-aeration cell-suspensions. *Biochimica Et Biophysica Acta*, 440, 429-447.

Niederman, R. A., Mallon, D. E. and Parks, L. C. (1979). Membranes of *Rhodopseudomonas sphaeroides* .6. Isolation of a fraction enriched in newly synthesized bacteriochlorophyll alpha-protein complexes. *Biochimica Et Biophysica Acta*, 555, 210-220.

Nori, A., Jensen, K. D., Tijerina, M., Kopeckova, P. and Kopecek, J. (2003). Tat-conjugated synthetic macromolecules facilitate cytoplasmic drug delivery to human ovarian carcinoma cells. *Bioconjugate Chemistry*, 14, 44-50.

Noy, D., Discher, B. M., Rubtsov, I. V., Hochstrasser, R. A. and Dutton, P. L. (2005). Design of amphiphilic protein maquettes: Enhancing maquette functionality through binding of extremely hydrophobic cofactors to lipophilic domains. *Biochemistry*, 44, 12344-12354.

Oates, J., Barrett, C. M., Barnett, J. P., Byrne, K. G., Bolhuis, A. and Robinson, C. (2005). The *Escherichia coli* twin-arginine translocation apparatus incorporates a distinct form of tatabc complex, spectrum of modular tata complexes and minor tatab complex. *Journal of Molecular Biology,* 346, 295-305.

Oesterhelt D. and Stoeckenius W. (1971). Rhodopsin-like protein from purple membrane of *Halobacterium halobium*. *Nature-New Biology*, 233, 149-.

Ogara, J. P. and Kaplan, S. (1997). Evidence for the role of redox carriers in photosynthesis gene expression and carotenoid biosynthesis in *Rhodobacter sphaeroides* 2.4.1. *Journal of Bacteriology*, 179, 1951-1961.

Ohhama, T., Seto, H. and Miyachi, S. (1985). C-13 nuclear magnetic-resonance studies on bacteriochlorophyll a biosynthesis in *Rhodopseudomonas spheroides* S. *Archives of Biochemistry and Biophysics*, 237, 72-79.

Olsen, J. D., Adams, P. G., Jackson, P. J., Dickman, M. J., Qian, P. and Hunter, C. N. (2014). Aberrant Assembly Complexes of the Reaction Center Light-harvesting 1 pufx (RC-LH1-pufx)

Core Complex of *Rhodobacter sphaeroides* Imaged by Atomic Force Microscopy. *Journal of Biological Chemistry*, 289, 29927-29936.

Oresnik, I. J., Ladner, C. L. and Turner, R. J. (2001). Identification of a twin-arginine leader-binding protein. *Molecular Microbiology*, 40, 323-331.

Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. and Remington, S. J. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science*, 273, 1392-1395.

Padan, E. (1979). Facultative anoxygenic photosynthesis in cyanobacteria. *Annual Review of Plant Physiology and Plant Molecular Biology*, 30, 27-40.

Palmer, T. and Berks, B. C. (2012). The twin-arginine translocation (Tat) protein export pathway. *Nature Reviews Microbiology*, 10, 483-496.

Palmer, T., Sargent, F. and Berks, B. C. (2005). Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends in Microbiology*, 13, 175-180.

Papiz, M. Z., Prince, S. M., Howard, T., Cogdell, R. J. and Isaacs, N. W. (2003). The structure and thermal motion of the B800-850 LH2 complex from *Rps. acidophila* at 2.0 Å over-circle resolution and 100 K: New structural features and functionally relevant motions. *Journal of Molecular Biology*, 326, 1523-1538.

Partridge, J. C. and Cummings, M. E. (1999). Adaptation of visual pigments to the aquatic environment. *Adaptive mechanisms in the ecology of vision*., 251-283.

Patel, R., Smith, S. M. and Robinson, C. (2014). Protein transport by the bacterial Tat pathway. *Biochimica Et Biophysica Acta-Molecular Cell Research*, 1843, 1620-1628.

Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R. and Piston, D. W. (1997). Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical Journal*, 73, 2782-2790.

