UNIVERSITY OF CALIFORNIA, SAN DIEGO

UTILIZING *CHLAMYDOMONAS REINHARDTII* AS A PLATFORM TO STUDY METABOLIC ENGINEERING

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Chemistry

by

Yuan Pu

Committee in charge:

Michael Burkart, Chair
Stephen Mayfield
Roberts Pomeroy

2016

The thesis of Yuan Pu is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____
Chair

The University of California, San Diego

2016

DEDICATION

I dedicate this thesis to my parents, Bingan Pu and Qing Bao, my aunt Jan Fang and her family for their support and love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all I want to thank my committee members including Dr. Burkart, Dr. Mayfield and Dr. Pomoroy. Thank you for everything you taught me and a lot of support in the past couple of years, also for your time in the summer for my defense.

Thanks Dr. Eva Sonnenshein and Dr. Joris Beld who led me to the world of microalgae from an undergraduate who knew nothing about algae.

Chapter 2, in full, is a reprint of the material as it appears in Journal of Applied Phycology, 2016. Sonnenschein, Eva C; Pu, Yuan; Beld, Joris; Burkart, Michael D 2016. The thesis author was the secondary author of this paper.

Thanks Dr. Suzanne Thomas from Salk Institute who gave a lot of support on my taxadiene project and helped a lot on my thesis writing and defense preparation.

Thanks Dr. Livia S Ferreira-Camargo who gave a lot helpful advice and suggestions on my projects and during my thesis writing.

Thanks Dr. Lorillee Tallorin for your help and support during my last phase.

Last but not least, thanks my family – my parents, my aunt Dr. Jan Fang and her family and my dear roommates and friends, thank you for your love and support. Thanks Kevin Luo for always being my first listener, reader and audience.

ABSTRACT OF THE THESIS

UTILIZING *CHLAMYDOMONAS REINHARDTII* AS A PLATFORM TO STUDY METABOLIC ENGINEERING

by

Yuan Pu

Master of Science in Chemistry

University of California, San Diego, 2016

Professor Michael Burkart, Chair

Paclitaxel, as known as taxol, is a member of a large family terpenoid and it is famous for its anti-cancer properties. It was originally isolated from a large amount of yew tree tissues and is hard to synthesize chemically. *Chlamydomonas reinhardtii* as a microalgae whose genome could be modified easily and be cultivating environmental friendly, is producing building blocks and starting materials for an important intermediate taxa-4, 11-diene in taxol biosynthesis. Here we introduced a new taxa-4,11-diene synthase which consists of geranylgeranyl pyrophosphate synthase and taxa-

4,11-diene synthase into *C. reinhardtii* and tested its expression via western blots. In addition, we identified two sfp type phosphopantetheinyl transferases (PPTases) in *C. reinhardtii* PptC1 and PptC2. Both of them activate cognate acyl carrier protein (ACP) of the type II fatty acid synthase (FAS) while only PptC2 is able to recognize type II FAS ACP in *E.coli*. Moreover, there is another single type I PKS existing in *C. reinhardtii* and we believe PptC2 is responsible for its activation. From the results of the above described studies, we conclude that *C. reinhardtii* has great potential for metabolic engineering and production of high-value natural products.

INTRODUCTION

Fossil fuels are still the largest resource of energy on this planet. In 2015, oil, natural gas and coal in total accounted for 85.9% of global primary energy consumption (British Petroleum, 2016). There is a large need for sustainable substitutes of fossil fuels based products as fossil fuels reserves dwindle and energy consumption continues to increase. Decades ago, scientists started to work on utilizing photosynthetic organisms as platforms to generate new energy sources and reduce dependence on fossil fuels. The promise of exploring such organisms as a source of energy includes less reliance on fossil fuel sources and reduced greenhouse gas output for a more sustainable future.

Algae is a general term for a large and diverse group of photosynthetic organisms, ranging from simple unicellular genera to complicated multicellular forms. The simple unicellular algae species are designated as microalgae. Microalgae are the primary source for third generation of biofuels, are known for easy genome modification, mature transformation technology and production scalability. (Dragone 2010, Rasala 2014). *Chlamydomonas reinhardtii* is species of microalgae that has been used as a model photosynthetic organism in the third generation biofuel research for years. It is also used as a platform to produce valuable products like antibodies (Tran 2009), immunotoxin fusions (Rasala, 2012) and furthermore, has high potential for expression of recombinant proteins. (Scranton 2015, Almaraz-Delgado 2014, Barrera 2013, Rasala 2014).

Chlamydomonas reinhardtii have a large central chloroplast occupying over 40% of its total cell volume (Franklin 2004). Given its large size and accessibility, the chloroplast of C. *reinhardtii* is amenable to genetic engineering via the use of high velocity microprojectiles or gene gun technique (Scranton 2015, Boynton 1988). While proving to be a useful technique, a major limitation of the use gene gun is that it is hard to

ensure that all genome copies of the target chloroplast will be successfully modified. If not all copies contain the target construct, conversion may occur. Colonies in a stable transformation state are described as homoplasmic colonies (Franklin 2004). Once a colony candidate is confirmed as homoplasmic, it will be rigid and considered for larger scale production.

In this thesis, two different studies in *C. reinhardtii* will be discussed in the following chapters. Project specific introductions will be given at the beginning of each chapter.

CHAPTER 1:

INSERTION OF TAXA-4,11-DIENE SYNTHASE INTO CHLAMYDOMONAS REINHARDTII

Terpenes and terpenoids (also known as isoprenoids) are the largest group of natural products, of which 55000 members are known to exist (Koksal 2011). They are made using the universal C_5 building blocks isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Further cyclization and rearrangement possibilities lead to an incredibly diverse set of compounds(Gershenzon 1993). Many other natural products such as alkaloids, phenolics, and vitamins contain terpenoid elements, indicating the importance of terpene biosynthesis study. Although terpenes are synthesized naturally, the yield of such natural synthesis is usually low and does not meet the rising demand of terpenes in industry (Wang 2015). Many groups have reported their success in engineering of terpene biosynthesis pathways into *E.coli*, yeast and other host organisms to increase production of terpenes and/or their precursors. In this huge family of compounds, paclitaxel (taxol) is famous for its excellent anti-cancer properties, which has led to increased demand for the compound. In nature, the production of taxol requires 750kg of dry bark from Pacific yew (*Taxus brevifolia*) to extract only 1mg of compound. Total synthesis of taxol requires a 35-51 step procedure and gives only a 0.4% yield. (Koepp 1995, Huang 2001, Ajikumar 2010, Nicolaou 1994, Holton 1994, Walji 2007). Decades ago, scientists started study alternative ways towards biosynthesis of this valuable compound.

During the whole taxol biosynthesis, there is an important intermediate diterpene called taxa-4,11-diene (Engels 2008), which is cyclized from geranylgeranyl pyrophosphate (GGPP). Scientists bioengineered taxa-4,11-diene synthase into *E.coli* and yeast years ago (Huang 2001, Engels 2008). Phytol, an important component in most kinds of chlorophyll structures including the most common and universal chlorophyll a,

is also derived from GGPP. As a green microalgae, *C. reinhardtii* itself generates GGPP for phytol production and eventually chlorophyll a production, which could be used as the starting material for taxa-4,11-diene production. The Toyomasu group reported isolation of a chimera fusicoccadiene synthase from plant-pathogenic fungus *Phomopsis amygdali* (Toyomasu 2007). Fusicoccadiene synthase consists of two domains: a prenyltransferase domain at the C-terminal end which converts isoprene units into GGPP and a terpene cyclase domain at the N-terminal which cyclizes GGPP into fusicoccadiene. By cutting off the terpene cyclase domain of fusicoccadiene synthases and ligating it with taxa-4,11-diene synthase and then transforming this fused taxa-4,11-diene synthase into *C. reinhardtii* we hope to achieve novel taxadiene biosynthesis in a photosynthetic organism.

Materials and Methods

Plasmid construction

DNA of fusicoccadiene synthase with two domains was obtained from Sapphire Energy. DNA of taxadiene synthase was synthesized by Life Technologies and then amplified by polymerase chain reaction (PCR) as nucleotide fragments. Taxadiene synthase and GGPPS from fusicoccadiene synthase were assembled into optimized vector D2 for chloroplast expression in *C. reinhardtii*.

Algal strains, transformation and culturing

Two algal strains were utilized for transformation: *C. reinhardtii* wild type (wt) strain 137c (Mt+)(CC-125) and *C. reinhardtii* psbA mutant strain (W1.1)(Manuell 2007). Both strains were grown to mid-log phase in tris-acetate-phosphate (TAP) media (Gorman and Levine 1965) at approximately 25°C and 70 μmol photons m^{-2} s^{-1} on a rotary shaker. They were harvested and resuspended in TAP medium at concentration of 3×10^7 cells/ml. Resuspended cultures were plated on TAP solid agar plates containing 150 mg/mL kanamycin. Approximately 2×10^7 cells were placed on a TAP plate for one transformation.

About 6-8μg of plasmid DNA were prepared to bind to gold particles (Seashell Technology) and about 0.4mL DNA/gold with solvent mixture was injected into special

tubing (Bio-Rad) for particle bombardment transformation. The transformation process was done by utilizing gene gun system (Bio-Rad) with helium as the carrier gas at approximately 275 psi. After transformation, plates were placed under low constant illumination overnight and then moved to high constant illumination for transformed colonies. The whole process of clearing non-transformed algae cells and forming transformed colonies took 7-15 days.

PCR screening

All colonies survived on transformation plates were picked and restreaked on fresh TAP plates containing 200mg/mL kanamycin and grown under high constant illumination for 1-2 weeks. Colonies were then screened by PCR using specific primers designed for taxa-4,11-diene synthase (TS forward: AACGCTTTAGGTGATGGCGA

TS reverse: TGTTGCACCTGGGTTGTCTT) and GGPPS (GGPPS forward: ATAATACATATGACACAATTAGAATGGATGCGTCAAATAATAACCGGT

GGPPS reverse: AACACGTAATAATTCTAATAATAAACGCAT) to verify the existence of transformed genes. Colonies confirmed to contain target gene were screened by PCR using general primers for D2 vector (16S forward: CCGAACTGAGGTTGGGTTTA; 16S reverse: GGGGGAGCGAATAGGATTAG; P3HB forward: CGTCCACTAAAATTTATTTACCCGAAGGGG;

P3HB reverse: GTTAAGGCTAGCTGCTAAGTCTTCTTTTCGC) to verify if they are homoplasmic. All PCR products were analyzed on 1-2% agarose gel with ethidium bromide.

Western blots

1L strain cultures were grown to concentration about 2×10^6 cells/mL under high constant illustration and then harvested by centrifugation and resuspened in 20mL lysis buffer (750mM Tris-Cl pH 8.0, 15% sucrose wt/vol, 100mM β-mercaptoethanol, 1mM phenylmethylsulfonylfluoride). Samples were lysed by sonication and separated by centrifugation. Fractions were saved and run on 12% SDS-PAGE gel and transfer to PVDF membranes. Membranes were blocked with 5% (wt/v) skim milk dissolved in 1× TBST buffer for 90 min at 4°C. Then 5 µl anti His epitope tag mouse monoclonal antibody (Rockland Inc.) in 10mL blocking buffer was added for incubation at 4°C overnight. After three times of washing with 50mL 1× TBST buffer for 10 min at room temperature, 3 µl Anti-Mouse IgG (H & L) AP Conjugate (Promega) in 10mL blocking buffer was added for incubation at room temperature for 1 hour. The membrane was washed three times with 50mL 1× TBST buffer for 10 min at room temperature and then developed with nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro-3'-indoylphosphate (BCIP) in 10mL miliQ water.

GCMS

1L strain cultures were grown to concentration of about 2×10^6 cells/mL under high constant illustration and then harvest by centrifugation and resuspened in 50mL lysis buffer (KOH :MeOH = 10:1) followed by sonication and incubation at 55°C for 30 min. Then all natural products were extracted by 1mL hexane. Samples were run and analyzed on GCMS from Agilent using program starting with 40°C and then ramped to 280°C at speed of 8°C/min and hold at 280°C for extra 5 min.

Protein expression in *E.coli*

The fused taxa-4,111-diene synthase was assembled into pET-28a(+) and pHis8 vectors and then transformed into *E.coli* BL21 cells. Four 5mL LB and kanamycin culture of both strains were grown on shaker at 37°C until OD_{600} reached 0.6, then they were induced by 0, 0.25, 0.5 and 1 mM IPTG and incubate for another 3h on shaker at 37°C. 1ml fraction of each culture was taken and cells were lysed by adding 250µl lysis buffer (50mM Tris, 150mM NaCl, 15% glycerol) and sonication for 1 min on ice. Insoluble proteins were separated from soluble fraction by centrifugation. All samples were run on 12% SDS-PAGE gels and stained by coomassie blue.

Taxadiene Synthase Activity Assay and GCMS

GGPP as substrate was added to purified taxadiene synthase in the system of 25mM HEPES pH 8.0, 10% glycerol and 1mM MgCl. Assay was incubated at room temperature

for 1h and then extracted by hexane. Hexane sample was analyzed on GCMS using the same program for *C. reinhardtii* sample.

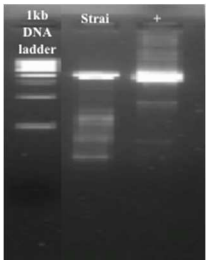
Results

The construct of the new Taxa-4,11-diene Synthase was assembled as shown in Figure 1 in optimized and kanamycin resistant vector D2 for chloroplast expression in *C. reinhardtii*. Wild type *C. reinhardtii* (both c137+ and W1.1) do not contain kanamycin resistant and do not survive on plates with antibiotic unless successfully transformed with the target construct after shooting. All surviving colonies on antibiotic containing plates were picked and transferred to new TAP plates with 150mg or 200mg kanamycin per plate for PCR screening.

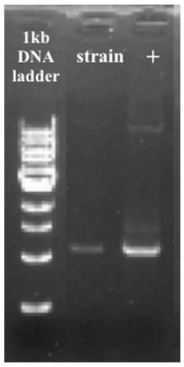
Figure 1: Diagram of construct transformed into C. reinhardtii

The presence of taxa-4,11-diene synthase and GGPPS domain from fusicoccadiene synthase in candidate colonies was verified by specific primers via PCR (Figure 2). In Figure 2, it is clearly shown that the selected strain contains the target genes by comparing the PCR product with algal lysate to gene positive PCR product. Homoplasmy, a term used to ensure that all 80 copies of chloroplast genomes contain the target genes, was verified through PCR screening with D2 vector specific primers (Figure 3). There are two sets of primers in one PCR reaction system: one set of primers amplifies 16S RNA in both wildtype and strain while the other set of primers amplifies the region of psbH gene which only works in wildtype since that part is replaced by the

target gene in the modified strain. In other words, only having the top band of 16S RNA but no lower band of psbH gene on the agarose gel shows that the strain is homoplasmic.



(a) Confirmation of taxadiene synthase on DNA gel



(b) Confirmation of GGPPS on DNA gel

Figure 2: Confirmation of presence of (a) taxadiene synthase and (b) GGPPS in strain via gene specific PCR

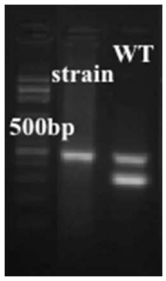


Figure 3: Confirmation of the homoplasmy of strain by showing the loss of bottom band that amplified the loci just in front of the psbH gene. The loss represents target gene fully replaced psbH gene in *C. reinhardtii* chloroplast genome.

When 1L colony candidate culture grew to middle logarithmic phase, they were harvested and added 30mL lysis buffer and then lysed by sonication. Samples from insoluble and soluble fraction were saved for western blots. Lysed biomass samples were analyzed on 12% SDS-PAGE gel and then transfer to PVDF membrane. PVDF membrane was soaked in blocking buffer first and then treated with anti His-tag antibody. Figure 4 shows the faint protein bands in strain: soluble fraction lane at approximately

120kDa and a more faint one at about 50kDa. The calculated size of taxadiene synthase domain alone should be about 90kDa, GGPPS should be about 40kDa and the fused gene of those two should be about 130kDa. Based on western blots result, the fused taxadiene synthase and GGPPS gene is not expressed well in this strain.

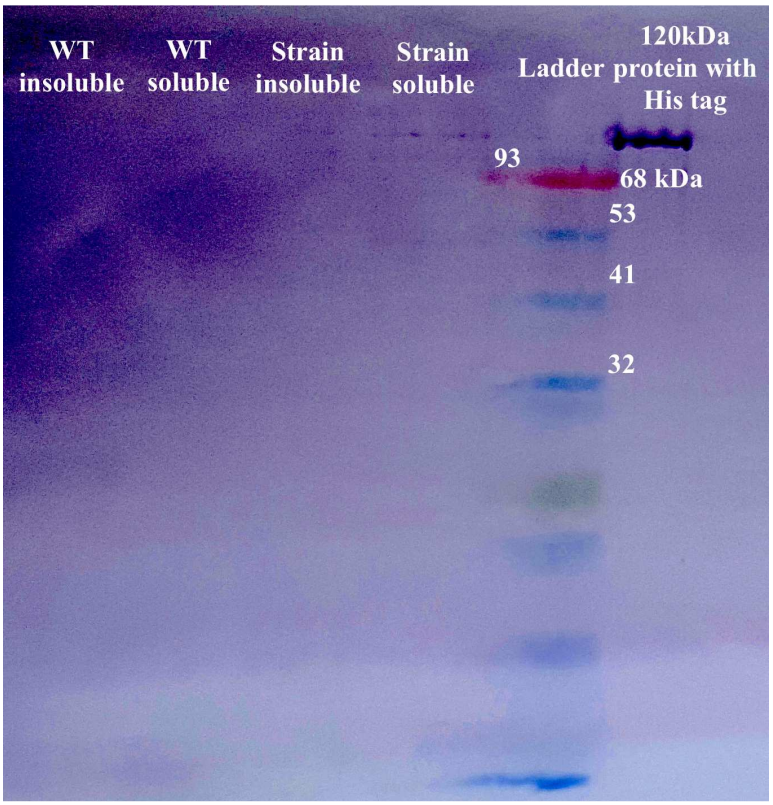
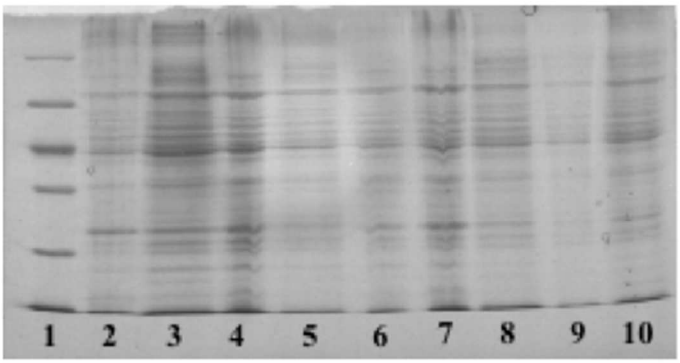


Figure 4: Western Blots of strain analyzed with anti His antibodies

To check if there is any taxa-4,11-diene produced in this *C. reinhardtii* strain, all natural products were extracted with hexane and analyzed on GCMS. Unfortunately, there was no taxa-4,11-diene product detected on chromatograph.

To test the general protein expression, the fused protein was introduced into pET28a(+) and pHis8 vectors and then expressed in *E.coli* BL21(DE3) cells. Proteins were induced by IPTG when cultures OD_{600} reached 0.6 to maximize expression levels. Induced cells were harvested, lysed, and analyzed on 12% SDS-PAGE gels. Figure 5

shows protein expression levels of construct pET28a taxadiene synthase and GGPPS in *E.coli* BL21 cells with different concentrations of IPTG induction. Lane 2 and 3 were induced with 0mM IPTG but compared to other lanes that were induced by 0.25-1mM IPTG there is no obvious differences between soluble or insoluble fractions. Moreover, there is no band at correct size corresponding to the size of fused GGPPS and taxadiene synthase. It indicates the fused taxa-4,11-diene synthase and GGPPS was not expressed in *E.coli*.