Peters, G. A. and Cellarius R. A. (1972). Photosynthetic membrane development in *Rhodopseudomonas spheroides* .2. Correlation of pigment incorporation with morphological aspects of thylakoid formation. *Journal of Bioenergetics*, 3, 345-and.

Pfleger, B. F., Pitera, D. J., D Smolke, C. and Keasling, J. D. (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nature Biotechnology*, 24, 1027-1032.

Pfleger, N., Lorch, M., Woerner, A. C., Shastri, S. and Glaubitz, C. (2008). Characterisation of Schiff base and chromophore in green proteorhodopsin by solid-state NMR. *Journal of Biomolecular NMR*, 40, 15-21.

Piston, D. W. and Kremers, G.-J. (2007). Fluorescent protein FRET: the good, the bad and the ugly. *Trends in Biochemical Sciences*, 32, 407-414.

Pitcher, R. S. and Watmough, N. J. (2004). The bacterial cytochrome cbb(3) oxidases. *Biochimica Et Biophysica Acta-Bioenergetics*, 1655, 388-399.

Punginelli, C., Maldonado, B., Grahl, S., Jack, R., Alami, M., Schroeder, J., Berks, B. C. and Palmer, T. (2007). Cysteine scanning mutagenesis and topological mapping of the *Escherichia coli* twin-arginine translocase tatc component. *Journal of Bacteriology*, 189, 5482-5494.

Qian, P., Bullough, P. A. and Hunter, C. N. (2008). Three-dimensional reconstruction of a membrane-bending complex - The RC-LH1-pufx core dimer of *Rhodobacter sphaeroides. Journal of Biological Chemistry*, 283, 14002-14011.

Qian, P., Hunter, C. N. and Bullough, P. A. (2005). The 8.5 angstrom projection structure of the core RC-LH1-pufx dimer of *Rhodobacter sphaeroides. Journal of Molecular Biology*, 349, 948-960.

Qian, P., Papiz, M. Z., Jackson, P. J., Brindley, A. A., Ng, I. W., Olsen, J. D., Dickman, M. J., Bullough, P. A. and Hunter, C. N. (2013). Three-Dimensional Structure of the *Rhodobacter sphaeroides* RC-LH1-pufx Complex: Dimerization and Quinone Channels Promoted by pufx. *Biochemistry*, 52, 7575-7585.

Quan, S., Hiniker, A., Collet, J.-F. and Bardwell, J. C. A. (2013). Isolation of bacteria envelope proteins. *Methods in molecular biology* (Clifton, N.J.), 966, 359-66.

Racker, E. and Stoeckenius W. (1974). Reconstitution of purple membrane-vesicles catalyzing light-driven proton uptake and adenosine-triphosphate formation. *Journal of Biological Chemistry*, 249, 662-663.

Ran, T., Ozorowski, G., Gao, Y., Sineshchekov, O. A., Wang, W., Spudich, J. L. and Luecke, H. (2013). Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes*. Acta Crystallographica Section D-Biological Crystallography*, 69, 1965-1980.

Randall, L. L. and Hardy, S. J. S. (1986). Correlation of competence for export with lack of tertiary structure of the mature species - a study invivo of maltose-binding protein in *Escherichia coli*. *Cell*, 46, 921-928.

Ratcliffe, E. C., Tunnicliffe, R. B., Ng, I. W., Adams, P. G., Qian, P., Holden-Dye, K., Jones, M. R., Williamson, M. P. and Hunter, C. N. (2011). Experimental evidence that the membrane-spanning helix of pufx adopts a bent conformation that facilitates dimerisation of the *Rhodobacter sphaeroides* RC-LH1 complex through N-terminal interactions. *Biochimica Et Biophysica Acta-Bioenergetics*, 1807, 95-107.

Ray, N., Oates, J., Turner, R. J. and Robinson, C. (2003). Dmsd is required for the biogenesis of DMSO reductase in *Escherichia coli* but not for the interaction of the dmsa signal peptide with the Tat apparatus. *Febs Letters*, 534, 156-160.

Reckel, S., Gottstein, D., Stehle, J., Loehr, F., Verhoefen, M.-K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., Bernhard, F., Schwalbe, H., Guentert, P. and Doetsch, V. (2011). Solution NMR Structure of Proteorhodopsin. *Angewandte Chemie-International Edition*, 50, 11942-11946.

Reddy, K. R., Jiang, J., Krayer, M., Harris, M. A., Springer, J. W., Yang, E., Jiao, J., Niedzwiedzki, D. M., Pandithavidana, D., Parkes-Loach, P. S., Kirmaier, C., Loach, P. A., Bocian, D. F., Holten, D. and Lindsey, J. S. (2013). Palette of lipophilic bioconjugatable bacteriochlorins for construction of biohybrid light-harvesting architectures. *Chemical Science*, 4, 2036-2053.

Regan, L. and DeGrado, W. F. (1988). Characterization of a helical protein designed from 1st principles. *Science*, 241, 976-978.

Reilly, P. A. and Niederman, R. A. (1986). Role of apparent membrane growth initiation sites during photosynthetic membrane-development in synchronously dividing *Rhodopseudomonas sphaeroides. Journal of Bacteriology*, 167, 153-159.

Remington, S. J. (2006). Fluorescent proteins: maturation, photochemistry and photophysics. *Current Opinion in Structural Biology*, 16, 714-721.

Ren, C., Patel, R. and Robinson, C. (2013). Exclusively membrane-inserted state of an uncleavable Tat precursor protein suggests lateral transfer into the bilayer from the translocon. *The FEBS journal*, 280, 3354-64.

Ricci, M., Bradforth, S. E., Jimenez, R. and Fleming, G. R. (1996). Internal conversion and energy transfer dynamics of spheroidene in solution and in the LH-1 and LH-2 light-harvesting complexes. *Chemical Physics Letters*, 259, 381-390.

Richter, S., Lindenstrauss, U., Luecke, C., Bayliss, R. and Brueser, T. (2007). Functional tat transport of unstructured, small, hydrophilic proteins. *Journal of Biological Chemistry*, 282, 33257-33264.

Riedel, T., Gomez-Consarnau, L., Tomasch, J., Martin, M., Jarek, M., Gonzalez, J. M., Spring, S., Rohlfs, M., Brinkhoff, T., Cypionka, H., Goeker, M., Fiebig, A., Klein, J., Goesmann, A., Fuhrman, J. A. and Wagner-Doebler, I. (2013). Genomics and Physiology of a Marine Flavobacterium Encoding a Proteorhodopsin and a Xanthorhodopsin-Like Protein. *Plos One*, 8.

Robertson, D. E., Farid, R. S., Moser, C. C., Urbauer, J. L., Mulholland, S. E., Pidikiti, R., Lear, J. D., Wand, A. J., DeGrado, W. F. and Dutton, P. L. (1994). Design and synthesis of multi-heme proteins. *Nature*, 368, 425-431.

Robinson, C. and Bolhuis, A. (2001). Protein targeting by the twin-arginine translocation pathway. *Nature Reviews Molecular Cell Biology*, 2, 350-356.

Rodrigue, A., Chanal, A., Beck, K., Muller, M. and Wu, L. F. (1999). Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial Tat pathway. *Journal of Biological Chemistry*, 274, 13223-13228.

Roszak, A. W., Howard, T. D., Southall, J., Gardiner, A. T., Law, C. J., Isaacs, N. W. and Cogdell, R. J. (2003). Crystal structure of the RC-LH1 core complex from *Rhodopseudomonas palustris*. *Science*, 302, 1969-1972.

Rothlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., dechancie, J., Betker, J., Gallaher, J. L., Althoff, E. A., Zanghellini, A., Dym, O., Albeck, S., Houk, K. N., Tawfik, D. S. and Baker, D. (2008). Kemp elimination catalysts by computational enzyme design. *Nature*, 453, 190-U4.

Ruch, S., Beyer, P., Ernst, H. and Al-Babili, S. (2005). Retinal biosynthesis in Eubacteria: in vitro characterization of a novel carotenoid oxygenase from *Synechocystis sp* PCC 6803. *Molecular Microbiology*, 55, 1015-1024.