Lane	
1	Ladder
2	Pellet 0mM IPTG
3	Soluble 0mM IPTG
4	Pellet 0.25mM IPTG
5	Soluble 0.25mM IPTG
6	Pellet 0.5mM IPTG
7	Pellet 1mM IPTG
8	Soluble 0.5mM IPTG
9	Soluble 0.5mM IPTG
10	Soluble 1mM IPTG

Figure 5: Fused taxa-4,11-diene synthase and GGPPS expression in *E.coli* BL21 cells induced by different concentrations of IPTG.

However, when expressing taxadiene synthase domain alone in *E.coli*, it expressed well (Fig.6). Therefore it was purified with Ni-resin and tested for its enzyme activity. After 1 hour assay incubation with GGPP as substrate, hexane was used to extract all possible products. Extracted sample was analyzed on GCMS using the same program for *C. reinhardtii* extraction sample. On this chromatograph there is a huge peak at retention time 16.244 min could be confirmed as taxa-4,11-diene (Figure 7). Comparing to chromatograph of C. reinhardtii extraction samples, there is no peak around 16.244 min that has identical or similar spectrum like figure 7.

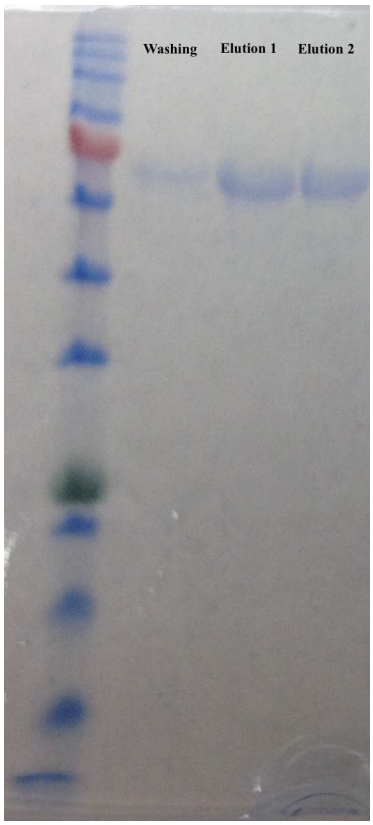


Figure 6: Taxadiene cyclase domain only expression in *E.coli* BL21 cells after purification

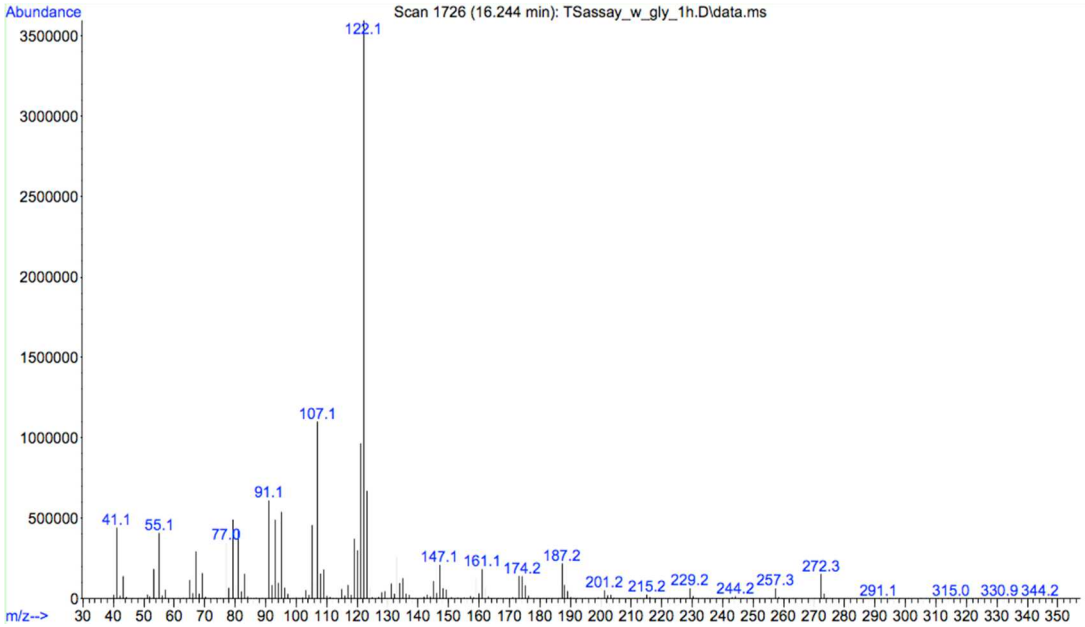


Figure 7: Spectrum of taxa-4,11-diene extracted from taxadiene synthase activity assay

Discussion

This taxa-4,11-diene synthase together with GGPPS is not the largest gene that scientists have attempted to transform into *C. reinhardtii* chloroplast genome, so it was unexpected that it did not express well. The PCR results clearly show that taxa-4,11-diene synthase was transformed into *C. reinhardtii* chloroplast genome and this strain is proved to be homoplasmic, but there was no taxa-4,11-diene as the product compound detected on GCMS from its natural products extraction as expected. Therefore, the general expression of this fused protein was tested in *E.coli* BL21 cells. In parallel, the expression level of taxadiene cyclase domain alone was tested as well as control. The protein expression level shows that the taxa-4,11-diene synthase fused with GGPPS is not expressed well in chloroplast of *C. reinhardtii* and did not express in *E.coli*. But the taxadiene cylase domain expressed well in *E.coli.* The following enzyme activity assay also showed that the cyclase domain is active and capable of producing detectable taxa-4,11-diene as product.

There are a couple of explanations for the poor expression of new taxa-4,11-diene synthase. One major reason could be that the D2 vector used has a weak *psbD* promoter and 5' to 3' untranslated region (UTR). In general, a D1 vector that contains *psbA* promoter and 5' UTR is the most effective vector to express recombinant proteins in chloroplasts of *C. reinhardtii* (Specht 2014). Also, when expressing the same foreign recombinant gene, the candidate strain with D1 system shows a higher protein expression level (Jones 2012). Transformation of D1 vector containing taxa-4,11-diene synthase and GGPPS into *C.reinhardtii* was attempted but did not succeed. Therefore, we decided to continue with this homoplasmic strain using D2 vector. In addition, since there is a His-

tag on the C-terminus end, protein purification with Ni-resin was also attempted several times, but did not work, most likely because of low protein expression level.

Because taxadiene cyclase domain is codon optimized for *C. reinhardtii* when it is synthesized, and GGPPS domain works well in C. reinhardtii according to the company it is from, genetic incompatibility should not be the reason for poor expression. Another possible reason for the poor expression of new taxa-4,11-diene synthase is poor fusion design of the two protein domains. Even though there is no stop codon between the two DNA sequences in the combined new taxa-4,11-diene synthase, a faint GGPPS band is visible in western blot analysis. There is no linker or spacer designed between the two domains. This unnatural fusion design may have inherent problems that would have resulted in enzyme function elimination or even abolition. It is also possible to cause the separation of two bands on the membrane, but only the GGPPS band would have been visualized through its His tag. We indeed visualized only the GGPPS band, so fusion design is a possibility.

A third reason is that the concentration of the sample was too low to be detected on GCMS. Considering the expression level of the fused protein is low, it follows that the production of taxa-4,11-diene would be low as well.

For future studies, it may be possible to directly transform only taxadiene cyclase into *C. reinhardtii,* considering that it has native ability GGPP production for phytol synthesis. This direction is not ideal however, because of lack of control over the use of GGPP in the cell for taxadiene over phytol synthesis. Assembly of recombinant protein into D1 vector might be one solution to address low protein expression. Other possible

solutions include: improving the fusion by inserting linkers between the two domains, running a double transformation of taxadiene synthase and GGPPS genes separately.

CHAPTER 2:

PHOSPHOPANTETHEINYLATION IN THE GREEN MICROALGAE CHLAMYDOMONAS REINHARDTII

Introduction

The eukaryotic green algae *Chlamydomonas reinhardtii* has been a model organism for many decades and has recently gained importance in the field of recombinant protein expression (Rochaix 1995; Griesbeck et al. 2006). *C. reinhardtii* can be used as production host for high-value products such as pharmaceutically relevant proteins including antibodies and immunotoxin fusions (Tran et al. 2009; Rasala et al. 2012; Jones et al. 2013). In biofuel research, *C. reinhardtii* further serves as reference for understanding algal fatty acid metabolism and accumulation and corresponding initial metabolic engineering experiments (Guschina and Harwood 2006; Blatti et al. 2012; Merchant et al. 2012; Wijffels et al. 2013; Li-Beisson et al. 2015; Scranton et al. 2015; Ahmad et al. 2015). The decline in fossil fuels has driven efforts in this area; however, there are still major gaps in our basic understanding of the pathways essential for the exploitation of microalgae as an energy source.

In fatty acid synthases (FASs), polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs), the central domain is the carrier protein that presents the growing product via a 4'-phosphopantetheine (4'-PPant) arm. This arm is posttranslationally attached by a phosphopantetheinyl transferase (PPTase) enzyme (Beld et al. 2014c). The PPTase catalyzes the transfer of 4'-PPant from coenzyme A (CoA) to a conserved serine residue of the carrier protein, thereby transforming the *apo*-synthase into its active *holo*-form. PPTases are thereby essential in primary and secondary metabolism as activators of these major synthases. PPTases have been grouped into three categories: the AcpS- (group I) and Sfp-type (group II) PPTases, each named after the corresponding prototype, and the FAS-integrated PPTases. First discovered in the 1960s,

the *holo*-ACP synthase (AcpS) from *Escherichia coli* represents the first category of small enzymes (approx. 120 aa) that are dedicated to its FAS (Alberts and Vagelos 1966). Sfp-type PPTases are named after Sfp from *Bacillus subtilis* (256 aa), a protein identified due to its importance in surfactin production (Nakano and Zuber 1990; Nakano et al. 1992). Sfp itself is very promiscuous and has been applied in various technologies and metabolic engineering efforts (Sunbul et al. 2009; Beld et al. 2014c). It is able to activate FASs, PKSs and NRPSs. Generally, organisms carry at least an AcpS-type PPTase and for some, further PPTases can be present to fulfil other functions than FAS activation. As an exception in the bacterial kingdom, cyanobacteria and a few other species have been shown to possess only Sfp-type PPTases (Copp and Neilan 2006). Humans carry a single Sfp-type PPTase, namely AASDHPPT, which has a broad substrate specificity and has evolved to serve novel functions (Joshi et al. 2003). Besides this classification into types, current sequence-based *in silico* analysis is not able to provide information on the functionalities of a PPTase.

In contrast to bacteria, eukaryotic cells possess two fatty acid biosynthetic systems: a type I FAS encoded on one polypeptide and a type II FAS presented on separate proteins, the latter resembling the bacterial synthase. Some eukaryotes such as the protozoans *Toxoplasma gondii* and *Dictyostelium discoideum* appear to have one PPTase committed to each FAS type (Cai et al. 2005; Nair et al. 2011). However, others like *Homo sapiens*, *Cryptosporidium parvum* and *Plasmodium falciparum* have only one PPTase, solely responsible for all phosphopantetheinylation reactions (Joshi et al. 2003; Cai et al. 2005). The fungal PPTases are metabolically distinct: one PPTase is encoded within the type I FAS (megasynthase) gene and a second one, Lys5, is involved in lysine

biosynthesis (Lambalot et al. 1996; Fichtlscherer et al. 2000). Additional PPTases can be present for posttranslational modification of proteins such as PKSs or NRPSs (Beld et al. 2014c). In plants and algae, fatty biosynthesis is located in the chloroplasts and mitochondria (Beld et al. 2014a), while the genes encoding these type II FASs are located in the nucleus. Thus, the question arises where and how the two nuclear-encoded ACPs are posttranslationally modified. In plants, two PPTases have been analysed so far (Elhussein et al. 1988; Guan et al. 2015). For the spinach PPTase, it was shown that both *apo*- and *holo*-ACP were efficiently transported into the chloroplast and no satisfying answer for the localization of phophopantetheinylation has been found (Fernandez and Lamppa 1990; Savage and Post-Beittenmiller 1994; Yang et al. 1994). For the mitochondrial PPTase (mtPPTase) from *Arabidopsis thaliana*, it was shown that it is only able to activate the mature *apo*-mtACP localized in the mitochondrion, thus the process of phosphopantetheinylation occurs in the organelle after import from the cytosol.

Finally, efforts in metabolic engineering of the fatty acid and secondary metabolism have demonstrated the essential nature of PPTase activation (Ku et al. 1997; Kealey et al. 1998; Amiri-Jami and Griffiths 2010). While heterologous expression of modular synthases in *Escherichia coli* and *Saccharomyces cerevisiae* required co-expression of Sfp (Kealey et al. 1998), a natural product synthase expressed in tobacco could be activated by an endogenous PPTase (Yalpani et al. 2001). Thus algae deserve increased attention as potential expression hosts for these biosynthetic pathways, including the ability to naturally produce complex, high-value fatty acids and lipids. However, contrary to their central function and importance for metabolic engineering, algal PPTases have not yet been described. Knowledge on the functionalities of these

enzymes will improve our ability to utilize microalgae for the production of lipid-based fuels and valuable molecules such as polyunsaturated fatty acids or antibiotics.

In this study, we have investigated two PPTase homologues from the green microalgae *C. reinhardtii*. Their phylogenetic position and predicted structure were analysed *in silico*, and the effect of known PPTase inhibitors on algal growth was monitored. The algal PPTases were then heterologously produced in *E. coli* in order to evaluate the substrate range with various CPs. Finally, we screened the available genome data for other microalgal PPTases and investigated other potential targets of PPTases besides fatty acid metabolism.

Experimental

Strains and materials:

Chlamydomonas reinhardtii c137 (mt+) was grown in TAP (Tris-acetate-phosphate) media (Gorman and Levine 1965) at 23°C and 70 μmol photons m^{-2} s^{-1} on a rotary shaker. RNA was purified from a culture in late logarithmic phase using PureLink Plant RNA Reagent (LifeTechnologies) following the manufacturer's protocol. cDNA was prepared with the Verso cDNA Kit (Thermo Scientific).

Sequence analysis, cloning, protein expression and purification:

The genome sequence of *C. reinhardtii* was screened for potential PPTases with *Escherichia coli* AcpS (GenBank: NP_417058), *Bacillus subtilis* Sfp (GenBank: CAA44858) or AASDHPPT (GenBank: NP_056238) as query using DELTA-BLAST (Altschul et al. 1990). The genes of the two best hits (GenBank: XP_001700873, herein referred to as PptC1, and XP_001689489 as PptC2) were amplified from cDNA with Phusion High Fidelity polymerase (New England Biolabs) in GC buffer, betaine (1 M final concentration), and in the case of pptC2, with 1.5% DMSO using the primers: pptC1forward: GAGACATATGATGCAGCGCTGTGCCGTCACTGCG, pptC1reverse: GAGACTCGAGCCCCTCCGCCACGCCCGC, pptC2forward: GAGACATATGATGCAGACTAGCTATGAAAGGAACGATGG, pptC2reverse: GAGACTCGAGGCATTGCTGGGCCAGCGC. After NdeI/XhoI restriction digest (New England Biolabs), pptC1 was cloned into pET29a, without and with stop codon into a modified pET24a containing an MBP-tag (Kosa et al. 2012) and pptC2 into pET29a

(Novagen). After Nde/KpnI-HF restriction digest, PptC1 was additionally inserted into a modified pET29a containing GFP (HindIII/EcoRI-inserted) (Kosa et al. 2012). For protein expression in *E. coli* BL21 (DE3), culture was grown at 37°C to an OD of 0.6 and induced with 1 mM IPTG at 37°C for 3 hours. The proteins were purified using Ni^{2+}-NTA (Novagen) affinity chromatography and amylose resin (New England Biolabs).

Protein modelling and phylogenetic analysis:

Protein structures were homology modeled with Swiss-Model (Arnold et al. 2006). PPTase sequences were aligned using ClustalW (Larkin et al. 2007) and a neighbor-joining tree was generated with MEGA6 (Tamura et al. 2013).

In vivo inhibition assay:

C. reinhardtii and *E. coli* were precultured in TAP or LB medium (Gerhardt et al. 1994) for four days or overnight, respectively. For the assay, precultures were diluted to an OD590 of 0.01 (*C. reinhardtii*) or OD590 of 0.0001 (*E. coli*) and aliquoted as 5 mL into glass tubes or 120 µl into wells of a microtiter plate. 10, 1 or 0.1 mM of the PPTase inhibitors 6-NOBP (Yasgar et al. 2010), ML267 (Foley et al. 2014) or AcpS7 (Joseph-McCarthy et al. 2005) were added. Growth was monitored using OD590 over seven days (*C. reinhardtii*) or 24 hours (*E. coli*).

In vitro activity assays:

In vitro phosphopantetheinylation activity of PptC1 and 2 was performed using the coenzyme A-biosynthetic enzymes CoaA, D and E, different CPs and a (7-nitro-2-1,3-

benzoxadiazol-4-yl)-modified pantetheine probe (Beld et al. 2014b) in a one-pot reaction (Worthington and Burkart 2006; Haushalter et al. 2008). In detail, the reaction was performed in 50 mM phosphate buffer (pH 8.0) with 12.5 mM $MgCl_2$, 8 mM ATP, 0.1 μg μl^{-1} CoaA, 0.1 μg μl^{-1} CoaD, 0.1 μg μl^{-1} CoaE, 0.1 μg μl^{-1} PPTase, 0.2% TritonX, 2 mM probe, ~1 mg mL^{-1} CP or synthase at a final volume of 50 μl for 3 hours at 37°C. 25 μl of reaction was separated on 13% urea PAGE and visualized under UV.

PptC2 activity by HPLC.

For the pH profile, reaction mixtures with 38 μM apoACP, 10 nM PptC2, 500 μM CoA, 10 mM $MgCl_2$, and 75 mM buffer were incubated at 37°C for 10 min and quenched with 50 mM EDTA. MES/NaAc buffers were used for pH 4.5–6.5 and Tris-HCl buffers for pH 7.0–8.5. For PptC2 activity test, reaction mixtures contained 38 μM apoACP, 1.2 nM PptC2, 75 mM Tris-HCl buffer (pH 7.5) and 10 mM $MgCl_2$ and the concentration of CoA was varied between 1 and 500 μM. Reactions were incubated at 37°C for 10 min and quenched with 50 mM EDTA. The peak area of holoACP was determined by HPLC.

In silico screen for algal PPTase targets and PPTases:

Microalgal genomic data was obtained either from NCBI or DOE Joint Genome Institute Genome Portal. PPTase genes in algal sequence data were identified using DELTA-BLAST (Altschul et al. 1990) with Sfp (GenBank: CAA44858), AcpS (GenBank: NP_417058) or AASDHPPT (GenBank: NP_056238) as query against Chlorophyta (taxid: 3041), Rhodophyta (taxid: 2763) and Bacillariophyceae (taxid: 33849), Phaeophyceae (taxid: 2870), Dinophyceae (taxid: 2864), Haptophyta (taxid: 2830), and

Chrysophyceae (taxid: 2825) and by blasting directly against the sequenced genomes at JGI. Non-microalgal sequences and those not containing the conserved residues (G105, D107, K150, E151, K155) (Lambalot et al. 1996) were removed. Alignments were generated using Muscle and a neighbour joining tree (bootstrap 100) was calculated using MEGA 6.0 (Tamura et al. 2013).