Ruiz-Gonzalez, M. X. and Marin, I. (2004). New insights into the evolutionary history of type 1 rhodopsins. *Journal of Molecular Evolution*, 58, 348-358.

Sabehi, G., Loy, A., Jung, K. H., Partha, R., Spudich, J. L., Isaacson, T., Hirschberg, J., Wagner, M. and Beja, O. (2005). New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *Plos Biology*, 3, 1409-1417.

Salih, A., Larkum, A., Cox, G., Kuhl, M. and Hoegh-Guldberg, O. (2000). Fluorescent pigments in corals are photoprotective. *Nature*, 408, 850-853.

Salis, H. M., Mirsky, E. A. and Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology*, 27, 946-U112.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning A Laboratory Manual Second Edition Vols. 1 2 and 3. *Cold Spring Harbor*, New York, USA. Illus. Paper.

Sanders, C., Wethkamp, N. and Lill, H. (2001). Transport of cytochrome c derivatives by the bacterial Tat protein translocation system. *Molecular Microbiology*, 41, 241-246.

Sargent, F., Stanley, N. R., Berks, B. C. and Palmer, T. (1999). Sec-independent protein translocation in *Escherichia coli* - A distinct and pivotal role for the tatb protein. *Journal of Biological Chemistry*, 274, 36073-36082.

Schachman, H. K., Pardee, A. B. and Staner, R. Y. (1952). Studies on the macromolecular organization of microbial cells. *Arch. Biochem. Biophys*. 38, 245.

Scheuring, S., Rigaud, J. L. and Sturgis, J. N. (2004). Variable LH2 stoichiometry and core clustering in native membranes of *Rhodospirillum photometricum. Embo Journal*, 23, 4127-4133.

Scheuring, S., Seguin, J., Marco, S., Levy, D., Breyton, C., Robert, B. and Rigaud, J. L. (2003). AFM characterization of tilt and intrinsic flexibility of *Rhodobacter sphaeroides* light harvesting complex 2 (LH2). *Journal of Molecular Biology*, 325, 569-580.

Schlichter, D., Fricke, H. W. and Weber, W. (1986). Light harvesting by wavelength transformation in a symbiotic coral of the red-sea twilight zone. *Marine Biology*, 91, 403-407.

Schmidt, K., (1978) Biosynthesis of carotenoids. Clayton, R.K. and Siström, W.R. (eds.), Photosynthetic bacteria. *Plenum Press*, New York, 729-750.

Schreiber, S., Stengel, R., Westermann, M., Volkmer-Engert, R., Pop, O. I. and Mueller, J. P. (2006). Affinity of tatc(d) for tata(d) elucidates its receptor function in the *Bacillus subtilis* twin arginine translocation (Tat) translocase system. *Journal of Biological Chemistry*, 281, 19977-19984.

Sener, M. K. and Schulten, K. (2009). From Atomic-Level Structure to Supramolecular Organization in the Photosynthetic Unit of Purple Bacteria. *Advances in Photosynthesis and Respiration*, 28, 275-294.

Sener, M. K., Olsen, J. D., Hunter, C. N. and Schulten, K. (2007). Atomic-level structural and functional model of a bacterial photosynthetic membrane vesicle. *Proceedings of the National Academy of Sciences of the United States of America,* 104, 15723-15728.

Sener, M., Hsin, J., Trabuco, L. G., Villa, E., Qian, P., Hunter, C. N. and Schulten, K. (2009). Structural model and excitonic properties of the dimeric RC-LH1-Pufx complex from *Rhodobacter sphaeroides*. *Chemical Physics*, 357, 188-197.

Shagin, D. A., Barsova, E. V., Yanushevich, Y. G., Fradkov, A. F., Lukyanov, K. A., Labas, Y. A., Semenova, T. N., Ugalde, J. A., Meyers, A., Nunez, J. M., Widder, E. A., Lukyanov, S. A. and Matz, M. V. (2004). GFP-like proteins as ubiquitous metazoan superfamily: Evolution of functional features and structural complexity. *Molecular Biology and Evolution*, 21, 841-850.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E. and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp red fluorescent protein. *Nature Biotechnology*, 22, 1567-1572.

Shaner, N. C., Lin, M. Z., Mckeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W. and Tsien, R. Y. (2008). Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nature Methods*, 5, 545-551.

Shaner, N. C., Patterson, G. H. and Davidson, M. W. (2007). Advances in fluorescent protein technology. *Journal of Cell Science*, 120, 4247-4260.

Shaner, N. C., Steinbach, P. A. and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nature Methods*, 2, 905-909.

Sharp, R. E., Moser, C. C., Rabanal, F. and Dutton, P. L. (1998). Design, synthesis, and characterization of a photoactivatable flavocytochrome molecular maquette. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 10465-10470.

Shastri, S., Vonck, J., Pfleger, N., Haase, W., Kuehlbrandt, W. and Glaubitz, C. (2007). Proteorhodopsin: Characterisation of 2D crystals by electron microscopy and solid state NMR. *Biochimica Et Biophysica Acta-Biomembranes*, 1768, 3012-3019.

Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M., Lukyanov, S., Zaraisky, A. G., Davidson, M. W. and Chudakov, D. M. (2009). Far-red fluorescent tags for protein imaging in living tissues. *Biochemical Journal*, 418, 567-574.

Shemin, D. (1956). The biosynthesis of porphyrins. Essays In Biochemistry (Graf, S., ed.) *Wiley*, New York, pp. 241-258

Shifman, J. M., Gibney, B. R., Sharp, R. E. and Dutton, P. L. (2000). Heme redox potential control in de novo designed four-alpha-helix bundle proteins. *Biochemistry*, 39, 14813-14821.

Shifman, J. M., Moser, C. C., Kalsbeck, W. A., Bocian, D. F. and Dutton, P. L. (1998). Functionalized de novo designed proteins: Mechanism of proton coupling to oxidation/reduction in heme protein maquettes. *Biochemistry*, 37, 16815-16827.

Shimomura, O., Johnson, F. H. and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from luminous hydromedusan, *Aequorea. Journal of Cellular and Comparative Physiology*, 59, 223-and.

Shneour, E. A. (1962). Source of oxygen in *Rhodopseudomonas spheroides* carotenoid pigment conversion. *Biochimica Et Biophysica Acta*, 65, 510-and.

Shreve, A. P., Trautman, J. K., Frank, H. A., Owens, T. G. and Albrecht, A. C. (1991). Femtosecond energy-transfer processes in the b800-850 light-harvesting complex of *Rhodobacter sphaeroides* 2.4.1. *Biochimica Et Biophysica Acta*, 1058, 280-288.

Siebert, C. A., Qian, P., Fotiadis, D., Engel, A., Hunter, C. N. and Bullough, P. A. (2004). Molecular architecture of photosynthetic membranes in *Rhodobacter sphaeroides*: the role of pufx. *Embo Journal*, 23, 690-700.

Siefert, E., Irgens, R. L. and Pfennig, N. (1978). Phototropic purple and green bacteria in a sewage-treatment plant. *Applied and Environmental Microbiology*, 35, 38-44.

Simon, R., Priefer, U. and Puhler, A. (1983). A broad host range mobilization system for in vivo genetic-engineering - transposon mutagenesis in gram-negative bacteria. *Bio-Technology*, 1, 784-791.

Sirevåg, R. (1995). Carbon metabolism in green bacteria. *Anoxygenic Photosynthetic Bacteria*, 871-883.

Sistrom, W. R., Griffiths, M. and Stanier, R. Y. (1956). The biology of a photosynthetic bacterium which lacks colored carotenoids. *Journal of Cellular and Comparative Physiology*, 48, 473-515

Skalicky, J. J., Gibney, B. R., Rabanal, F., Urbauer, R. J. B., Dutton, P. L. and Wand, A. J. (1999). Solution structure of a designed four-alpha-helix bundle maquette scaffold. *Journal of the American Chemical Society*, 121, 4941-4951.