Results

Phylogeny and predicted structure of *C. reinhardtii* PPTases

Using DELTA-BLAST with *E. coli* AcpS, *B. subtilis* Sfp and the human AASDHPPT against the *C. reinhardtii* genome produced two dominant hits (Table S1), both *C. reinhardtii* gene products, herein referred to as PptC1 (Genbank acc. no. XP_001700873, 256 aa) and PptC2 (Genbank acc. no. XP_001689489, 510 aa), carry the active site residues of Sfp (Beld et al. 2014c): D107 and E109 (replaced by an A in PptC1) coordinating the Mg^{2+} ion, P76 and K75 supporting coenzyme A, and E151 that deprotonates the serine of the incoming ACP to facilitate PPant transfer (Fig. S1). PptC1 has a higher identity to the query sequences (23.7 to 29.6%) in comparison to PptC1 (19.2 to 27.8%). Overall, PptC1 has the highest identity to AASDHPPT and PptC2 to AcpS.

Prediction by sequence alignment was confirmed by modelling of the tertiary structure. The best accuracy of model prediction using SWISS-MODEL (Kiefer et al. 2009) was obtained with AASDHPPT (PDB ID 2CG5) as template for both PptC1 (GMQE 0.58) and 2 (GMQE 0.23), while accuracy using Sfp (4MRT) or AcpS (1FTH) was lower (GMQE 0.41 and 0.20 or 0.23 and 0.07) (Fig. 1). Predicted structures of PptC1 with either template possess folds characteristic of Sfp-type PPTases, however, in case of PptC2, only the Sfp-based model demonstrates resemblance (Fig. S2).

A phylogenetic tree of the described PPTases reveals that indeed the two *C. reinhardtii* PPTases group within the Sfp-type enzymes (Fig. S3). PptC1 appears as

outgroup within the human PPTase-like enzyme cluster, PptC2 is placed with the PPTases MtaA and MupN (Silakowski et al. 1999; Shields et al. 2010). The algal PPTases do no cluster with mtPPTase of *Arabidopsis*, however, share an sequence identity of 22% (PptC1) and 29% (PptC2).

Further *in silico* characterization was performed to predict protein localization. The protein sequences were submitted to PredAlgo (Tardif et al. 2012), and no specific targeting sequence was identified for PptC1, while PptC2 was predicted to be targeted to the mitochondria, albeit with a rather weak score (Mscore = 0.80).

While we wished to study knockouts of these genes, generation of knockout strains and gene silencing in *C. reinhardtii* remain challenging techniques. In order to complement our bioinformatic analysis of the PPTases *in vivo*, we supplemented cultures with known PPTase inhibitors and analysed growth. Here, we fed ML267 (Foley et al. 2014) (specific for Sfp-type PPTases), AcpS7 (Joseph-McCarthy et al. 2005) (specific for AcpS-type PPTases) and 6-NOBP (Yasgar et al. 2010) (general, but poor PPTase inhibitor) to wild-type *C. reinhardtii* and evaluated the resultant phenotype (Fig. S4). Growth of *C. reinhardtii* was completely abolished by 10 mM ML267, while AcpS7 and 6-NOBP appeared to have no strong effect. In comparison, growth of *E. coli* (carrying both type of PPTases) was only inhibited by addition of 6-NOBP. It is possible that 6-NOBP does not enter the algal cells or has no inhibition of algal PPTase activity.

Protein expression and *in vitro* activity

PptC2 was cloned from cDNA into an *E. coli* expression vector and solubly expressed with an N-terminal His-tag. The same procedure was applied for PptC1, but it expressed as an insoluble aggregate. Although we obtained soluble protein from N-terminal MBP and C-terminal 6xHis-tag fusions, the protein showed no activity (data not shown). Finally, PptC1 was expressed as a C-terminal GFP-His fusion. This complex was soluble and active in a one-pot PPTase activity assay (Fig. 1) described below.

PPTases have two substrates, coenzyme A and a carrier protein, and often show promiscuous behaviour for both. The substrate range of the algal PPTases was evaluated *in vitro* using a one-pot carrier protein (CP) labeling reaction (Worthington and Burkart 2006). Therein, a fluorescent pantetheine derivative is converted into a coenzyme A analog using heterologous *E. coli* CoaA, CoaD, CoaE. In the same reaction, the PPTase catalyzes labeling of *apo*-CP with the *in situ* generated coenzyme A analog. The resulting fluorescent crypto-CP is visualized via PAGE and UV analysis (Fig. 2). The following CPs from various carrier protein-dependent pathways were analysed for substrate selectivity for both algal PPTases: type I and II FAS (*E. coli*, *C. reinhardtii* and human ACP); type I and II PKS (mycocerosic acid synthase from *Mycobacterium tuberculosis* and actinorhodin ACP from *Streptomyces coelicolor* A3(2)); and NRPS (BpsA from *Streptomyces lavendulae*) (Table 2). While PptC2 catalyzed the labeling reaction for *E. coli* ACP, *C. reinhardtii* cACP and actinorhodin ACP, PptC1 interacted with *C. reinhardtii* cACP and demonstrated weak catalysis with *E. coli* ACP.

Activity of PptC2 was separately measured through formation of *E. coli holo*-ACP by HPLC (Fig. 3). The pH optimum of PptC2 was observed at pH 7.5 (Fig. 3 A), which is similar to the optimum of the cyanobacterial PPTase Sppt, but higher than those

from other bacterial or the human PPTase (Finking et al. 2002; Joshi et al. 2003; Roberts et al. 2009). Using 38 μM *E. coli apo*-ACP and 1.2 nM PptC2 and CoA concentrations from 1 to 500 μM, PptC2 activity increased with CoA concentration and maximum activity was not reached at 500 μM (Fig. 3 B).

Algal PPTases

Following the criterion of the conserved residues G105, D107, K150, E151, K155 (Sfp) (Lambalot et al. 1996) after DELTA-BLAST (Fig. S5), PPTases were identified in most of the available algal genome data. The green algal PPTases formed two distinct clusters, each containing one of the *C. reinhardtii* PPTases (Fig. 4). Cluster 1 contained PptC1 and cluster 2 contained PptC2 and Sfp. Two PPTases were found in most chlorophyta and the single cryptophyte identified. Only one PPtase per species was identified for the rhodophyta and bacillariophyta strains. Only the PPTase of the heterkont *Aureococcus anophagefferens*, the chlorarachniophyte *Bigelowiella natans*, one PPTase of the cryptophyte *Guillardia theta* and the single red algal PPTase fall within the two clusters of the green algal PPTases, while all other PPTases show less similarity to the PPTases of green microalgae.

Wang *et al.* proposed that group II PPTases can be sub-classified into two- (D-X-E) and three-magnesium-binding-residue-PPTases (D-E-E) (Wang et al. 2014). Using three examples, they grouped algal PPTases into the first group, with A as second residue. While PptC1 follows this pattern, PptC2 is a three-magnesium-binding-residue-PPTase. Correspondingly, all other PPTases of cluster 1 and 2 (Fig. 4) follow this classification

(Fig. S4). The other PPTases identified carry either E (*Ectocarpus siliculosus* CBJ49162.1, *Emiliania huxleyi* XP_005760789.1), M or V or as second residue (Fig. S4). M and V as second residue also occur in animals (M), plants (V) and fungi (M, V) (Wang et al. 2014)

Discussion

Two PPTases of the green microalgae *C. reinhardtii* were identified and heterologously produced in *E. coli* for *in vitro* activity analysis. Due to sequence similarity, structural modelling and inhibition assay, both PPTases were assigned to the Sfp-type, which agrees with their evolutionary origin. Green microalgae originated from primary endosymbiosis of a photosynthetic cyanobacterium-like prokaryote, and in cyanobacteria, only Sfp-type PPTase were identified independent of the presence or absence of natural product gene clusters (Copp and Neilan 2006; Roberts et al. 2009). As proposed by Finking *et al.*, the Sfp-type PPTases might have developed from duplication of an ancestral AcpS PPTase (Finking et al. 2002). Due to their broad activity range, the AcpS-type PPTase might have been lost over time. The high level of conservation within cyanobacteria suggests that this happened in an early, common cyanobacterial ancestor. Also other bacteria such as *Pseudomonas aeruginosa* PAO1 carry only an Sfp-type PPTase, which raises the question of whether the Sfp-type could have been the ancestral prototype (Copp and Neilan 2006).

No algal PPTases have been described previously. One PPTase identified in plants was isolated from spinach (Elhussein et al. 1988); however, at that time the different types of PPTases were unknown. The mtPPTase of *A. thaliana* has been classified as Sfp-type PPTase (Guan et al. 2015). We propose that the conservation of the Sfp-type PPTases in cyanobacteria was carried on through evolution to algae and plants in contrast to the other kingdoms of life that generally have at least one AcpS-type PPTase. Interestingly, the study of spinach found the major PPTase activity in the cytosol

and only minor activity in the chloroplast and mitochondrion. The authors proposed that the phosphopantetheinylation of the chloroplast ACP takes place in the cytosol and that this "pre"-*holo*-ACP is recognized by the plastid, taken up and further processed to *holo*-ACP. This could also be the case in *C. reinhartii*, as no specific targeting sequencing was identified with PredAlgo.

The PPTase PptC2 was able to recognize cognate as well as non-cognate CPs of the type II FAS and type II PKS systems, however, none of the other tested CPs from type I synthases or NRPS. *C. reinhardtii* uses a type II FAS system that is encoded in the nucleus, but located in the chloroplast, as in spinach (Beld et al. 2014c). Additionally, a type I PKS of 21,004 amino acids is encoded in the nuclear genome (Sasso et al. 2012). Neither the PKS, nor its product have been detected or described so far. Thus, it remains unknown if green microalgae are capable of producing complex polyketides, as found in dinoflagellates (Rein and Borrone 1999), or what kind of molecules result from this machinery. Microalgae in general possess a wide range of PKS and NRPS genes, which would require activation by a PPTase (Shelest et al. 2015; Kohli et al. 2016). Green microalgae are not known to produce toxins such as those characteristic of dinoflagellate algal blooms. Therefore, we hypothesize that PptC2 is dedicated to modifying the plastid-localized FAS CP, and that it also activates the type I PKS, if produced. Besides sharing less sequence similarity than PptC2 with the *Arabidopsis* mtPPT, PptC1 could be the PPTase devoted to the mitochondrial type II FAS. It appears to have a more narrow activity range.