Slamovits, C. H., Okamoto, N., Burri, L., James, E. R. and Keeling, P. J. (2011). A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nature Communications*, 2.

Slouf, V., Chabera, P., Olsen, J. D., Martin, E. C., Qian, P., Hunter, C. N. and Polivka, T. (2012). Photoprotection in a purple phototrophic bacterium mediated by oxygen-dependent alteration of carotenoid excited-state properties. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 8570-8575.

Spencer, M. E. and Guest, J. R. (1974). Proteins of inner membrane of *Escherichia coli* - identification of succinate-dehydrogenase by polyacrylamide-gel electrophoresis with sdh amber mutants. *Journal of Bacteriology*, 117, 947-953.

Springer, J. W., Parkes-Loach, P. S., Reddy, K. R., Krayer, M., Jiao, J., Lee, G. M., Niedzwiedzki, D. M., Harris, M. A., Kirmaier, C., Bocian, D. F., Lindsey, J. S., Holten, D. and Loach, P. A. (2012). Biohybrid Photosynthetic Antenna Complexes for Enhanced Light-Harvesting. *Journal of the American Chemical Society*, 134, 4589-4599.

Spudich, J. L. and Jung, K.-H. (2005). Microbial Rhodopsins: Phylogenetic and Functional Diversity. *Handbook of Photosensory Receptors*, 1-23.

Spudich, J. L., Yang, C. S., Jung, K. H. and Spudich, E. N. (2000). Retinylidene proteins: Structures and functions from archaea to humans. *Annual Review of Cell and Developmental Biology*, 16, 365-+.

Stanley, N. R., Palmer, T. and Berks, B. C. (2000). The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli. Journal of Biological Chemistry*, 275, 11591-11596.

Strandberg, E. and Killian, J. A. (2003). Snorkeling of lysine side chains in transmembrane helices: how easy can it get? *Febs Letters*, 544, 69-73.

Sturgis, J. N. and Niederman, R. A. (2009). Organization and Assembly of Light-Harvesting Complexes in the Purple Bacterial Membrane. *Advances in Photosynthesis and Respiration*, 253-273.

Swainsbury, D. J. K., Friebe, V. M., Frese, R. N. and Jones, M. R. (2014a). Evaluation of a biohybrid photoelectrochemical cell employing the purple bacterial reaction centre as a biosensor for herbicides. *Biosensors and Bioelectronics*, 58, 172-178.

Swainsbury, D. J. K., Scheidelaar, S., Van Grondelle, R., Killian, J. A. and Jones, M. R. (2014b). Bacterial Reaction Centers Purified with Styrene Maleic Acid Copolymer Retain Native Membrane Functional Properties and Display Enhanced Stability. *Angewandte Chemie-International Edition*, 53, 11803-11807.

Tabita, F. R. (1995). The biochemistry and metabolic regulation of carbon metabolism and CO_2 fixation in purple bacteria. *Anoxygenic Photosynthetic Bacteria*, 885-914.

Tan, S. C., Yan, F., Crouch, L. I., Robertson, J., Jones, M. R. and Welland, M. E. (2013). Superhydrophobic carbon nanotube electrode produces a near-symmetrical alternating current from photosynthetic protein-based photoelectrochemical cells. *Advanced Functional Materials*, 23, 5556-5563.

Teter, S. A. and Theg, S. M. (1998). Energy-transducing thylakoid membranes remain highly impermeable to ions during protein translocation. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 1590-1594.

Thomas, J. D., Daniel, R. A., Errington, J. and Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli. Molecular Microbiology*, 39, 47-53.

Tichy, H. V., Oberle, B., Stiehle, H., Schiltz, E. and Drews, G. (1989). Genes downstream from *pucB* and *pucA* are essential for formation of the B800-850 complex of *Rhodobacter capsulatus. Journal of Bacteriology*, 171, 4914-4922.

Tikh, I. B., Held, M. and Schmidt-Dannert, C. (2014). Biobrick(TM) compatible vector system for protein expression in *Rhodobacter sphaeroides. Applied Microbiology and Biotechnology*, 98, 3111-3119.