Interestingly, genome sequenced green microalgae to date appear to carry two distinct PPTases. PPTases are often classified into the KEA and KES subfamilies

depending on the residue 152 (based upon Sfp sequence) (Copp and Neilan 2006; Beld et al. 2014c). No subclustering is identified due to this feature in the *C. reinhardtii* PPTases. They demonstrate the KEA pattern, as do most other identified algal PPTases. A few algal PPTases demonstrate a KES pattern, but these cluster independent of phylogeny (Fig. S5). Currently, only model species of the other algal phyla have been sequenced, and therefore we are lacking sufficient genomic data for further analysis. Nevertheless, the microalgal PPTases identified herein extend the proposed sub-classification of group II PPTases by Wang *et al*. (Wang et al. 2014). The conservation of the second-residue (A and E) within the three magnesium-binding residues indicates an early evolutionary separation of the two PPTase versions within the green microalgae. PPTases of the other algal phyla show different second-residue conservation than green microalgae, underlining the complex phylogeny and evolutionary history of photosynthetic eukaryotes.

PPTases are central activators in primary and secondary metabolism found in all kingdoms of life. Herein, we characterized two PPTases from green microalgae. We were able to demonstrate that PptC1 and PptC2 from the green alga *C. reinhardtii* have different substrate specificities and that these two types of PPTases appear conserved throughout the phylum of chlorophyta. We identified possible CP substrates of these PPTases from FAS, PKS, NRPS, and hybrid pathways. Microalgae represent an efficient, sustainable production host for future bio-based products, including biofuels and other high-value products. A more complete understanding of these organisms, including the metabolic processes of fatty acid and secondary metabolism, is vital to progress in this

area, informing future metabolic engineering efforts and improving to access algae-derived products.

Tables

Table 2. Substrate range of algal PPTases. + and - indicating the activity of each *C. reinhardtii* PPTase with ACP or synthase in one pot reaction as detected by fluorescent labelling.

Enzyme type	Enzyme	PptC1	PptC2
Type I FAS	human ACP	-	-
Type II FAS	*E. coli* ACP	(+)	+
Type II FAS	*C. reinhardtii* ACP	+	+
Type I PKS	mycocerosic acid synthase	-	-
Type II PKS	actinorhodin ACP	-	+
NRPS	BpsA	-	-

Figures

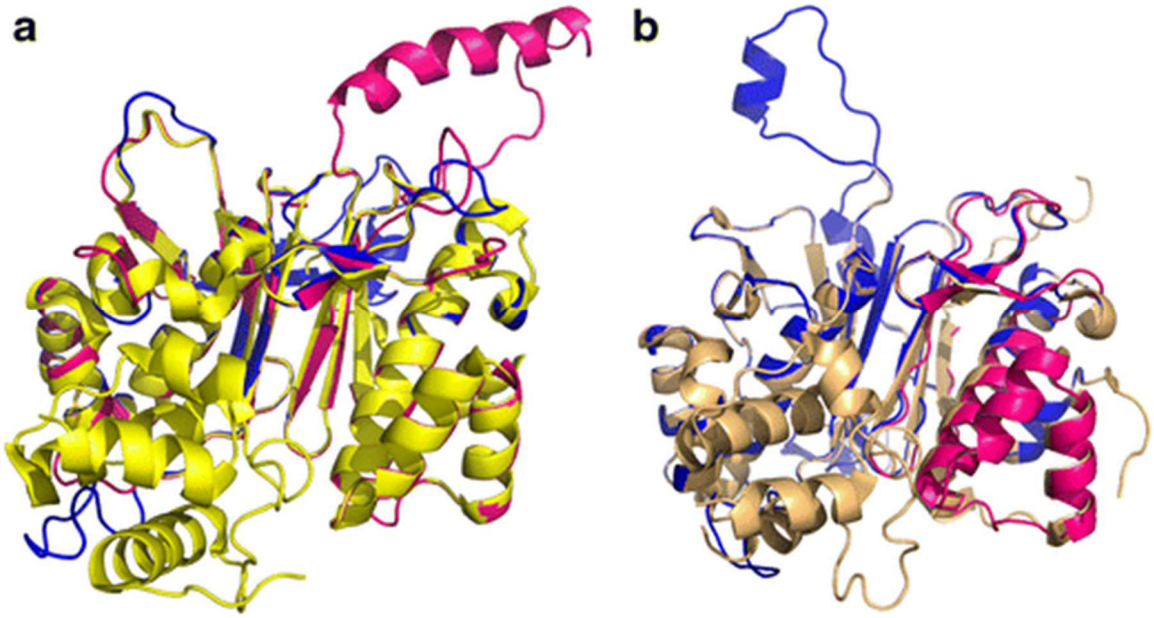


Figure 8: Structure prediction of PptC1 (blue) and PptC2 (pink) onto **a** Sfp (4MRT) (yellow) and **b** AASDHPPT (2CG5) (orange). GMQE: PptC1/Sfp = 0.41, PptC2/Sfp = 0.20, PptC1/AASDHPPT = 0.58, PptC2/AASDHPPT = 0.23.

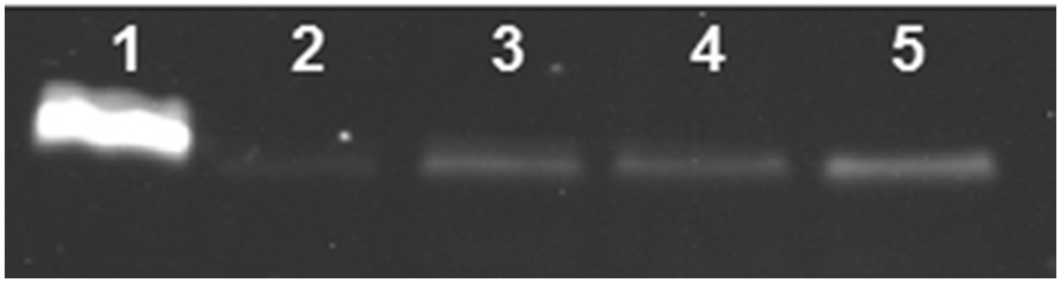


Figure 9: Fluorescent image of one pot reaction of 1: *E. coli* ACP with Sfp, 2: *C. reinhardtii* ACP without PPTase, 3: with Sfp; 4: with PptC1, 5: with PptC2.

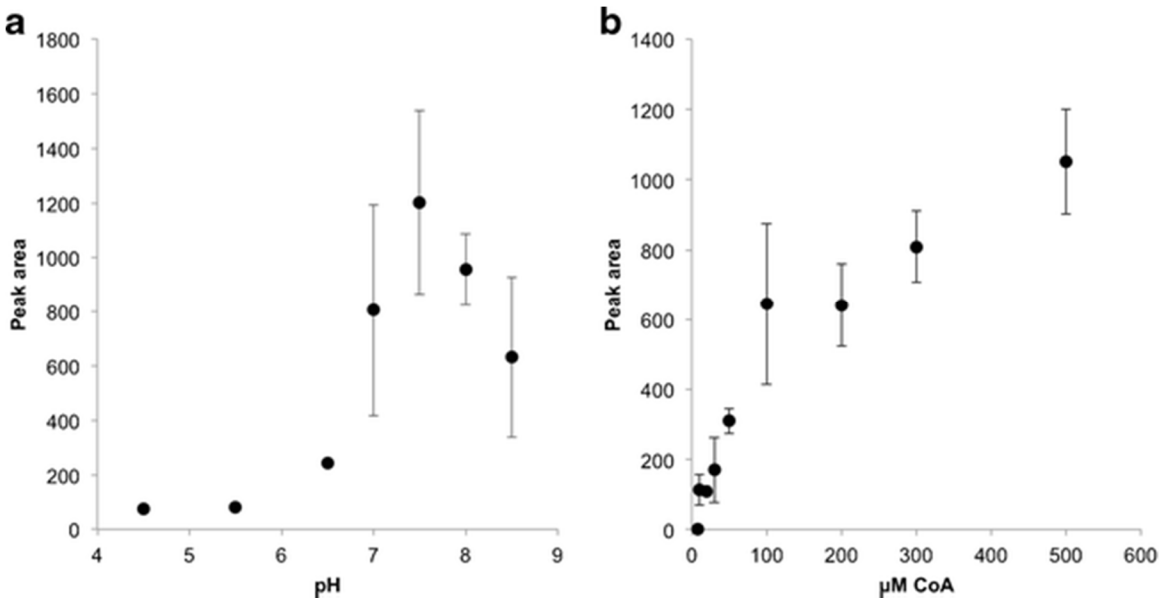


Figure 10: **a** pH profile of PptC2. PPTase activity was measured through formation of holoACP by HPLC at different pH values.(Finking et al. 2002) Reaction mixtures including 38 µM apoACP, 10 nM PptC2, 500 µM CoA, and 75 mM buffer were incubated at 37°C for 10 min and quenched with 50 mM EDTA. MES/NaAc buffers were used for pH 4.5–6.5 and Tris-HCl buffers for pH 7.0–8.5. **b** Reaction mixtures contained 38 µM apoACP, 1.2 nM PptC2, 75 mM Tris-HCl buffer (pH 7.5) and 10 mM $MgCl_2$ and the concentration of CoA was varied between 1 and 500 µM. Reactions were incubated at 37°C for 10 min and quenched with 50 mM EDTA. The peak area of holoACP was determined by HPLC.