Tsien, R. Y. (1998). The green fluorescent protein. *Annual Review of Biochemistry*, 67, 509-544.

Tsien, R. Y., Bacskai, B. J. and Adams, S. R. (1993). FRET for studying intracellular signalling. *Trends in Cell Biology*, 3, 242-245.

Tunggal, J., Wartenberg, M., Paulsson, M. and Smyth, N. (2003). Expression of the nidogen-binding site of the laminin gamma 1 chain disturbs basement membrane formation and maintenance in F9 embryoid bodies. *Journal of Cell Science*, 116, 803-812.

Van Grondelle, R., Dekker, J. P., Gillbro, T. and Sundstrom, V. (1994). Energy-transfer and trapping in photosynthesis. *Biochimica Et Biophysica Acta-Bioenergetics*, 1187, 1-65.

Van Niel, C. B. (1941). The bacterial photosyntheses and their importance for the general problem of photosynthesis. *Advances in Enzymology and Related Subjects of Biochemistry*, 1, 263-328.

Varo, G., Brown, L. S., Lakatos, M. And Lanyi, J. K. (20030. Characterization of the photochemical reaction cycle of proteorhodopsin. *Biophysical Journal*, 84, 1202-1207

Vatter, A. E. And Wolfe, R. S. (1958). The structure of photosynthetic bacteria. *Journal of Bacteriology*, 75, 480-488.

Visscher, K. J., Bergstrom, H., Sundstrom, V., Hunter, C. N. and van Grondelle, R. (1989). Temperature-dependence of energy-transfer from the long wavelength antenna BChl-896 to the reaction center in *Rhodospirillum rubrum, Rhodobacter sphaeroides* (wt and m21 mutant) from 77 to 177k, studied by picosecond absorption-spectroscopy. *Photosynthesis Research*, 22, 211-217.

Von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane-proteins correlates with the trans-membrane topology. *Embo Journal*, 5, 3021-3027.

Von Lintig, J. and Vogt, K. (2000). Filling the gap in vitamin A research - Molecular identification of an enzyme cleaving beta-carotene to retinal. *Journal of Biological Chemistry*, 275, 11915-11920.

Wachter, R. M., Elsliger, M. A., Kallio, K., Hanson, G. T. and Remington, S. J. (1998). Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure with Folding and Design*, 6, 1267-1277.

Walter, J. M., Greenfield, D., Bustamante, C. and Liphardt, J. (2007). Light-powering *Escherichia coli* with proteorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 2408-2412.

Walther, T. H., Gottselig, C., Grage, S. L., Wolf, M., Vargiu, A. V., Klein, M. J., Vollmer, S., Prock, S., Hartmann, M., Afonin, S., Stockwald, E., Heinzmann, H., Nolandt, O. V., Wenzel, W., Ruggerone, P. and Ulrich, A. S. (2013). Folding and Self-Assembly of the tata Translocation Pore Based on a Charge Zipper Mechanism. *Cell*, 152, 316-326.

Walther, T. H., Grage, S. L., Roth, N. and Ulrich, A. S. (2010). Membrane Alignment of the Pore-Forming Component tata(d) of the Twin-Arginine Translocase from *Bacillus subtilis* Resolved by Solid-State NMR Spectroscopy. *Journal of the American Chemical Society*, 132, 15945-15956.

Walz, T., Jamieson, S. J., Bowers, C. M., Bullough, P. A. and Hunter, C. N. (1998). Projection structures of three photosynthetic complexes from *Rhodobacter sphaeroides*: LH2 at 6 angstrom LH1 and RC-LH1 at 25 angstrom. *Journal of Molecular Biology*, 282, 833-845.

Wang, W. W., Sineshchekov, O. A., Spudich, E. N. and Spudich, J. L. (2003). Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. *Journal of Biological Chemistry*, 278, 33985-33991.

Watkins, D. W., Armstrong, C. T. and Anderson, J. L. R. (2014). De novo protein components for oxidoreductase assembly and bio logical integration. *Current Opinion in Chemical Biology*, 19, 90-98.

Watrob, H. M., Pan, C. P. and Barkley, M. D. (2003). Two-step FRET as a structural tool. *Journal of the American Chemical Society*, 125, 7336-7343.

White, G. F., Schermann, S. M., Bradley, J., Roberts, A., Greene, N. P., Berks, B. C. and Thomson, A. J. (2010). Subunit Organization in the tata Complex of the Twin Arginine Protein Translocase a site-directed EPR spin labeling study. *Journal of Biological Chemistry*, 285, 2294-2301.

Willem, M., Miosge, N., Halfter, W., Smyth, N., Jannetti, I., Burghart, E., Timpl, R. and Mayer, U. (2002). Specific ablation of the nidogen-binding site in the laminin gamma 1 chain interferes with kidney and lung development. *Development*, 129, 2711-2722.

Williams, J. (1988). Construction of specific mutations in the photosystem II photosynthetic reaction centre by genetic engineering methods in the cyanobacterium Synechocystis 6803. *In*: The Cyanobacteria (Glazer, A. And Packer, L., eds.) , pp. 809-812.

Woodbury, N. W. and Allen, J. P. (1995). The pathway, kinetics and thermodynamics of electron transfer in wild type and mutant reaction centers of purple nonsulfur bacteria. *Anoxygenic Photosynthetic Bacteria*, 527-557.

Yan, B., Spudich, J. L., Mazur, P., Vunnam, S., Derguini, F. And Nakanishi, K. (1995). Spectral tuning in bacteriorhodopsin in the absence of counterion and coplanarization effects. *Journal of Biological Chemistry*, 270, 29668-29670.

Ye, S. X., Strzalka, J. W., Discher, B. M., Noy, D., Zheng, S. Y., Dutton, P. L. and Blasie, J. K. (2004). Amphiphilic 4-helix bundles designed for biomolecular materials applications. *Langmuir*, 20, 5897-5904.

Yoshitsugu, M., Shibata, M., Ikeda, D., Furutani, Y. and Kandori, H. (2008). Color change of proteorhodopsin by a single amino acid replacement at a distant cytoplasmic loop. *Angewandte Chemie-International Edition*, 47, 3923-3926.

Yoshitsugu, M., Yamada, J. and Kandori, H. (2009). Color-Changing Mutation in the E-F Loop of Proteorhodopsin. *Biochemistry*, 48, 4324-4330.

Yutin, N. and Koonin, E. V. (2012). Proteorhodopsin genes in giant viruses. *Biology Direct*, 7.

Zeng, X. H., Choudhary, M. and Kaplan, S. (2003). A second and unusual *pucba* operon of *Rhodobacter sphaeroides* 2.4.1: Genetics and function of the encoded polypeptides. *Journal of Bacteriology*, 185, 6171-6184.

Zhu, Y. S. and Hearst, J. E. (1986). Regulation of expression of genes for light-harvesting antenna proteins LH1-i and LH-ii - reaction center polypeptides rc-l, rc-m, and rc-h - and enzymes of bacteriochlorophyll and carotenoid biosynthesis in *Rhodobacter capsulatus* by light and oxygen. *Proceedings of the National Academy of Sciences of the United States of America*, 83, 7613-7617.

Zhu, Y. S. and Kaplan, S. (1985). Effects of light oxygen and substrates on steady-state levels of messenger RNA coding for ribulose-1 5-bisphosphate carboxylase and light-harvesting and reaction center polypeptides in *Rhodopseudomonas sphaeroides*. *Journal of Bacteriology*, 162, 925-932.

Zuber, H. And Brunisholz, R. (1991). Structure and function of antenna polypeptides and chlorophyll–protein complexes: Principles and variability. Chlorophylls (Scheer, H., ed.), *CRC Press*, Boca Raton, Florida, 627–703

Zubkov, M. V. (2009). Photoheterotrophy in marine prokaryotes. *Journal of Plankton Research*, 31, 933-938.

Zurdo, J., Fernandezcabrera, C. and Ramirez, J. M. (1993). A structural role of the carotenoid in the light-harvesting ii-protein of *Rhodobacter capsulatus*. *Biochemical Journal*, 290, 531-537.