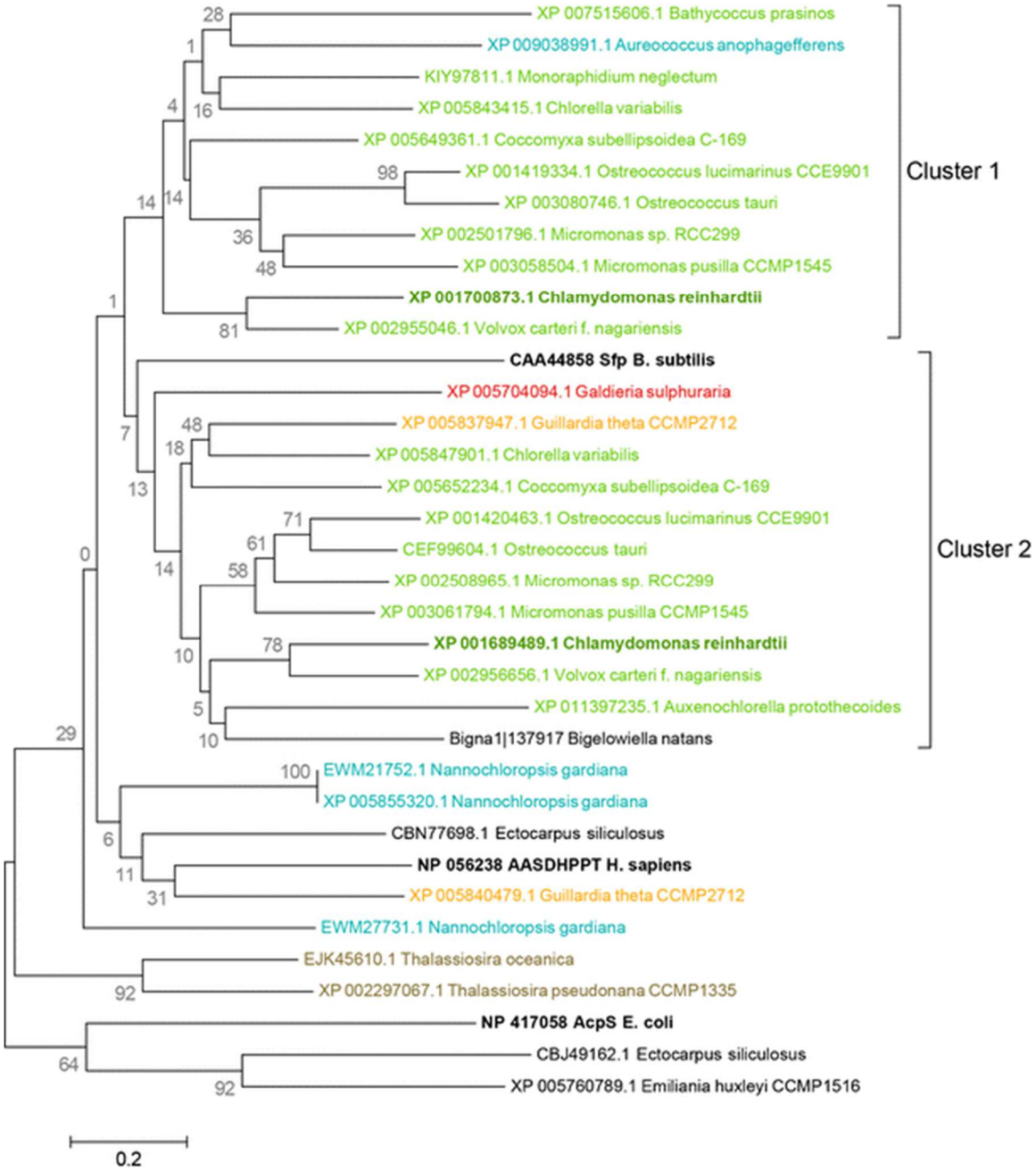


Figure 11: Phylogenetic analysis of algal PPTases by neighbour joining method (100 bootstrap replicates) (Tamura et al. 2013). Bacillariophyta were marked in brown, Cryptophyta in orange, Chlorophyta in green, Heterokontophyta in blue and Rhodophyta in red.

References

Ahmad I, Sharma AK, Daniell H, Kumar S (2015) Altered lipid composition and enhanced lipid production in green microalga by introduction of brassica diacylglycerol acyltransferase 2. Plant Biotechnol J 13:540–550. doi: 10.1111/pbi.12278

Alberts WA, Vagelos PR (1966) Acyl carrier protein. J Biol Chem 241:5201–5204.

Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol 215:403–410. doi: 10.1016/S0022-2836(05)80360-2

Amiri-Jami M, Griffiths MW (2010) Recombinant production of omega-3 fatty acids in Escherichia coli using a gene cluster isolated from *Shewanella baltica* MAC1. J Appl Microbiol 109:1897–1905. doi: 10.1111/j.1365-2672.2010.04817.x

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201. doi: 10.1093/bioinformatics/bti770

Beld J, Blatti JL, Behnke C, et al (2014a) Evolution of acyl-ACP thioesterases and β-ketoacyl-ACP synthases revealed by protein–protein interactions. J Appl Phycol 26:1619–1629. doi: 10.1007/s10811-013-0203-4

Beld J, Cang H, Burkart MD (2014b) Visualizing the Chain-Flipping Mechanism in Fatty-Acid Biosynthesis. Angew Chemie Int Ed 53:14456–14461. doi: 10.1002/anie.201408576

Beld J, Sonnenschein EC, Vickery CR, et al (2014c) The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. Nat Prod Rep 31:61–108. doi: 10.1039/c3np70054b

Blatti JL, Beld J, Behnke C a, et al (2012) Manipulating fatty acid biosynthesis in microalgae for biofuel through protein-protein interactions. PLoS One 7:e42949. doi: 10.1371/journal.pone.0042949

Cai X, Herschap D, Zhu G (2005) Functional Characterization of an Evolutionarily Distinct Phosphopantetheinyl Transferase in the Apicomplexan *Cryptosporidium parvum*. Eukaryot Cell 4:1211–1220. doi: 10.1128/EC.4.7.1211

Copp JN, Neilan BA (2006) The Phosphopantetheinyl Transferase Superfamily : Phylogenetic Analysis and Functional Implications in Cyanobacteria. Appl Environ Microbiol 72:2298–2305. doi: 10.1128/AEM.72.4.2298

Elhussein SA, Miernyk JA, Ohlrogge JB (1988) Plant holo-(acyl carrier protein)

synthase. Biochem J 252:39–45.

Fernandez MD, Lamppa GK (1990) Acyl Carrier Protein (ACP) lmport into Chloroplasts Does not Require the Phosphopantetheine : Evidence for a Chloroplast Holo-ACP Synthase. Plant Cell 2:195–206.

Fichtlscherer F, Wellein C, Mittag M, Schweizer E (2000) A novel function of yeast fatty acid synthase. Subunit alpha is capable of self-pantetheinylation. Eur J Biochem 267:2666–71.

Finking R, Solsbacher J, Konz D, et al (2002) Characterization of a new type of phosphopantetheinyl transferase for fatty acid and siderophore synthesis in *Pseudomonas aeruginosa*. J Biol Chem 277:50293–50302. doi: 10.1074/jbc.M205042200

Foley TL, Rai G, Yasgar A, et al (2014) 4-(3-Chloro-5-(trifluoromethyl)pyridin-2-yl)-N-(4-methoxypyridin-2-yl)piperazine-1-carbothioamide (ML267), a potent inhibitor of bacterial phosphopantetheinyl transferase that attenuates secondary metabolism and thwarts bacterial growth. J Med Chem 57:1063–1078. doi: 10.1021/jm401752p

Gerhardt P, Murray R, Wood W, Krieg N (1994) Methods for general and molecular bacteriology. ASM Press, Washington D.C.

Gorman DS, Levine RP (1965) Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of Chlamydomonas reinhardi. Proc Natl Acad Sci U S A 54:1665–9.

Griesbeck C, Kobl I, Heitzer M (2006) *Chlamydomonas reinhardtii*: a protein expression system for pharmaceutical and biotechnological proteins. Mol Biotechnol 34:213–223. doi: 10.1385/MB:34:2:213

Guan X, Chen H, Abramson A, et al (2015) A phosphopantetheinyl transferase that is essential for mitochondrial fatty acid biosynthesis. Plant J n/a–n/a. doi: 10.1111/tpj.13034

Guschina IA, Harwood JL (2006) Lipids and lipid metabolism in eukaryotic algae. Prog Lipid Res 45:160–186. doi: 10.1016/j.plipres.2006.01.001

Haushalter RW, Worthington AS, Hur GH, Burkart MD (2008) An orthogonal purification strategy for isolating crosslinked domains of modular synthases. Bioorg Med Chem Lett 18:3039–3042. doi: 10.1016/j.bmcl.2008.01.026

Jones CS, Luong T, Hannon M, et al (2013) Heterologous expression of the C-terminal antigenic domain of the malaria vaccine candidate Pfs48/45 in the green algae *Chlamydomonas reinhardtii*. Appl Microbiol Biotechnol 97:1987–1995. doi:

10.1007/s00253-012-4071-7

Joseph-McCarthy D, Parris K, Huang A, et al (2005) Use of structure-based drug design approaches to obtain novel anthranilic acid acyl carrier protein synthase inhibitors. J Med Chem 48:7960–7969. doi: 10.1021/jm050523n

Joshi AK, Zhang L, Rangan VS, Smith S (2003) Cloning, expression, and characterization of a human 4'-phosphopantetheinyl transferase with broad substrate specificity. J Biol Chem 278:33142–33149. doi: 10.1074/jbc.M305459200

Kealey JT, Liu L, Santi D V, et al (1998) Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. Proc Natl Acad Sci U S A 95:505–509.

Kiefer F, Arnold K, Künzli M, et al (2009) The SWISS-MODEL Repository and associated resources. Nucleic Acids Res 37:D387–392. doi: 10.1093/nar/gkn750

Kohli GS, John U, Van Dolah FM, Murray SA (2016) Evolutionary distinctiveness of fatty acid and polyketide synthesis in eukaryotes. ISME J 1–14. doi: 10.1038/ismej.2015.263

Kosa NM, Haushalter RW, Smith AR, Burkart MD (2012) Reversible labeling of native and fusion-protein motifs. Nat Methods 9:981–987. doi: 10.1038/NMETH.2175

Ku J, Mirmira RG, Liu L, Santi D V (1997) Expression of a functional non-ribosomal peptide synthetase module in *Escherichia coli* by coexpression with a phosphopantetheinyl transferase. Chem Biol 4:203–207.

Lambalot R, Gehring AM, Fluge RS, et al (1996) A new enzyme transferases superfamily - the phosphopantetheinyl transferases. Chem Biol 3:923–936.

Larkin MA, Blackshields G, Brown NP, et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. doi: 10.1093/bioinformatics/btm404

Li-Beisson Y, Beisson F, Riekhof W (2015) Metabolism of acyl-lipids in *Chlamydomonas reinhardtii*. Plant J 82:504–522. doi: 10.1111/tpj.12787

Merchant SS, Kropat J, Liu B, et al (2012) TAG, you're it! *Chlamydomonas* as a reference organism for understanding algal triacylglycerol accumulation. Curr Opin Biotechnol 23:352–363. doi: 10.1016/j.copbio.2011.12.001

Nair DR, Ghosh R, Manocha A, et al (2011) Two functionally distinctive phosphopantetheinyl transferases from amoeba *Dictyostelium discoideum*. PLoS One 6:e24262. doi: 10.1371/journal.pone.0024262

Nakano MM, Corbell N, Besson J, Zuber P (1992) Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. Mol Gen Genet 232:313–321.

Nakano MM, Zuber P (1990) Molecular biology of antibiotic production in Bacillus. Crit Rev Biotechnol 10:223–240. doi: 10.3109/07388559009038209

Rasala B, Lee P, Shen Z, et al (2012) Robust expression and secretion of Xylanase1 in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. PLoS One 7:e43349. doi: 10.1371/journal.pone.0043349

Rein KS, Borrone J (1999) Polyketides from dinoflagellates: origins, pharmacology and biosynthesis. Comp Biochem Physiol B Biochem Mol Biol 124:117–131.

Roberts A a, Copp JN, Marahiel M a, Neilan B a (2009) The *Synechocystis* sp. PCC6803 Sfp-type phosphopantetheinyl transferase does not possess characteristic broad-range activity. Chembiochem 10:1869–1877. doi: 10.1002/cbic.200900095

Rochaix J-D (1995) *Chlamydomonas reinhardtii* as the photosynthetic yeast. Annu Rev Genet 29:209–230.

Sasso S, Pohnert G, Lohr M, et al (2012) Microalgae in the postgenomic era: a blooming reservoir for new natural products. FEMS Microbiol Rev 36:761–785. doi: 10.1111/j.1574-6976.2011.00304.x

Savage LJ, Post-Beittenmiller D (1994) Phosphopantethenylated Precursor Acyl Carrier Protein Is Imported into Spinach (*Spinacia oleracea*) Chloroplasts. Plant Physiol 104:989–995.

Scranton MA, Ostrand JT, Fields FJ, Mayfield SP (2015) *Chlamydomonas* as a model for biofuels and bio-products production. Plant J 82:523–531. doi: 10.1111/tpj.12780

Shelest E, Heimerl N, Fichtner M, Sasso S (2015) Multimodular type I polyketide synthases in algae evolve by module duplications and displacement of AT domains in trans. BMC Genomics 16:1015. doi: 10.1186/s12864-015-2222-9

Shields JA, Rahman AS, Arthur CJ, et al (2010) Phosphopantetheinylation and specificity of acyl carrier proteins in the mupirocin biosynthetic cluster. ChemBioChem 11:248–255. doi: 10.1002/cbic.200900565

Silakowski B, Schairer HU, Ehret H, et al (1999) New lessons for combinatorial biosynthesis from myxobacteria. J Biol Chem 274:37391–37399. doi: 10.1074/jbc.274.52.37391

Sunbul M, Zhang K, Yin J (2009) Using phosphopantetheinyl transferases for enzyme

posttranslational activation, site specific protein labeling and identification of natural product biosynthetic gene clusters from bacterial genomes. In: Methods in enzymology, 1st edn. Elsevier Inc., pp 255–275

Tamura K, Stecher G, Peterson D, et al (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. doi: 10.1093/molbev/mst197

Tardif M, Atteia A, Specht M, et al (2012) PredAlgo: a new subcellular localization prediction tool dedicated to green algae. Mol Biol Evol 29:3625–3639. doi: 10.1093/molbev/mss178

Tran M, Zhou B, Pettersson PL, et al (2009) Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. Biotechnol Bioeng 104:663–673. doi: 10.1002/bit.22446

Wang YY, Li YD, Liu JB, et al (2014) Characterization and evolutionary implications of the triad Asp-Xxx-Glu in group II phosphopantetheinyl transferases. PLoS One 9:1–11. doi: 10.1371/journal.pone.0103031

Wijffels RH, Kruse O, Hellingwerf KJ (2013) Potential of industrial biotechnology with cyanobacteria and eukaryotic microalgae. Curr Opin Biotechnol 24:405–413. doi: 10.1016/j.copbio.2013.04.004

Worthington AS, Burkart MD (2006) One-pot chemo-enzymatic synthesis of reporter-modified proteins. Org Biomol Chem 4:44–46. doi: 10.1039/b512735a

Yalpani N, Altier DJ, Barbour E, et al (2001) Production of 6-methylsalicylic acid by expression of a fungal polyketide synthase activates disease resistance in tobacco. Plant Cell 13:1401–1409.

Yang LM, Fernandez MD, Lamppa GK (1994) Acyl carrier protein (ACP) import into chloroplasts. Covalent modification by a stromal holoACP synthase is stimulated by exogenously added CoA and inhibited by adenosine 3’,5'-bisphosphate. Eur J Biochem 224:743–750.

Yasgar A, Foley TL, Jadhav A, et al (2010) A strategy to discover inhibitors of *Bacillus subtilis* surfactin-type phosphopantetheinyl transferase. Mol Biosyst 6:365–375. doi: 10.1039/b913291k

Acknowledgement

Chapter 2, in full, is a reprint of the material as it appears in Journal of Applied Phycology, 2016. Sonnenschein, Eva C; Pu, Yuan; Beld, Joris; Burkart, Michael D 2016. The thesis author was the secondary author of this paper.

References

Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., ... & Stephanopoulos, G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science*, *330*(6000), 70-74.

Boynton, J. E., Gillham, N. W., Harris, E. H., Hosler, J. P., & Johnson, A. M. (1988). Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. *Science*, *240*(4858), 1534.

British Petroleum (2016) *BP Statistical Review of World Energy June 2016*. Available at www.bp.com/statisticalreview.

Dragone, G., Fernandes, B. D., Vicente, A. A., & Teixeira, J. A. (2010). Third generation biofuels from microalgae.

Engels, B., Dahm, P., & Jennewein, S. (2008). Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic engineering*, *10*(3), 201-206.

Franklin, S. E., & Mayfield, S. P. (2004). Prospects for molecular farming in the green alga Chlamydomonas reinhardtii. *Current opinion in plant biology*,*7*(2), 159-165.

Gershenzon, J., & Croteau, R. B. (1993). Terpenoid biosynthesis: The basic pathway and formation of monoterpenes, sesqiuterpenes, an diterpenes.*Terpenoid Biosynthesis*.

Holton, R. A., Kim, H. B., Somoza, C., Liang, F., Biediger, R. J., Boatman, P. D., ... & Kim, S. (1994). First total synthesis of taxol. 2. Completion of the C and D rings. *Journal of the American Chemical society*, *116*(4), 1599-1600.

Huang, Q., Roessner, C. A., Croteau, R., & Scott, A. I. (2001). Engineering Escherichia coli for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorganic & medicinal chemistry*, *9*(9), 2237-2242

Jones, C. S., Luong, T., Hannon, M., Tran, M., Gregory, J. A., Shen, Z., ... & Mayfield, S. P. (2013). Heterologous expression of the C-terminal antigenic domain of the malaria vaccine candidate Pfs48/45 in the green algae Chlamydomonas reinhardtii. *Applied microbiology and biotechnology*, *97*(5), 1987-1995.

Koepp, A. E., Hezari, M., Zajicek, J., Vogel, B. S., LaFever, R. E., Lewis, N. G., & Croteau, R. (1995). Cyclization of Geranylgeranyl Diphosphate to Taxa-4 (5), 11 (12)-diene Is the Committed Step of Taxol Biosynthesis in Pacific Yew.*Journal of Biological Chemistry*, *270*(15), 8686-8690.

Köksal, M., Jin, Y., Coates, R. M., Croteau, R., & Christianson, D. W. (2011). Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature*, *469*(7328), 116-120.

Nicolaou, K. C., Yang, Z., Liu, J. J., Ueno, H., Nantermet, P. G., Guy, R. K., ... & Sorensen, E. J. (1994). Total synthesis of taxol. *Nature*, *367*(6464), 630-634.

Rasala, B. A., & Mayfield, S. P. (2014). Photosynthetic biomanufacturing in green algae; production of recombinant proteins for industrial, nutritional, and medical uses. *Photosynthesis research*, *123*(3), 227-239.

Specht, E. A. (2014). *Improving the genetic tractability of the green alga Chlamydomonas reinhardtii* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Database.

Tran, M., Zhou, B., Pettersson, P. L., Gonzalez, M. J., & Mayfield, S. P. (2009). Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnology and bioengineering*, *104*(4), 663-673.

Walji, A. M., & MacMillan, D. W. (2007). Strategies to bypass the taxol problem. Enantioselective cascade catalysis, a new approach for the efficient construction of molecular complexity. *Synlett*, *2007*(10), 1477-1489.

Wang, X., Ort, D. R., & Yuan, J. S. (2015). Photosynthetic terpene hydrocarbon production for fuels and chemicals. *Plant biotechnology journal*, *13*(2), 137-146